© IWA Publishing 2008 Journal of Hydroinformatics | 10.2 | 2008

Single-reservoir operating rules for a year using multiobjective genetic algorithm

Taesoon Kim, Jun-Haeng Heo, Deg-Hyo Bae and Jin-Hoon Kim

ABSTRACT

A monthly operating rule for single-reservoir operation is developed in this study. Synthetic inflow data over 100 years are generated by using a time series model, AR(1), and piecewise-linear operating rules consisting of 4 and 5 linear lines are found using the implicit stochastic optimization method. In order to consider multiobjective functions in reservoir system operation, a multiobjective genetic algorithm (NSGA-II) is adopted to obtain the optimization results. The search space of NSGA-II is carefully refined using frequency analysis of historical data, and the relationship between inflow and constraints is also investigated. It is determined that 4 and 5 segments are the optimal number of segments for the piecewise-linear operating rule, and the effect of random number seeding on NSGA-II is evaluated. Six years of historical inflow data are used for the simulation model and the results show that the developed operating rule would handle various inflow series. As a result, probabilistic reservoir storage forecasts can be provided to a system operator so as to enable the operator to evaluate the current status of a reservoir quantitatively.

Key words | frequency analysis, multiobjective genetic algorithm, operating rules, probabilistic reservoir storage forecast

Taesoon Kim
Jun-Haeng Heo (corresponding author)
School of Civil and Environmental Engineering,
Yonsei University,
Seoul 120-749
South Korea
E-mail: *jhheo@yonsei.ac.kr*

Deg-Hyo Bae
Jin-Hoon Kim*
Department of Civil and Environmental Engineering,
Sejong University,
Seoul 143-747,
South Korea
*Current address: Youngsan River Control Center,
78-7 Daesonggil,
Naju-City,
Jeollanam-Do 520-200,
South Korea

INTRODUCTION

The reservoir operating rule has been extensively studied over the last four decades after Young (1967) first introduced a method to obtain general operating rules from deterministic optimization results. In a single-reservoir operation, the operating rule is found to provide reservoir system operators with long-term reservoir operating policies that might satisfy multiple objectives and constraints of reservoir operation. Since the operating rule is used as a sort of reference, system operators often deviate from these policies for satisfying specific constraints or objectives such as for navigation, recreation, etc.

On the other hand, an optimization model for reservoir operation would be able to find an optimal solution for the particular objectives without any constraint violations. Even though this solution is obviously optimal for the optimization model, it is difficult for a system operator to accept the optimal solution without any changes since reservoir operation involves very complex decision-making processes; thus, no optimization model could consider all aspects of reservoir operation. Moreover, the optimization results cannot be published for everyone to examine and debate prior to implementation (Oliveira & Loucks 1997). Therefore, a predefined operating rule that is computed from the optimization model is essential for effective reservoir operation.

In general, the reservoir operating rule has been developed by two approaches. The explicit stochastic optimization method (Loucks *et al.* 1981; Tejada-Guibert *et al.* 1993) is used to develop the operating rule based on the transition probabilities of decision variables between two serial periods like June and July; as a result, the release during July can be computed using that during June by using the transition probability. Although the explicit stochastic optimization method is intuitively appealing and appears to be more

doi: 10.2166/hydro.2008.019

mathematically based, it suffers from the lack of memory since the decision variables are required to be divided into many discretized values with small intervals, and this could cause the memory requirement to increase exponentially.

Another approach is the implicit stochastic optimization method (Karamouz *et al.* 1992). Here, time series modeling is usually applied to generate synthetic inflow and optimization results of reservoir operation are obtained. Once the optimization results are obtained, the implicit stochastic method can easily find the operating rule using regression analysis. However, the regression relationship between the independent and dependent variables could weaken because, even though the independent variables are almost the same, the corresponding dependent variables might be completely different and this would result in the poor performance of the regression equation.

Of these two methods, the implicit stochastic optimization method is applied in this study. This is because the case study reservoir in this study has the largest storage capacity in South Korea. Thus the explicit stochastic approach is not the proper one for this large storage reservoir. Karamouz & Houck (1987) compared deterministic dynamic programming with regression (DPR) with stochastic dynamic programming (SDP) and concluded that, for small reservoirs, the performance of SDP (capacity of 20% of the mean annual flow) is better than that of DPR; in contrast, DPR performs better for large reservoirs (capacity exceeding 50% of the mean annual flow).

A piecewise-linear operating rule whose end points are computed directly by the optimization model is employed because of the above-mentioned shortcomings of regression analysis. The piecewise-linear operating rule consists of several linear segments connected with each other by two end points. The nondominated sorting genetic algorithm-II (NSGA-II) developed by Deb *et al.* (2002) is adopted in this study. NSGA-II is one of the most imitating multiobjective genetic algorithms (MOGA) (Van Veldhuizen & Lamont 2000) and it has been applied to a few reservoir system optimization problems (Liong *et al.* 2004; Kim *et al.* 2006; Reddy & Kumar 2006).

Kim *et al.* (2006) showed the optimization results by NSGA-II when it is applied to a multireservoir system and presented the method how decision-makers analyze a Pareto-front in order to discriminate true decision variables. In contrast, this study suggests a predefined operating rule and provides probabilistic long-term storage forecasting to a decision-maker of reservoir operation. Unfortunately, there is no predefined operating rule for reservoirs in South Korea: therefore, the results of the current work would be more effective tools than the previous work.

Several past research works have used the piecewise-linear operating rule and GAs. Oliveira & Loucks (1997) suggested a piecewise-linear operating rule for a multi-reservoir system comprising two hypothetical systems. They computed a target system release using the available water and inflow and then calculated the releases from each single reservoir based on the target system release. Ahmed & Sarma (2005) applied a piecewise-linear operating rule approach and a GA to single-reservoir optimization in India. They used 4, 6, 7 and 11 segmented piecewise-linear operating rules and generated synthetic monthly inflow data over 100 years using an artificial neural network (ANN). The optimization results from GA are compared with those from SDP, and because supplying water resources without shortage is usually a more important objective than hydropower production, they suggested that GA-derived policies are more efficient than SDP-derived ones.

Recently, some research reported that the operating rule curves derived from a GA are more adequate, effective and robust than those derived from the current M5 rule of the Shihmen reservoir in Taiwan (Chang & Chang 2001; Chang *et al.* 2005*a*,*b*). In addition, a new multiobjective GA implementation called the macro-evolutionary multiobjective GA (MMGA) was applied to develop the rule curves for the Fei-Tsui reservoir system in Taiwan (Chen *et al.* 2007). However, all these studies use only the water level to obtain the release plan for 10 days of operation, and the obtained release plan cannot provide the specific release amount to a reservoir system operator.

The objective of this study is to suggest an effective formulation method for the piecewise-linear operating rule when the multiobjective GA, NSGA-II, is applied to develop the reservoir operating rule. The characteristics of multiobjective optimization are inherent in reservoir operation; these include minimizing water shortage, maximizing hydropower production, preventing flooding of the downstream area, etc. Thus, the multiobjective optimization method would be essential for developing the reservoir operating rule.

In addition, detailed approaches regarding the formulation of NSGA-II for developing the reservoir operating rule are suggested. The upper and lower limits of the first and last end points of the piecewise-linear operating rule are investigated since these two limits determine the search space of NSGA-II. The number of segments to be used is also examined to select the appropriate number. In order to provide a reservoir system operator with useful information that is necessary for evaluating the current status of reservoir storage, probabilistic long-term reservoir storage forecasts are computed using the developed piecewise-linear operating rule and 100 years of synthetic generated inflow data.

GENETIC ALGORITHMS AND NONDOMINATED SORTING GENETIC ALGORITHM-II

GAs are based on the mechanics of natural selection and natural genetics (Goldberg 1989). They were developed in the 1960s and refined throughout the 1970s by John Holland and his coworkers at the University of Michigan to explain and model the adaptability of natural systems. Holland's monograph (Holland 1975), *Adaptation in Natural and Artificial Systems*, is the seminal book in this field, and Goldberg's book, *Genetic Algorithms in Search, Optimization and Machine Learning* (Goldberg 1989), deals with the most practical optimization problems in evolutionary computation.

GAs are a very effective optimization method for solving multiobjective problems (MOPs) because they exploit not only a single solution but also a set of many possible solutions, namely a population, for searching the global optimum. Hence, they can easily determine which solution is superior to the others, even though only one run of the optimization model is performed; this is a major advantage of GAs when "Pareto-optimality", which was originally introduced by Edgeworth (1881) and later generalized by Pareto (1896), is applied to achieve global or near-global optimum solutions for MOPs. Readers who are interested in GAs and Pareto-optimality can refer to the textbooks written by Mitchell (1996) and Deb (2001), and a number of technical papers, e.g. Van Veldhuizen & Lamont (2000), are good references.

NSGA-II (Deb *et al.* 2002) is widely used to solve MOPs in the field of engineering (Deb & Jain 2003; Deb & Raji 2003; ISI 2004), and recently a few researchers have examined water resources engineering including reservoir system optimization (Kuo *et al.* 2003; Prasad & Park 2004; Prasad *et al.* 2004; Reed & Minsker 2004; Kim *et al.* 2006). The key features of NSGA-II that are different from those of the former NSGA (Srinivas & Deb 1994) are a reduction in time complexity, a parameterless sharing procedure that uses crowding distance for ensuring diversity in a population, and elitism that can speed up the performance of GAs and can also help in preventing the loss of good solutions once they are found (Deb *et al.* 2002). In this study, since our attention would be confined to the application method of NSGA-II to develop the reservoir operating rule, interested readers can find a detailed explanation regarding NSGA-II in Deb *et al.* (2002) and Deb & Jain (2003).

CASE STUDY AREA AND TIME SERIES MODELING

The case study area is the Soyanggang dam basin. The Soyanggang dam is located upstream of the North Han River in the Korea Peninsula (Figure 1). The Han River consists of two tributaries, namely the North and South Han Rivers, and three large storage reservoirs: the Soyanggang and Hwacheon reservoirs on the North Han River and the Choongju reservoir on the South Han River. These reservoirs supply most of the domestic, industrial and irrigation water resources to the Seoul Metropolitan Area, which has a population of 23.7 million.

Among these three storage reservoirs, the Soyanggang and Choongju reservoirs are multipurpose reservoirs. They are required to release the designated water supply downstream through turbines and lower their water level to the regulated water level in order to secure sufficient flood control capacity during the flood season between 21 June and 20 September. In this study, the Soyanggang reservoir basin (Table 1) is selected as the case study area because it has a longer historical record than the Choongju reservoir.

Figure 2 illustrates the monthly historical inflows of the Soyanggang reservoir between 1974 and 2005. A significantly large inflow usually occurs during summer because of

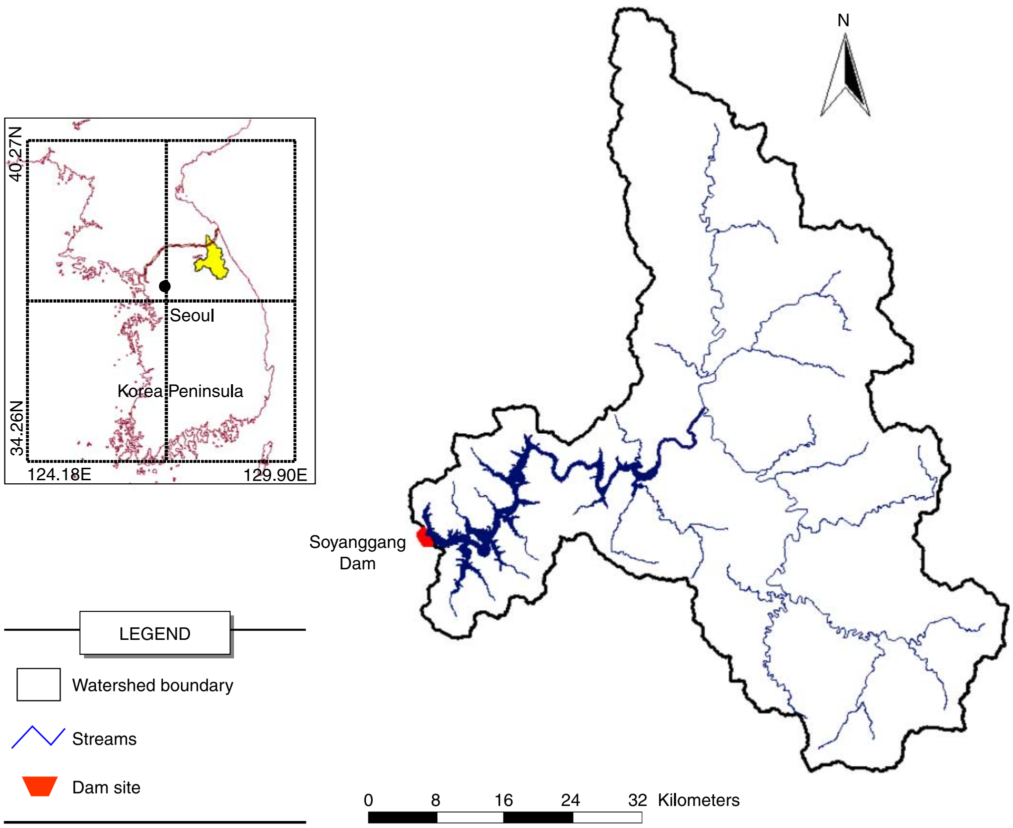

Figure 1 | The Soyanggang reservoir basin in the Korea Peninsula.

the intensive heavy rainfall from July to September that is more than approximately 65% of the annual inflow. A time-series model with periodic coefficients can be used because of this within-year periodicity. However, the number of periodic parameters, for example, 12 different parameters if the order 1 autoregressive time series model, AR(1), is used, can be very large and the periodic autocorrelation function (ACF) of the standardized inflow at the Soyanggang reservoir shows no periodic characteristics; therefore, a

Table 1 | Description of the Soyanggang reservoir

Basin area	2703 km^2
Mean annual inflow (1974–2005)	2192.85 MCM
Annual designated water supply demand to downstream	1456.96 MCM
Design flood level (maximum Storage capacity)	EL.198.00 m (2900 MCM)
Normal pool level	EL.193.50 m (2543.75 MCM)
Regulated water level for flood season	EL.190.30 m (2345.52 MCM)
Low water level	EL.150.00 m (686.82 MCM)
Total storage capacity	2900 MCM
Flood control capacity	500 MCM
Design discharge capacity	5625 m^3 s^{-1}

MCM: million cubic meters

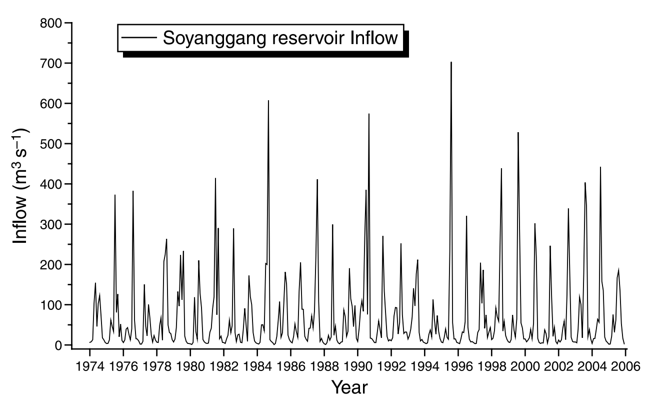

Figure 2 | Monthly inflow at the Soyanggang reservoir from January 1974 to December 2005.

periodic time-series model with constant coefficients (Salas *et al.* 1980) is applied.

In order to identify the time-series model, the serial ACF and partial ACF (PACF) are computed. Lag 1, 2, 3 and 4 serial ACF show larger values than 95% confidence limits, and PACF, which is one of the favorable measures for identifying the time-series model, shows only lag 1 PACF outside the 95% confidence limits. Hence, in this study, AR (1) is applied in order to generate the synthetic inflow of the Soyanggang reservoir.

MODEL FORMULATION

Selection of probable limits for operating rule using frequency analysis

The piecewise-linear operating rule consists of several lines connecting each end point. The dependent variables (release from the reservoir during a month) are computed by using the corresponding independent variables such as the sum of initial storage and monthly inflow (Oliveira & Loucks 1997; Ahmed & Sarma 2005; Chang *et al.* 2005*a*).

Figure 3 is an example of the piecewise-linear operating rule applied in this study for a month when 7 end points are used. The X axis is the sum of the initial storage and the monthly synthetic generated inflow and the Y axis is the monthly release corresponding to the independent variable on the X axis. The chromosome is coded by using

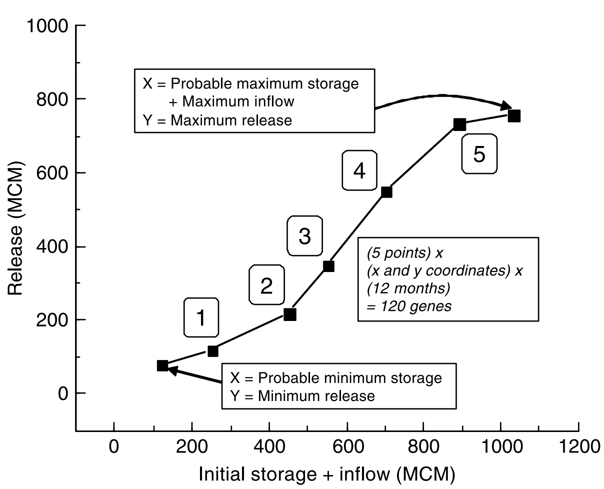

Figure 3 | Chromosome formulation when the 6-segmented piecewise-linear operating rule is used.

real numbers since the use of binary coding causes the length of chromosome to become very long. A chromosome is composed of the x and y coordinates of end points corresponding to 12 months, as shown in Equation (1):

$$[x_{1,1}y_{1,1}x_{1,2}y_{1,2}x_{1,3}y_{1,3}x_{1,4}y_{1,4}x_{1,5}y_{1,5}\vdots\cdots\vdots x_{12,1}y_{12,1}x_{12,2} y_{12,2}x_{12,3}y_{12,3}x_{12,4}y_{12,4}x_{12,5}y_{12,5}] \quad (1)$$

where $x_{m,n}$ is the x coordinate of the nth end point of month m, and $y_{m,n}$ is the y coordinate.

In Figure 3, the operating rule has 6 piecewise lines and 7 end points. The first and last end points can be calculated by the storage or release limits, and they are treated as known values before optimization is performed; thus, only 5 inner end points are used for formulating the chromosome of NSGA-II. Because each end point has x and y coordinates and the time horizon is a year, 120 genes are necessary for a chromosome in this 6-segmented piecewise-linear operating rule.

The known x and y coordinates of the first and last end points should be determined carefully since they are the limits of the possible search space of NSGA-II. If this search space is very large, the operation rule may be improper for application to real-time reservoir operation. Past researches regarding the piecewise-linear operating rule developed by GAs have used dead storage for computing the x coordinate of the first point and an arbitrary point far from the origin for computing that of the last point (Oliveira & Loucks 1997; Ahmed & Sarma 2005). And, for the y coordinate (monthly release), these researchers have used the zero value for the minimum release and a specific release for the maximum that makes full reservoir storage at the start of the next month in order to prevent the violation of storage limits.

However, if the above approach is applied to the case study of the Soyanggang reservoir, appropriate reservoir operating rules cannot be developed since the difference between the maximum and minimum storages is much larger, (at most 1856.93 MCM); thus, the search space becomes very large. Moreover, because the historical maximum and minimum storage records of the Soyang-gang reservoir show strong within-year periodicity, if the same maximum and minimum storage limits are used over 12 months, it is difficult to develop an operating rule effectively.

Therefore, the frequency analysis approach is used to determine the maximum and minimum storage limits. Historical maximum and minimum storage records at the start of each month from 1974 to 2005 are used for computing the maximum and minimum storages corresponding to 100-year quantiles. For the frequency analysis, FARD2002 developed by NIDP of South Korea (NIDP 2002) is applied. It is a powerful and advanced frequency analysis tool. Fourteen of the most frequently used distributions are applied along with four effective goodness-of-fit tests including the probability plot correlation coefficient test (PPCC test).

According to the results obtained by applying FARD2002 to the historical maximum and minimum storage records of the Soyanggang, the generalized extreme value (GEV) distribution is the recommended probability distribution function (PDF) for the maximum and minimum storage at the start of each month. The 100-year quantiles can be calculated by using the cumulative distribution function (CDF) of GEV. The maximum and minimum storages for each month correspond to quantiles that have a 1% exceedance probability of the maximum historical storage data and 1% nonexceedance probability of the minimum historical data, respectively.

Figure 4 shows the probable maximum and minimum storages and historical storages at the start of each month.

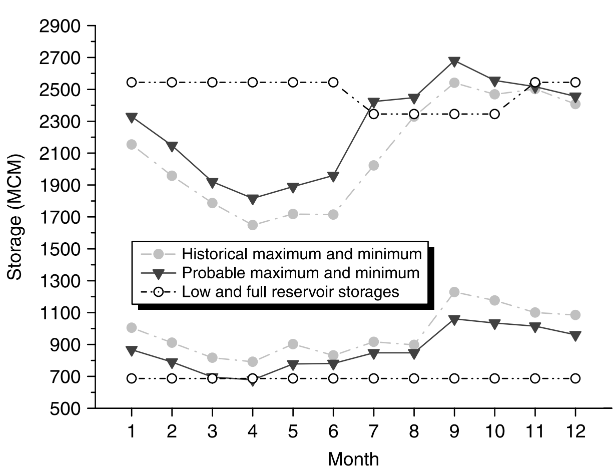

Figure 4 | Historical and probable maximum and minimum storages. The probable minimum and maximum storages corresponding to 100-year quantiles of CDF of GEV distribution.

From July to October, the probable maximum storages are greater than the regulated reservoir storage for the flood season because these probable maximum storages are not averaged values over a month but a kind of instantaneous value at the start of the month. In particular, the probable maxima for September and October are larger than the full reservoir storage corresponding to the normal pool level in Table 1, but they are smaller than the design flood level storage that is the practical maximum storage capacity based on safety considerations. Therefore, it can be said that all the computed maximum and minimum storages are reasonably estimated, and they are slightly larger than the historical values.

In addition to determining the x coordinates of the first and last end points, the y coordinate corresponding to the monthly release is also calculated. The minimum release of the dam, zero, cannot be used because the Soyanggang reservoir releases a designated amount of water downstream; thus, 50% of the designated release downstream is used to secure the minimum water supply. The maximum release in the non-flood season is set to the maximum monthly turbine discharge; this is because, in this season, water is usually released through turbines for hydropower production. Hence, it is impossible in practice to release more water in a month than the maximum turbine discharge.

In the flood season, release from the spillway occurs and the maximum release is calculated by Equation (2). Here, the maximum number of days when spillway release occurs is at most 10, according to the historical records; thus, 1/3 of the A value is selected for a month. Therefore, a sum of two thirds of the monthly maximum turbine discharge and a third of monthly maximum water release through spillway is the maximum total release from reservoir in the flood season:

$$[(1 - A) \times \text{maximum turbine discharge} + A \times \text{spillway}] \quad (2)$$

Objective function, state equation, constraints and genetic operators

Two multiobjective functions are used. The first objective function is to minimize the shortage index that is equal to the sum of the squares of the annual shortages over a 100-

year period when each annual shortage is expressed as a ratio of the annual requirements (USACE 1997). This function is defined as

$$SI = \frac{100\sum_{i=1}^{N}\left[\frac{S_A}{D_A}\right]}{N} \quad (3)$$

where SI is shortage index; N, the number of years; S_A, the annual shortage (annual demand volume minus annual volume supplied) and D_A, the annual demand volume.

The second objective function is to maximize the sum of hydroelectric power production, and hydroelectric power is defined as

$$\sum_{t=1}^{N\times 12} 9.81 Q_t^T H_t e/3600\,(\mathrm{GWh}) \quad (4)$$

where Q_t^T is the total flow through the turbines during period t (million cubic meters per month); H_t, the falling height (m); and e, the plant efficiency (Loucks *et al.* 1981).

When NSGA-II is applied, it is required that the multiobjective functions are in conflict with each other. For example, minimizing treatment cost and maximizing the removal efficiency of wastewater are evidently conflicting multiobjectives functions. However, in reservoir operation problems, because of nonlinearity among the decision and state variables, it is difficult to formulate well-defined conflicting multiobjective functions (Kim & Heo 2004). Moreover, if an objective function is computed by the decision or state variables for a specific time and not for the entire time horizon, a well-distributed Pareto-optimal front is usually difficult to obtain, and only a few concentrated Pareto-optimal points that are not useful in terms of the trade-off relationship among the multiobjective functions can be obtained (Kim 2005). The best way to avoid this useless Pareto-optimal front is to find well-defined conflicting multiobjective functions. However, if it is difficult to find such functions, the second best way is to formulate multiobjective functions with regard to the entire time horizon of the optimization model, such as Equations (3) and (4).

The state equation is the water balance equation between the start and end of a month; the storage and release limits of the Soyanggang dam are used as the constraints of NSGA-II. Equation (5) represents the state equation and Equations (6) and (7) represent the constraints with respect to storage and release, respectively:

$$S_{t+1} = S_t + I_t - R_t - Eva_t \quad (5)$$

$$S_{\min} \le S_t \le S_{\max} \quad (6)$$

$$R_{\min} \le R_t \le R_{\max} \quad (7)$$

where S_t is the storage at the start of month t; I_t the monthly inflow during month t; R_t the release; Eva_t, the evaporation from the water surface of the reservoir. Furthermore, $S_{\min}$ and $S_{\max}$ are the lower and upper limits of storages, respectively; $R_{\min}$ and $R_{\max}$ the lower and upper limits of release, respectively.

In addition, another constraint for the end points of the piecewise-linear relationship is essential to formulate a useful operating rule. If the x coordinate of any $(i+1)$th end point is smaller than that of the i th end point, the relationship between the independent and dependent variables has no meaning. Moreover, if the y coordinate of any $(i+1)$th end point is smaller than that of the (i)th end point, the release decreases although the sum of storage and inflow increases. Therefore, the x and y coordinates of the $(i+1)$th end point should be larger than those of the (i)th end point; these constraints are shown as follows:

$$\begin{aligned} x_i &< x_{i+1} \\ y_i &< y_{i+1} \end{aligned} \quad (8)$$

where x_i is the x coordinate of the (i)th end point and y_i is the y coordinate.

Equations (9) and (10) represent the constraints for the terminal storage at the end of the last month (S_{13}) and the designated release from the Soyanggang reservoir for a month. The maximum terminal storage is set to 110% of the initial storage and the minimum is $\alpha \times S_1$, and each month's release R_t is required to be greater than $\beta \times R_{\mathrm{duty}}$. R_{duty} is the monthly designated water supply demand to downstream (Table 1):

$$\alpha \times S_1 \le S_{13} \le 1.1 \times S_1 \quad (9)$$

$$\beta \times R_{\mathrm{duty}} \le R_t. \quad (10)$$

The population size is 1,000 and the number of generations is 500 in NSGA-II. The crossover probability is set to 0.7 and the mutation probability is the reciprocal of the number of chromosomes. The distribution index of simulated binary crossover (SBX) in NSGA-II determines the variance of the offspring from the parent. The smaller the distribution index, the larger would be the investigated search space because the offspring might have a very different value from the parent. In contrast, if a large distribution index is used, NSGA-II converges very quickly, but the storages and releases of the operational results calculated by the optimal solutions are too irregular and unsuitable to be applied to real-time reservoir operation (Kim & Heo 2004). Therefore, even though the convergence rate would be slow, a small distribution index of 5.0 is applied to the SBX of NSGA-II.

Effects of inflow, water supply, and terminal storage to piecewise-linear operating rule

In the reservoir operations conducted in South Korea, one of the most important objectives is to supply water resources without the occurrence of shortages. For this reason, stable water supply up to a year is essential, and the terminal storage that is the reservoir storage value at the end of a year should be sufficient to supply water resources for the next several months prior to the flood season that is the water recharging period in South Korea.

In the optimization model, the constraint for terminal storage, Equation (9), is used and this constraint is very closely related to the constraint for water supply, Equation (10). However, because the available water resources in a year are limited, if α is very large, then β cannot be larger than a specific value, and vice versa. In addition, if the annual inflow is smaller than the annual water supply demand of 1456.96 MCM, as shown in Table 1, β cannot be 100% although α is much smaller than 100%. Therefore, in this section, the relationship between the water supply demand, terminal storage and annual inflow is examined and the best way to formulate the piecewise-linear operating rule properly is found.

Two simulations are performed. The first one uses 24 pairs of α and β. All the state equations, objective functions, constraints from Equations (3)–(10) and 100-year synthetic inflow data are used, and the number of segments of the operating rule is 6.

Since α is the minimum terminal storage and β is the minimum water supply limit downstream, larger values of these two parameters result in better performance of the reservoir operation. However, if β is 95%, no Pareto-optimal solution can be found. If β is 90%, 33 Pareto-optimal solutions can be found when α is 60%. If β is 85%, 30 Pareto-optimal solutions can be found when α is 60%, and if β is 80%, 198 Pareto-optimal solutions can be found when α is 60%, 116 Pareto-optimal solutions when α is 65% and 22 Pareto-optimal solutions when α is 70%. As a result, although the minimum water supply demand is lowered to 80%, at most 70% of the initial storage, which is only 1124.6 MCM or 44.2% of the normal pool reservoir storage shown in Table 1, is guaranteed for the next several months. Note that only 44.2% water storage is insufficient to supply water resources in winter and next spring.

The reason for obtaining this result could be that only one set of the 12 piecewise-linear operating rules is used, whereas 100 synthetic generated inflows are used as input data for NSGA-II. If the regression technique is applied, 100 different optimization results are computed using 100 different synthetic generated inflows. Furthermore, an operating rule having the minimum least squared errors for these 100 different optimization results is developed; thus, it becomes easier to satisfy the constraints of the optimization model. However, the piecewise-linear operating rule for one month has only one fixed form for all the synthetic generated inflows, and all the constraints should be satisfied using this one operating rule. Hence, the most conservative operating rule that satisfies all the constraints would be formulated, thereby resulting in relatively small values of the terminal storage constraint (α) and water supply constraint (β) such as 70% and 80%, respectively, as shown in Table 2.

Table 3 shows the results obtained without using the constraint of terminal storage. Since the terminal storage constraint is not used, it would be expected that Pareto-

Table 2 | Pareto-optimal solutions number using 24 differenent α and β

α/β	60%	65%	70%	75%	80%	85%
80%	198	116	22	×	×	×
85%	30	×	×	×	×	×
90%	33	×	×	×	×	×
95%	×	×	×	×	×	×

Table 3 | Number of Pareto-optimal solutions obtained considering only the constraints of water supply (β)

β (%)	105	100	95	90	85	80
No. of Pareto-optimal solutions	×	×	9	20	16	20

optimal solutions can be found, although β is 100% or more. However, no Pareto-optimal solution can be found when β is 105% or even 100%. This result shows that, even if there is no constraint associated with terminal storage, 100% of the water resources cannot be supplied downstream when the annual inflow is smaller than the annual water supply demand. As a result, it can be said that the piecewise-linear operating rule cannot satisfy the constraint of water supply demand if an insufficient annual inflow that is smaller than the annual water supply demand is used in the optimization model. Therefore, the operating rule is developed using only sufficient inflow and the performance of the developed piecewise-linear operating rule is examined.

If the reservoir operating rule is developed using only sufficient annual inflow, the simulation result obtained using the operating rule should be checked to determine whether or not the operating rule is appropriate for real-time reservoir operation. Figure 5 indicates the storage and turbine release resulting from the operating rule that is developed using only sufficient synthetic inflow. The situation wherein the historical annual inflow is larger than the water supply demand can be called as "sufficient historical inflow." In this situation, typical storage and turbine release patterns are observed; the storages and turbine releases are high during summer and relatively low during other seasons.

In contrast, if insufficient inflows are adopted, the storage values achieved are smaller than those computed using sufficient historical inflows. Moreover, the pattern of turbine release is very different from that obtained using sufficient historical inflow and the storage values are at most slightly larger than the designated water supply demand. It is noted that all these results are achieved by using the same piecewise-linear operating rule developed for only sufficient synthetic inflow data. Therefore, it can be said that the piecewise-linear operating rule developed from only sufficient synthetic inflow data can be applied to the reservoir operation, even though the inflow is insufficient.

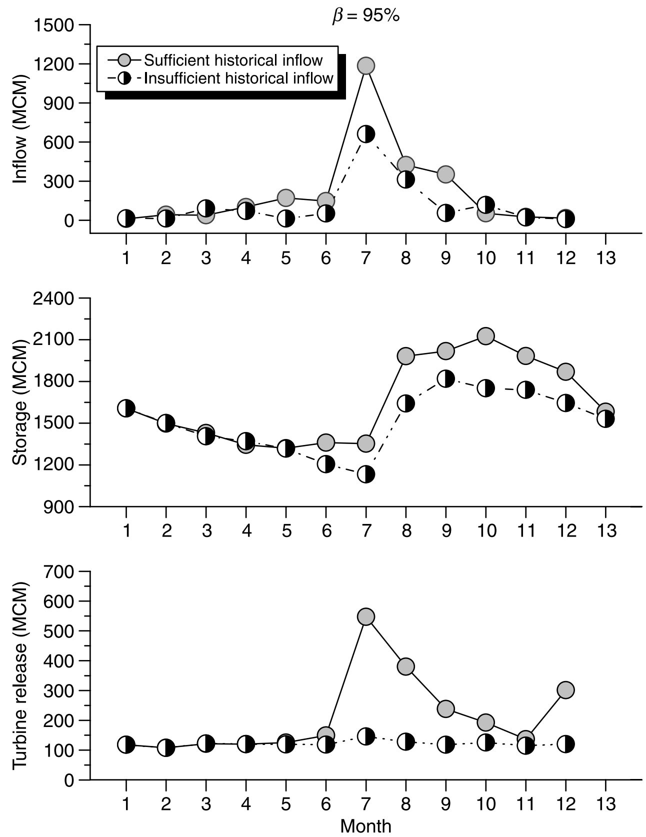

Figure 5 | Inflow, storage and turbine release obtained from the developed piecewise-linear operating rule.

RESULTS AND ANALYSIS

Determination of optimal segment number

It is important to determine the optimal segment number of the piecewise-linear operating rule. If the segment number is larger than required, the chromosome would have many genes, resulting in longer computation times. In contrast, if the segment number is smaller, the number of constraint violations would increase because the operating rule is too simple to satisfy the complicated constraints of reservoir operation.

Ahmed & Sarma (2005) provided an interesting fact regarding the piecewise-linear operating rule: the more the number of segments, the worse the performance of the piecewise-linear operating rule if the inflow data used were different from those used for the optimization model while developing the operating rule. This could be because the GA becomes more specific to the data of the optimization model if the number of segments increases.

Therefore, a small number of segments, namely, 3, 4, 5 and 6 segments, are investigated to determine the optimal segment number. Each case has three different constraints for the water supply in Equation (10): $\beta = 95\%$, 90% and 85%. Pareto-fronts for all 12 cases are computed by NSGA-II with the same seed number. No Pareto-optimal solutions without constraint violations are found for the case in which 3 segments are used with $\beta = 95\%$ and that in which 6 segments are used with $\beta = 95\%$; thus, the solutions in the last generation are infeasible.

With the exception of these two cases, the remaining 10 cases can be divided into two groups according to hydroelectric power generation. The minimum values of the shortage index are almost the same; however, the hydropower generation values, which are functions of the effective head and release, as shown in Equation (4), are different in the Pareto-fronts.

Figure 6 shows 5 Pareto-fronts that have relatively larger hydropower production. Among these 5 Pareto-fronts, those computed using 5 segments with $\beta = 95\%$ and 6 segments with $\beta = 95\%$ show worse performances than the other three cases because they have larger shortage indexes and smaller hydropower production. In the other three cases, it can be said that the Pareto-front that is computed using 4 segments with $\beta = 85\%$ is the best solution because the overall hydroelectric power generation is large and the shortage index values are generally smaller than those of the Pareto-optimal solutions computed using 3 segments with $\beta = 85\%$ and 5 segments with $\beta = 85\%$.

In Figure 7 Pareto-fronts those from 3, 4, and 5 segments are very similar to each other and the best Pareto-optimal solutions can be computed for 5 segments with $\beta = 90\%$. Even though the hydropower production in this case is slightly smaller than that in the other cases, the shortage index values are more important because the Soyanggang reservoir is a multipurpose reservoir that should release a designated amount of water downstream. Therefore, in this study, the optimal segment numbers selected for the piecewise-linear operating rule are 4 segments with $\beta = 85\%$ and 5 segments with $\beta = 90\%$.

Seed number effects

A GA uses a stochastic search procedure based on random number generation. All pseudo-random number generators have a seed number to initiate the calculation of a random number. When different seed numbers are used, if the results calculated are very different then it can be considered that the optimization process is not properly formulated and the optimization results are not the optimal solutions.

Ten different seed numbers are used for evaluating the effect of the seed number on the optimization results. After obtaining the Pareto-optimal solutions, a compromise Pareto-optimal solution that has no preference for any two multiobjective functions is selected and the operating rule is developed using this compromise Pareto-optimal

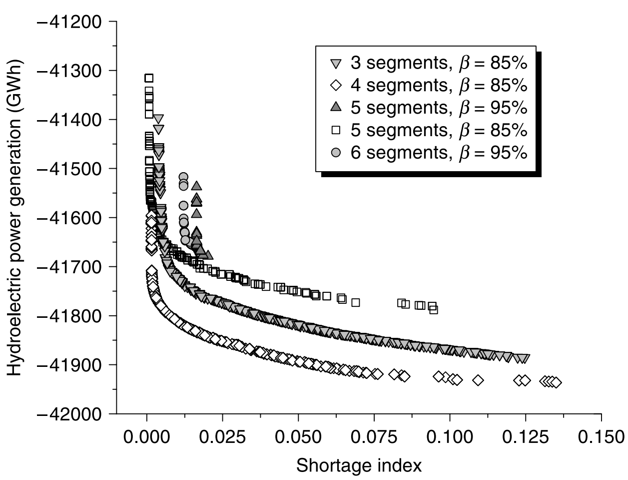

Figure 6 | Five Pareto-fronts with relatively large hydroelectric power generation.

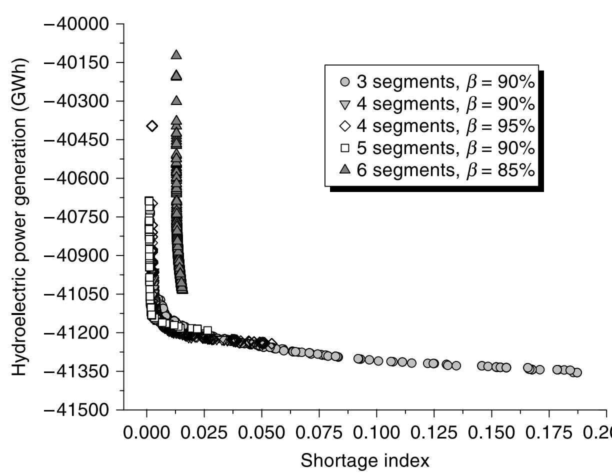

Figure 7 | Five Pareto-fronts with relatively small hydroelectric power generation.

solution. Figure 8 shows 10 piecewise-linear operating rules for February. Each of them is different but the basic forms of these 10 rules are similar, and the gradients of the lines of most of the operating rules significantly change at about 1500–1700 MCM. These results show that, although different seed numbers are used, the operating rules developed by using the optimization results are not significantly different.

The seed number effect could be eliminated if many of the optimization results are computed and then averaged values are used for the operating rules. However, the optimization of reservoir operation is very complex. A large number of constraints are required. For example, 110 constraints are required if the 5-segmented operating rule is applied. In addition, the water balance equation given as Equation (5), which is necessary for representing the state of the optimization problem, is serially connected with the time periods before and after the event. Therefore, it would be impossible to use averaged values for the reservoir operation, and in this study our attention is confined to the achievement of appropriate parameter settings and the results for a single random seed number.

Reservoir operation results using historical inflow

In order to evaluate the applicability of the developed piecewise-linear operating rule to the real-time monthly single-reservoir operation of the Soyanggang reservoir, a simulation model is formulated using the historical inflow over 6 years from 2000 to 2005 and the historical initial storages, which are the storage values on 1 January for each year. The two developed piecewise-linear operating rules, which have 4 and 5 segments, are adopted in the simulation model.

The most important factor for reservoir operation is definitely the inflow because it represents the available water resources, which can be used by the decision-maker of any reservoir for making beneficial decisions in the future. All historical annual inflows are greater than the annual designated water supply to downstream except for that of 2001 which is 1437.54 MCM, and annual inflows reach to at most 3405.28 MCM and at least 1437.54 MCM. In addition, the historical initial storages vary significantly; these storages constitute the water resources and can be used before the water recharging season, namely the flood season in South Korea. The historical initial storages are within 45.3–74.3% of the maximum storage capacity of 2900 MCM for the Soyanggang reservoir, as shown in Table 4.

Figure 9 indicates the historical monthly inflows of the Soyanggang reservoir over 6 years. It can be said that the inflow patterns within a year show large variability. For example, in 2001, when the smallest annual inflow occurred, all the 12 monthly inflows are relatively smaller than those of the other years, while in 2003, when the largest annual inflow occurred, two peaks are observed, first in April and then in August. On the other hand, in 2004, every monthly inflow is relatively smaller than those of the other years with the exception of one peak in July, which has the largest peak inflow value. Therefore, if the developed

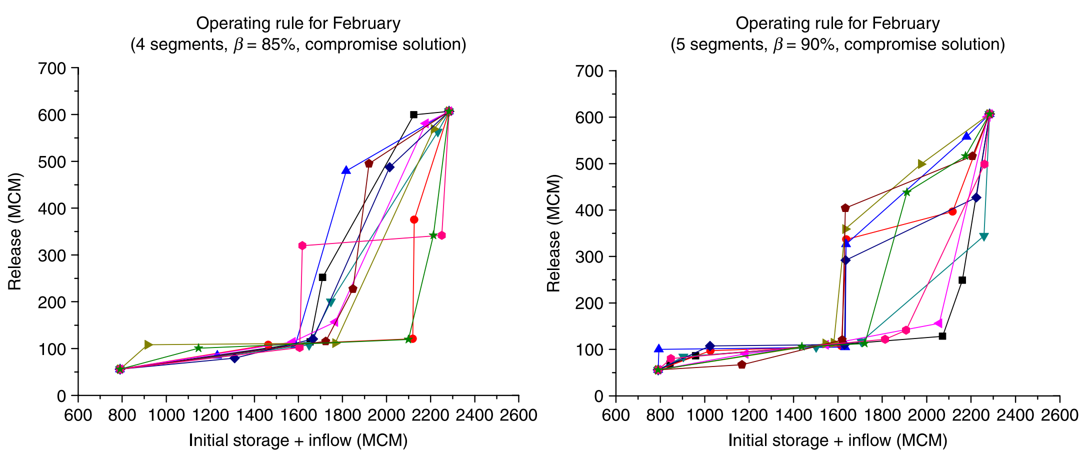

Figure 8 | Operating rules developed by 10 different seed numbers for February. The left operating rules have 4 segments with $\beta = 85\%$ and the right ones have 5 segments with $\beta = 90\%$.

Table 4 | Historical initial storages and annual sums of inflow from 2000 and 2005

Year	2000	2001	2002	2003	2004	2005
Initial storage (MCM)	2154.0 (74.3%*)	1452.3 (50.1%)	1314.6 (45.3%)	1510.1 (52.1%)	1585.0 (54.7%)	1526.0 (52.6%)
Annual sum of inflow (MCM)	1870.33	1437.54	2153.55	3405.28	2564.00	1990.00

*Percentile is the ratio of initial storage to maximum storage capacity, 2900 MCM.

piecewise-linear operating rule can handle these 6 years of historical inflows, which have significantly variable annual inflows and diversity within the annual inflow patterns, it could be applied to the real-time operation of the Soyanggang reservoir.

The simulation is performed using the historical initial storages, historical inflows and developed 4- and 5-segmented piecewise-linear operating rules. Figure 10 shows the computed reservoir storage results for the release that is computed using the developed piecewise-linear operating rule with the historical inflows and historical initial storages. Here, "hist" represents the historical reservoir storage; "water", the storage when a Pareto-optimal solution that has much better objective function values for water shortage than for hydroelectric power production is selected and the corresponding operating rule is employed; "power", the operational result when a Pareto-optimal solution is selected for hydropower production; and "comp", the reservoir storage when a compromise Pareto-optimal solution is used between two multiobjective functions.

All the simulation results provide reasonable reservoir operation results in comparison with the historical operation of the Soyanggang reservoir. Of course, historical operation is not the global optimum, but it should be noted

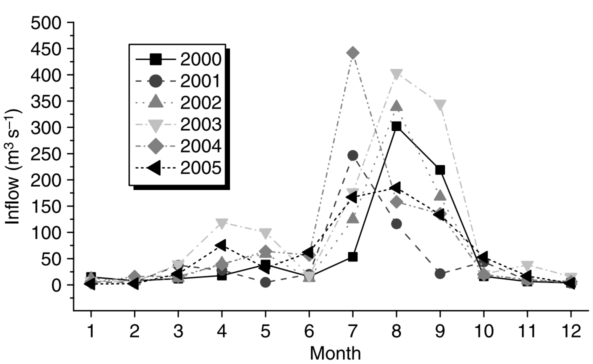

Figure 9 | Historical monthly inflows of the Soyanggang reservoir from 2000 to 2005.

that, for applying the developed operating rule to real-time reservoir operation, the operational results obtained by the developed operating rule should not be significantly different from historical operations.

In addition to the resemblance between the historical and simulation results, the results from the developed operating rules fully satisfy the constraints of terminal storage, Equation (9), even though high or low initial storages are used in the simulation model. For example, the initial storage in 2000 is 2154.0 MCM, which is 74.3% of the maximum storage capacity of the Soyanggang reservoir. Even though this value is much larger than the value obtained by the optimization model, which is 1606.6 MCM, the terminal storages obtained from the three operating rules of "water", "comp" and "power" are approximately 1526 MCM, which is approximately 95% of 1606.6 MCM; thus, the obtained terminal storages satisfy the constraints of terminal storage in the optimization model. In contrast, even if the initial storage is small, such as 1314.6 MCM in 2002, all the three computed terminal storages satisfy the constraint of terminal storage in Equation (9).

Figure 10 graphically shows the reservoir storage results when the developed operating rules are employed, and Table 5 indicates the comparison results when two developed operating rules are used. Because the lower limit of water supply demand that is 85% of β in Equation (10) is used for the 4-segmented operating rules, the shortage indexes are larger than those of the 5-segmented operating rules; as a result, the average water shortages are more severe. On the other hand, the "power" column provides the results computed using a Pareto-optimal solution for hydropower production; thus, the hydroelectric power sum is slightly larger than the sums of the "water" and "comp" columns. The average water shortages of the "water" column also have the smallest value when three average shortage values having the same β value are compared. All these simulation results reveal that the developed piece-

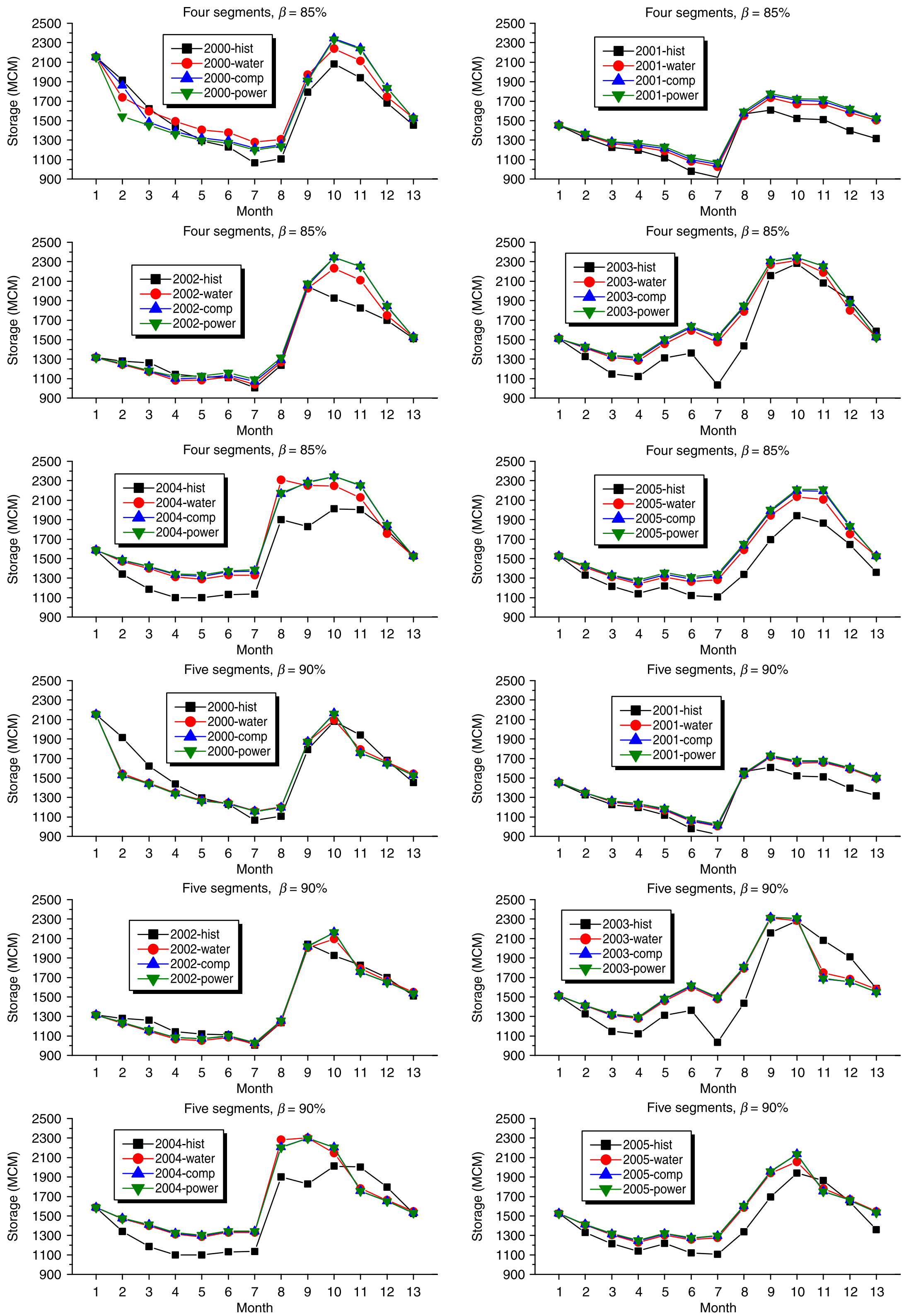

Figure 10 | Reservoir operation results obtained using the developed 4- and 5-segmented piecewise-linear operating rules.

Table 5 | Reservoir operation results obtained using piecewise-linear operating rules

	Four segments, β = 85%			Five segments, β = 90%		
	Water	**Comp**	**Power**	**Water**	**Comp**	**Power**
Shortage index	0.23	0.37	0.48	0.17	0.21	0.23
Hydroelectric power sum (GWh)	3159.64	3164.09	3169.90	3090.02	3119.34	3120.65
Average turbine discharge (MCM)	2220.28	2215.00	2214.96	2187.28	2204.27	2203.91
Average water shortage (MCM)	− 10.49	− 20.65	− 33.88	− 8.94	− 14.79	− 23.77
Average hydroelectric power (GWh)	526.61	527.3	528.32	515.00	519.89	520.11
Average number of water shortages	4.50	5.83	6.16	4.50	6.17	6.17

Release sum and all the averaged values are calculated using results over 6 years from 2000 and 2005.

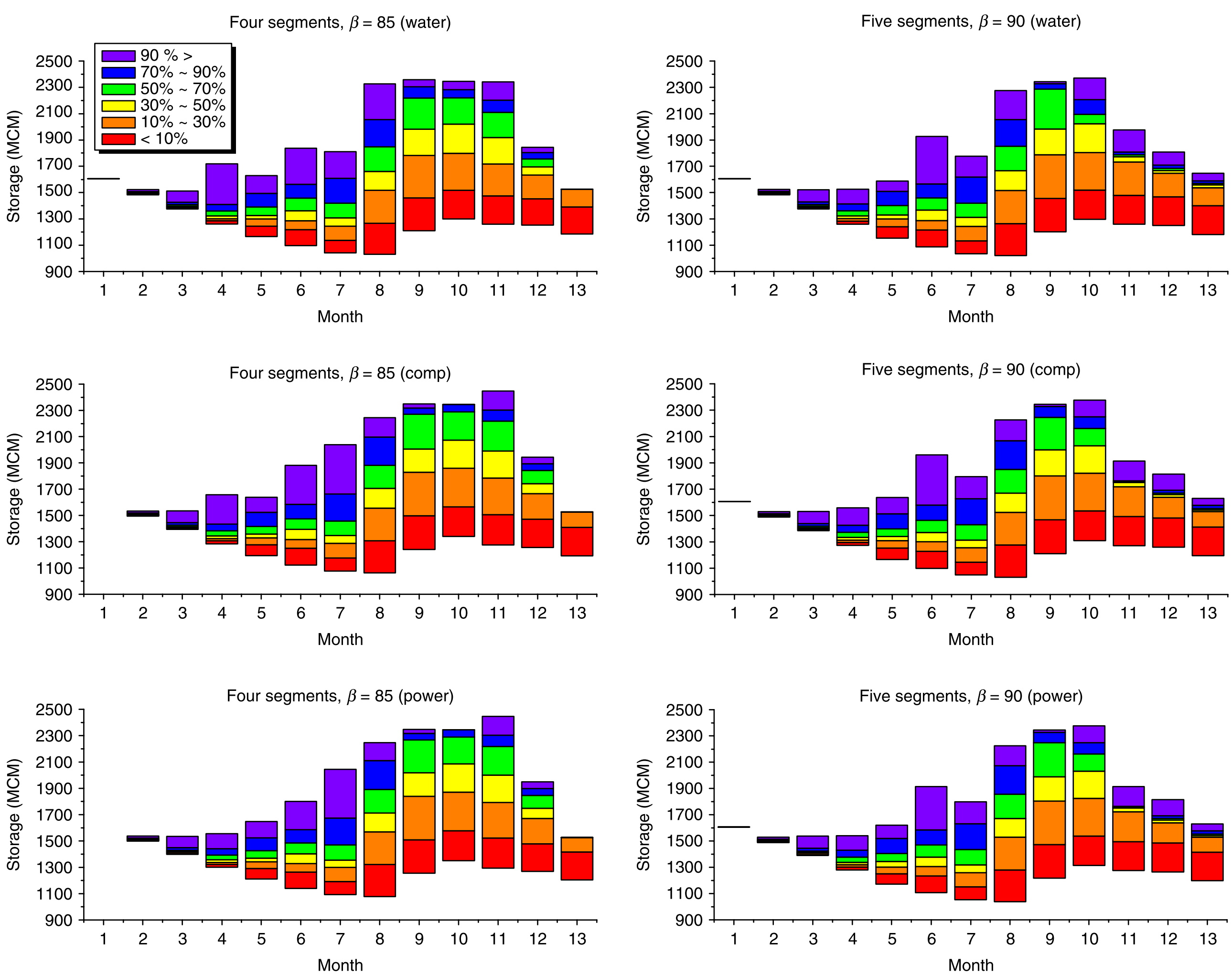

Figure 11 | Probabilistic long-term reservoir storage forecast by piecewise-linear operating rule.

wise-linear operating rule would be able to provide appropriate reservoir operation results that fully consider the effects of multiobjective functions in the optimization model used for developing the operating rule.

Probabilistic long-term reservoir storage forecasting

Probabilistic monthly storage forecasting is performed using the developed piecewise-linear operating rule. First, 100 storage forecast results are calculated by using the possible releases computed from 100 years of synthetic inflow data as input for the operating rule. Then, 100 storage values for each month are sorted in ascending order in order to compute quantiles having 10%, 30%, 50%, 70% and 90% nonexceedance probabilities.

Figure 11 plots 6 probabilistic reservoir storage forecast results. The left column shows the results for the 4-segmented operating rule with $\beta = 85\%$ and the right column shows those for the 5-segmented operating rule with $\beta = 90\%$. If three storage forecasts in only the left or right column are compared, the "water" graph usually has the lowest storage values, implying that more water resources are supplied downstream. It is interesting to note that the extent of the nonexceedance probability between 10% and 30% is generally the largest, and then this would mean that the probable storage of the Soyanggang reservoir has more chance to be low value, which is relatively small.

Figure 11 can provide quantitative information regarding the overall reservoir status to the decision-maker. In general, the decision-maker uses the average storage values over the past reservoir operation period to evaluate the reservoir status so that, if the current storage is larger than the average, it can be said that there is "enough" water in the reservoir and vice versa. However, this information just provides a qualitative status of the reservoir. If the decision-maker wants to know the reservoir water storage status quantitatively, the probabilistic reservoir forecast results are used. Thus, instead of using words such as "enough" to describe the reservoir storage, the decision-maker would be able to provide more quantitative information, e.g. the reservoir storage is within 10% and 30%; therefore, more water resources should be secured for preventing water shortage in future.

CONCLUSIONS

In this study, the implicit stochastic optimization approach is used and, instead of using regression analysis for the optimization results, the developed reservoir operating rule is found directly from the optimization model. The piecewise-linear operating rule for the Soyanggang reservoir was developed using a multiobjective genetic algorithm (NSGA-II) and the synthetic inflow that was generated by time-series modeling. The following are used in the optimization model: two multiobjective functions with regard to minimizing shortage index and maximizing the sum of hydropower production; constraints of reservoir capacity, water supply demand, terminal storage, designated water release and coordinates of each end point; and water balance equations for water storage, release and evaporation.

In order to formulate the operating rule effectively, two aspects of the piecewise-linear operating rule are examined in detail: search space determination and effects of inflow and constraints. First, the upper and lower limits of the first and last end points are determined by frequency analysis. If the upper and lower limits are simply set to the storage values corresponding to normal pool and low water levels, respectively, the search space of NSGA-II would become very large, particularly for the non-flood season. Therefore, each quantile having 1% exceedance and nonexceedance probabilities is computed by the frequency analysis of the historical storage record on the first day of a month; these quantities are used as the upper and lower limits of the optimization model. Furthermore, these limits could be reasonable estimations with slight variations from the historical maximum and minimum values.

On the other hand, the relationships between the inflow and the constraints with regard to water supply demand and terminal storage are also investigated. If the annual inflow is smaller than the annual designated water supply demand, both constraints cannot be simultaneously satisfied. Thus, a piecewise-linear operating rule is developed using only a "sufficient inflow," and the simulation results computed by the developed operating rule show favorable storage and turbine release patterns, regardless of whether the inflow is sufficient or insufficient.

In the case study, the simulation results are obtained by using the developed piecewise-linear operating rule.

Four- and five-segmented operating rules are adopted along with 6 years of historical inflow data of the Soyanggang reservoir. The reservoir operation results show that the developed piecewise-linear operating rule can handle various inflow series that have different characteristics and can generally satisfy the constraints defined in the optimization model, including the constraint of terminal storage. In addition, a probabilistic long-term reservoir storage forecast is provided. This storage forecast would be useful information since a system operator is able to evaluate the current status of the reservoir quantitatively, but not qualitatively.

Some improvements can be made to this study. The multiobjective functions used in this study are with regard to water shortage and hydropower production. However, recently, other objectives such as navigation, recreation and environment have become more important. It is difficult to measure all these objectives; further, it is difficult to develop two or three conflicting multiobjective functions that include these new objectives. Moreover, an inflow forecasting model should be included in the optimization model for developing the reservoir operating rule. While the deterministic synthetic generated inflow simply reflects the characteristics of past inflows, forecasting models such as ensemble streamflow forecasting (ESP) can improve the performance of the reservoir operating rule.

ACKNOWLEDGEMENTS

This research was supported by a grant (1-6-3) from the Sustainable Water Resources Research Center of 21st Century Frontier Research Program.

REFERENCES

Ahmed, J. A. & Sarma, A. K. 2005 Genetic algorithm for optimal operating policy of a multipurpose reservoir. *Wat. Res. Manage.* **19** (2), 145–161.

Chang, L. -C. & Chang, F. -J. 2001 Intelligent control for modelling of real-time reservoir operation. *Hydrol. Process.* **15**, 1621–1634.

Chang, F. -J., Chen, L. & Chang, L. -C. 2005*a* Optimizing the reservoir operating rule curves by genetic algorithms. *Hydrol. Process.* **19** (11), 2277–2289.

Chang, Y. -T., Chang, L. -C. & Chang, F. -J. 2005*b* Intelligent control for modeling of real-time reservoir operation, Part II: artificial neural network with operating rule curves. *Hydrol. Process.* **19** (7), 1431–1444.

Chen, L., McPhee, J. & Yeh, W. W. G. 2007 A diversified multiobjective GA for optimizing reservoir rule curves. *Adv. Wat. Res.* **30** (5), 1082–1093.

Deb, K. 2001 *Multi-objective Optimization Using Evolutionary Algorithms*. John Wiley & Sons, Chichester.

Deb, K. & Jain, S. 2003 Multi-speed gearbox design using multi-objective evolutionary algorithms. *J. Mech. Design* **125** (3), 609–619.

Deb, K. & Raji, A. R. 2003 Reliable classification of two-class cancer data using evolutionary algorithms. *Biosystems* **72**(1–2), 111–129.

Deb, K., Pratap, A., Agarwal, S. & Meyarivan, T. 2002 A fast and elitist multiobjective genetic algorithm: NSGA-II. *Evol. Comput. IEEE Trans.* **6** (2), 182–197.

Edgeworth, F. Y. 1881 *Mathematical Physics*. Keagan Paul, London.

Goldberg, D. E. 1989 *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison Wesley, Reading, MA.

Holland, J. H. 1975 *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI.

ISI 2004 Fast breaking papers. http://www.esi-topics.com/fbp/ ISI.

Karamouz, M. & Houck, M. H. 1987 Comparison of stochastic and deterministic dynamic programming for reservoir operating rule generation. *Wat. Res. Bull.* **23** (1), 1–9.

Karamouz, M., Houck, M. H. & Delleur, J. W. 1992 Optimization and simulation of multiple reservoir systems. *J. Wat. Res. Plann. Manage.* **118** (1), 71–81.

Kim, T. 2005 *Multireservoir System Optimization Using Multi-objective Genetic Algorithms*. PhD dissertation, Yonsei University.

Kim, T. & Heo, J.-H. 2004 Multireservoir system optimization using multi-objective genetic algorithms. In *Proceedings of the 2004 World Water and Environmental Resources Congress, Salt Lake City, UT*. ASCE, Virginia, doi: 10.1061/40737 (2004)244.

Kim, T., Heo, J. -H. & Jeong, C. -S. 2006 Multireservoir system optimization in the Han River basin using multi-objective genetic algorithms. *Hydrol. Process.* **20** (9), 2057–2075.

Kuo, J. -T., Chen, W. -C. & Chen, L. 2003 Multiobjective water resources systems analysis using genetic algorithms– application to Chou-shou River Basin, Taiwan. *Wat. Sci. Technol.* **48** (10), 71–77.

Liong, S. -Y., Tariq, A. A. -F. & Lee, K. S. 2004 Application of evolutionary algorithm in reservoir operations. *J. Inst. Engrs., Singapore* **44** (1), 39–54.

Loucks, D. P., Stedinger, J. R. & Haith, D. A. 1981 *Water Resource Systems Planning and Analysis*. Prentice-Hall, Englewood Cliffs, NJ.

Mitchell, M. 1996 *An Introduction to Genetic Algorithms*. MIT Press, Cambridge, MA.

NIDP 2002 FARD 2002 User Manual (in Korean). Hydrology and Water Resources System Laboratory, Yonsei University.

Oliveira, R. & Loucks, D. P. 1997 Operating rules for multireservoir systems. *Wat. Res.* **33** (4), 839–852.

Pareto, V. 1896 Cours d'economic politique. F. Rough, Lausanne.

Prasad, T. D. & Park, N. -S. 2004 Multiobjective genetic algorithms for design of water distribution networks. *J. Wat. Res. Plann. Manage.* **130** (1), 73–82.
Prasad, T. D., Walters, G. A. & Savic, D. A. 2004 Booster disinfection of water supply networks: multiobjective approach. *J. Wat. Res. Plann. Manage.* **130** (5), 367–376.
Reddy, M. J. & Kumar, D. N. 2006 Optimal reservoir operation using multi-objective evolutionary algorithm. *Wat. Res. Manage.* **20** (6), 861–878.
Reed, P. & Minsker, B. 2004 Striking the balance: long-term groundwater monitoring design for conflicting objectives. *J. Wat. Res. Plann. Manage.* **130** (2), 140–149.
Salas, J. D., Delleur, J. W., Yevjevich, V. & Lane, W. L. 1980 *Applied modeling of hydrologic time series*. Water Resources Publications, Littleton, Co.
Srinivas, N. & Deb, K. 1994 Multiobjective optimization using nondominated sorting in genetic algorithms. *Evol. Comput.* **2** (3), 221–248.
Tejada-Guibert, J. A., Johnson, S. A. & Stedinger, J. R. 1993 Comparison of two approaches for implementing multireservoir operating policies derived using stochastic dynamic programming. *Wat. Res.* **29** (12), 3969–3980.
USACE 1997 *Hydrologic Engineering Requirements for Reservoirs*. US Army Corps of Engineers, CA.
Van Veldhuizen, D. A. & Lamont, G. B. 2000 Multiobjective evolutionary algorithms: analyzing the state-of-the-art. *Evol. Comput.* **8** (2), 125–147.
Young, G. K. 1967 Finding reservoir operating rules. *J. Hydraul. Div.* **93** (HY6), 297–321.

First received 4 June 2007; accepted in revised form 23 November 2007