

Developing reservoir operational decision rule by genetic programming

E. Fallah-Mehdipour, O. Bozorg Haddad and M. A. Mariño

ABSTRACT

The reservoir operational decision rule is an equation that can balance reservoir system parameters in each period by considering previous experiences of the system. That equation includes variables such as inflow, volume storage and released water from the reservoir that are commonly related to each other by some constant coefficients in predefined linear and nonlinear patterns. Although optimization tools have been extensively applied to develop an optimal operational decision rule, only optimal constant coefficients have been derived and the operational patterns are assumed to be fixed in that operational rule curve. Genetic programming (GP) is an evolutionary algorithm (EA), based on genetic algorithm (GA), which is capable of calculating an operational rule curve by considering optimal operational undefined patterns. In this paper, GP is used to extract optimal operational decision rules in two case studies by meeting downstream water demands and hydropower energy generation. The extracted rules are compared with common linear and nonlinear decision rules, LDR and NLDR, determined by a software package for interactive general optimization (LINGO) and GA. The GP rule improves the objective functions in the training and testing data sets by 2.48 and 8.53%, respectively, compared to the best rule by LINGO and GA in supplying downstream demand. Similarly, the hydropower energy generation improves by 48.03 and 44.21% in the training and testing data sets, respectively. Results show that the obtained objective function value is enhanced significantly for both the training and testing data using GP. They also indicate that the proposed rule, based on GP, is effective in determining optimal rule curves for reservoirs.

Key words | decision rule, genetic programming, reservoir system

E. Fallah-Mehdipour

O. Bozorg Haddad (corresponding author)
Department of Irrigation & Reclamation
Engineering,
Faculty of Agricultural Engineering & Technology,
College of Agriculture & Natural Resources,
University of Tehran,
Karaj, Tehran,
Iran
E-mail: Haddad@iust.ac.ir

M. A. Mariño

Department of Land, Air & Water Resources,
Department of Civil & Environmental Engineering,
and
Department of Biological & Agricultural
Engineering,
University of California,
139 Veihmeyer Hall,
University of California,
Davis, CA 95616-8628,
USA

INTRODUCTION

Reservoirs are important structures that can store and release water based on decisions made by operators of the system. Those decisions directly affect the purpose of the operation, such as supplying downstream demands, generating hydropower energy and controlling floods. Prior experience helps the operator to make an appropriate decision to calculate how much (amount) and when (time) to release water from the reservoir.

In recent decades, different types of rules have been widely used to extract operational policies from long-term operational experiences. Linear decision rules (LDRs), standard operation policy (SOP), hedging rules (HRs) and nonlinear decision rules (NLDRs) are common rules that

use linear and nonlinear equations to identify operation policies. There are simulation and optimization techniques that can be used to extract operation policies. Software packages that can simulate reservoir conditions are used to extract operation policies. Although less precise than software packages, trial-and-error can be used to determine optimal/near-optimal solutions. While it is possible to calculate optimal/near-optimal solutions by trial-and-error, the probability of success is directly related to the number of times one executes trial-and-error calculations, which can be time-consuming. Thus, use of an optimization method together with a simulation model is recommended to determine optimal reservoir operation policies. Linear

programming (LP) and nonlinear programming (NLP) are two general optimization techniques that can be used to calculate optimal solutions (Fallah-Mehdipour *et al.* 2011) by using software packages such as LINGO and a general algebraic modeling system (GAMS). However, those software packages are sometimes not capable of calculating an optimal solution in complex problems with a large number of decision variables and linear and nonlinear constraints (e.g. Bozorg Haddad *et al.* 2008a, b). In recent years, evolutionary algorithms (EAs) have been used as optimization tools in the determination of optimal solutions in complex problems. Although the use of those algorithms does not guarantee a global optimal solution, EAs are good candidates to determine optimal/near-optimal solutions.

Various EAs have been extensively used to solve reservoir operation problems, especially in the determination of an operational decision rule. Genetic algorithm (GA) is a random-based algorithm that searches the decision space by using techniques inspired by natural evolution. First used by Goldberg (1989), GA has been extensively applied in science and engineering. In the water resources field, GA with its modifications have been used in reservoir operation by Oliveira & Loucks (1997), Wardlaw & Sharif (1999), Sharif & Wardlaw (2000), Cai *et al.* (2001), Chang & Chang (2001), Reis *et al.* (2005) and Chang *et al.* (2005a, b). In all the aforementioned investigations, the decision variables are numerical values. In those studies, formulation of operational decision rules have been assigned and their coefficients determined by GA. For instance, released water from a reservoir is related by linear or nonlinear equations to water storage and inflow at each period of reservoir operation. Although the coefficients of operational rule curves are optimized by GA, the operational patterns are dictated to the system.

Genetic programming (GP) is one of the EAs based on GA, in which mathematical operators and functions are added to the numerical values as decision variables. Thus, GP is capable of presenting a mathematical equation as a result, which involves different variables. This equation has been applied in many water resource problems to estimate or predict a variable which is directly/indirectly dependent on the other variable(s). Savic *et al.* (1999) applied GP to flow prediction for the Kirkton catchment in Scotland. The results obtained were compared to optimally calibrated conceptual models and an artificial neural network (ANN). Results

showed that data-driven approaches (GP and ANN) gave acceptable predicted values, considering the relative size of the models and the number of variables included. Khu *et al.* (2001) used GP to forecast runoff for the Orgeval catchment in France. GP functions as an error-updating procedure complemented the rainfall-runoff model, MIKE11/NAM. Results indicated that the proposed methodology was able to forecast accurate storm events for different updating intervals. Rabunal *et al.* (2007) determined the unit hydrograph of a typical urban basin using GP and ANN. Results showed that there is no considerable difference between a hydrograph of a conceptual model and one resulting from GP and ANN. Sivapragasam *et al.* (2008) applied GP for flood routing in natural channels. Results showed that hydrograph peaks are accurately predicted and there is no time lag in the occurrence of the peak, unlike with the nonlinear Muskingum model. Guven & Gunal (2008) used GP for prediction of local scour downstream of hydraulic structures. The GP-based formulation results were compared with experimental results and other equations and were found to be more accurate. By using GP, Sivapragasam *et al.* (2009) modeled evaporation-seepage losses for reservoir water balance in semi-arid regions. Results of GP and Penman's model for both evaporation loss estimation and reservoir scheduling were compared. While GP and Penman's combination model performed equally well for estimating evaporation losses, GP was also able to model seepage losses (or other losses from a reservoir) to a much better degree. Kisi & Guven (2010) used linear GP (LGP) to estimate evaporation using data from three climatology stations in California. They compared results of the LGP with ANN and support vector machine (SVR), with the LGP showing to be a superior alternative to the SVR and ANN techniques. Guven & Kisi (2010) successfully used LGP for the estimation of suspended sediment yield in natural rivers. Izadifar & Elshorbagy (2010) employed GP, ANN and statistical regression models to estimate actual evapotranspiration using meteorological variables. Results showed that GP and regression models performed better than ANN in the estimation of actual evapotranspiration.

In this paper, reservoir operational LDR and NLDR are developed using LINGO as LP and NLP solvers and GA. The applied rules have a specific linear and nonlinear form which can affect the efficiency of the operation of a reservoir system by applying the high-order equations. To compare the

capability of these tools to determine reservoir operational rules with an optimization technique, GP is then used to extract a decision rule for the operation of a reservoir system supplying downstream demands and generating hydropower energy in two case studies. Results show the efficiency of the GP rule compared to those of LINGO and GA by applying traditional rules with a specified pattern.

RESERVOIR SYSTEM SIMULATION

Reservoirs are artificial lakes to balance the flow in highly managed systems, taking in water during high flows and releasing it during low flows. This system operates for several purposes, such as supplying downstream demands, generating hydropower energy and flood control. There are several investigations in the short, long and integrating short and long term (e.g. [Batista Celeste *et al.* 2008](#)) reservoir operation without considering any operational decision rules. In these studies, the main variable which is commonly identified as the decision variable is released water. Operational decision rules help the operator to calculate the value of release in each period of operation. To determine an operational rule, a general mathematical equation is usually presented in the model equations:

$$R_t = f(S_t, Q_t) \quad (1)$$

in which R_t = release from the reservoir at period t ; S_t = storage volume of the reservoir at the beginning of the t th period; Q_t = inflow to the reservoir during period t ; and f = linear or nonlinear function for transferring storage volume and inflow to the released water from the reservoir.

The common pattern of Equation (1), which is a linear decision rule, is identified as (e.g. [Mousavi *et al.* 2007](#); [Bozorg Haddad *et al.* 2008a](#))

$$R_t = a \times Q_t + b \times S_t + c \quad (2)$$

In Equation (2), a linear pattern is assumed for reservoir operation and the coefficient values (decision variables) are calculated. Thus, it may be possible to determine a more effective decision rule compared to those extracted by Equation (2).

In this paper, a reservoir operational rule curve without any assumed pattern will be extracted by GP and compared with commonly used LDR and NLDR by LINGO and GA. Thus, all the variables, mathematical operators and coefficient values comprise the decision variables. Because the GP-developed rule is not bounded by a predefined set of rules, GP is capable of yielding more flexible rules which may help the operators to meet more targets. The capability of the methodology will be tested by applying it to two reservoir systems with two purposes: supplying downstream demands and generating hydropower energy. The capability of the methodology in generating GP decision rules in convex and non-convex problems is discussed.

Supplying downstream demand purpose

For the purpose of meeting downstream demand, the objective function is considered to be the minimization of the total squared deviation of the released water:

$$\text{Min. } Z_1 = \sum_{t=1}^T \left(\frac{R_t - D_t}{D_t} \right)^2 \quad (3)$$

in which Z_1 = objective function of the supplying downstream demand purpose; T = number of operating periods; and D_t = downstream demand of reservoir at period t .

The main equation in the reservoir operation is the continuity equation:

$$S_{t+1} = S_t + Q_t - R_t - SP_t - \text{Loss}_t \quad (4)$$

in which S_{t+1} = storage volume of the reservoir at the beginning of the $(t+1)$ th period; SP_t = volume of spilled water from reservoir at period t ; and Loss_t = volume of water lost from the reservoir at period t .

The model's formulation is constrained by the following relations:

$$\text{Loss}_t = F_1(\text{Ev}_t, \bar{A}_t) \quad (5)$$

$$\bar{A}_t = (A_t + A_{t+1})/2 \quad (6)$$

$$A_t = F_2(S_t) \quad (7)$$

$$R_{\text{Min}} \leq R_t \leq R_{\text{Max}} \quad (8)$$

$$S_{\text{Min}} \leq S_t \leq S_{\text{Max}} \quad (9)$$

where F_1 = function for calculating volume of lost water considering evaporation rate; Ev_t = evaporation depth at period t ; \bar{A}_t = average surface at period t ; A_t = water surface at the start of period t ; F_2 = linear function for transferring storage volume to water surface; R_{Min} , R_{Max} = minimum and maximum allowable capacity for release from reservoir; and S_{Min} , S_{Max} = minimum and maximum storage of reservoir.

Generating hydropower energy purpose

The objective of the generating hydropower energy purpose is to make the power generation as close to the installed capacity as possible. Mathematically, the objective function may be written as

$$\text{Min. } Z_2 = \sum_{t=1}^T \left(1 - \frac{P_t}{\text{PPC}} \right) \quad (10)$$

in which Z_2 = objective function of the generating hydropower energy purpose; P_t = generated power during period t ; and PPC = installed capacity of the power plant.

For hydropower energy generation, Equations (11)–(15) are added to the optimization:

$$P_t = F_3(\gamma, e, RP_t, PF, \bar{H}_t, TW_t) \quad (11)$$

$$\bar{H}_t = (H_t + H_{t+1})/2 \quad (12)$$

$$H_t = F_4[(S_t)^3] \quad (13)$$

$$TW_t = F_5[(R_t)^3] \quad (14)$$

$$RPS_t = R_t - RP_t \quad (15)$$

in which F_3 = function of hydropower energy generated; γ = specific weight of water; e = efficiency of the power plant;

RP_t = release from power plant for generated power of the reservoir at period t ; PF = plant factor; \bar{H}_t = average head of the reservoir during period t ; H_t , H_{t+1} = water elevation at the start of period t and $t + 1$; TW_t = tail water elevation at period t ; F_4 = nonlinear function for transferring storage volume to water elevation; F_5 = nonlinear function for transferring hydropower release to tail water elevation; and RPS_t = spilled water from power plant at period t .

In this paper, a dynamic penalty function, which is a linear relation, is used to tackle constraints. This penalty function is added to the minimization objective functions, Equations (3) and (10), as follows:

$$\text{DPF} = A \{ |\text{Min.}[(S_t - S_{\text{Min}}), (S_{\text{Max}} - S_t), 0]| \} + B \{ |\text{Min.}[(R_t - R_{\text{Min}}), (R_{\text{Max}} - R_t), 0]| \} + C \quad (16)$$

where DPF = dynamic penalty function and A , B , and C = positive constants of the linear relation of the penalty factor. Both Equations (8) and (9) are considered in the DPF.

EVOLUTIONARY ALGORITHMS

In artificial intelligence, EAs are a subset of evolutionary computations that can yield optimal/near-optimal solutions in all types of problems such as: linear/nonlinear, discrete/continuous and convex/nonconvex, using validated experimental theories of biological evolution and natural processes. In this paper, two EAs, GA and GP, are employed to develop optimal operation decision rules.

Genetic algorithm

GA belongs to the larger class of EAs, which generates solutions in optimization problems using techniques inspired by natural evolution, such as: inheritance, selection, mutation, and crossover. In GA, a population of strings called chromosomes, which encode candidate solutions to an optimization problem, evolves towards better solutions. The evolution usually starts from a random population of chromosomes. In each generation, the fitness of every chromosome in the population is evaluated, multiple chromosomes are stochastically selected and modified

from the current population by considering their fitness to form a new population. The new population is then used in the next generation of the algorithm. Commonly, the algorithm terminates when either a maximum number of generations or iterations is reached. More information about the GA is contained in Goldberg (1989) and Deb (2001).

Genetic programming

As with GA, GP is an artificial intelligence method with a random search basis to determine an optimal solution. In GA, each decision variable is called a gene and a set of genes is identified as a chromosome. Thus, all chromosomes have the same length in each generation. However, in GP, the chromosomes have a tree structure which can include different numbers of decision variables and can produce an expression as a solution. This expression with a tree structure involves terminals as leaves and functions as nodes. All the numerical and non-numerical variables are assumed to be the terminal set. In contrast, the arithmetic operators (\pm, \times, \div), mathematical functions (e.g. sin, cos), Boolean operators (e.g. and or), logical expressions (e.g. if-then-else) and other user-defined functions are identified as the function set. Figure 1 shows two examples of tree-structure chromosomes in GP. As is shown, $\{x, 1, 5\}$ and $\{x, 12\}$ are respectively the terminal sets of $y(x) = x^2 - x + 5$ and $y(x) = \sin(x) + 12$ expressions. A random set of trees is generated as the initial population in the first generation of the GP searching process. The trees are then compared by considering the calculated fitness

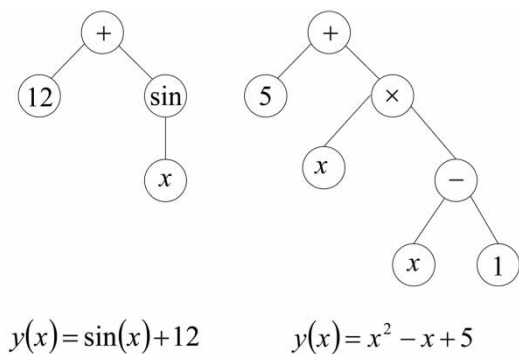


Figure 1 | Examples of GP expressions.

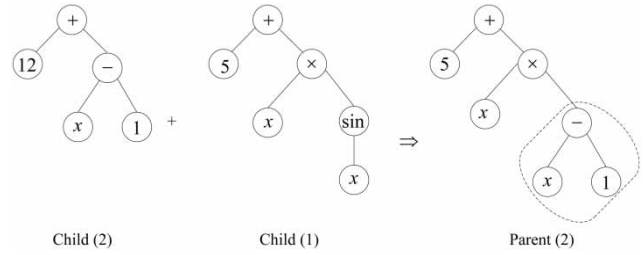


Figure 2 | Crossover between parents and creating children in GP.

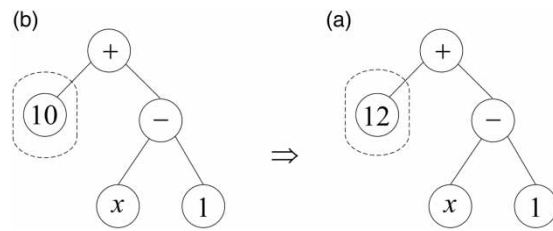


Figure 3 | Tree structures (a) before and (b) after mutation.

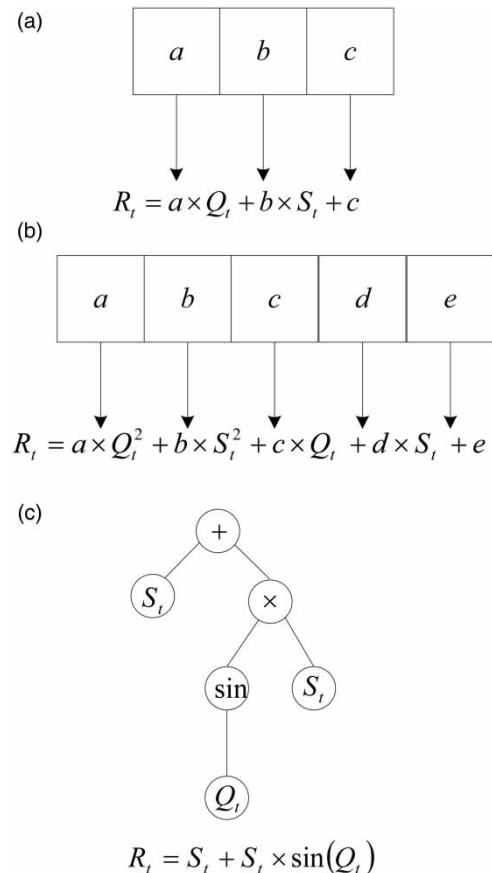


Figure 4 | Schematic structure of solution for GA in (a) LDR, (b) NLDR and (c) GP.

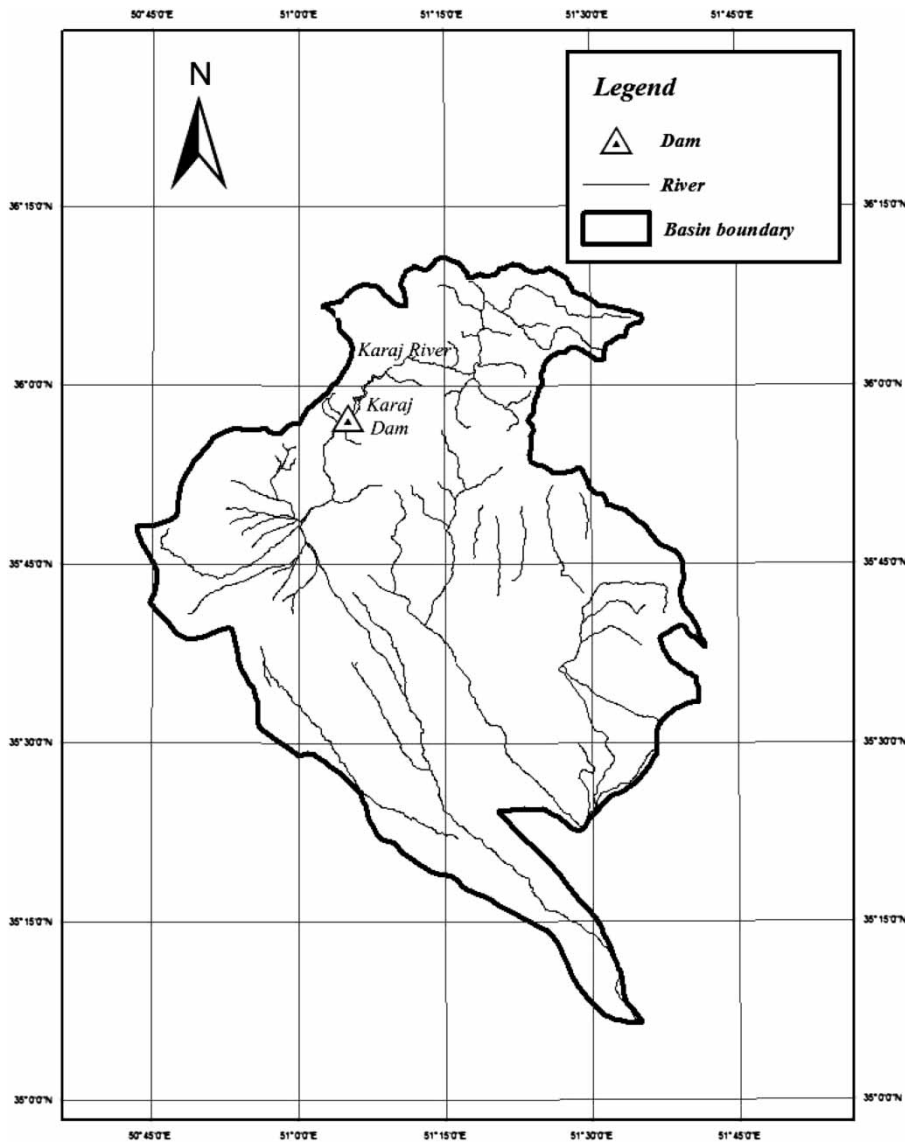


Figure 5 | Location of first case study in the Karaj Basin.

function for each tree. The trees with the better fitness values are selected using techniques such as roulette wheel, tournament or ranking method. The next generation

is prepared using two genetic operators: crossover and mutation. In the GP crossover operator, two trees are assigned as the parents and two children are produced by

Table 1 | Results and statistical measures of five runs of the GA rule curves for the first case study

Rule type	Number of runs					Statistical measures				
	1	2	3	4	5	Minimum	Average	Maximum	Standard deviation	Coefficient of variation
LDR	35.763	35.769	36.110	36.327	36.328	35.763	36.059	36.328	0.282	0.008
NLDR	33.770	34.591	35.698	35.689	34.167	33.770	34.783	35.698	0.881	0.025

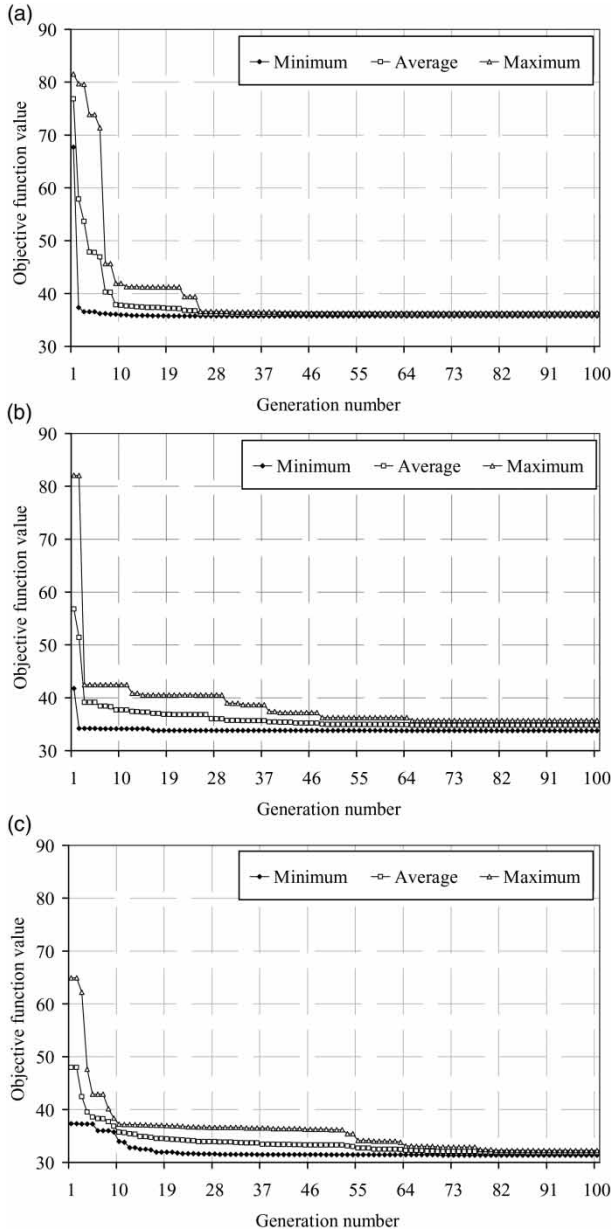


Figure 6 | Trend of minimum, average, and maximum of objective function in five runs for (a) LDR and (b) NLDR by GA and (c) GP in the first case study.

Table 2 | Results of LINGO and GA rule curves for the first case study

Rule type	Rule curve formulation	Method	Coefficient values					Objective function
			a	b	c	d	e	
LDR	$R_t = a \times Q_t + b \times S_t + c$	LINGO	2.1E-03	1.5E-01	1.4E + 01	-	-	34.270
		GA	7.6E-02	2.8E-01	9.6E-01	-	-	35.763
NLDR	$R_t = a \times Q_t^2 + b \times S_t^2 + c \times Q_t + d \times S_t + e$	LINGO	1.0E-05	1.2E-04	1.2E-01	1.5E-01	1.3E + 01	32.130
		GA	2.8E-06	-3.9E-09	7.2E-02	2.4E-01	8.4E + 00	33.770

swapping sub-trees from parents. Figure 2 illustrates crossover between parents and creating children in the GP. The next genetic operator is mutation, which randomly exchanges a decision variable in a node with another random variable. Figure 3 shows the structure of two trees before and after mutation. The produced trees using genetic operators are the input for the next generation. This process continues up to the maximum number of generations or iterations.

APPLICATION

This paper considers two case studies to illustrate the capability of GP to extract reservoir operational decision rules and compare the algorithm’s efficiency with common LDR and NLDR developed by LINGO as LP and NLP solvers and GA. The first and second case studies consider reservoir operation for the purpose of supplying downstream demands and generating hydropower energy, respectively. Thus, the capabilities of different rules are compared in convex and non-convex problems. Those rules have linear and nonlinear forms which respectively coincide with the first and second order of inflow and volume storage. Results are then compared with the rule developed by GP in each case study.

GA and GP were coded in the software package Matlab7.0 and run on a PC/WindowsXP, 256 MB RAM, 2 GHz computer. The execution time of each run was less than 1 minute. In addition, four arithmetic operators, including \pm , \times and \div and four mathematical functions involving \sin , \cos , power (x^y) and square root ($\sqrt{\quad}$) were considered as the function set in GP. The schematic structure of the chromosome and tree for the GA and GP rules are

presented in Figure 4. As is shown in Figure 4(a), the LDR has a linear equation for which GA finds three constant coefficients. Thus, these coefficients are settled as the

genes in a GA chromosome. If a second-order NLDR is extracted by GA, each chromosome would have five coefficients as the decision variables in each chromosome

Table 3 | Results and statistical measures of five runs of GP rule curve for the first case study

Number of runs					Statistical measures				
1	2	3	4	5	Minimum	Average	Maximum	Standard deviation	Coefficient of variation
31.333	32.234	31.930	31.833	31.948	31.333	31.859	32.234	0.329	0.010

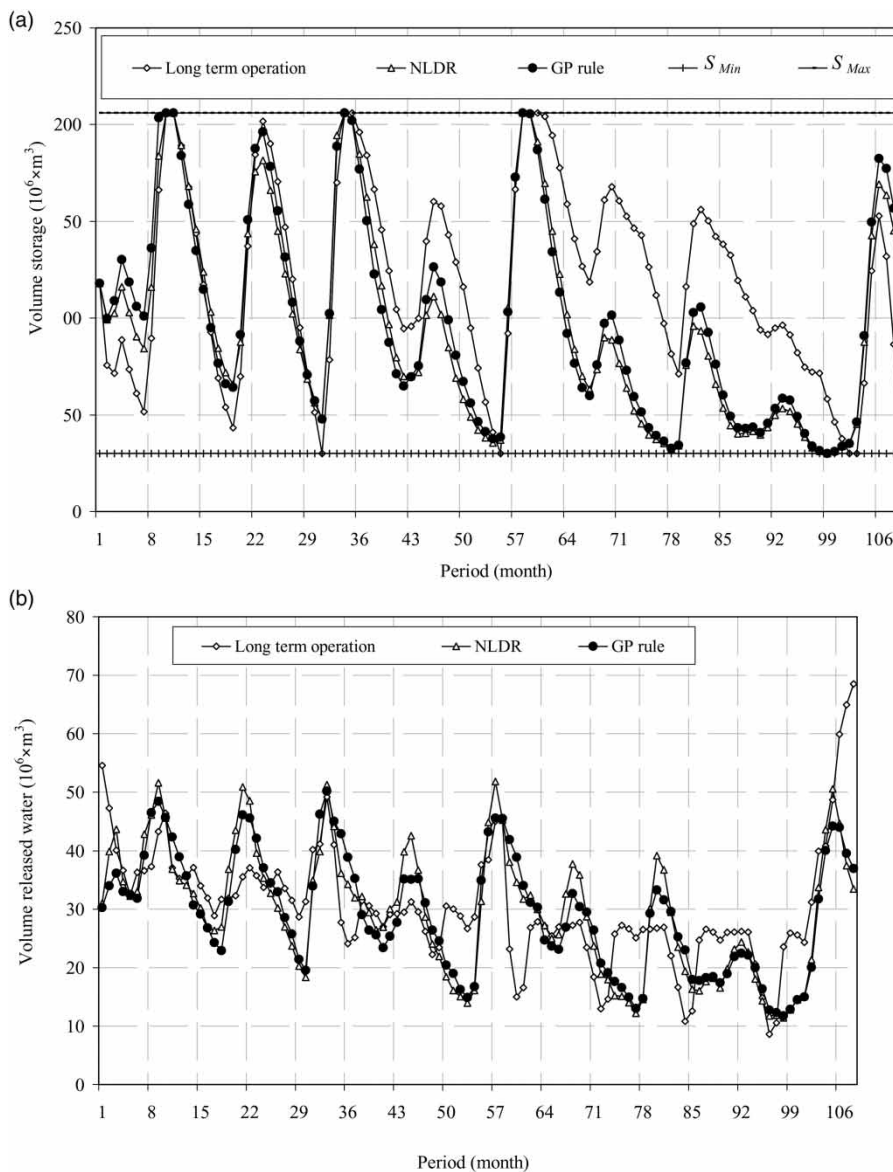


Figure 7 | Volume of (a) storage and (b) released water for the training data of the first case study.

(Figure 4(b)). In the GP developed rule, each tree shows a mathematical equation without a predefined structure. In this tree, each node has arithmetic operators, mathematical functions and variables including inflow and volume storage. Figure 4(c) shows a schematic of tree structure for a GP rule.

First case study

The first case study considers the Karaj reservoir, with an active volume of $176 \times 10^6 \text{ m}^3$, located on the Karaj River with an annual average inflow of $415.23 \times 10^6 \text{ m}^3$ in Iran (Figure 5). The reservoir supplies part of the Karaj and Tehran demands while the other part is supplied by groundwater. Input data including inflow and demand were divided into two training (9 years) and testing (3 years) sets.

LDR and NLDR by LINGO: first case study

As a first step, the optimal objective function was determined for the training data set without any rule. The best (minimum) value of the obtained objective was 27.69 by LINGO. The result of the long-term operation was only determined for the applied time series. Thus, to operate a

reservoir system in real-time, an operational rule curve should be used in reservoir modeling.

At the second step, the linear and nonlinear rules were determined by LINGO. The best results (objective functions) for LDR and NLDR were 34.27 and 32.13, respectively. The best value of objective function without any operation rules was respectively 19.20 and 13.82% smaller (better) than the best results of LDR and NLDR.

LDR and NLDR by GA: first case study

To compare the GA tool as an EA with LINGO as a gradient-based optimization tool, the LDR and NLDR were extracted by GA, using 50 chromosomes and 100 generations. Table 1 shows results of five different runs of the random-based GA algorithm for LDR and NLDR and their statistical measures. It should be noted that the greater the number of runs conducted, the higher the probability to obtain a better solution. According to the results, the coefficients of variation of both rules are acceptable (small value). The minimum value of the NLDR objective function is 5.57% better (smaller) than the minimum value of the LDR. This result shows that more precision is attained in the estimation of released water in the NLDR than in the linear one, even though the results of the different runs are

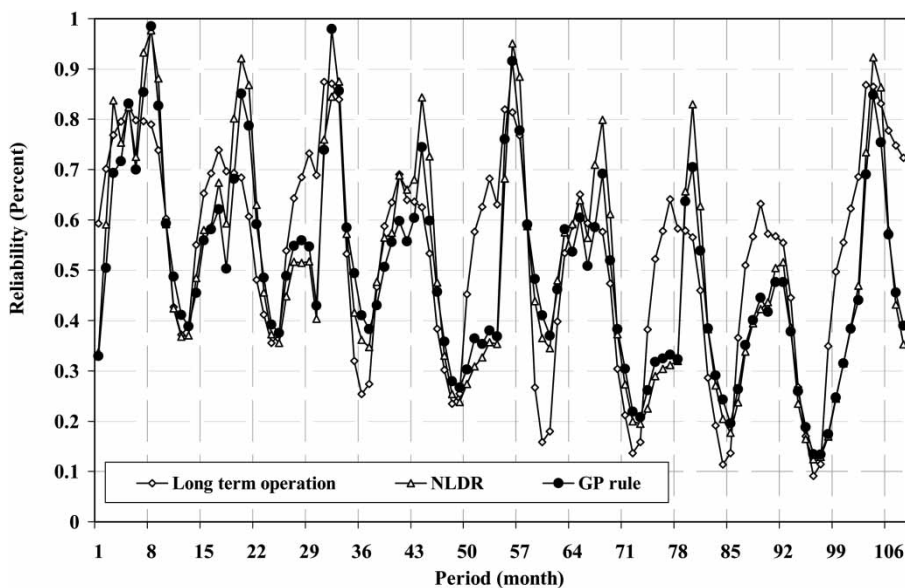


Figure 8 | The reliability percentage of meeting demand for the training data of the first case study.

independent. The minimum value of the obtained objective function for the linear and nonlinear rules is 1.56 and 5.40% better than the maximum value of the obtained objective functions, respectively. It shows that the range of results obtained in different runs is not high using GA, although GA search is an independent search process used in those runs. The similar trend of minimum, average and maximum of these independent objective functions at the end of each generation indicates the high probability of obtaining a better solution even by a single run. This trend of objective

function convergence was also experienced in other investigations in which EAs were used as the optimization tool (e.g. Bozorg Haddad *et al.* 2008a, b). Figures 6(a) and (b) show the convergence trends of the objective function for the LDR and NLDR, respectively. As is shown, there is a decreasing trend of the minimum, average and maximum values of the objective function for five runs. For an appropriate comparison, the best values of objective functions using LINGO and GA are presented in Table 2. As is shown, the best (minimum) value of the objective function

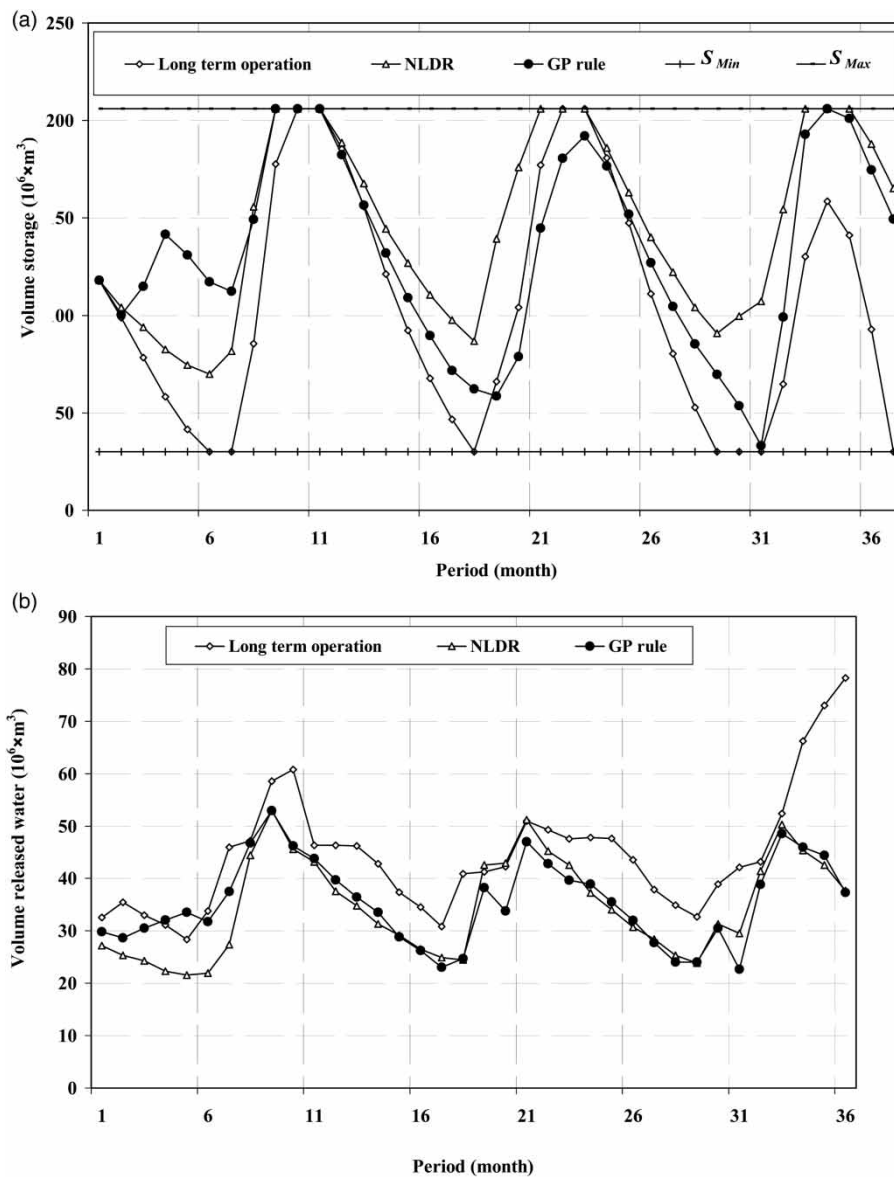


Figure 9 | Volume of (a) storage and (b) released water for the test data of the first case study.

using GA (33.77) is 4.86% better (smaller) than the corresponding value using LINGO (32.13). Thus, GA is capable of determining a near-optimal solution compared to LINGO as the gradient-based solver.

Developed rule by GP: first case study

In EAs, the number of computational efforts needed to achieve optimal/near-optimal solution is equal to the number of function evaluations. This value in each generation or iteration of GA and GP is equal to the number of

chromosomes and trees, respectively. Thus, the total number of function evaluations in the search process for GA and GP is respectively equal to the number of chromosomes and trees which is multiplied by the number of generations or iterations. At this step, GP was used for five different runs involving 50 trees and 100 generations, the same as the GA function evaluation (50 chromosomes \times 100 generations). Figure 6(c) illustrates the decreasing trend of the objective function for GP, as in the case of the GA trend. Table 3 shows objective function values and their statistical measures for five different runs. The best

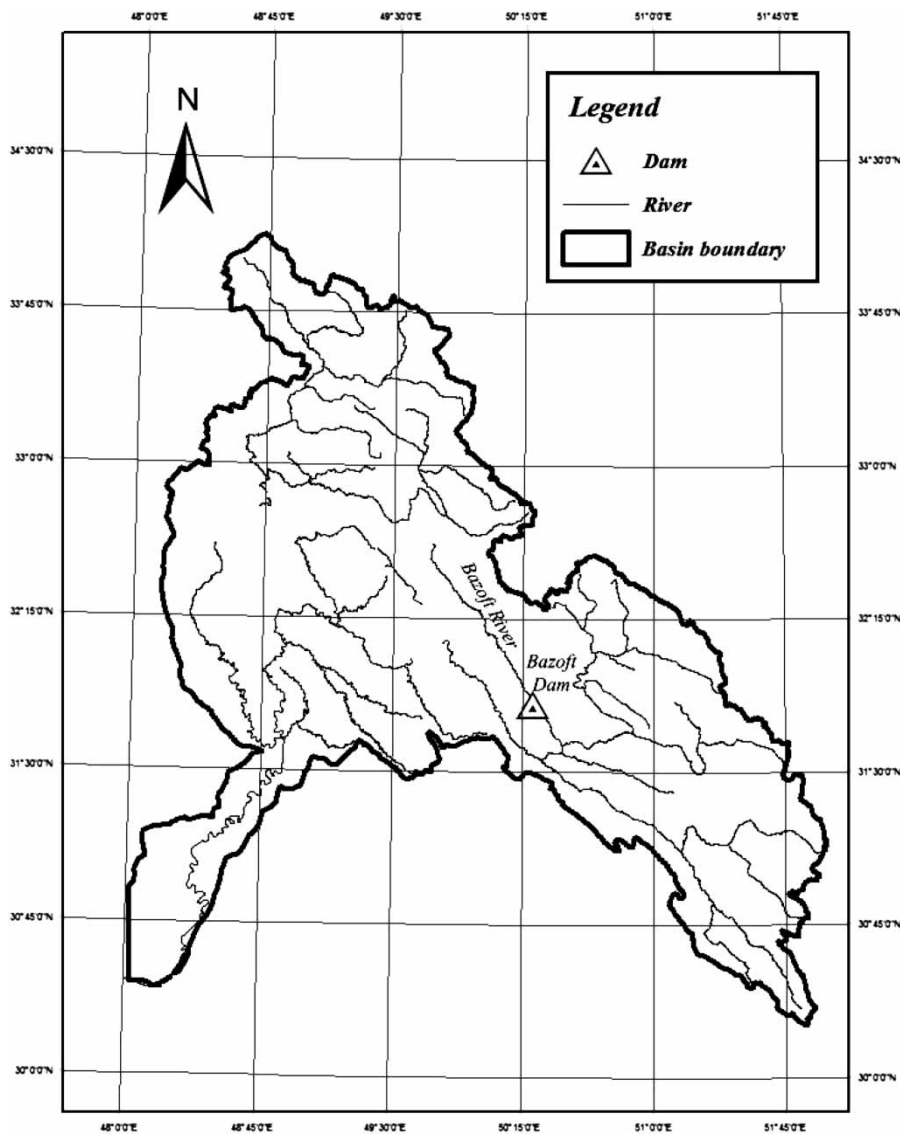


Figure 10 | Location of second case study in the Karoon Basin.

(minimum) value of the objective function is 7.22% better (smaller) than the best value obtained by GA. Table 3 also shows that the coefficient of variation is equal to 0.01 for five different runs. Thus, the deviation of calculated objective functions compared to their average value is small. It follows that there is a good probability to determine an appropriate solution even by one run. Equation (17) is the developed rule with a minimum value of the objective function:

$$R_t = \sqrt{11.351Q_t + (8.693 + \sin(S_t))S_t + 6.7042} \quad (17)$$

This equation has a nonlinear form, including both arithmetic operators and mathematical functions.

To further compare the results, Figure 7 shows the volume storage and released water of the long-term operation by LINGO as the best possible result, NLDR by LINGO as the best result of traditional rules and the GP rule. As is shown, the volume storage of the 3 final years is lower than in previous years. These years are dry periods and the

operator should use more water for volume storage to supply demand. Figure 8 shows the reliability percentage of supplying demand for the long-term operation and NLDR by LINGO and the GP rule even though the reliabilities of NLDR and GP rules are less than the maximum value of reliability in the same periods by long-term operation. It is more than the minimum value in other periods.

At the final step, the NLDR by LINGO as the best result of a predefined pattern rule and GP-developed rule were tested for 3 years (36 months). The objective functions of these two rules were calculated as 7.39 and 6.76 for the LINGO and GP, respectively. According to these results, GP is more capable to determine an appropriate operational rule curve. Figure 9 shows storage volume and released water for the test data set. As is shown, the rule curves use less water from volume storage to supply demands than the long-term operation model without considering any rule curves. Thus, their volume storage is more than the minimum allowable volume storage. It should be noted the

Table 4 | Results and statistical measures of five runs of GA rule curves for the second case study

Rule type	Number of runs					Statistical measures				
	1	2	3	4	5	Minimum	Average	Maximum	Standard deviation	Coefficient of variation
LDR	38.748	38.524	34.558	35.418	35.600	34.558	36.569	38.748	1.929	0.053
NLDR	32.005	32.416	33.280	35.032	32.934	32.005	33.133	35.032	1.168	0.035

Table 5 | Results of LINGO and GA rule curves for the second case study

Rule type	Rule curve formulation	Method	Coefficient values					Objective function
			a	b	c	d	e	
LDR	$R_t = a \times Q_t + b \times S_t + c$	LINGO	1.3E - 01	2.0E - 01	7.3E + 00	-	-	33.595
		GA	1.2E - 01	2.5E - 01	9.8E - 01	-	-	34.558
NLDR	$R_t = a \times Q_t^2 + b \times S_t^2 + c \times Q_t + d \times S_t + e$	LINGO	1.1E - 04	9.7E - 04	1.0E - 04	1.0E - 03	1.0E + 00	31.280
		GA	8.6E - 06	5.7E - 06	7.7E - 02	3.3E - 01	9.9E - 01	32.005

Table 6 | Results and statistical measures of five runs of GP rule curve for the second case study

1	Number of runs				Statistical measures				
	2	3	4	5	Minimum	Average	Maximum	Standard deviation	Coefficient of variation
16.257	20.976	21.424	25.245	21.520	16.718	21.176	25.245	3.027	0.143

best possible objective function for long-term operation of test data sets without considering any rule curves was 3.38 by LINGO.

Second case study

The second case study considers the Bazoft reservoir, located on one of the main rivers, the Bazoft River in Iran (Figure 10). The annual average inflow to the Bazoft reservoir with an active volume of $308.15 \times 10^6 \text{ m}^3$ is about $2012.60 \times 10^6 \text{ m}^3$. This reservoir with an installed capacity of $290 \times 10^6 \text{ W}$ has been designed and operated for hydro-power energy generation. Data sets of inflow and demand were divided into two sets, 25 and 9 years, for the training and test processes, respectively.

LDR and NLDR by LINGO: second case study

The developed rule steps in this case study are the same as in the first case study. Generally, gradient-based optimization tools considering LP and NLP methods are capable of finding best (optimum) solution. In this paper, the best possible optimal solution has been found by the LINGO as a gradient-based tool.

A fitness value of 3.84 was obtained by considering the optimal long-term operation of the reservoir without any operational rule using LINGO. Although this obtained objective relates to the best result for training times series, it is not possible to use them directly to make decisions for the real-time operation of a reservoir. Addition of embedded operation rules as a constraint in the simulation relations is one of the methods used to extract a decision rule in the operating model. Objective functions of the embedded LDR and NLDR were respectively calculated as 33.595 and 31.280 using LINGO.

LDR and NLDR by GA: second case study

In LINGO as the LP and NLP solver, the rule's mathematical pattern should be identified in the simulation model. Thus, LINGO's capability is limited in determining a decision rule. In this paper, to develop the rule pattern, GP will be used. Thus, at first, the capability of GA to determine a traditional rule should be verified. To find LDR and

NLDR by GA, 50 chromosomes with 100 generations were used. Thus, the total number of function evaluations for GA is equal to the number of chromosomes which is multiplied to the number of generations. Table 4 shows results of five different runs for the linear and nonlinear rules and their statistical measures. These results show no considerable

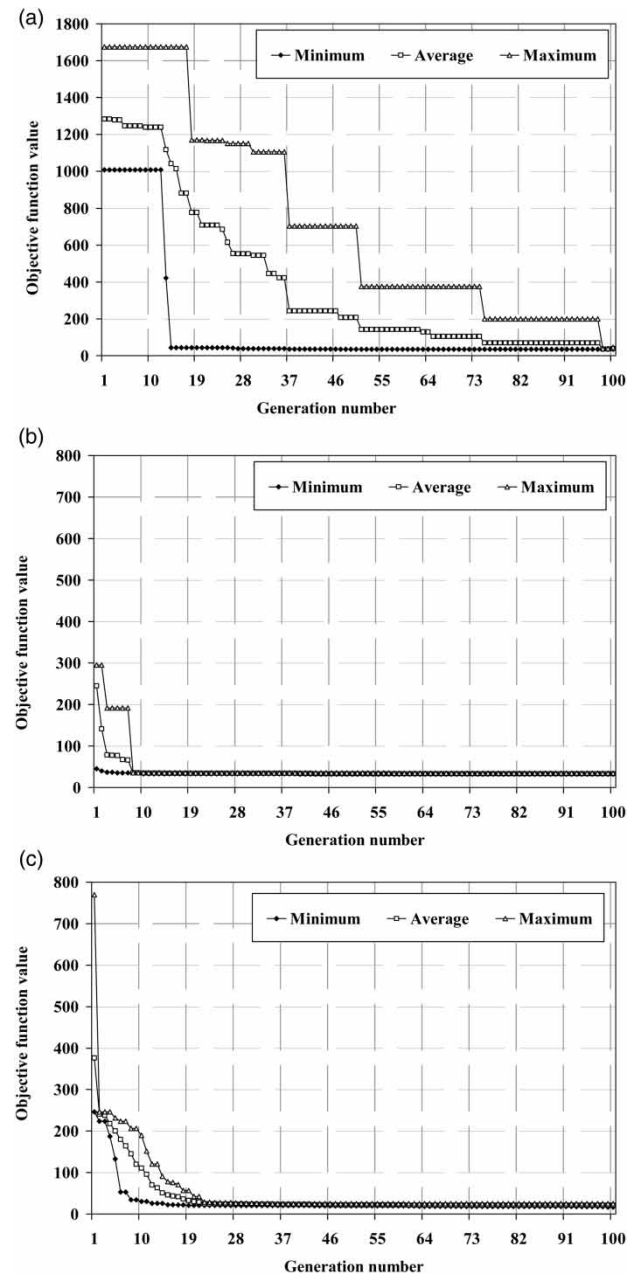


Figure 11 | Trend of minimum, average and maximum of objective function in five runs for (a) linear and (b) nonlinear rule curve by GA, and (c) GP in the second case study.

difference between the objective functions of five runs. As shown in Table 4, all the minimum, average and maximum values of the NLDR are less than the corresponding values of the LDR. Thus, the NLDR estimated released water with more precision with the same number of function

evaluations compared to the linear rule curve. Table 5 shows a comparison of the best values of the rules by GA and LINGO. The best value of the objective function (31.280) using LINGO was 2.27% better than the best value obtained by GA (32.005).

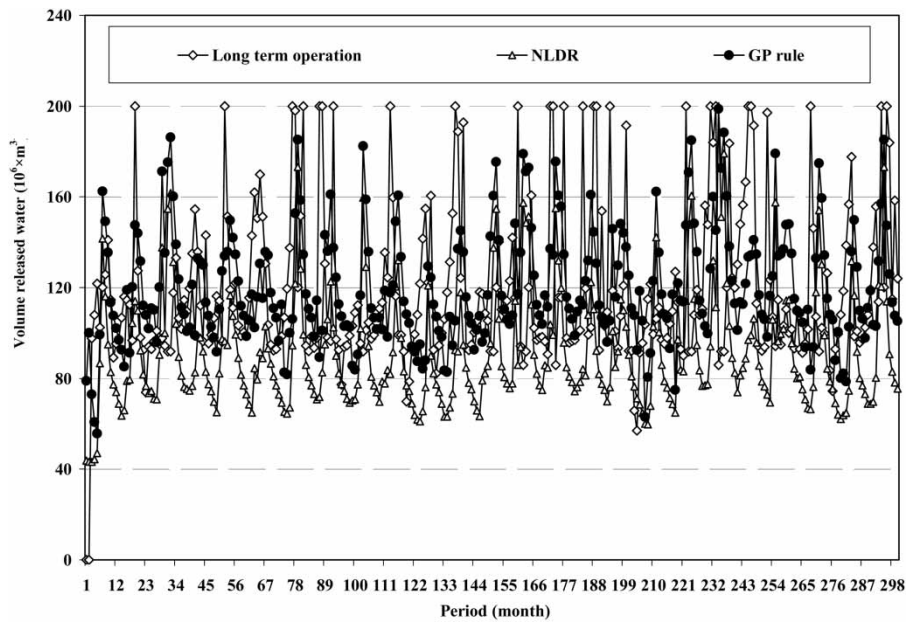


Figure 12 | Volume of released water for the training data in the second case study.

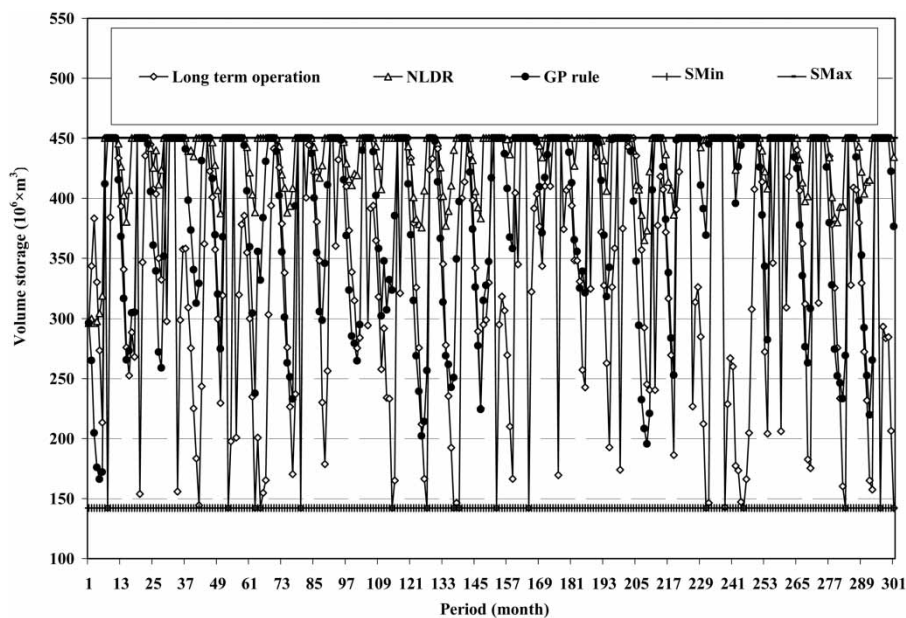


Figure 13 | Volume of storage for the training data in the second case study.

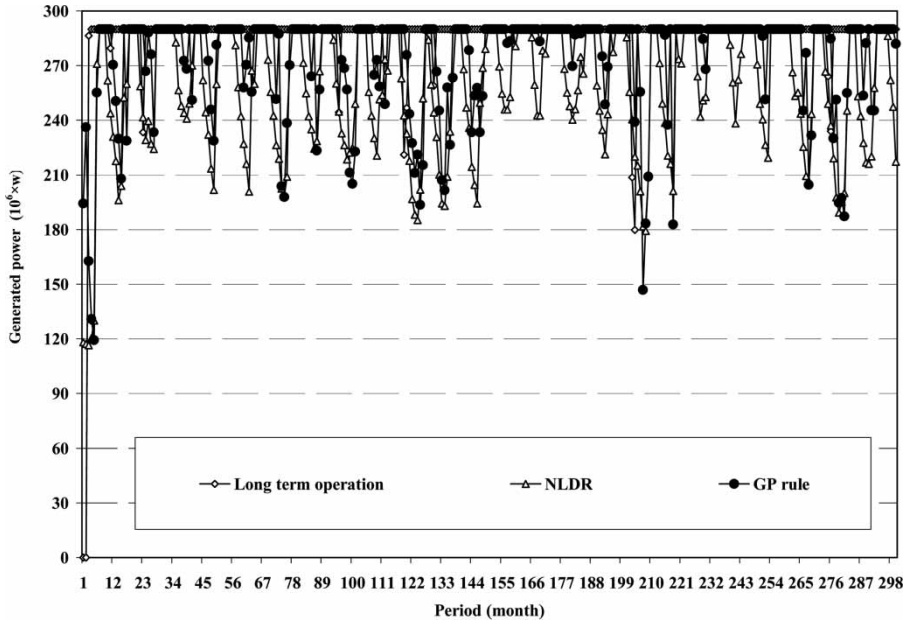


Figure 14 | Generated power for the training data in the second case study.

Developed rule by GP: second case study

As shown in previous sections, the difference between LDR and NLDR results and long-term operation is high. Thus, use of a decision rule that is capable and flexible to derive an operating rule with a smaller difference from a long-term result is the goal of operators. In this subsection, a developed rule by GP is extracted and compared with traditional rules. GP was used to determine optimal rules with the same specifications of the GP employed in the first case study. Table 6 shows the results of five runs and their statistical measures. The best objective function value from GP is 48.03 and 49.20% less than the best results by LINGO and GA. As presented in Table 6, the coefficient

variation of the calculated objective function for five different runs was 0.143, a very small value. Thus, it is possible to achieve the minimum value of the objective function even by one run.

The decreasing trends of the statistical measures for GA and GP are shown in Figure 11. The rule obtained by using GP, having a nonlinear structure with mathematical functions, is as follows:

Released water, storage volume and generated power of the long-term operation without any decision rule, NLDR by LINGO as the best result of the traditional rules and the GP rule are shown in Figures 12–14, respectively. As is shown, the long-term operation just releases water equal to the maximum allowable value in some periods. The GP rule

$$\begin{aligned}
 R_t = & 2(\sqrt{Q_t} + \sqrt{S_t}) \\
 & + \sqrt{\sqrt{2Q_t - S_t + (S_t^2 + S_t Q_t)^{\sin(\sqrt{S_t})}} + \sqrt{(S_t Q_t - Q_t + S_t)^{\sin(\sqrt{Q_t})} + (S_t^3) \sqrt{\sin(Q_t^{S_t})}} + \sqrt{Q_t} + Q_t} \\
 & + \sqrt[4]{2Q_t - S_t + (2S_t^2 + 2S_t Q_t)^{\sin(\sqrt{S_t})}} + \sqrt{2S_t} + \sqrt{Q_t + S_t}
 \end{aligned}
 \tag{18}$$

releases more water in different periods and generates more power in those periods. Thus, the obtained objective function of GP is less than the same value by NLDR.

Finally, the best result of a common predefined pattern rule (NLDR) which was found by LINGO and the nonlinear rule by GP were tested. The objective functions of these two rules were respectively calculated as 23.30 and 13.00 for rule curves by LINGO and GP for 9 years (108 months). This operational period was modeled by LINGO without any rules, yielding an objective function value of 2.84. Figure 15 shows storage volume, released water and generated power for the test data set in the second case study. As is shown, released water from the GP rule used more water from volume storage in most of the operational periods. Thus, the storage volume of the GP rule is less than in the NLDR in the operational periods.

CONCLUDING REMARKS

Application of GP, for developing an optimal existing relation between input and output data in water resources, resulted in a solution with performance measures much better than other data-developed methods. In this paper, GP was used to develop the best relation between inflow, volume storage and release as the operational rule curve for two different operational purposes: supplying downstream demand and generating hydropower energy. Minimization of the total deficit rate, which is the total difference between released water and demand divided by the obtained demand during the planning horizon, and minimization of the total difference between power generation and installed capacity divided by the installed capacity during the planning horizon were respectively considered as the objective functions. Note that the effects of equal water excess or shortage are the same for each period in the supplying downstream demand purpose. The objective function of the supplying downstream demand purpose is the minimization of the total squared deviation of the released water in the planning horizon. Thus, to minimize the objective function and decrease the sum of differences between released water and demand in all periods in the planning horizon, the value of released water is nearly the same as the demand in each period. By using LINGO as LP and NLP

solvers and GA in two case studies, LDR and NLDR were calculated and compared even though the obtained objective function was better (smaller) by LINGO. GA was capable of achieving near-optimal yield rule curves by 3.67% average of difference with LINGO in the LDR and NLDR, respectively.

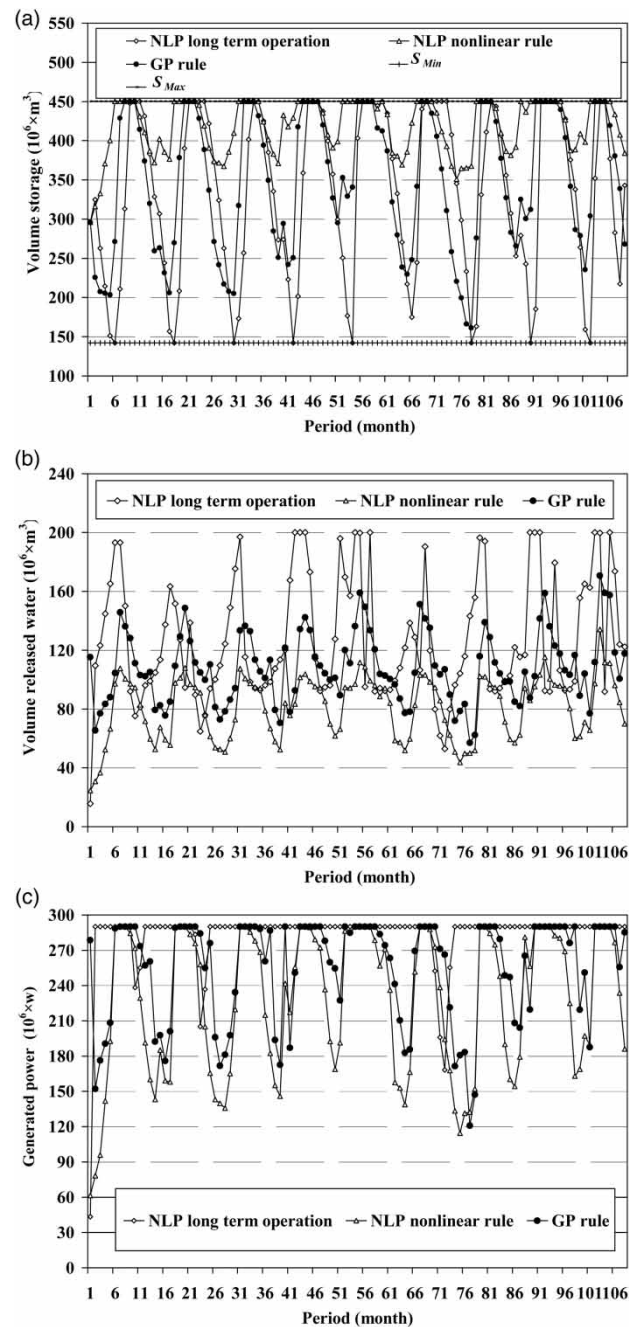


Figure 15 | Volume of (a) storage, (b) released water and (c) generated power for the test data in the second case study.

The rules determined by LINGO and GP were then compared in the training and testing data sets. The GP rule was more effective with a high performance for both of the case studies with the same number of function evaluations. To determine the improvement rate of the obtained objection function by GP, the difference between the best (minimum) calculated objective function by GP and LINGO was divided by the obtained objection function by LINGO. The GP objective functions were respectively 2.48 and 48.03% better (smaller) than those obtained from LINGO for the training data in the first and second case studies. By using GP, testing-data results showed 8.53 and 44.21% improvement for the first and second case studies, respectively. These results indicated that the efficiency of the optimal rule curves for reservoirs improved by the proposed rule based on GP. Application of GP in the optimal operation of a reservoir with other operation purposes, such as flood control, is recommended for future studies.

REFERENCES

- Batista Celeste, A., Suzuki, K. & Kadota, A. 2008 Integrating long- and short-term reservoir operation models via stochastic and deterministic optimization: case study in Japan. *J. Water Res. Plann. Mngmnt.* **134** (5), 440–448.
- Bozorg Haddad, O., Afshar, A. & Mariño, M. A. 2008a Honey-bee mating optimization (HBMO) algorithm in deriving optimal operation rules for reservoirs. *J. Hydroinf.* **10** (3), 257–264.
- Bozorg Haddad, O., Afshar, A. & Mariño, M. A. 2008b Design-operation of multi-hydropower reservoirs: HBMO approach. *Water Res. Mngmnt.* **22** (12), 1709–1722.
- Cai, X., McKinney, D. & Lasdon, L. S. 2001 Solving nonlinear water management models using a combined genetic algorithm and linear programming approach. *Adv. Water Res.* **24** (6), 667–676.
- Chang, F.-J., Chen, L. & Chang, L.-C. 2005a Optimizing the reservoir operating rule curves by genetic algorithms. *Hydrol. Process.* **19** (11), 2277–2289.
- Chang, Y.-T., Chang, L.-C. & Chang, F.-J. 2005b Intelligent control for modeling of real-time reservoir operation, Part II: artificial neural network with operating rule curves. *Hydrol. Process.* **19** (7), 1431–1444.
- Chang, L.-C. & Chang, F.-J. 2001 Intelligent control for modeling of real-time reservoir operation. *Hydrol. Process.* **15** (9), 1621–1634.
- Deb, K. 2001 *Multi-objective Optimization Using Evolutionary Algorithms*. John Wiley & Sons, New York.
- Fallah-Mehdipour, E., Bozorg Haddad, O., Beygi, S. & Mariño, M. A. 2011 Effect of utility function curvature of Young's bargaining method on the design of WDNs. *Water Res. Mngmnt.* **25** (9), 2197–2218.
- Goldberg, D.E. 1989 *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison Wesley, New York.
- Güven, A. & Günel, M. 2008 Genetic programming approach for prediction of local scour downstream of hydraulic structure. *J. Irrig. Drainage Engng.* **134** (2), 241–249.
- Güven, A. & Kisi, O. 2010 Estimation of suspended sediment yield in natural rivers using machine-coded linear genetic programming. *Water Res. Mngmnt.* **25** (2), 691–704.
- Izadifar, Z. & Elshorbagy, A. 2010 Prediction of hourly actual evapotranspiration using neural network, genetic programming, and statistical models. *Hydrol. Process.* **24** (23), 3413–3425.
- Khu, S.T., Liang, S.-Y., Babovic, V., Madsen, H. & Muttill, N. 2001 Genetic programming and its application in real-time runoff forecasting. *J. AWRA* **37** (2), 439–451.
- Kisi, O. & Güven, A. 2010 Evapotranspiration modeling using linear genetic programming technique. *J. Irrig. Drainage Engng.* **136** (10), 715–723.
- Mousavi, S. J., Ponnambalam, K. & Karray, F. 2007 Inferring operating rules for reservoir operations using fuzzy regression and ANFIS. *Fuzzy Sets Syst.* **158** (10), 1064–1082.
- Oliveira, R. & Loucks, D. P. 1997 Operating rules for multireservoir systems. *Water Resour. Res.* **33** (4), 839–852.
- Rabunal, J. R., Puertas, J., Suarez, J. & Rivero, D. 2007 Determination of the unit hydrograph of a typical urban basin genetic programming and artificial neural networks. *Hydrol. Process.* **21** (4), 476–485.
- Reis, L. F. R., Walters, G. A., Savic, D. & Chaudhry, F. H. 2005 Multi-reservoir operation planning using hybrid genetic algorithm and linear programming (GA-LP): an alternative stochastic approach. *Water Res. Mngmnt.* **19** (6), 831–848.
- Savic, D. A., Walters, G. A. & Davidson, J. W. 1999 A genetic programming approach to rainfall-runoff modeling. *Water Res. Mngmnt.* **13** (3), 219–231.
- Sharif, M. & Wardlaw, R. 2000 Multi-reservoir system optimization using genetic algorithms: case study. *J. Comput. Civ. Engng.* **14** (4), 225–263.
- Sivapragasam, C., Maheswaran, R. & Venkatesh, V. 2008 Genetic programming approach for flood routing in natural channels. *Hydrol. Process.* **22** (5), 623–628.
- Sivapragasam, C., Vasudevan, G., Maran, J., Bose, C., Kaza, S. & Ganesh, N. 2009 Modeling evaporation-seepage losses for reservoir water balance in semi-arid regions. *Water Res. Mngmnt.* **23** (5), 853–867.
- Wardlaw, R. & Sharif, M. 1999 Evaluation of genetic algorithm for optimal reservoir system operation. *J. Water Res. Plann. Mngmnt.* **125** (1), 25–33.