A Genetic Algorithm Parallel Strategy for Optimizing the Operation of Reservoirs with Multiple Eco-environmental Objectives

Duan Chen[1,2*], Qiuwen Chen[3], Arturo S. Leon[2], Ruonan Li[4]

[1]Changjiang River Scientific Research Institute, Wuhan 430010, China

[2]Department of Civil and Construction Engineering, Oregon State University, Corvallis, Oregon 97331, United States of America

[3] CEER Nanjing Hydraulic Research Institute, Nanjing 210029, China

[4] Research Center for Eco-environmental Sciences, Chinese Academy of Sciences, Beijing 100081, China

Abstract:

Optimizing the operation of reservoir involving ecological and environmental (eco-environmental) objectives is challenging due to the often competing social-economic objectives. Non-dominated Sorting Genetic Algorithm-II is a popular method for solving multi-objective optimization problems. However, within a complex search space, the NSGA-II population (i.e., a group of candidate solutions) may be trapped in local optima as the population diversity is progressively reduced. This study proposes a computational strategy that operates several parallel populations to maintain the diversity of the candidate solutions. An improved version of the NSGA-II, called c-NSGA-II is implemented by incorporating multiple recombination operators from the Borg-MOEA, a state-of-the-art optimization algorithm that has been recently proposed. The parallel strategy is then coupled into the routine of the c-NSGA-II and applied to the operation of the Qingshitan reservoir (Southwest of China) which includes three eco-environmental and two social-economic objectives. Three metrics (convergence, diversity, and hyper volume index) are used for evaluating the optimization performance. The results show that the proposed parallel strategy significantly improves the solution quality in both convergence and diversity. Two characteristic schemes are identified for the operation of the Qingshitan reservoir for tradeoff between eco-environmental and social-economic objectives.

Keywords: Reservoir Operation; Ecological and Environmental Objectives; NSGA-II; Parallel Strategy

1 Introduction

Maximizing yield of reservoir has been one of the most important aspects for integrated water resource management (White, 1998; Cardwell et al., 2006). In most integrated management of reservoirs, the primary objectives are the maximization of social-economic objectives(e.g., power generation, flood control etc.). However, objectives and constraints involving eco-environmental aspects (e.g., fish community, environmental flow etc.) are being gradually included in reservoir operation due to growing concerns on river ecosystem. Most of the early studies use a constant minimum flow discharge as a fixed constraint (Homa et al., 2005). This “minimum flow” operation has minor impact on the primary social-economic objectives (Castellettl et al., 2008), however this operation is commonly criticized for its limitations on ecosystem benefits (Whiting, 2002; Petts, 2009). Some other studies focus on implementing objectives that involves fish habitat and fish population (Sale et al., 1982; Cardwell et al., 1996) or a time-varying constraint based on dynamic flow requirement of a target fish (Chen et al., 2015; Li, et al., 2015). Recently, a concept based on river system integrity has been further advocated (Richter et al., 1996, 1997, Poff et al., 2010, Yin et al., 2011) instead of giving priority to the fish species. Operation of the reservoir that combines social-economic and eco-environmental objectives, some of which may be highly nonlinear and discontinuous (Dorn and Ranjithan, 2003; Smith et al., 2007), present a complex multi-objective optimization problem within a high dimension solution space that is often non-differentiable and non-convex (Yeh, 1985; Wurbs, 1993; Labadie, 2004). Finding optimal solutions to the problem is challenging and requires a sophisticated optimization algorithm.

Non-dominated Sorting Genetic Algorithm-II (Deb et al., 2002), also known as NSGA-II, is one of the most popular multi-objective evolutionary algorithms (MOEAs). The NSGA-II is recognized for its superior performance in multi-objective optimization and has been receiving increasing attention for practical applications on reservoir operation (Prasad and Park, 2004; Atiquzzaman et al., 2006; Yin et al., 2011). However, premature convergence (i.e., little or no improvement of solutions) is often a problem for the NSGA-II (Leung et al., 1997; Hrstka and Kučerová., 2004). In a complex search space, it is difficult for the algorithm to find better solutions at every iteration (i.e., generation). The 'best' gene in the present generation will duplicate itself repeatedly by recombination operator unless better gene is found. The population of the NSGA-II is likely dominated by the 'best' gene after a few iterations. This premature convergence drastically reduces the diversity of the candidate solutions (i.e., population) and leads to convergence to a local optima. The NSGA-II is found difficult to obtain the global optimal solutions to the optimization problem within complex or high-dimensional search spaces (Hrstka and Kučerová., 2004, Hadka and Reed 2013) e.g., the optimization of reservoir operation that involves multiple eco-environmental objectives.

To maintain the diversity of the candidate solutions in the NSGA-II, the present study proposes a computing strategy that simultaneously operates multiple parallel populations, instead of a single population in the serial NSGA-II. To improve the performance of the optimization, an adaptive version of the original NSGA-II called c-NSGA-II is firstly prepared. The c-NSGA-II employ multiple recombination operators that are from the Borg-MOEA(Hadka and Reed 2013), a recently developed state-of-the-art optimization framework. The proposed parallel strategy is then incorporated into the routine of the c-NSGA-II. We apply both parallel NSGA-II (p-NSGA-II) and the c-NSGA-II to the operation of the Qingshitan Reservoir in the southwest of China, which comprise three eco-environmental and two social-economic objectives. The convergence, diversity and quality of solutions of the two optimization techniques are compared using three indexes which are commonly used in the literatures. The objective of this study is to: (1) develop a state-of-the-art optimization method combining parallel strategy and the c-NSGA-II; and (2) apply the method in the operation of the Qingshitan reservoir to test the performance.

2 Methodology

2.1 c-NSGA-II

Recombination is one of the most important operators/process in the GA that produces child solutions from parent solutions. Hadka and Reed (2013) compared various state-of-the-art multi-objective evolutionary algorithms (MOEA) and developed a unified optimization framework called Borg-MOEA to overcome critical issues of the MOEA such as dominance resistance and deterioration of the optimization algorithm. One of the major features of the Borg MOEA is the use of multiple recombination operators to generate offspring from the parent population.

The c-NSGA-II is developed by incorporating the multiple recombination operators from the Borg-MOEA (Hadka and Reed, 2013) on the basis of the NSGA-II (Deb et al., 2002). Including the original simulated binary crossover (SBX, Deb and Agrawal, 1994), four other popular recombination operators are incorporated, i.e., Differential Evolution operator (DE, Storn and Price, 1997), Parent-Centric Crossover (PCX, Deb et al., 2011), Unimodal Normal Distribution Crossover (UNDX, Kita et al., 1999) and Simplex Crossover (SPX, Tsutsui et al., 1999). The c-NSGA-II uses primary principles of the original NSGA-II by mimicking evolution process of genes using selection, recombination and mutation operators. The procedure of the c-NSGA-II optimization is briefly described below (see also Figure 1). For a multi-objective problem (MOP), a set of non-dominated solutions is obtained according to the concept of non-dominance. A solution is called Pareto-optimal when there is no solution that would improve at least one objective function value without worsening at least one other objective function value. The set of all the Pareto-optimal solutions of the MOP is then defined as Pareto-optimal front or Pareto front.

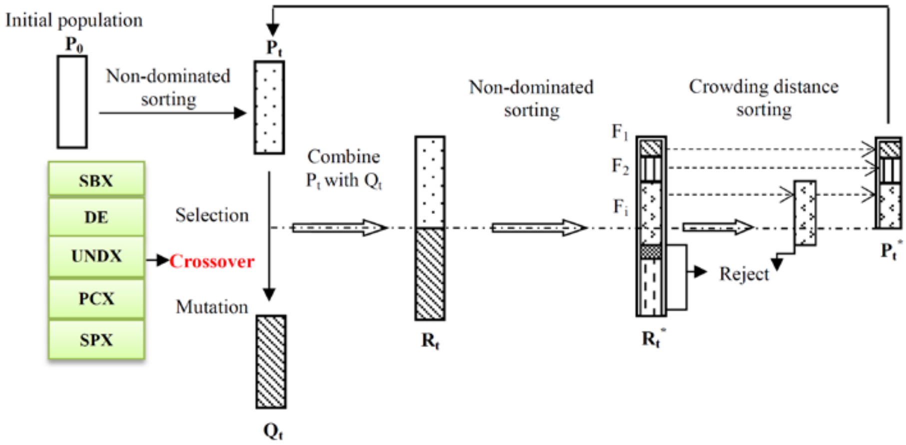

Figure 1. Major process of the c-NSGA-II optimization (Adapted from Deb et al. 2002)

2.2 p-NSGA-II

Master-slave model (MS), multi-population model (MP) and diffusion model (DM) are three major parallel paradigms that are often implemented in MOEAs. However, their implementation varies significantly in the literature. It is more common to use the MS and MP paradigm for small scale clusters i.e., less than 50 processors. The MS is probably the easiest implementation for paralleling MOEAs. For the MS paradigm, the master processor stores the populations and performs all critical computational processes (selection, recombination and mutation). The slave processor is used only to evaluate fitness functions and return objective values. It is usual to assign approximately equal workload to all the slaves. The MS paradigm has only one parameter (i.e., the number of processors) and can be easily implemented. However, the relatively large computational costs of communication are often criticized (Cantu-Paz, 2000; Coello et al., 2001). Moreover, the MS do not improve serial algorithm’s performance due to its parallelized scheme.

The MP, often referred to as “island model”, is another popular parallel paradigm. The MP generates multiple subpopulations and distributes them on different processors. Each processor starts with a separate population and exchanges information with each other. This exchange is called migration and requires the specification of frequency of migration, numbers of the migrants as well as allocations of the migrants (Cantu-Paz, 2007). Early studies showed that the MP has a potential of 'superlinear' speed-up (Coello Coello et al., 2002), however they require a sophisticated design and careful tuning of the aforementioned parameters (Tang et al., 2006; Cantu-Paz, 2007).

In the present study, a parallel strategy that combines the MP and MS paradigms is incorporated in the c-NSGA-II and form a parallel method called p-NSGA-II. The main steps of the p-NSGA-II are described as follows (also in Figure 2). First, a pre-determined number of subgroups (N_g) with the same population size (P_g) are allocated to CPU processors. Each of these subgroups randomly generates the initial population. Each subgroup simultaneously employs the c-NSGA-II to evolve populations until the predetermined numbers of generations (Gg) are reached or a certain number of feasible solutions are found. A gene pool i.e., the combined group (P_c) is set up to store the optimized solutions from each subgroup. Only a given ratio (R_m) of the solutions from each subgroup is allowed to migrate to the gene pool for maintaining the diversity of the P_c . The c-NSGA-II is employed again for the P_c to continue the optimization until the stopping criteria is satisfied. The stopping criteria used in this study is the number of generations (G_c) for the combined group.

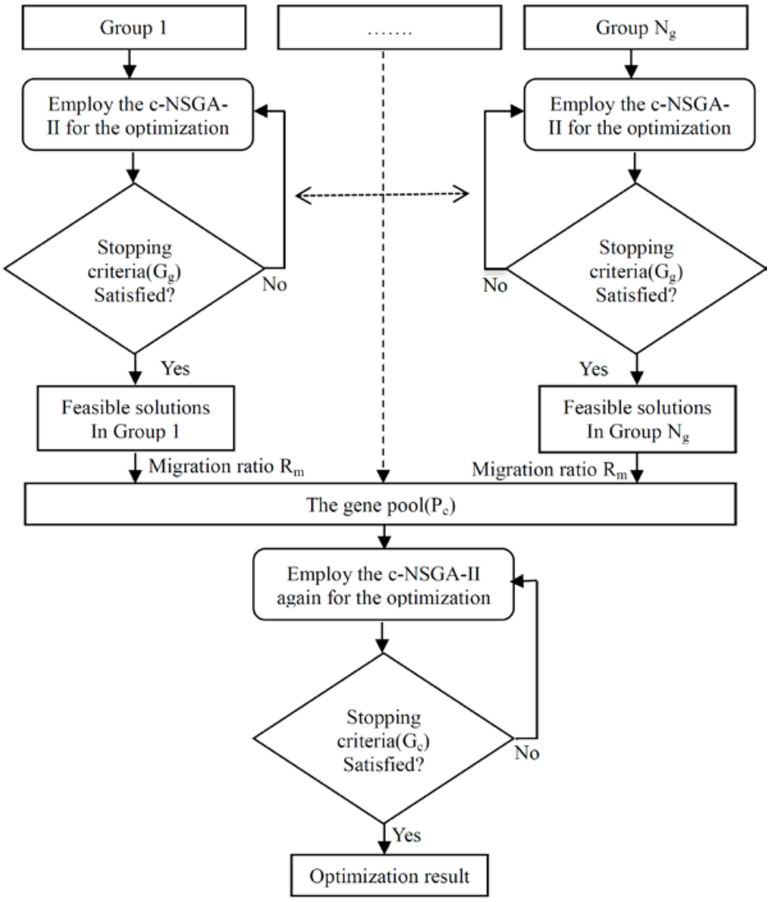

Figure 2. Flow chart of the parallel strategy for the c-NSGA-II (p-NSGA-II)

As can be observed in Figure 2, the proposed parallel strategy is a combination of the MP and MS models. This strategy has multiple subpopulations at the beginning to allow for more diversity. Then the optimization evolves in a similar way to the MP model. Unlike in the MP model where the information exchange is between the subpopulations, in the proposed strategy the information exchange is between the subpopulations and the gene pool only. The information exchange is similar to the MS model, which reduces the intensive communication costs of the MP model. Compared to the c-NSGA-II, only two additional parameters, namely, number of subgroups (N_g) and migration ratio (R_m), need to be specified in the p-NSGA-II. P_c is determined by N_g and R_m and is computed within the optimization process.

2.3 Selected performance metrics

For a multi-objective optimization, preferred results include: 1) convergence to the true Pareto-optimal front, and 2) diversity of solutions in the Pareto-optimal front (Deb, et al., 2002).

2.3.1 Convergence index

There are various metrics to quantify the convergence of MOEAs, such as error ratio (Veldhuizen, 1999), set coverage metric (Zltzler, 1999), etc. Herein a commonly-used generational distance metric (Veldhuizen, 1999, Deb et al., 2002) is adopted as the convergence index. This metric measures Euclidean Distance in the objective space between two sets of solutions. The generational distance index (C_{index}) between two sets of solutions is given by

$$C_{index} = \frac{\sum_{i=1}^{n} (\min_{j=1}^{n^*} \sqrt{\sum_{m=1}^{M} (f_m^{(i)} - f_m^{*(j)})^2})}{n} \tag{1}$$

where n and n^* are the number of solutions in these two sets, respectively, and these numbers are not necessarily the same. In Equation (1), M is the number of objectives; $f_m^{(i)}$ and $f_m^{*(j)}$ are the *m-th* objective function value of the *i-th* and *j-th* member of the two sets. The smaller C_{index} represents the better convergence toward the Pareto-optimal front.

2.3.2 Diversity index

Besides convergence, spread of the solutions is also important for decision making. It is often desirable to find solutions that span the entire Pareto-optimal region. Herein the metric proposed by Deb et al. (2001, 2002) is used for quantifying the diversity of the solutions. This metric is given by:

$$D_{index} = \frac{d_f + d_l + \sum_{i=1}^{N-1} \left| d_i - \overline{d} \right|}{d_f + d_l + (N-1) \cdot \overline{d}} \tag{2}$$

where D_{index} is the diversity index. N is the number of the solutions in the obtained nondominated set. d_f and d_l are the Euclidean distances between the extreme solutions and the boundary solutions, respectively. d_i is the Euclidean distance between consecutive solutions in the obtained nondominated set. $\overline{d}$ is the average of all the distance d_i, $i=1,2,\ldots,N-1$. Similarly as Convergence index, the smaller diversity index represents better diversity of solutions.

2.3.3 Hypervolume index

Hypervolume indicator is found to be a good metric for evaluating the performance of multi-objective optimization (Zitzler et al., 2003; Reed et al., 2013) due to its property of combining the convergence and diversity metrics into a single index. The hypervolume indicator basically measures the volume of objective space covered by a set of non-dominated solutions. The hypervolume indicator (I_h) (Zitzler et al., 2003; Reed et al., 2013) is defined as:

$$I_h(A) = \int_{\underbrace{(0,0,\ldots,0)}_{n}}^{\overbrace{(1,1,\ldots,1)}^{n}} \alpha_A(Z)\,dz \tag{3}$$

where A is any objective vector set, Z is the space (0,1), n for the normalized objectives (n = 5 in our test case), $\alpha_A(Z)$ is the attainment function which will have a value of 1 if A is a weakly dominated solution set in Z. The hypervolume indicator calculates the volume of the objective space enclosed by the attainment function and the axes. A higher hypervolume indicator denotes better quality of the solutions in terms of convergence and diversity.

3. Case study

The Qingshitan reservoir is a major hydraulic facility of the Lijiang River basin in the southwest of China and provides water for multiple purposes. It is located at the upstream of the Gantang River, the largest tributary of the Lijiang River. The reservoir releases water through turbines for hydropower production and through a bypass channel for irrigation. Due to soil drainage, part of the irrigation flow returns to the Gantang River through irrigation channels. The total flow immediately downstream of the Qingshitan reservoir joins the Lijiang

River at Guilin city, which was primarily used for domestic water consumption. The location of the case study and the schematic of water distribution in the Qingshitan Reservoir are presented in Figure 3.

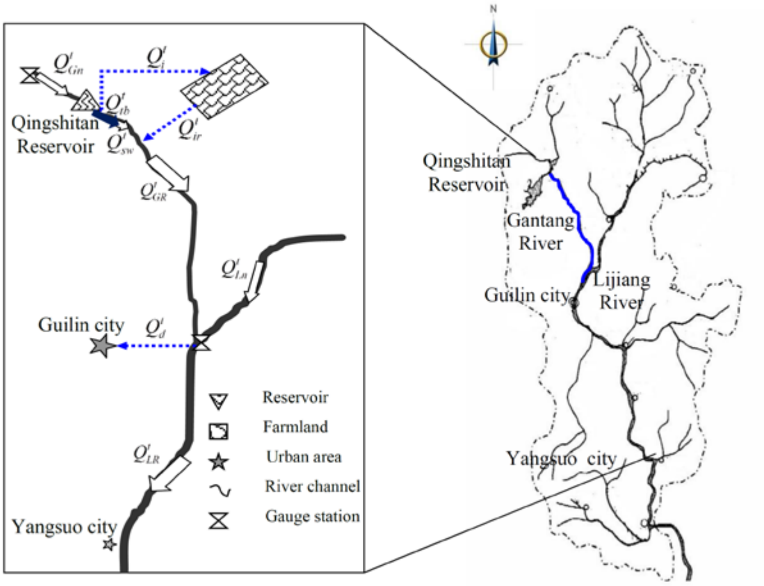

Figure 3. Location of case study and schematic of water distribution in the Qingshitan Reservoir. Q^t_{Gn} and Q^t_{Ln} denote the upstream natural flow of the Gantang River and the Lijiang River at time *t*, respectively; Q^t_i denotes the operational flow for irrigation at time t; Q^t_{tb} and Q^t_{sw} denotes the operational flow through turbines and spillways at time *t*, respectively; Q^t_{ir} denotes the returned flow from irrigation at time *t* (due to soil drainage); Q^t_d denotes the flow for domestic use in the Guilin city at time *t*; Q^t_{GR} and Q^t_{LR} denote the total flow of the Gantang and Lijiang Rivers at time *t*, respectively.

The Qingshitan reservoir used to be operated only for agricultural irrigation, power generation and domestic water supply for the city of Guilin. Recently, the Lijiang River has been receiving increasing attention from tourists due to its unique karstic landscape. For this recreational purpose, the reservoir needs to discharge a certain flow for maintaining navigation of cruise boats in the Lijiang River. On the other hand, a few previous studies showed that the present operational scheme of the Qingshitan reservoir has strongly negative impacts on the aquatic habitat (Ye et al., 2011, Li et al., 2011). Therefore, new operational schemes for the Qingshitan reservoir are being developed for including additional objectives. Based on a previous study on the Qingshitan reservoir (Chen et al., 2012), minimizing natural flow alteration in the two rivers are proposed as ecological objectives to restore the ecosystem in the Gantang and Lijiang rivers. Maintaining an acceptable water quality in the Lijiang River, by conveying a minimum amount of water in this river is another added objective, which is both environmental and social-economic preferred.

The proposed objectives and constraints of the reservoir operation are described in Sections 3.1 and 3.2. Since the reservoir is annually regulated, the optimization period is chosen to be one year. The daily flow releases for hydropower production, irrigation and domestic water supply are the decision variables.

3.1 Objective function

(1) Maximize general water supply (GWS)

This objective is composed of three sub-objectives, namely agricultural irrigation, domestic water supply and tourism cruise navigation. These three sub-objectives are all expected to be maximized in the optimization. The objective function is expressed as follows:

$$\max(GWS) = \max\frac{1}{n}\cdot\sum_{t=1}^{n}\left(w_1\cdot\min(\frac{Q_i^t}{Q_{i\_tar}^t},1) + w_2\cdot\min(\frac{Q_{tb}^t + Q_{sw}^t + Q_{ir}^t + Q_{Ln}^t}{Q_{d\_tar}^t},1) + w_3\cdot\min(\frac{Q_{tb}^t + Q_{sw}^t + Q_{Ln}^t - Q_d^t}{Q_{w\_tar}^t},1)\right) \quad (4)$$

where GWS is average rate of general water supply during the optimization period; n is the total number of time steps; w_1, w_2 and w_3 are weights of the irrigation, water supply and navigation objectives, respectively. According to “the Operation Handbook of the Qingshitan Reservoir”, these three objectives are equally important; therefore, all weights have a value of 1/3; $Q_{i\_tar}^t$, $Q_{d\_tar}^t$ and $Q_{w\_tar}^t$ (m^3/s) are discharge requirements at time t for irrigation, domestic water supply and recreation (cruise navigation), respectively. Other notations were explained in Figure 3.

(2) Maximize power generation (PG)

$$\max PG = \max\sum_{t=1}^{n}\left(\eta\times g\times\overline{H_r^t}\times Q_{tb}^t\times\Delta t\right)/PG_{\max} \quad (5)$$

where PG is the percentage to the $PG_{\max}$ (kWh), which is the maximum electricity (i.e., generated at full capacity) during the optimization period; η is efficiency of turbine; g is the gravitational acceleration (m/s^2). $\overline{H_r^t}$ is average water head (m) at time t; $Q_{d\max}^t$ (m^3/s) is allowed maximum discharge through turbines; Δt is the operational time of the turbine during each time step.

(3) Minimize flow alteration in the Gantang River (FAGR)

$$\min FAGR = \min\frac{1}{n}\cdot\sum_{t=1}^{n}\left(\frac{\left|Q_{tb}^t + Q_{sw}^t + Q_{ir}^t - Q_{Gn}^t\right|}{Q_{Gn}^t}\right) \quad (6)$$

where *FAGR* is average flow alteration rate of the Gantang River from its natural flow;

(4) Minimize the flow alteration in the Lijiang River (FALR)

$$\min FALR = \min\frac{1}{n}\cdot\sum_{t=1}^{n}\left(\frac{\left|(Q_{tb}^t + Q_{sw}^t + Q_{ir}^t - Q_d^t + Q_{Ln}^t) - (Q_{Gn}^t + Q_{Ln}^t)\right|}{Q_{Ln}^t + Q_{Gn}^t}\right) \quad (7)$$

where *FALR* is average flow alteration rate of the Lijiang River from its natural flow

(5) Maximize Water Quality in the Lijiang River (WQ)

$$\max WQ = \max\frac{1}{n}\cdot\sum_{t=1}^{n}\left((Q_{tb}^t + Q_{sw}^t + Q_{ir}^t - Q_d^t + Q_{Ln}^t)/Q_{wq}^t\right) \quad (8)$$

where *WQ* is water quality index in the Lijiang River and Q_{wq}^t is the required discharge to maintain an acceptable water quality in the Lijiang River.

3.2 Constraints

The optimization model satisfies the mass balance equation (or continuity equation) and constraints as follows:

(1) Continuity equation

$$V^{t+1} - V^t = (Q_{Gn}^t - Q_f^t - Q_i^t)\cdot\Delta t \quad (9)$$

where V^{t+1} and V^t are reservoir storage at the t and t+1 time. There is no tributary between the Gantang River and the Lijiang River. In the present numerical tests, water leakage and evaporation losses are not considered.

(2) Reservoir water level constraints

$$H_{r\_\min} \le H_r^t \le H_{r\_\max} \quad (10)$$

where H_r^t is the reservoir water level at time t; $H_{r\_\min}$ and $H_{r\_\max}$ are allowed minimum and maximum reservoir water levels, respectively. The maximum reservoir water level becomes normal water surface elevation (WSE) during flood season.

(3) Irrigation discharge constraints

$$0 \le Q_i^t \le Q_{i\_allow}^t \tag{11}$$

where $Q_{i\_allow}^t$ is allowed maximum discharge in the irrigation channel at time *t*.

(4) Domestic water supply discharge constraints

$$Q_{d\_minium}^t \le Q_d^t \le Q_{d\_allow}^t \tag{12}$$

where $Q_{d\_minium}^t$ is minimum daily water consumption, determined by the lowest record in a year. $Q_{d\_allow}^t$ is allowed maximum discharge in the water supply channel at time *t*.

(5) Turbine flow constraints

$$0 \le Q_{tb}^t \le Q_{tb\_\max} \tag{13}$$

where $Q_{tb\_\max}$ is allowed maximum discharge of Q_{tb}^t;

(6) spill flow constraints

$$0 \le Q_{sw}^t \le Q_{sw\_\max}^t \tag{14}$$

where $Q_{sw\_\max}^t$ is allowed maximum discharge of Q_{sw}^t at time *t*.

(7) Output constraints

$$N_{d\_\min} \le N_d^t \le N_{d\_\max} \tag{15}$$

where N_d^t is power output at time *t*. $N_{d\_\min}$ and $N_{d\_\max}$ are firm output and output capacity, respectively.

3.3 Parameters setting

The c-NSGA-II and p-NSGA-II are both applied to the operation of the Qingshitan Reservoir. For the c-NSGA-II, the population is set to 320. The number of generations is set to a relatively large number (4000) to investigate the convergence of the solutions. Other parameters e.g., mutation rate are set to the same typical values used in the literature (Deb et al., 2002; Yandamuri et al., 2006; Yin et al., 2011; Sindhya et al., 2011).

For the p-NSGA-II, three experiments are conducted with different numbers of parallel groups. The number of subgroups (*Ng*) is set to 8 in Experiment I, which means that 8 processors are used for the parallel computation. The population for each subgroup (P_g)is set to 80. The population of the combined group (P_c) is determined by the product $Ng \times P_g \times R_m$. For matching the population of the c-NSGA-II and the combined group of the p-NSGA-II, we intentionally set R_m to 0.5 in the first experiment. Likewise, for experiments II (Ng = 16) and III (Ng = 32), R_m is set to 0.25 and 0.125, respectively. For all the three experiments, the number of generations for the subgroups (G_g) is set to 3000. The number of generations for the combined group (P_c) is set to 1000, which results in a total of 4000 generations, the same as the number of generations for the c-NSGA-II.

Because of the random nature of Genetic Algorithms, optimization results may have some difference for different runs, like other random-based search algorithms. Therefore, 30 random-seed replicate runs are used in each experiment and the average value as well as the corresponding variance is reported.

4 Results

4.1 Results of NSGA-II and p-NSGA-II

The performance of the c-NSGA-II and p-NSGA-II with different subgroups are compared for the three aforementioned metrics. The convergence and diversity indexes are calculated at each generation, whose results are presented in Figure 4. For the c-NSGA-II, the results of the two indexes are based on a constant population of 320. For the p-NSGA-II, the convergence and diversity indexes for the first 3000 generations are average values of each subgroup (e.g., Ng = 8), which is 80 populations. The results for the last 1000 generations are based on the combined populations, which is of 320, the same size as the population of the c-NSGA-II.

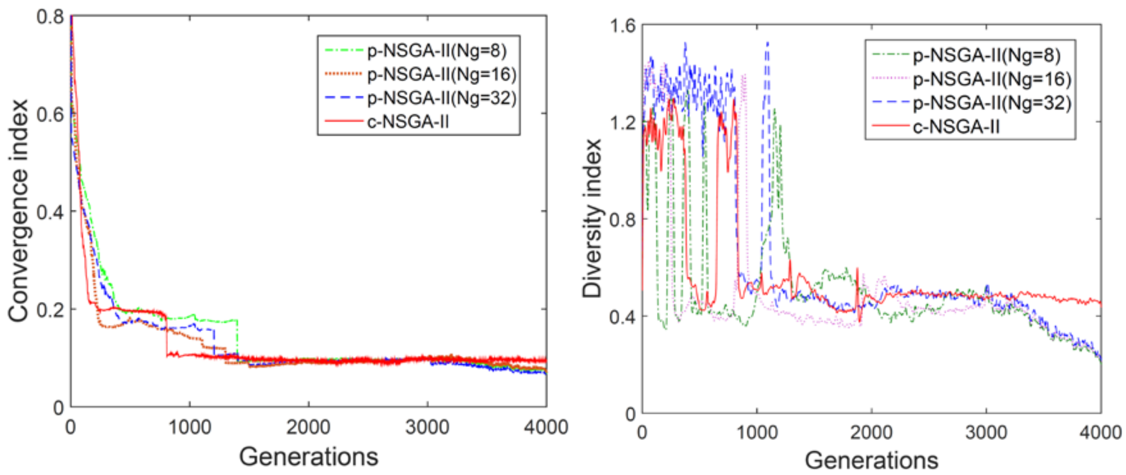

Figure 4. Convergence index (left) and diversity index (right) as a function of generation for the c-NSGA-II and p-NSGA-II with different parallel groups

During the first few hundred generations (e.g., 500), the convergence index for both the c-NSGA-II and p-NSGA-II decreases rapidly. Between 500 and 1500 generations, there is a discontinuity in the convergence rate for the c-NSGA-II and p-NSGA-II. After this discontinuity all experiments converge to approximately the same value (i.e., 0.12). For generations between 1500 and 3000, the convergence index for the c-NSGA-II and p-NSGA-II oscillates slightly around 0.12. For generations larger than 3000, the convergence index for the c-NSGA-II still oscillates slightly around 0.12. However, the convergence index for the p-NSGA-II is reduced gradually after the parallel groups are combined. Between the generations 3000 and 4000, the convergence index for the p-NSGA-II is reduced about 30% (from 0.12 to about 0.076). Compared to the convergence index that decreases more or less monotonically, the diversity index presents large oscillations especially during the first 1500 generations. For generations between 1500 and 3000, the diversity index oscillates slightly around 0.41 for the c-NSGA-II and p-NSGA-II. For generations larger than 3000, the diversity index for the c-NSGA-II still oscillates slightly around 0.41. However, the diversity index for the p-NSGA-II is reduced more than 50% (from 0.48 to about 0.21) after the parallel groups are combined.

The convergence, diversity at the last generation i.e., the final non-dominated solutions, is presented in Table 1. The hypervolume index is also included in the table. The mean value and variance for each experiment is shown in bold and in parentheses, respectively. As observed in Table 1, the convergence and diversity indexes of the p-NSGA-II are smaller than that of the c-NSGA-II, which means better performance on the two indexes. The results of Table 1 also show that the hypervolume index of the p-NSGA-II is higher than that of the c-NSGA-II The hypervolume value of the p-NSGA-II increases nearly 10% than the one of the c-NSGA-II, meaning better quality of the solutions from the p-NSGA-II. On the other hand, the different experiments of the p-NSGA-II show similar results on the performance indexes. The p-NSGA-II with 16 subgroups i.e., Ng=16 obtained slightly better mean values on convergence and hypervolume. The best diversity index i.e., smallest value is obtained for Ng=32.

Table 1. Performance index of the NSGA-II and p-NSGA-II with different subgroups

Method	Convergence index	Diversity index	Hypervolume index
c-NSGA-II	**0.1168** (0.0029)	**0.4816** (0.023)	**0.3085** (0.0018)
p-NSGA-II (Ng=8)	**0.0778** (0.0031)	**0.2142** (0.017)	**0.3359** (0.0021)
p-NSGA-II (Ng=16)	**0.0748** (0.0018)	**0.2128** (0.021)	**0.3384** (0.0016)
p-NSGA-II (Ng=32)	**0.0752** (0.0017)	**0.2125** (0.025)	**0.3367** (0.0019)

4.2 Optimal operational solutions for the Qingshitan Reservoir

Unlike some mathematical functions with theoretically Pareto-optimal solutions, there is no "known" true Pareto-optimal front for the real-world reservoir operation. Hence, a reference set of solutions is assumed to be the “true” Pareto-optimal front. The reference set is generated by applying the non-dominated sorting to the combined best solutions from all optimization runs (Kollat et al., 2008). This reference set is displayed in Figure 5 as a parallel line plot, which is a common way of showing results for more than three objectives (Fu et al., 2012). The Pareto solutions for the c-NSGA-II and p-NSGA-II (N_g=16) at the last generation are also presented in Figure 5. Each line in Figure 5 represents a solution for the five objectives and the arrow shows the direction in which the objective function is improving. Compared to the c-NSGA-II (Figure 5, left), the lines of the p-NSGA-II (Figure 5, right) cover a wider range for each objective, meaning that the solutions are more spread. The results of p-NSGA-II are also found more similar with the reference set (Figure 5, middle).

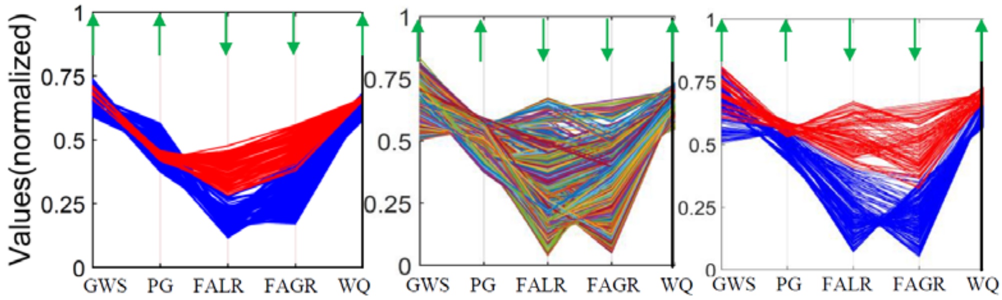

Figure 5. Parallel line plot for optimal solutions for the c-NSGA-II (left), p-NSGA-II (right) and reference set (middle)

The lines in Figure 5 also represent trade-off between the different objectives. For low-dimension objectives such as bi-objective, the trade-off between the objectives is easy to be recognized because of the one-to-one competing relation. For high dimension objectives, the trade-off are more complex due to the many-to-many competing relations. To identify the trade-off relations between the five objectives, a cluster analysis (Duran & Odell, 2013), which group similar objects, are applied to cluster the lines for both the c-NSGA-II and p-NSGA-II (N_g=16). Two different solution groups (A and B) are identified through the cluster analysis. In general, Scheme A (red line) has lower values than Scheme B (blue line) for all the five objectives for the c-NSGA-II and p-NSGA-II (N_g = 16). As can be observed in Figure 5, the result of the p-NSGA-II is more obvious in such grouping relations.

The range (lower and upper bounds) of the objective values can be also compared from Figure 5. All the objective values are normalized (i.e., bounded between 0 and 1) for a fair comparison. For all objectives, larger ranges are observed for the p-NSGA-II (N_g=16) compared to the c-NSGA-II. The comparison also show that the p-NSGA-II (N_g=16) solutions are closer to the reference set than those of the c-NSGA-II.

5 Discussions

Overall, the population and generation parameters influence the performance of Genetic Algorithms. In general, the larger the population and number of generations, the better the quality of the solutions (De Jong, 2007). For a fair comparison between the c-NSGA-II and p-NSGA-II experiments, the generation number is the same (i.e., 4000). For the population, the p-NSGA-II has smaller population (i.e., 80 in each subgroup) than the population of the c-NSGA-II (i.e., 320) during the first 3000 generations. However, after 3000 generations (i.e., the last 1000 generations), the populations (in the combined group) of the p-NSGA-II are the same as the c-NSGA-II. As can be observed in Figure 4, the c-NSGA-II converges faster than the p-NSGA-II at the beginning. The result is expected because a larger random population has a higher probability for achieving better solutions. However, after an initial fast convergence, the c-NSGA-II failed to improve its performance in terms of convergence and diversity. This is the so-called premature convergence which results from some "super genes" dominating the search space (Leung et al., 1997; Hrstka and Kučerová., 2004). In a premature convergence, the population is trapped in local optima. Premature convergence is less of an issue for the p-NSGA-II as this method gains diversity after the parallel subgroups are combined (at 3000 generations in this case study). It can be noted in Figure 4, that the performance of the p-NSGA-II, in terms of convergence and diversity, improves after the parallel subgroups are combined (i.e., at 3000 generations). Also, as shown in Table 1, compared to the c-NSGA-II, all the p-NSGA-II experiments achieved lower values for the convergence and diversity indexes, and a slightly higher value for the hypervolume index. These comparisons indicate that the p-NSGA-II has a better performance than the c-NSGA-II.

Among the different experiments of the p-NSGA-II, the p-NSGA-II with 16 subgroups obtain the smallest value of the convergence index (0.0748), meaning the best performance in convergence. On the other hand, the p-NSGA-II with 32 subgroups achieve the smallest value of the diversity index (0.2125), meaning the best performance in diversity. Enhancing convergence and maintaining diversity of the solutions are dual-goals for designing a sophisticated algorithm for multi-objective optimization (Deb, et al., 2002). The two goals are difficult to achieve at the same time because of the complex interactions between convergence and diversity in MOEA (Laumanns & Deb, 2001; Goel & Stander, 2010. In our case study, the p-NSGA-II with an intermediate number of parallel groups (i.e., Ng=16) appears to be the best choice as it achieves the largest hypervolume index (0.3384). It should be noted, however, that the differences of the results for the three experiments of the p-NSGA-II are very small (Table 1).

The results in Table 2 show that the p-NSGA-II with Ng=16 has a better spread of solutions than the c-NSGA-II for most of the objectives. It is worth mentioning that a better spread of solutions is desirable because it provides more flexibility in the reservoir operation. In addition, most of p-NSGA-II results are closer to the reference set indicating a better accuracy. Through the aforementioned cluster analysis two solutions groups are identified (i.e., Schemes A and Schemes B). These groups show different characteristics on balancing the five objectives of the reservoir operation. The solutions in Schemes A are found to have less deviation from the natural flow regime for both rivers (smaller values of the FAGR and FALR objectives), however, less water supply and power generation (smaller values of the GWS and PG objectives). This is because the FAGR and FALR objectives favour a natural flow regime (e.g., less flow regulation by the reservoir). As expected, the solutions for Schemes A have a lower water quality (smaller values of WQ) because of the lesser flow regulation compared to the solutions for the Schemes B. On the contrary, the Schemes B has better solutions for social-economic interests (i.e., larger values for the GWS and PG objectives) at expenses of the eco-environmental objectives. These results show strong conflicts between the social-economic and eco-environmental objectives.

6 Conclusions

Growing concerns on river ecosystem and environment increases the complexity of reservoir operation and challenges the optimization algorithms. This study shows that a parallel strategy which employs multiple parallel groups of populations help to improve the performance of the NSGA-II. This improvement is due to the increase of diversity provided by the parallel groups. However, the performance does not improve monotonically with the number of parallel groups. In the case of the Qingshitan reservoir operation, 16 parallel groups achieve the best overall performance. The parallel strategy is developed based on the combination of two commonly-used parallel computing paradigms, i.e., the master-slave and the island models. The implementation of the proposed strategy is straightforward and can be easily incorporated into other random-based search algorithms.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (51109012, 51479188), the Chinese Ministry of Water Resources special founds on public causes (201101005), and the Bonneville Power Administration through cooperative agreement # TIP258.

Reference

Atiquzzaman, M., S. Liong, and X. Yu (2006), Alternative Decision Making in Water Distribution Network with NSGA-II, *J. Water Resour. Plann. Manage., 132(2)*, 122-126.

Cantu-Paz E. (2000), Efficient and accurate parallel genetic algorithms. Norwell, MA: Kluwer Academic Publishers

Cantu-Paz E. (2007), Parallel setting in parallel genetic. F. Lobo, C. Lima, Z. Michalewicz (Eds.), Parameter Setting in Evolutionary Algorithms. Studies in Computational Intelligence, Springer , pp. 259–276.

Cardwell, H., H.I. Jager, and M.J. Sale (1996), Designing instream flows to satisfy fish and human water needs, *J.Water Res. PL.-ASCE 122(5).*,356-363.

Cardwell, H., R. Cole, L. Cartwright, and L. Martin(2006), Integrated water resource management: Definitions and conceptual musings, *J. Contemp. Water Res. Educ.*, 135, 8-18.

Castelletti, A., F. Pianosi, R. and Soncini-Sessa(2008), Water reservoir control under economic, social and environmental constraints, *Automatica, 44*: 1595-1607.

Chen, D., Li, R., Chen, Q., & Cai, D. (2015). Deriving Optimal Daily Reservoir Operation Scheme with Consideration of Downstream Ecological Hydrograph Through A Time-Nested Approach. Water Resources Management, 1-16.Chen, Q., D. Chen, et al (2013), Adapting the operation of two cascaded reservoirs for ecological flow requirement of a de-watered river channel due to diversion-type hydropower stations, *Ecol. Model.*,252(3),266-272, dx.doi.org/10.1016/j.ecolmodel.2012.03.008.

Chen, Q., D. Chen, et al.,(2012), Optimizing the operation of the Qingshitan reservoir in the Lijiang River for multiple human interests and quasi-natural flow maintenance, *Journal of Environmental Sciences.,24(11),* doi:10.1016/S1001-0742(11)61029-2.

Chen, Q., Yang, Q., Lin, Y(2011), Development and application of a hybrid model to analyze spatial distribution of macro-invertebrates under flow regulation in the Lijiang River. Ecological Informatics, 6: 407-413.

Coello Coello C, Pulido GT(2001), Multiobjective optimization using a micro-genetic algorithm. In: Spector L et al., editors. Proceedings of the genetic and evolutionary computation conference (GECCO 2001). San Francisco, CA: Morgan Kaufmann Publishers;. p.274–82.

Coello Coello C, Van Veldhuizen DA, Lamont GB (2002), Evolutionary algorithms for solving multi-objective problems. New York, NY: Kluwer Academic Publishers;.

De Jong K(2007), Parameter setting in EAs: a 30 year perspective. In: Lobo F G, Lima G F, Michalewicz Z, eds. Parameter Setting in Evolutionary Algorithms. Berlin: Springer-Verlag, 1-18.

Deb K, Zope P, Jain A(2002), Distributed computing of pareto-optimal solutions using multi-objective evolutionary algorithms, Technical report KanGAL Report No. 2002008, Indian Institute of Technology.

Deb, K. and Agrawal, R. B. (1994). Simulated binary crossover for continuous search space. Technical Report IITK/ME/SMD-94027, Indian Institute of Technology, Kanpur, UP, India.

Deb, K. (2001), Multi-Objective Optimization Using Evolutionary Algorithms, Wiley.

Deb, K., & Jain, H. (2011).Self-Adaptive Parent to Mean-Centric Recombination for Real-Parameter Optimization. Indian Institute of Technology Kanpur, Tech. Rep.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan(2002), A fast and elitist multi-objective genetic algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation 6 (2)*, 182-197.

Dorn, J.L., and S. Ranjithan(2003), Evolutionary Multiobjective Optimization in Watershed Water Quality Management, in: Evolutionary Multi-Criteria Optimization, *Lecture Notes in Computer Science* (LNCS), vol. 2632, Springer-Verlag, Berlin, pp. 692-706.

Duran, B. S., & Odell, P. L. (2013). Cluster analysis: a survey (Vol. 100). Springer Science & Business Media.

Fu, G., Kapelan, Z., Kasprzyk, J. R., & Reed, P. (2012). Optimal design of water distribution systems using many-objective visual analytics. Journal of Water Resources Planning and Management.

Goel, T. S., & Stander, N. (2010). A study of convergence characteristics of multi-objective evolutionary algorithms. In Proceedings of the 13th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Paper AIAA-2010 (Vol. 9233, pp. 13-15).

Hadka, D., & Reed, P. (2013). Borg: An auto-adaptive many-objective evolutionary computing framework. Evolutionary Computation, 21(2), 231-259.

Homa, E.S., R.M. Vogel, M.P. Smith, C.D. Apse, A. Huber-Lee, and J. Sieber (2005), An Optimization Approach for Balancing Human and Ecological Flow Needs. In: EWRI 2005: Impacts of Global Climate Change, *Proceedings of the 2005 World Water and Environmental Resources Congress Anchorage*, Alaska.

Hrstka, O., & Kučerová, A. (2004). Improvements of real coded genetic algorithms based on differential operators preventing premature convergence. Advances in Engineering Software, 35(3), 237-246.

Jager, H.I., and K.A. Rose(2003), Designing optimal flow patterns for fall Chinook salmon in a Central Valley river, *N. Am. J. Fish Manage. ,23:* 1-21.

Kita, H., Ono, I., and Kobayashi, S. (1999). Multi-parental extension of the unimodal normal distribution crossover for real-coded genetic algorithms. In Congress on Evolutionary Computation, pages

Kollat, J. B., Reed, P. M. & Kasprzyk, J. R(2008), A new epsilon-dominance hierarchical Bayesian optimization algorithm for large multiobjective monitoring network design problems. Advances in Water Resources, 31, 828-845.

Labadie, J.W.(2004), Optimal operation of multi-reservoir systems: State-of-the-art review, *J.Water Resour. Plann. Manage., 130(2)*: 93-111.

Laumanns, M., & Deb, K. (2001). On the convergence and diversity-preservation properties of multi-objective evolutionary algorithms. Eidgenössische Technische Hochschule Zürich (ETH), Institut für Technische Informatik und Kommunikationsnetze (TIK).

Leung, Y., Gao, Y., & Xu, Z. B. (1997). Degree of population diversity-a perspective on premature convergence in genetic algorithms and its markov chain analysis. Neural Networks, IEEE Transactions on, 8(5), 1165-1176.

Li, R., Chen, Q., Ye, F.(2011), Modelling the impacts of reservoir operations on the downstream riparian vegetation and fish habitats in the Lijiang River. Journal of Hydroinformatics, 13(2): 229-244

Li, R., Q. Chen, and D. Chen(2012), Ecological hydrograph based on Schiozothorax chongi habitat conservation in the dewatered river channel between Jinping cascaded dams, Science China (Technological Sciences), doi:10.1007/s11431-011-4614-7.

Li, R., Chen, Q., Tonina, D., & Cai, D. (2015). Effects of upstream reservoir regulation on the hydrological regime and fish habitats of the Lijiang River, China. Ecological Engineering, 76, 75-83.Petts, G.E.(2009), Instream flow science for sustainable river management, J.Water Resour.Asssoc.,45:1071-1076.

Poff, N. L., Richter, B. D., Arthington, A. H., Bunn, S. E., Naiman, R. J., Kendy, E., ... & Warner, A. (2010). The ecological limits of hydrologic alteration (ELOHA): a new framework for developing regional environmental flow standards. Freshwater Biology, 55(1), 147-170.

Prasad, T. and N. Park (2004), Multiobjective Genetic Algorithms for Design of Water Distribution Networks, J. Water Resour. Plann. Manage., 130(1), 73-82.

Reed P, Yamaguchi S(2004), Making it easier to use the multiple population parallelization scheme for evolutionary algorithms. In: Sehlke J et al.,editors. 2004 EWRI world water and environmental resources congress. Salt Lake City, Utah: ASCE.

Reed, P.M., Hadka, D., Herman, J., Kasprzyk, J., and Kollat, J.(2013), "Evolutionary Multiobjective Optimization in Water Resources: The Past, Present, and Future." (Editor Invited Submission to 35th Anniversary Special Issue), Advances in Water Resources, v51, 438-456.

Richter, B.D., J.V. Baumgartner, J. Powell, and D.P. Braun(1996), A method for assessing hydrologic alteration within ecosystems. *Conserv. Biol., 10*: 1163-1174.

Richter, B.D., J.V. Baumgartner, R. Wigington, and D.P. Braun (1997), How much water does a river need? *Freshwater Biol., 37*: 231-249.

Sale, M. J., E. D. Brill, and E. E. Herricks(1982), An approach to optimizing reservoir operation for downstream aquatic resources ,*Water Resources Research,* 18:4.

Shiau, J.T., and F.C.Wu(2004), Feasible diversion and instream flow release using range of variability approach, *J. Water Res. Pl.-ASCE.130(5)*: 395-404.
Shiau, J.T., and F.C.Wu(2006), Compromise programming methodology for determining instream flow under multiobjective water allocation criteria, *J. Am. Water Resour. As. 42(5)*: 1179-1191.
Sindhya, Karthik and Deb, Kalyanmoy and Miettinen, Kaisa(2011), Improving convergence of evolutionary multi-objective optimization with local search: a concurrent-hybrid algorithm, Natural Computing,10(4), 10.1007/s11047-011-9250-4.
Smith, B.T., H.I. Jager, M.J. Sale(2007), *Optimization of Hydropower Resources: Current Practices and Opportunities for Improvement*, ORNL/TM-2006/91,Oak Ridge National Laboratory.
Storn, R. and Price, K. (1997). Differential evolution — a simple and efficient heuristic for global optimization over continuous spaces. Journal of Global Optimization, 11(4):341–359.
Suen, J. P., and J.W. Eheart(2006), Reservoir Management to Balance Ecosystem and Human Needs: Incorporating the Paradigm of the Ecological Flow Regime, *Water Resour. Res. 42*,W03417.
Tsutsui, S., Yamamura, M., and Higuchi, T. (1999). Multi-parent recombination with simplex crossover in real coded genetic algorithms. In Genetic and Evolutionary Computation Conference(GECCO 2999).
Van Veldhuizen DA, Zydallis JB, Lambont GB(2003). Considerations in engineering parallel multiobjective evolutionary algorithms. IEEETrans Eolut Comput,7(2):144–73.
White, G.F.(1998),Reflections on the 50-year International Search for Integrated Water Management, *Water Policy., 1(1)*:21-27.
Whiting, P.J.(2002), Stream flow necessary for environmental maintenance, *Annu Rev Earth Pl. Sc., 30*,181-206.
Wurbs, R.A.(1993), Reservoir-System simulation and optimization models, *J. Water Resour. Plann. Manage. 119(4)*: 455-472.
Yandamuri, S., K. Srinivasan, and S. M. Bhallamudi(2006), Multiobjective Optimal Waste Load Allocation Models for Rivers Using Nondominated Sorting Genetic Algorithm-II.” *J. Water Resour. Plann. Manage.*
Ye, F., Chen, Q., Li, R.(2010), Modelling the riparian vegetation evolution due to flow regulation of Lijiang River by unstructured cellular automata. Ecological informatics, 5: 108-114
Yeh, W.(1985), Reservoir management and operation models: A state-of-the-art review, *Water Resour. Res., l21*: 1797–1818.
Yin, X., and Z. Yang(2011), Development of a coupled reservoir operation and water diversion model: Balancing human and environmental flow requirements, *Ecol. Model. 222(2)*: 224-231.
Zitzler, E., L. Thiele , M. Laumanns , C. M. Fonseca and V. G. da Fonseca(2003), "Performance assessment of multiobjective optimizers: An analysis and review", IEEE Trans. Evol. Comput., vol. 7, no. 2, pp.117 -132.