Improving multi-objective reservoir operation optimization with sensitivity-informed problem decomposition

Jinggang Chu[1], Chi Zhang[1], Guangtao Fu[2], Yu Li[1], Huicheng Zhou[1]

[1] School of Hydraulic Engineering, Dalian University of Technology, Dalian 116024, China

[2] Centre for Water Systems, College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Harrison Building, Exeter EX4 4QF, UK

*Corresponding author:

Dr. Chi Zhang

School of Hydraulic Engineering, Dalian University of Technology

Dalian 116024, China

Tel.: 86-411-8470-8517

Fax: 86-411-8470-8517

E-mail: czhang@dlut.edu.cn

Abstract

This study investigates the effectiveness of a sensitivity-informed method for multi-objective operation of reservoir systems, which uses global sensitivity analysis as a screening tool to reduce the computational demands. Sobol′'s method is used to screen insensitive decision variables and guide the formulation of the optimization problems with a significantly reduced number of decision variables. This sensitivity-informed problem decomposition dramatically reduces the computational demands required for attaining high quality approximations of optimal tradeoff relationships between conflicting design objectives. The search results obtained from the reduced complexity multi-objective reservoir operation problems are then used to pre-condition the full search of the original optimization problem. In two case studies, the Dahuofang reservoir and the inter-basin multi-reservoir system in Liaoning province, China, sensitivity analysis results show that reservoir performance is strongly controlled by a small proportion of decision variables. Sensitivity-informed problem decomposition and pre-conditioning are evaluated in their ability to improve the efficiency and effectiveness of multi-objective evolutionary optimization. Overall, this study illustrates the efficiency and effectiveness of the sensitivity-informed method and the use of global sensitivity analysis to inform problem decomposition when solving the complex multi-objective reservoir operation problems.

Keywords water supply; complexity reduction; multi-objective optimization; preconditioning; sensitivity analysis; reservoir operation

1 Introduction

Reservoirs are often operated considering a number of conflicting objectives (such as different water uses) related to environmental, economic and public services. The optimization of Reservoir Operation Systems (ROS) has attracted substantial attention over the past several decades. In China and many other countries, reservoirs are operated according to reservoir operation rule curves which are established at the planning/design stage to provide long-term operation guidelines for reservoir management to meet expected water demands. Reservoir operation rule curves usually consist of a series of storage volumes or levels at different periods (Liu et al., 2011a and 2011b). For the optimal ROS problem, the values of storage volumes or levels are optimized to achieve one or more objectives. Quite often, there are multiple curves, related to different purposes of reservoir operation. The dimension of a ROS problem depends on the number of the curves and the number of time periods. For a cascaded reservoir system, the dimension can be very large, which increases the complexity and problem difficulty and poses a significant challenge for most search tools currently available (Labadie, 2004; Draper and Lund, 2004; Sadegh et al., 2010; Zhao et al., 2014).

In the context of multi-objective optimal operation of ROS, there is not one single operating policy that improves simultaneously all the objectives and a set of non-dominating Pareto optimal solutions are normally obtained. The traditional approach to multi-objective optimal reservoir operation is to reformulate the multi-objective problem as a single objective problem through the use of some scalarization methods, such as the weighted sum method (Tu et al., 2003 and 2008; Shiau, 2011). This method has been developed to repeatedly solve the single objective problem using different sets of weights so that a set of Pareto-optimal solutions to the original multi-objective problem could be obtained (Srinivasan and Philipose, 1998; Shiau and Lee, 2005). Another well-known method is the ε-constraint method (Ko et al., 1997; Mousavi and Ramamurthy, 2000; Shirangi et al., 2008): all the objectives but one are converted into constraints and the level of satisfaction of the constraints is optimized to obtain a set of Pareto-optimal solutions. However, with the increase in problem complexity (i.e., the number of objectives or decision variables), both approaches become inefficient and ineffective in deriving the Pareto-optimal solutions.

In the last several decades, bio-inspired algorithms and tools have been developed to directly solve multi-objective optimization problems by simultaneously handling all the objectives (Nicklow et al., 2010). In particular, multi-objective evolutionary algorithms (MOEA) have been increasingly applied to the optimal reservoir operation problems, with intent of revealing tradeoff relationships between conflicting objectives. Suen and Eheart (2006) used the non-dominated sorting genetic algorithm (NSGAII) to find the Pareto set of operating rules that provides decision makers with the optimal trade-off between human demands and ecological flow requirements. Zhang et al. (2013b) used a multi-objective adaptive differential evolution combined with chaotic neuron networks to provide optimal trade-offs for multi-objective long-term reservoir operation problems, balancing hydropower operation and the requirement of reservoir ecological environment. Chang et al. (2013) used an adjustable particle swarm optimization – genetic algorithm (PSO-GA) hybrid algorithm to minimize water shortages and maximize hydro-power production in management of Tao River water resources.

However, significant challenges remain for using MOEAs in large, real-world ROS applications. The high dimensionality of ROS problems makes it very difficult for MOEAs to identify ‘optimal or near optimal’ solutions with the computing resources that are typically available in practice. Thus the primary aim of this study is to investigate the effectiveness of a sensitivity-informed optimization methodology for multi-objective reservoir operation, which uses sensitivity analysis results to reduce the dimension of the optimization problems, and thus improves the search efficiency in solving these problems. This framework is based on the previous study by Fu et al. (2012), which developed a problem decomposition framework that can dramatically reduce the computational demands required to obtain high quality solutions for optimal design of water distribution systems. The ROS case studies used to demonstrate this framework consider the optimal design of reservoir water supply operation policies. Storage volumes at different time periods on the operation rule curves are used as decision variables. It has been widely recognized that the determination of these decision variables requires a balance among different ROS objectives. Sobol'’s sensitivity analysis results are used to form simplified optimization problems considering a small number of sensitive decision variables, which can be solved with a dramatically reduced number of model evaluations to obtain Pareto approximate solutions. These Pareto approximate solutions are then used to pre-condition a full search by serving as starting points for the multi-objective evolutionary algorithm. The results from the Dahuofang reservoir and inter-basin multi-reservoir system case studies in Liaoning province, China, whose conflicting objectives are minimization of industry water shortage and minimization of agriculture water shortage, illustrate that sensitivity-informed problem decomposition and pre-conditioning provide clear advantages to solve large-scale multi-objective ROS problems effectively.

2 Problem formulation

Most reservoirs in China are operated according to rule curves, i.e., reservoir water supply operation rule curves. Because they are based on actual water storage volumes, they are simple to use. Fig. 1 shows typical water supply operation rule curves from Dahuofang reservoir based on 36 10-day periods.

As we know that water demand could be fully satisfied only when there is sufficient water in reservoir. Water supply operation rule curve, which is used to operate most reservoirs in China, represents the limited storage volume for water supply in each period of a year. In detail, water demand will be fully satisfied when the reservoir storage volume is higher than water supply operation rule curve, whereas water demand need to be rationed when the reservoir storage volume is lower than water supply operation rule curve. In general, a reservoir has more than one water supply target, and there is one to one correspondence between water supply rule curve and water supply target. The water supply with lower priority will be limited prior to the water supply with higher priority when the reservoir storage volume is lower. To reflect the phenomenon that different water demands can have different reliability requirements and different levels of priority in practice, the operation rule curve for the water supply with the lower priority is located above the operation rule curve for the water supply with the higher priority.

Fig. 1 shows water supply operation rule curves for agriculture and industry where the maximum storage is smaller in the middle due to the flood control requirements in wet seasons. In Fig. 1, the red line with circle represent water supply rule curve for agriculture, the green line with triangle represent water supply rule curve for industry, and the water supply rule curve for agriculture with lower priority is located above the water supply rule curve for industry with higher priority. The water storage available between the minimum and maximum storages is divided into three parts: zone 1, zone 2 and zone 3 by the water supply rule curves for agriculture and industry.

Specifically, both the agricultural demand D_1 and the industrial demand D_2 could be fully supplied when the actual water storage is in zone 1, which is above the water supply rule curve for agriculture; when the actual water storage is in zone 2, which is above the water supply rule curve for industry and below the water supply rule curve for agriculture, the industrial demand D_2 could be fully supplied, and the agricultural demand D_1 has to be rationed; both the agricultural demand D_1 and the industrial demand D_2 have to be rationed when the actual water storage is in zone 3, which is below the water supply rule curve for industry. The water supply rule for a specific water user consists of one water supply rule curve and rationing factors that indicate the reliability and priority of the water user. Assuming that the specified water rationing factor α_1 is applied to the water supply rule curve for agriculture in Fig. 1, the agricultural demand D_1 could be fully supplied without rationing when the actual water storage is in zone 1, however, when the water storage is in zone 2 or zone 3, the agricultural demand has to be rationed, i.e., $\alpha_1 * D_1$. Similarly, assuming that the specified water rationing factor α_2 is applied to the water supply rule curve for industry in Fig. 1, the industrial demand D_2 could be fully supplied without rationing when the actual water storage is in zone 1 or zone 2, however, when the water storage is in zone 3, the industrial demand has to be rationed, i.e., $\alpha_2 * D_2$.

Because it could be assumed that the historical inflow into the reservoir would be repeated in the future, to provide long-term operation guidelines for reservoir management to meet expected water demands in a future planning year, the water demands in the future planning year and long-term historical inflow are used. The optimization objectives for water supply operation rule curves are to minimize water shortages during the long-term historical period. The ROS design problem is formulated as a multi-objective optimization problem, i.e., minimizing multiple objectives simultaneously. In this paper, the objectives are to minimize industry and agriculture water shortages:

$$\min f_i(\boldsymbol{x}) = SI_i = \frac{100}{N}\sum_{j=1}^{N}\left(\frac{D_{i,j}-W_{i,j}(\boldsymbol{x})}{D_{i,j}}\right)^2 \quad (1)$$

where $\boldsymbol{x}$ is the vector of decision variables, i.e., the water storages at different periods on a water-supply rule curve; SI_i is the shortage index for water demand i

(industrial water demand when i = 1, agricultural water demand when i = 2), which measures the frequency and magnitude of annual shortages occurred during N years, and is used as an indicator to reflect water supply efficiency; N is the total number of years simulated; $D_{i,j}$ is the sum of target demands for water demand i during the jth year; $W_{i,j}(x)$ is the sum of delivered water for water demand i during the jth year.

For the ROS optimization problem, the mass balance equations are:

$$S_{t+1} - S_t = I_t - R_t - SU_t - E_t \tag{2}$$

$$R_t = g(\boldsymbol{x}),\ SU_t = k(\boldsymbol{x}), E_t = e(\boldsymbol{x}) \tag{3}$$

$$ST_t^{\min} \leq S_t \leq ST_t^{\max}, ST_t^{\min} \leq \boldsymbol{x} \leq ST_t^{\max} \tag{4}$$

where S_t is the initial water storage at the beginning of period t; S_{t+1} is the ending water storage at the end of period t; I_t, R_t, SU_t and E_t are inflow, delivery for water use, spill and evapotranspiration loss, respectively; and $ST_t^{\max}$ and $ST_t^{\min}$ are the maximum and minimum storage, respectively. Additionally, because $W_{i,j}(x)$ in Equation (1) is the sum of delivered water for water demand i during the jth year, the sum value of R during the jth year equals to $W_{1,j}(x) + W_{2,j}(x)$.

3 Methodology

Pre-conditioning is a technique that uses a set of known good solutions as starting points to improve the search process of optimization problems (Nicklow et al., 2010). It is very challenging in determining good initial solutions, and different techniques including the domain knowledge can be used. This study utilizes a sensitivity-informed problem decomposition to develop simpler search problems that consider only a small number of highly sensitive decisions. The results from these simplified search problems can be used to successively pre-condition search for larger, more complex formulations of ROS design problems. The ε-NSGAII, a popular multi-objective evolutionary algorithm, is chosen as it has been shown effective for many engineering optimization problems (Kollat and Reed, 2006; Tang et al., 2006; Kollat and Reed, 2007). For the two-objectives (ε_{SI_1} and ε_{SI_2}) considered in this paper, their epsilon values in ε-NSGAII were chosen based on reasonable and practical requirements and were both set to 0.01. According to the study by Fu et al. (2012), the sensitivity-informed methodology, as shown in Fig. 2, has the following steps:

1. Perform a sensitivity analysis using Sobol''s method to calculate the sensitivity indices of all decision variables regarding the ROS performance measure;

2. Define a simplified problem that considers only the most sensitive decision variables by imposing a user specified threshold (or classification) of sensitivity;

3. Solve the simplified problem using ε-NSGAII with a small number of model simulations;

4. Solve the original problem using ε-NSGAII with the Pareto optimal solutions from the simplified problem fed into the initial population.

3.1 Sobol''s sensitivity analysis

Sobol''s method was chosen for sensitivity analysis because it can provide a detailed description of how individual variables and their interactions impact model performance (Tang et al., 2007b; Zhang et al., 2013a). A model could be represented in the following functional form:

$$y = f(\boldsymbol{x}) = f\left(x_1, \cdots, x_p\right) \tag{5}$$

where y is the goodness-of-fit metric of model output, and $\boldsymbol{x} = \left(x_1, \cdots, x_p\right)$ is the parameter set. Sobol''s method is a variance based method, in which the total variance of model output, $D(y)$, is decomposed into component variances from individual variables and their interactions:

$$D(y) = \sum_i D_i + \sum_{i<j} D_{ij} + \sum_{i<j<k} D_{ijk} + \cdots + D_{12\cdots m} \tag{6}$$

where D_i is the amount of variance due to the ith variable x_i, and D_{ij} is the amount of variance from the interaction between x_i and x_j. The model sensitivity resulting from each variable can be measured using the Sobol''s sensitivity indices of different orders:

$$\text{First-order index: } S_i = \frac{D_i}{D} \tag{7}$$

$$\text{Second-order index: } S_{ij} = \frac{D_{ij}}{D} \tag{8}$$

$$\text{Total-order index: } S_{Ti} = 1 - \frac{D_{\sim i}}{D} \tag{9}$$

where $D_{\sim i}$ is the amount of variance from all the variables except for x_i, the first-order index S_i measures the sensitivity from the main effect of x_i, the second-order index S_{ij} measures the sensitivity resulting from the interactions between x_i and x_j, and the total-order index S_{Ti} represents the main effect of x_i and its interactions with all the other variables.

3.2 Performance metrics

Since MOEA search is stochastic, performance metrics are used in this study to compare the quality of the approximation sets derived from replicate multi-objective evolutionary algorithm runs. Three indicators were selected: the generational distance (Veldhuizen and Lamont, 1998), the additive ε-indicator (Zitzler et al., 2003), and the hypervolume indicator (Zitzler and Thiele, 1998).

The generational distance measures the average Euclidean distance from solutions in an approximation set to the nearest solution in the reference set, and indicates perfect performance with zero. The additive ε-indicator measures the smallest distance that a solution set need to be translated to completely dominate the reference set. Again, smaller values of this indicator are desirable as this indicates a closer approximation to the reference set.

The hypervolume indicator, also known as the S metric or the Lebesgue measure, measures the size of the region of objective space dominated by a set of solutions. The hypervolume not only indicates the closeness of the solutions to the optimal set, but also captures the spread of the solutions over the objective space. The indicator is normally calculated as the volume difference between a solution set derived from an optimization algorithm and a reference solution set. In this study, the worst case solution is chosen as reference. For example, the worst solution is (1, 1) for two minimization objectives in the normalized objective space. Thus larger hypervolume indicator values indicate improved solution quality and imply a larger distance from the worst solution.

4 Case study

Two case studies of increasing complexity are used to demonstrate the advantages of the sensitivity-informed methodology: (1) the Dahuofang reservoir, and (2) the inter-basin multi-reservoir system in Liaoning province, China. The inter-basin multi-reservoir system test case is a more complex ROS problem with Dahuofang, Guanyinge and Shenwo reservoirs. In the two ROS problems, the reference sets were obtained from all the Pareto optimal solutions across a total of 10 random seed trials, each of which was run for a maximum number of function evaluations (NFE) of 500,000. Additionally, the industrial and agricultural water demands in the future planning year, i.e., 2030, and the historical inflow from 1956 to 2006 were used to optimize reservoir operation and meet future expected water demands in the two case studies.

4.1 Dahuofang reservoir

The Dahuofang reservoir is located in the main stream of Hun River, in Liaoning province, Northeast China. The Dahuofang reservoir basin drains an area of 5437km^2, and within the basin the total length of Hun River is approximately 169km. The main purposes of the Dahuofang reservoir are industrial water supply and agricultural water supply to central cities in Liaoning province. The reservoir characteristics and yearly average inflow are illustrated in Table 1.

The Dahuofang ROS problem is formulated as follows: the objectives are minimization of industrial shortage index and minimization of agricultural shortage index as described in Equation (1); the decision variables include storage volumes on the industrial and agricultural curves. For the industrial curve, a year is divided into

24 time periods (with ten days as scheduling horizon from April to September, and a month as scheduling horizon in the remaining months). Thus there are twenty-four decision variables for industrial water supply. The agricultural water supply occurs only in the periods from the second ten-day of April to the first ten-day of September, thus there are fifteen decision variables for agricultural water supply. In total, there are thirty-nine decision variables.

4.2 Inter-basin multi-reservoir system

As shown in Fig. 3, Dahuofang, Guanyinge and Shenwo reservoirs compose the inter-basin multi-reservoir system in Liaoning province, China.

Liaoning province in China covers an area of 1.46×10^5 km^2 with an extremely uneven distribution of rainfall in space. The average amount of annual precipitation decreases from 1100 mm in east to 600 mm in west (WMR-PRC, 2008). However, the population, industries, and agricultural areas mainly concentrate in the western parts. Therefore, it is critical to develop the best water supply rules for the inter-basin multi-reservoir system to decrease the risk of water shortages caused by the mismatch of water supplies and water demands in both water deficit regions and water surplus regions. Developing inter-basin multi-reservoir water supply operation rules has been promoted as a long-term strategy for Liaoning province to meet the increasing water demands in water shortage areas. In the inter-basin multi-reservoir system of Liaoning province, the abundant water in Dahuofang, Guanyinge and Shenwo reservoirs is diverted downstream to meet the water demands in water shortage areas, especially the region between Daliaohekou and Sanhekou hydrological stations.

The main purposes of the inter-basin multi-reservoir system are industrial water supply and agricultural water supply to eight cities (Shenyang, Fushun, Anshan, Liaoyang, Panjin, Yingkou, Benxi and Dalian) of Liaoning province, and environmental water demands need to be satisfied fully. The characteristics of each reservoir in the inter-basin multi-reservoir system are illustrated in Table 2.

The flood season runs from July to September, during which the inflow takes up a large part of the annual inflow. The active storage capacities of Dahuofang and Shenwo reservoirs reduce significantly during flood season for the flood control.

The inter-basin multi-reservoir operation system problem is formulated as follows: the objectives are minimization of industrial shortage index and minimization of agricultural shortage index as described in Equation (1). Regarding Shenwo reservoir, which has the same water supply operation rule curve features as Dahuofang reservoir, the decision variables include storage volumes on the industrial and agricultural curves and there are thirty-nine decision variables. Regarding Guanyinge reservoir, the decision variables include storage volumes on the industrial curve and water transferring curve due to the requirement of exporting water from Guanyinge reservoir to Shenwo reservoir in the inter-basin multi-reservoir system, which is similar to the water supply operation rule curve for industrial water demand, and there are forty-eight decision variables. Therefore, the inter-basin multi-reservoir system has six rule curves and $39 \times 2 + 48 = 126$ decision variables in total.

5 Results and discussions

5.1 Dahuofang reservoir

In the Dahuofang reservoir case study, a set of 2000 Latin Hypercube samples were used per decision variable yielding a total number of $2000 \times (39 + 2) = 82000$ model simulations used to compute Sobol''s indices. Following the recommendations of Tang et al. (2007a, b) boot-strapping the Sobol'' indices showed that 2000 samples per decision variable were sufficient to attain stable rankings of global sensitivity.

The first-order indices representing the individual contributions of each variable to the variance of the objectives are shown in blue in Fig. 4. The total-order indices representing individual and interactive impacts on the variance of the objectives are represented by the total height of bars. Agr4_2 represents decision variable responding to water storage volume on the agricultural curve at the second ten days of April and ind3_3 represents decision variable responding to water storage volume on the industrial curve at the last ten days of March, and so on. Considering the shortage index for the industrial water demand, the water storages at time periods ind1, ind2, ind3, ind10, ind11, and ind12, i.e., the water storages at time periods 1, 2, 3, 10, 11, and 12 of water supply operation rule curves for industrial water demand are the most sensitive variables, accounting for almost 100% of the total variance. However, the interactive effects from variables are not noticeable due to the characteristics of industrial water supply and the influences of rules for industrial water demand. Considering the agricultural shortage index, the water storages at time periods from agr4-2 to agr5-3, i.e., the water storages at the first five time periods of water supply operation rule curves for agricultural water demand are the most sensitive variables. This is explained by the characteristics of agricultural water supply and the influences of water supply operation rule curves for agricultural water demand, implying that the interactive effects from variables are noticeable because the agricultural water supply is limited in the whole year if the agricultural water supply in one time period is limited and the largest agricultural water demand occurs in the second and last ten days of May.

5.1.1 Simplified problems

Building on the sensitivity results shown in Fig. 4, one simplified version of the Dahuofang ROS problem is formulated: only 11-periods are considered for optimization, i.e., time periods ind1, ind2, ind3, ind10, ind11, and ind12 for industrial curve and agr4-2, agr4-3, agr5-1, agr5-2, and agr5-3 for agricultural curve based on a total-order Sobol''s index threshold of greater than 10%. The threshold is subjective and its ease-of-satisfaction decreases with increasing numbers of parameters or parameter interactions. In all of the results for the Sobol''s method, parameters classified as the most sensitive contribute, on average, at least 10 percent of the overall model variance (Tang et al., 2007a, b). The full search 39-period problem serves as the performance baseline relative to the reduced complexity problem.

5.1.2 Pre-conditioned optimization

In this section, the pre-conditioning methodology is demonstrated using the 11-period simplification of the Dahuofang ROS test case from the prior section, while the insensitive decision variables are set randomly first and kept constant during the solution of the simplified problem.

Using the sensitivity-informed methodology, the 11-period case was first solved using ε-NSGAII with a maximum NFE of 2000, and the Pareto optimal solutions combined with the constant insensitive decision variables were then used as starting points to start a complete new search with a maximum NFE of 498,000. The standard search using ε-NSGAII was set to a maximum NFE of 500,000 so that the two methods have the same NFE used for search. In this case, 10 random seed trials were used given the computing resources available. The search traces in Fig. 5 show for all three metrics (generational distance, additive epsilon indicator, and hypervolume) that the complexity-reduced case can reliably approximate their portions of the industrial and agricultural water shortage tradeoff given their dramatically reduced search periods. All three metrics show diminishing returns at the end of the reduced search periods. The pre-conditioning results are shown in Fig. 5 in red search traces continuing from the blue reduced complexity search results.

Fig. 5 clearly highlight that the sensitivity-informed pre-condition problems dramatically enhance search efficiency in terms of the generational distance, additive epsilon indicator, and hypervolume metrics. Overall, sensitivity-informed problem decomposition and pre-conditioning yield strong efficiency gains and more reliable search (i.e., narrower band widths on search traces) for the Dahuofang ROS test case.

Fig. 6(a) shows Pareto fronts from a NFE of 3000, 5000 and 8000 in the evolution process of one random seed trial. In the case of the pre-conditioned search, the solutions from 3000, 5000 and 8000 evaluations are much better than the corresponding solutions in the case of standard baseline search. The results show that the Pareto approximate front of the pre-conditioned search is much wider than that of the standard search, and clearly dominates that of the standard search in all the regions across the entire objective space.

Fig. 6(b) shows the best and worst Pareto fronts from a NFE of 500,000 and 8000 in the evolution process of ten seed trials. In the case of the pre-conditioned search, the best solutions from 500,000 evaluations are better than the corresponding solutions in the case of standard baseline search. Although it is obvious that there are not many differences between solutions obtained from pre-conditioned search and solutions from standard baseline search due to the complexity of the problem, the best Pareto fronts from a NFE of 8000 in the case of the pre-condition search are approximate the same as the best Pareto fronts from a NFE of 500,000 in the case of the standard baseline search.

Fig. 7 shows the computational savings for two thresholds of hypervolume values 0.80 and 0.85 in the evolution process of each seed trial. In both cases of the thresholds of hypervolume values 0.80 and 0.85, NFE of the pre-conditioned search is less than standard baseline search for each seed. In the case of the threshold of hypervolume value 0.80, the average NFEs of full search and pre-conditioned full search are approximately 94,564 and 25,083 for one seed run respectively, and the computation is saved by 73.48%. Although the NFE of Sobol''s analysis is 82,000, the average NFEs of pre-conditioned full search is approximately $25{,}083 + 82{,}000/10 = 33{,}283$ for each seed run, and the computational saving is 64.80%.

Similarly, in the case of the threshold of hypervolume value 0.85, which is extremely difficult to achieve, the average NFEs of full search and pre-conditioned full search are approximately 214,049 and 105,060 for each seed run respectively, and the computation is saved by 50.92%. When the computation demand by Sobol''s analysis is considered, the computational saving is still 47.09%.

5.2 Inter-basin multi-reservoir system

5.2.1 Sensitivity analysis

Similar to the Dahuofang case study, a set of 2000 Latin Hypercube samples were used per decision variable yielding a total number of $2000 \times (126 + 2) = 256{,}000$ model simulations to compute Sobol''s indices in this case study.

The first-order and total-order indices for 126 decision variables are shown in Fig. 8. Similar to the results obtained from the Dahuofang ROS Problem in Fig. 4, the variance in the two objectives, i.e., industrial and agricultural shortage indices, are largely controlled by the water storages at time periods from agr4-2 to agr5-3 of Shenwo reservoir water supply operation rule curves for agricultural water demand, the water storages at time periods from agr4-2 to agr5-3 of Dahuofang reservoir water supply operation rule curves for agricultural water demand, the water storages at time periods ind1, ind2, ind3, ind7-1, ind10, ind11, and ind12 of Dahuofang reservoir water supply operation rule curves for industrial water demand based on a total-order Sobol''s index threshold of greater than 3%, which is subjective and its ease-of-satisfaction decreases with increasing numbers of parameters or parameter interactions. These 17 time periods are obvious candidates for decomposing the original optimization problem and formulating a pre-conditioning problem. Therefore, the simplified problem is defined from the original design problem with the 109 intensive time periods removed, i.e., the insensitive decision variables are set randomly first and kept constant during the solution of the simplified problem. It should be noted that the increased interactions across sensitive time periods in this test case. These interactions verify that this problem represents a far more challenging search problem.

5.2.2 Pre-conditioned optimization

Using the sensitivity-informed methodology, the simplified problem was first solved using ε-NSGAII with a maximum NFE of 5000, and the Pareto optimal solutions combined with the constant insensitive decision variables were then used as starting points to start a complete new search with a maximum NFE of 495,000. The standard search using ε-NSGAII was set to a maximum NFE of 500,000 so that the two methods have the same NFE used for search. In this case, 10 random seed trials are used given the computing resources available. Similar to the results obtained from the Dahuofang ROS problem in Fig. 5, the search traces in Fig. 9 show all three metrics (generational distance, additive epsilon indicator, and hypervolume) that represent performance metrics for the inter-basin multi-reservoir water supply operation system problem. Similarly, the pre-conditioning results are shown in Fig. 9 in red search traces continuing from the blue reduced complexity search results. It is clear that the sensitivity-informed pre-condition problems enhance search efficiency in terms of the generational distance, additive epsilon indicator, and hypervolume metrics. However, with the increase in problem complexity in comparison to the first case study (i.e., the number of decision variables from 39 to 126), the search of ROS optimization problem becomes more difficult, and so the metrics obtained from pre-conditioned search are not improved greatly compared with the standard baseline search and the pre-conditioning results shown in Fig. 9 are as good as the results shown in Fig. 5.

Fig. 10(a) shows Pareto fronts from a NFE of 6000, 8000 and 10,000 in the evolution process of one random seed trial. In the case of the pre-conditioned search, the solutions from the three NFE snapshots are much better than those from standard baseline search. Similar to Fig. 6(a), the results show that the Pareto approximate front of the pre-conditioned search is much wider than that of the standard search, and clearly dominates that of the standard search in all the regions across the entire objective space. Additionally, in the case of the pre-conditioned search, the solutions from 6000 evaluations are as good as those from 8000 evaluations and 10,000 evaluations. And they are much better than the solutions from the standard baseline search. It should be noted that the slow progress in the Pareto approximate fronts from 6000 to 10,000 evaluations reveals the difficulty of the inter-basin multi-reservoir operation system problem.

Fig. 10(b) shows the best and worst Pareto fronts from a NFE of 500,000 in the evolution process of ten seeds trials. Although it is obvious that the best Pareto approximate front of the pre-conditioned is as good as that of the standard search in all the regions across the entire objective space approximately, the Pareto solutions from 10 trials of the pre-conditioned search have significantly reduced variation, indicating a more reliable performance of the pre-conditioned method. In other words, the results show that the Pareto solution from one random seed trial of the pre-conditioned search is as good as the best solution from ten random seed trials of the standard search. That is to say, in the case of the pre-conditioned search, one random seed trial with a NFE of 500,000 is sufficient to obtain the best set of Pareto solutions, however, in the case of the standard search, ten seed trials with a total of $500{,}000 * 10 = 5{,}000{,}000$ NFE are required to obtain the Pareto solutions. Note that the NFE of Sobol'’s analysis is 256,000, which is about half of the NFE of one random seed trial. Thus, an improvement in search reliability can significantly reduce the computational demand for a complex search problem such as the multi-reservoir case study, even when the computation required by sensitivity analysis is included.

5.3 Discussions

For a very large and computationally intensive ROS problem, the full search problem is likely to be difficult so that it could not be optimized sufficiently in practice. The simplified problems can be used to generate high quality pre-conditioning solutions and thus dramatically improve the computational tractability of complex problems. This, however, requires using suitable optimization algorithms like ε-NSGAII which are capable of overcoming the risks for pre-mature convergence when pre-conditioning search (Fu et al., 2012).

The methodology tested in this study aims to reduce the number of decision variables through sensitivity-guided decomposition to form simplified problems. The optimization results from the two ROS problems show the reduction in decision space can make an impact on the reliability and efficiency of the search algorithm. For the Dahuofang ROS problem, recall that the original optimization problem has 39 decision variables, and the simplified problem has 11 decision variables based on Sobol'’s analysis. In the case of the inter-basin multi-reservoir operation system, the original optimization problem has 126 decision variables, and the simplified problem has a significantly reduced number of decision variables, i.e., 17. Searching in such significantly reduced space formed by sensitive decision variables makes it much easier to reach good solutions.

Although Sobol'’s global sensitivity analysis is computationally expensive, it captures the important sensitive information between a large number of variables for ROS models. This is critical for correctly screening insensitive decision variables and guiding the formulation of ROS optimization problems of reduced complexity (i.e., fewer decision variables). For example, in the Dahuofang ROS problem, accounting for the sensitive information, i.e., using total-order or first-order indices, result in a simplified problem for threshold of 10% as shown in Fig. 4. Compared with the standard search, this sensitivity-informed problem decomposition dramatically reduces the computational demands required for attaining high quality approximations of optimal ROS tradeoffs relationships between conflicting objectives, i.e., the best Pareto fronts from a NFE of 8000 in the case of the pre-condition search are approximately the same as the best Pareto front from a NFE of 500,000 in the case of the standard baseline search.

It should be noted that the sensitivity-informed problem decomposition framework is completely independent of multi-objective optimization algorithms, that is, any multi-objective algorithms could be embedded in the framework, including AMALGAM (Vrugt and Robinson, 2007). When dealing with three or more objectives, the formulation of the optimization problems with a significantly reduced number of decision variables will dramatically reduce the computational demands required to attain Pareto approximate solutions in a similar way to the two-objective optimization case studies considered in this paper.

6 Conclusions

This study investigates the effectiveness of a sensitivity-informed optimization method for the ROS multi-objective optimization problems. The method uses a global sensitivity analysis method to screen out insensitive decision variables and thus forms simplified problems with a significantly reduced number of decision variables. The simplified problems dramatically reduce the computational demands required to attain Pareto approximate solutions, which themselves can then be used to pre-condition and solve the original (i.e., full) optimization problem. This methodology has been tested on two case studies with different levels of complexity- the Dahuofang reservoir and the inter-basin multi-reservoir system in Liaoning province, China. The results obtained demonstrate the following:

1. The sensitivity-informed optimization problem decomposition dramatically increases both the computational efficiency and effectiveness of the optimization

process when compared to the conventional, full search approach. This is demonstrated in both case studies for both MOEA efficiency (i.e., the NFE required to attain high quality tradeoffs) and effectiveness (i.e., the quality approximations of optimal ROS tradeoffs relationships between conflicting design objectives).

2. The Sobol'’s method can be used to successfully identify important sensitive information between different decision variables in the ROS optimization problem and it is important to account for interactions between variables when formulating simplified problems.

Overall, this study illustrates the efficiency and effectiveness of the sensitivity-informed method and the use of global sensitivity analysis to inform problem decomposition. This method can be used for solving the complex multi-objective optimization problems with a large number of decision variables, such as optimal design of water distribution and urban drainage systems, distributed hydrological model calibration, multi-reservoir optimal operation and many other engineering optimization problems.

Acknowledgements

This study was supported by the China Postdoctoral Science Foundation Grants (2014M561231), and National Natural Science Foundation of China Grants (51320105010, 51279021, and 51409043).

References

Chang, J.X., Bai, T., Huang, Q., Yang, D.W. 2013. Optimization of water resources utilization by PSO-GA. Water Resources Management 27 (10), 3525-3540.

Draper, A.J., Lund, J.R. 2004. Optimal hedging and carryover storage value. Journal of Water Resources Planning and Management 130 (1), 83-87.

Fu, G.T., Kapelan, Z., Reed, P. 2012. Reducing the complexity of multi-objective water distribution system optimization through global sensitivity analysis. Journal of Water Resources Planning and Management 138 (3), 196-207.

Ko, S.K., Oh, M.H., Fontane, D.G. 1997. Multiobjective analysis of service-water-transmission systems. Journal of Water Resources Planning and Management 132 (2), 78-83.

Kollat, J.B., Reed, P.M. 2006. Comparing state-of-the-art evolutionary multi-objective algorithms for long-term groundwater monitoring design. Advances in Water Resources 29, 792-807.

Kollat, J.B., Reed, P.M. 2007. A computational scaling analysis of multiobjective evolutionary algorithms in long-term groundwater monitoring applications. Advances in Water Resources 30 (3), 408-419.

Labadie, J.W. 2004. Optimal operation of multireservoir systems: state-of-the-art review. Journal of Water Resources Planning and Management 130 (2), 93-111.

Liu, P., Guo, S.L., Xu, X.W., Chen, J.H. 2011a. Derivation of aggregation-based joint operating rule curves for cascade hydropower reservoirs. Water Resources Management 25, 3177-3200.

Liu, P., Cai, X.M., Guo, S.L. 2011b. Deriving multiple near-optimal solutions to deterministic reservoir operation problems. Water Resources Research 47, W08506, doi:10.1029/2011WR010998.

Mousavi, H., Ramamurthy, A.S. 2000. Optimal design of multi-reservoir systems for water supply. Advances in Water Resources 23, 613-624.

MWR-PRC (Ministry of Water Resources of PRC), 2008. China Water Resources Bulletin. Beijing: China Water Resources and Hydropower Press, (in Chinese).

Nicklow, J., Reed, P.M., Savic, D., Dessalegne, T., Harrell, L., Chan-Hilton, A., Karamouz, M., Minsker, B., Ostfeld, A., Singh, A., Zechman, E. 2010. State of the art for genetic algorithms and beyond in water resources planning and management. Journal of Water Resources Planning and Management 136 (4), 412-432.

Sadegh, M., Mahjouri, H., Kerachian, R. 2010. Optimal inter-basin water allocation using crisp and fuzzy shapely games. Water Resources Management 24, 2291-2310.

Shiau, J.-T. 2011. Analytical optimal hedging with explicit incorporation of reservoir release and carryover storage targets. Water Resources Research 47, W01515, doi:10.1029/2010WR009166.

Shiau, J.-T., Lee, H.-C. 2005. Derivation of optimal hedging rules for a water-supply reservoir through compromise programming. Water Resources Management 19 (2), 111-132.

Shirangi, E., Kerachian, R., Bajestan, M.S. 2008. A simplified model for reservoir operation considering the water quality issues: Application of the Young conflict resolution theory. Environmental Monitoring and Assessment 146, 77-89.

Srinivasan, K., Philipose, M.C. 1998. Effect of hedging on over-year reservoir performance. Water Resources Management 12, 95-120.

Suen, J.P., Eheart, J.W. 2006. Reservoir management to balance ecosystem and human needs: Incorporating the paradigm of the ecological flow regime. Water Resources Research 42, W03417, doi:10.1029/2005WR004314.

Tang, Y., Reed, P. M., van Werkhoven, K., Wagener, T. 2007a. Advancing the identification and evaluation of distributed rainfall-runoff models using global sensitivity analysis. Water Resources Research 43, W06415, doi:10.1029/2006WR005813.

Tang, Y., Reed, P. M., Wagener, T. 2006. How effective and efficient are multiobjective evolutionary algorithms at hydrologic model calibration? Hydrology and Earth System Sciences 10, 289-307.

Tang, Y., Reed, P. M., Wagener, T., van Werkhoven, K. 2007b. Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. Hydrology and Earth System Sciences 11, 793-817.

Tu, M.-Y., Hsu, N.-S., Tsai, F.T.-C., Yeh, W.W.-G. 2008. Optimization of hedging rules for reservoir operations. Journal of Water Resources Planning and Management 134 (1), 3-13.

Tu, M.-Y., Hsu, N.-S., Yeh, W.W.-G. 2003. Optimization of reservoir management and operation with hedging rules. Journal of Water Resources Planning and

Management 129 (2), 86-97.

Veldhuizen, D.A.V., Lamont, G.B. 1998. Evolutionary Computation and Convergence to a Pareto Front. In: John, R., Koza, editors. Late Breaking Papers at the Genetic Programming Conference; p. 221-228.

Vrugt, J.A., Robinson, B.A. 2007. Improved evolutionary optimization from genetically adaptive multimethod search. Proceedings of the National Academy of Sciences of the United States of America 104, 708-711.

Zhang, C., Chu, J.G., Fu, G.T. 2013a. Sobol''s sensitivity analysis for a distributed hydrological model of Yichun River Basin, China. Journal of Hydrology, 480, 58-68.

Zhang, H.F., Zhou, J.Z., Fang, N., Zhang, R., Zhang, Y.C. 2013b. An efficient multi-objective adaptive differential evolution with chaotic neuron network and its application on long-term hydropower operation with considering ecological environment problem. Electrical Power and Energy Systems 45 (1), 60-70.

Zhao, T.T.G., Zhao, J.S., Yang, D.W. 2014. Improved dynamic programming for hydropower reservoir operation. Journal of Water Resources Planning and Management 140 (3), 365-374.

Zitzler, E., Thiele, L. 1998. Multiobjective optimization using evolutionary algorithms-a comparative case study. In: Eiben, A., Back, T., Schoenauer, M., Schwefel, H.-P., editors. Parallel problem solving from nature (PPSN V). Lecture notes in computer science. Berlin, Amsterdam, the Netherlands: Springer-Verlag; p. 292-301.

Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C.M., Fonseca, V.G. 2003. Performance assessment of multiobjective optimizers: An analysis and review. IEEE Transactions on Evolutionary Computation 7 (2), 117-132.

Table 1 Reservoir characteristics and yearly average inflow (10^8 m^3)

Reservoir name	Minimum capacity	Utilizable capacity	Flood control capacity	Yearly average inflow
Dahuofang	1.34	14.30	10.00	15.70

Table 2 Characteristics of each reservoir in the inter-basin multi-reservoir system

Reservoir	Active storage (10^8 m^3)		**Role in water supply project**
	Flood season	Non-flood season	
Dahuofang	10.00	14.30	Supplying water
Guanyinge	14.20	14.20	Supplying water and exporting water to Shenwo
Shenwo	2.14	5.43	Supplying water and importing water from Guanyinge

List of Figure Captions

Fig. 1 Reservoir operational rule curves

Fig. 2 Flowchart of the sensitivity-informed methodology

Fig. 3 Layout of the inter-basin multi-reservoir system

Fig. 4 First-order and total-order indices for the Dahuofang ROS problem regarding (a) industrial shortage index and (b) agricultural shortage index. The x-axis labels represent decision variables (water storage volumes on the industrial and agricultural curves)

Fig. 5 Performance metrics for the Dahuofang ROS problem - (a) Generational Distance; (b) Additive Epsilon Indicator; (c) Hypervolume

Fig. 6 Pareto fronts derived from pre-conditioned and standard full searches for the Dahuofang ROS problem. (a) Sample Pareto fronts with different numbers of function evaluations for one random seed trial. (b) The best and worst Pareto fronts of ten seed trials.

Fig. 7 Computational savings for two hypervolume values - (a) $hypervolume = 0.80$; (b) $hypervolume = 0.85$

Fig. 8 First-order and total-order indices for the inter-basin multi-reservoir operation problem regarding industrial shortage index and agricultural shortage index. The x-axis labels represent decision variables (water storage volumes on the industrial, agricultural and water transferring curves)

Fig. 9 Performance metrics for the inter-basin multi-reservoir water supply operation problem - (a) Generation Distance; (b) Additive Epsilon Indicator; (c) Hypervolume

Fig. 10 Pareto fronts derived from pre-conditioned and standard full searches for the inter-basin multi-reservoir operation problem. (a) Sample Pareto fronts with different numbers of function evaluations for one random seed trial. (b) The best and worst Pareto fronts of ten seed trials.