Journal of Hydrology 520 (2015) 356–366

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

A preference-based multi-objective model for the optimization of best management practices

Lei Chen, Jiali Qiu, Guoyuan Wei, Zhenyao Shen *

State Key Laboratory of Water Environment Simulation, School of Environment, Beijing Normal University, Beijing 100875, PR China

ARTICLE INFO

Article history:
Received 11 July 2014
Received in revised form 5 November 2014
Accepted 7 November 2014
Available online 18 November 2014
This manuscript was handled by Geoff Syme, Editor-in-Chief, with the assistance of Paul Jeffrey, Associate Editor

Keywords:
Nonpoint source pollution
Multi-objective optimization
Preference
Reference point
Best management practices

SUMMARY

The optimization of best management practices (BMPs) at the watershed scale is notably complex because of the social nature of decision process, which incorporates information that reflects the preferences of decision makers. In this study, a preference-based multi-objective model was designed by modifying the commonly-used Non-dominated Sorting Genetic Algorithm (NSGA-II). The reference points, achievement scalarizing functions and an indicator-based optimization principle were integrated for searching a set of preferred Pareto-optimality solutions. Pareto preference ordering was also used for reducing objective numbers in the final decision-making process. This proposed model was then tested in a typical watershed in the Three Gorges Region, China. The results indicated that more desirable solutions were generated, which reduced the burden of decision effort of watershed managers. Compare to traditional Genetic Algorithm (GA), those preferred solutions were concentrated in a narrow region close to the projection point instead of the entire Pareto-front. Based on Pareto preference ordering, the solutions with the best objective function values were often the more desirable solutions (i.e., the minimum cost solution and the minimum pollutant load solution). In the authors' view, this new model provides a useful tool for optimizing BMPs at watershed scale and is therefore of great benefit to watershed managers.

1. Introduction

Nonpoint source (NPS) pollutants are major causes of water quality impairment. Best management practices (BMPs), as inherent aspects of watershed management, have been proven to effectively reduce NPS pollution (Laurent and Ruelland, 2011; Van Wie et al., 2013). Therefore, it is desirable to select a preferred set of BMPs for the watershed that would result in the greatest reduction in pollutant loads for the least cost. The selection and implementation of BMPs is constrained by objective conditions, including environmental characteristics and economic and social factors (Maringanti et al., 2011; Shen et al., 2013; Zare et al., 2012). A major challenge for multi-objective optimization is the incorporation of information that reflects the preferences of watershed managers that is useful in the decision-making processes, especially given limited funds and human resources.

The traditional Genetic Algorithms (GA) that have been applied to various multi-objective problem solving tasks have received a great deal of attention from decision makers (Deb et al., 2002a; Hadka and Reed, 2013). The goal of GA is to uncover the interactions between conflicting objectives to find a well-converged and well-distributed set of Pareto-optimality solutions (Deb, 2001; Deb et al., 2002b). The optimization methods have been widely applied in the water resources planning and management (Nicklow et al., 2010; Shafiee and Zechman, 2011). The Non-dominated Sorting GA (NSGA-II) has become one of the most widely used multi-objective GA for the selection and placement of BMPs at the watershed scale, considering the efficiency and cost of BMPs as well as farm income (Hsieh et al., 2010; Maringanti et al., 2011; Panagopoulos et al., 2013). However, NSGA-II faces difficulties when applied to problems without any information from decision makers (Deb, 1999a). The difficulties include: (1) the emphasis on all of the non-dominated solutions may produce selection pressure for the decision makers, and (2) a set of Pareto-optimality solutions does not provide a good presentation of the properties of the solutions near the desired region. Instead, it is a time-consuming search process based on the entire Pareto-front. Therefore, it is necessary to identify methods that can find a preferred and smaller set of Pareto-optimality solutions (Nedjah et al., 2012; Park and Ok, 2012).

Numerous past studies have attempted to provide decision makers with a set of preferred solutions near the desired region

* Corresponding author. Tel./fax: +86 10 58800398.
E-mail addresses: chenlei1982bnu@bnu.edu.cn (L. Chen), zyshen@bnu.edu.cn (Z. Shen).

http://dx.doi.org/10.1016/j.jhydrol.2014.11.032
0022-1694/

of interest. In these studies, the cost-effectiveness-operability objectives were weighed by the decision makers' preferences and aggregated into a single objective function. This simple approach seems reasonable because the weights for the objectives were directly derived from the levels of satisfaction of the local decision makers. However, the weights were subjective to the different purposes of the local governments. Alternatively, Deb (1999b) modified NSGA by incorporating the goal programming idea while Phelps and Köksalan (2003) used pairwise comparisons to incorporate the preferences into the fitness function. Woodruff et al. (2013) proposed the many-objective visual analytics (MOVA) framework to demonstrate how decision biases arise for lower dimensional, highly aggregated problem formulations. Currently, one of the main tools for expressing information that reflects the preferences of decision makers is the use of reference points, which consist of aspiration levels that reflect desirable values for the objective functions and allow the decision makers to search for the most preferred solutions (Deb and Sundar, 2006; Molina et al., 2009).

Agriculture is the main economic activity in the Three Gorges Reservoir Region (TGRR). The massive application of fertilizers releases much nitrogen (N) and phosphorus (P), which results in serious eutrophication. To cost-effectively control pollution, it is extremely important to identify the locations for various BMPs to minimize cost. The preferences of the decision maker have a large influence on the selection of watershed management programs, especially in developing countries like China. In this study, the NSGA-II was modified to directly obtain a region around the reference points based on the methods previously proposed by Thiele et al. (2009). In addition, the concept of Pareto preference ordering is used to prioritize more desirable solutions for further analysis. The Daning River watershed was used as a case study to verify an interactive, multi-objective optimization method. The objectives of this paper are to (1) encourage the participation of the decision makers in the entire multi-objective optimization process for searching for the most preferred areas, (2) avoid the extremely optimistic or pessimistic expectations for the watershed aquatic environment, and (3) identify the solutions with the best objective function values.

2. The framework for a preference-based NSGA-II

The framework for the preference-based optimization is shown in Fig. 1. Firstly, the initial Pareto-optimality front was generated by incorporating the Soil and Water Assessment Tool (SWAT) and the traditional NSGA-II (Part I) to provide the decision space for decision makers. Secondly, a preference-based algorithm was used to find a smaller and preferred set of solutions in the Pareto-front based on the information that reflects the preferences of decision makers (Part II). Finally, the concept of Pareto preference ordering was used to reduce the number of eligible solutions (Part III).

2.1. The generation of the initial Pareto-front

In this study, the Pareto-optimality fronts were generated using the NSGA-II method. The NSGA-II has gained popularity in many fields because it has been developed to overcome issues of high computational complexity and lack of elitism and a need for specifying the sharing parameter (Deb et al., 2002a). A major difference between the NSGA-II and other GAs is the method of operator selection. The NSGA-II uses the non-dominated sorting and ranking selection with the crowded comparison operator. The innovative aspects of this algorithm are as follow (Panda and Yegireddy, 2013; Zare et al., 2012):

1. Fast non-dominated sorting: The fast non-dominated sorting approach has a better book-keeping strategy to speed up the non-dominated sorting process and reduce the computation complexity.
2. Crowding distance: The NSGA-II adopts a crowding distance to measure the density of individuals in the same front. The overall crowding distance is calculated as the sum of individual distance values based on their m objectives in the n-dimensional space. Behind the non-domination rank, the crowding distance of each individual is also calculated by the average Euclidean distance between it and each individual in a front. Then, the selection is performed using a crowded comparison operator.
3. Elitist crowded comparison operator: This operator guides the selection process at various stages toward a uniformly spread-out Pareto-optimality front. The crowding distance is applied to select one with a greater crowding distance from two individuals in the same front. The elitist crowded comparison operator combines offspring population members with parent population in the selection process.

In this study, the operator for the selection process is based on individual fitness and sorted quickly by a non-dominated sorting method. The excellent individuals were selected from the chromosome, which were regarded as the parents to generate offspring. To ensure that the good gene can be inherited by the next generation, the better adapted individuals from each generation were given a larger virtual fitness score. For the selection mechanism, this study used a stochastic tournament selection method. The basic idea was that n individuals were randomly selected from the chromosomes of every generation, and the best individual among the n individuals was retained and directly inherited by the next generation, which was consistent with the idea of the elite reservation strategy. Mutation and crossover were used to create a new population for the next generation. The model terminated when the maximum generation was reached, which was the stopping condition that provided a range of optimized solutions for the multi-objective functions. More NSGA-II information can be obtained from Deb et al. (2002a).

The optimization results of the NSGA-II were very sensitive to the parameters, including the number of generations, the initial population size and the mutation probability. In this study, a sensitivity analysis was conducted on the NSGA-II parameters using the Morse Classification Screening Method (One factor At a Time or OAT), which only changed a certain parameter and observed its influence on the results. The Pareto-front located as close to the origin of coordinate system as possible is desired, and the corresponding parameter values were those that resulted in the least sum of distances from each solution on the Pareto-front to the origin. For the specific sensitivity analysis procedure, refer to Maringanti et al. (2009) and Maringanti et al. (2011).

To avoid the subjectivity of the decision maker in the parameter sensitivity analysis, this study introduced the convergence index and the distribution index (representing diversity) for the Pareto-optimality set, which were proposed by Deb et al. (2002a). The convergence index γ measures the convergence of Pareto-optimality solutions based on the minimum Euclidean distance from the solutions $x = (x_1, x_2, x_3, \ldots, x_n)$ of each generation to the chosen solutions on the Pareto-front $y = (y_1, y_2, y_3, \ldots, y_n)$. The calculation follows:

$$d_i = \min_{i=1}^{n} \sqrt{\sum_{m=1}^{k} (f_m(\vec{x}_i) - f_m(\vec{y}_i))^2} \tag{1}$$

where $f_m(\vec{x})$ is the m-th objective function of solution $\vec{x}$, $\vec{x}_i$ and $\vec{y}_i$ are the k-dimensional vectors of the i-th solution from $x = (x_1, x_2, x_3,$