Journal of Hydrology 511 (2014) 242–253

Contents lists available at ScienceDirect

Journal of Hydrology

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

Multi-objective optimization of empirical hydrological model for streamflow prediction

Jun Guo [a,*], Jianzhong Zhou [b], Jiazheng Lu [a], Qiang Zou [c], Huajie Zhang [b], Sheng Bi [b]

[a] State Grid Hunan Electric Power Corporation Research Institute, Changsha 410007, China
[b] School of Hydropower and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China
[c] Changjiang Water Resources Commission, Changjiang Institute of Survey, Planning, Design and Research, Wuhan 430010, China

ARTICLE INFO

Article history:
Received 16 April 2013
Received in revised form 19 January 2014
Accepted 22 January 2014
Available online 31 January 2014
This manuscript was handled by Corrado Corradini, Editor-in-Chief, with the assistance of Gokmen Tayfur, Associate Editor

Keywords:
Streamflow forecasting
Hydrological model
Multi-objective optimization
Model calibration

SUMMARY

Traditional calibration of hydrological models is performed with a single objective function. Practical experience with the calibration of hydrologic models reveals that single objective functions are often inadequate to properly measure all of the characteristics of the hydrologic system. To circumvent this problem, in recent years, a lot of studies have looked into the automatic calibration of hydrological models with multi-objective functions. In this paper, the multi-objective evolution algorithm MODE-ACM is introduced to solve the multi-objective optimization of hydrologic models. Moreover, to improve the performance of the MODE-ACM, an Enhanced Pareto Multi-Objective Differential Evolution algorithm named EPMODE is proposed in this research. The efficacy of the MODE-ACM and EPMODE are compared with two state-of-the-art algorithms NSGA-II and SPEA2 on two case studies. Five test problems are used as the first case study to generate the true Pareto front. Then this approach is tested on a typical empirical hydrological model for monthly streamflow forecasting. The results of these case studies show that the EPMODE, as well as MODE-ACM, is effective in solving multi-objective problems and has great potential as an efficient and reliable algorithm for water resources applications.

1. Introduction

To make precise streamflow predictions, hydrological model calibration is always needed and is one of the most important issues in the field of hydrology. Model calibration is performed to find the optimal parameters which produce the best fit between simulated and observed hydrologic characteristics. Model calibration is normally performed with a single objective function. However, practical experience with the calibration of hydrologic models reveals that single objective functions are often inadequate to properly measure all of the characteristics of the hydrologic system (Vrugt et al., 2003). Therefore, it stands to reason that the calibration of hydrologic model should be performed with multiple objective functions, which represent different dynamic aspects inherent to the hydrological system. Gupta et al. (1998) have discussed the advantages of model calibration with multiple objective functions. Different from single objective optimization, the result of multi-objective optimization will not be a single best solution but consists of a set of non-dominated, or Pareto optimal, solutions. Generally, there are two ways of solving this multi-objective calibration problem. One is to convert the multi-objective problem to a single objective problem by assigning weights to different objective functions. This can only generate one Pareto solution at a time, making this method inefficient and time-consuming. The other way is to solve the multi-objective optimization problem directly based on the paradigm of non-dominated sorting. Non-dominated sorting is performed with the concept of Pareto dominance. If all objective functions of individual P are superior to those of individual Q, it means that P dominates Q and the Pareto rank of P is set higher (the rank value is smaller) than that of Q. If one or more objective functions of individuals P are superior to those of Q, while other objective functions of P are inferior to those of Q, this indicates that P is non-dominated with Q and P is set the same Pareto rank as Q for this situation. With this sorting method, each individual can be assigned a Pareto rank. This sorting scheme can take into account all different optimization objective functions simultaneously. A complete multi-objective optimization allows analysis of the trade-offs among the different objective functions and enables hydrologists to better understand the limitations of the current hydrologic model structure (Gupta et al., 1998). In recent years, much research has been devoted to developing or introducing multi-objective algorithms for optimization of hydrologic models (Yapo et al., 1998; Vrugt et al., 2003; Gupta et al., 1998, 1999; Boyle et al., 2000; Wagener et al., 2001; Xia et al., 2002; Leplastrier

* Corresponding author. Tel.: +86 18973102063.
E-mail addresses: guojunhust@gmail.com, guojunhust@126.com, guo@hust.edu.cn (J. Guo).

http://dx.doi.org/10.1016/j.jhydrol.2014.01.047
0022-1694/

et al., 2002; Reed et al., 2003; Reed and Minsker, 2004; Khu and Madsen, 2005; Gill et al., 2006; de Vos and Rientjes, 2007, 2008; Price et al., 2012; Guo et al., 2013). The most famous algorithms applied in multi-objective calibration of hydrological models are Multi-Objective Complex Evolution (MOCOM-UA) method (Yapo et al., 1998), Multiobjective Shuffled Complex Evolution Metropolis (MOSCEM-UA) algorithm (Vrugt et al., 2003; de Vos and Rientjes, 2008; Guo et al., 2013), Nondominated Sorting Genetic Algorithm II (NSGA-II) (Tang et al., 2005; Khu and Madsen, 2005; de Vos and Rientjes, 2007, 2008; Zhang et al., 2010; Guo et al., 2013), and the improved Strength Pareto Evolutionary Algorithm (SPEA2) (Tang et al., 2005; Zhang et al., 2010; Guo et al., 2013). MOSCEM-UA is an improved version of MOCOM-UA. And Tang et al. (2005) and Guo et al. (2013) have shown that SPEA2 performs better than MOSCEM-UA.

The multi-objective evolution algorithm Multi-Objective Differential Evolution with Adaptive Cauchy Mutation (MODE-ACM), which is designed to solve the short-term multi-objective optimal hydrothermal scheduling problem (Qin et al., 2010), is introduced for parameter optimization of a hydrological model. Moreover, to improve the performance of the MODE-ACM, an Enhanced Pareto Multi-Objective Differential Evolution algorithm (EPMODE) is proposed in this research. The features and capabilities of the EPMODE algorithm are illustrated using two case studies, and the results are compared with NSGA-II, SPEA2 and MODE-ACM.

The remainder of this paper is organized as follows: Section 2 presents the details of the proposed EPMODE algorithm and simple introductions to NSGA-II, SPEA2 and MODE-ACM are also given; Section 3 describes the performance metrics used in this research; Section 4 demonstrates the methods about two case studies; Section 5 presents comparisons of the results generated by the above mentioned four algorithms through two case studies; and conclusions are made in Section 6.

2. Multi-objective optimization algorithms

2.1. NSGA-II

NSGA-II, which is an improvement over the NSGA, was first proposed by Deb et al. (2002). It is one of the most effective and efficient algorithms for solving multi-objective problems. NSGA-II is an extension of Genetic Algorithm (GA) to a multi-objective optimization algorithm. The main evolving mechanism of NSGA-II is based on GA. However, as each individual of the population is related to two or more different objective functions, we cannot determine which individual is better with the selection operator of GA. The selection operator of NSGA-II has to be done under the concept of Pareto dominance.

NSGA-II overcomes limitations of the original version of the algorithm by employing a fast non-dominated sorting approach and a diversity preservation strategy. Suppose the population size is N and the number of objective functions is M, the fast non-dominated sorting method can reduce the computing time complexity to $O(MN^2)$. In addition, NSGA-II uses a Crowded-Comparison Operator to preserve the diversity of the optimal non-dominated frontier. With the Crowded-Comparison Operator, each point of the frontier is assigned a crowded value, and each crowded value is calculated according to its rank obtained from the previous fast non-dominated sorting. If the size of frontier is larger than the predefined size, then the point with highest rank and smallest crowded value will be deleted from the set of frontier. This diversity preservation strategy can improve the diversity of the non-dominated solutions set. Besides, in the algorithm NSGA, the share parameter is introduced to maintain the diversity of solutions set, and this parameter must be predefined. It is usually cumbersome to determine the proper parameter value. To circumvent this problem, NSGA-II proposes a parameter-less diversity preservation mechanism. For more details about NSGA-II, readers are encouraged to refer to Deb et al. (2002).

2.2. SPEA2

SPEA2 is an improved version over the SPEA (Zitzler and Thiele, 1999; Zitzler et al., 2001). SPEA2, like NSGA-II, can be treated as an extension of GA to a multi-objective optimization framework, but the operators of SPEA2 are designed more suitable and simpler for computation than NSGA-II (as NSGA-II must predefine a proper share parameter). It overcomes limitations of the original version of the algorithm by using an improved fitness assignment and diversity preservation using k-means clustering. Similar to SPEA, SPEA2 also contains a regular population and an archive (used for retaining the non-dominated solutions). The fitness of SPEA2 is determined by the strengths of its dominators in both archive and population, as opposed to SPEA where only archive members are taken into account. The archive size is set in advance. When the size of archive is larger than the predefined size, the diversity preservation operator is activated. It promotes diversity by iteratively removing the individual that has the minimum distance from its neighboring solutions. For more detailed descriptions of SPEA2, readers can refer to Zitzler and Thiele (1999) and Zitzler et al. (2001).

2.3. MODE-ACM

The evolution scheme of MODE-ACM is based on the differential evolution algorithm (Storn and Price, 1995, 1997). It includes four operators: crowded-comparison operator, archive update operator, differential evolving operator and adaptive Cauchy mutation operator. The crowded-comparison operator is the same as NSGA-II. The diversity preservation strategy of NSGA-II is adopted to updating the archive set. In each generation, the first rank solutions of the regular population are added to the archive individually, while the dominated solutions are removed from the archive. If the archive size is larger than predefined size, the crowded-comparison operator is activated to remove solutions with smaller crowded distance. Although the differential evolving algorithm converges fast and identifies the optimal solutions in the feasible space, it also suffers from the premature convergence problem. In essence, the cause of this problem is that the diversity of the population decreases along with the evolutionary process, and the algorithm cannot jump out of the local optimum and search new feasible space. From the definition of the differential evolving algorithm, it can be noted that if the differences between each individual are very small, the differential evolving algorithm mutation and crossover operators nearly cannot generate new individuals. If the algorithm has not converged to the global optimum, this indicates that a local optimum has constrained the algorithm. The adaptive Cauchy mutation operator is used to avoid premature convergence of the algorithm. In every h generations (h is usually set to 5 or 10), the diversity of each dimension in decision variable space is calculated. If the diversity of ith-dimension is smaller than the predefined threshold, the adaptive Cauchy mutation operator will be activated to add a small disturbance to that dimension. Generally, intense mutation (intense mutation means large range of magnitude shifts) can increase the probability of escaping the local optimum; however, overly intense mutation may affect the convergence performance of the algorithm. Therefore, the mutation operator is designed to be adaptive, allowing more intense mutation at the beginning while less intense mutation at the end. For more detailed descriptions of MODE-ACM, readers can refer to Qin et al. (2010).