Environmental Modelling & Software 69 (2015) 292–307

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Investigating the run-time searching behavior of the differential evolution algorithm applied to water distribution system optimization☆

Feifei Zheng*, Aaron C. Zecchin, Angus R. Simpson

School of Civil, Environmental and Mining Engineering, University of Adelaide, Adelaide, South Australia, 5005, Australia

ARTICLE INFO

Article history:
Received 28 March 2014
Received in revised form
15 September 2014
Accepted 22 September 2014
Available online 16 October 2014

Keywords:
Differential evolution algorithm
Evolutionary algorithms
Search behavior
Water distribution systems
Optimization

ABSTRACT

In recent years, the differential evolution algorithm (DEA) has frequently been used to tackle various water resource problems due to its powerful search ability. However, one challenge of using the DEA is the tedious effort required to fine-tune parameter values due to a lack of theoretical understanding of what governs its searching behavior. This study investigates DEA's search behavior as a function of its parameter values. A range of behavioral metrics are developed to measure run-time statistics about DEA's performance, with primary focus on the search quality, convergence properties and solution generation statistics. Water distribution system design problems are utilized to enable investigation of the behavioral analysis using the developed metrics. Results obtained offer an improved knowledge on how the control parameter values affect DEA's search behavior, thereby providing guidance for parameter-tuning and hence hopefully increasing appropriate take-up of the DEA within the industry in tackling water resource optimization problems.

1. Introduction

In the water resource community, researchers and engineers often have to deal with various optimization problems. These include hydrological model calibration, the planning, design and operation of water resource systems (Nicklow et al., 2010). The optimization process tries to find the best solution of the given problem within the specified constraints. Optimizing water resource problems are often extremely difficult due to the highly nonlinear and complex decision spaces (Razavi et al., 2012). Although traditional deterministic optimization techniques have been attempted to solve these problems, the results have often been unsatisfactory (Zheng et al., 2011a).

Over the past two decades, there has been a move towards developing or applying various evolutionary algorithms (EAs) to deal with water resource optimization problems (Maier et al., 2014). The differential evolution algorithm (DEA), first proposed by Storn and Price (1995) as one type of EA, has especially received a deal of attention in recent years (details of DEA are given in Section 3). For example, Vasan and Raju (2007) introduced the DEA to optimize flow allocations of an irrigation system; Reddy and Kumar (2007), applied the DEA to reservoir system optimization; Vasan and Simonovic (2010), and Zheng et al. (2013) employed the DEA to optimize the design of water distribution systems (WDSs). More recently, Chichakly et al. (2013) designed watershed-based stormwater management plans using the DEA, and Joseph and Guillaume (2013) used the DEA to calibrate hydrological models. It has been reported in these studies that the DEA exhibited better performance in efficiently finding optimal solutions compared to other types of evolutionary algorithms (EAs), such as genetic algorithms (GAs) and ant colony optimization (ACO) algorithms. This shows that the DEA is promising for dealing with a broad array of water resource optimization problems.

Previous studies have also shown that DEA's search behavior is heavily dependent on the values of the control parameters F (the differential weight used in the mutation operator) and CR (the crossover probability used in the crossover operator), while it is not significantly affected by the varying population size N (Qin and Suganthan, 2005; Das and Suganthan, 2011). Using the same WDS case studies, Suribabu (2010) concluded that the performance of DEA is significantly better than GAs, while Marchi et al. (2014)

☆ Thematic Issue on the Evolutionary Algorithms.

* Corresponding author. N124, Engineering North, School of Civil, Environmental and Mining Engineering, University of Adelaide, North Terrace Campus, Adelaide, South Australia, 5005, Australia. Tel.: +618 8313 1113.

E-mail addresses: feifei.zheng@adelaide.edu.au (F. Zheng), aaron.zecchin@adelaide.edu.au (A.C. Zecchin), angus.simpson@adelaide.edu.au (A.R. Simpson).

http://dx.doi.org/10.1016/j.envsoft.2014.09.022
1364-8152/

stated that GAs gave better results overall than the DEA. This contradiction can be explained by the fact that different parameter values, including *F* and *CR*, were used in these DE applications. Zheng et al. (2011b) performed a sensitivity analysis of DEA's parameters (*F* and *CR*) in terms of affecting the final solutions based on two WDS case studies. Results in their study showed that the DEA was unable to solve the optimization problems effectively if inappropriate parameter values were used. This suggests that a set of appropriate parameter values is very critical in obtaining the satisfactory performance of the DEA, which is similar as other types of EAs such as genetic algorithms and the harmony search algorithm (Savic and Walters, 1997; Geem, 2006).

Suitable parameter values when using an EA approach are normally optimization problem-dependent due to the variation of fitness landscapes associated with different problems and problem types (Tolson et al., 2009). Typically, a trial-and-error approach is used to calibrate the parameter values for the DEA applied to given optimization problems in water resources (Reddy and Kumar, 2007). This results in a large computational overhead especially when dealing with real world optimization problems, for which a large number of decision variables are normally involved. The tedious effort required for tuning parameter values has been frequently claimed by practitioners as one of the main reasons for their reluctance to embrace EAs in practice (Geem and Sim, 2010).

In order to address this issue, two potential research directions have been adopted. The first direction is the development of parameter-free EAs. Wu and Walski (2005), for example, proposed a self-adaptive penalty approach within a GA to remove the presetting of the penalty multiplier parameter for pipeline optimization. Geem and Sim (2010) proposed a parameter-setting-free harmony search algorithm to optimize the design of WDSs. More recently, Zheng et al. (2013) proposed a self-adaptive differential evolution algorithm (SADE) to optimize the design of WDSs, in which the two control parameter values *F* and *CR* were adapted along with the evolution of the solutions rather than being prespecified to fixed values in advance.

The second research direction is the characterization of EA's run-time search properties as a function of the varying control parameter values, thereby providing guidance for fine-tuning parameter values. Traditionally, an EA's search performance is typically assessed based on end-of-run performance measures (i.e. statistics describing the least-cost solution found, and the time taken to find the least-cost solution, see discussions in the position paper by Maier et al. (2014)). A state-of-the-art example of the end-of-run performance analysis is the work by Hadka and Reed (2012), in which a diagnostic assessment framework was developed for evaluating the effectiveness, reliability, efficiency and controllability of multi-objective evolutionary algorithms (MOEAs). In contrast to the extensive research on the end-of run performance assessment, there has been few investigations into characterizing an algorithm's properties from the point of view of the underlying run-time searching performance. The only example of the run-time performance analysis is the work of Zecchin et al. (2012), who investigated the run-time search behavior of various ACO operators applied to WDS optimization problems.

In the context of a parametric study, the end-of-run statistics enable the determination of a direct relationship between an algorithm's parameter settings and overall performance. However, a consideration of the run-time behavioral statistics can provide more insight as to how the different values of the control parameters affect an EA's searching behavior in terms of the exploration (the ability to broadly explore the whole search space) and exploitation (the ability to intensively exploit the promising regions) within the search process (Maier et al., 2014). This insight should provide guidance not only for practitioners to select appropriate parameter values of EAs based on an available computational budget, but also for algorithm developers to understand more deeply the direct and measured impact of parameter variations on the search behavior. For example, for applications with a limited computational budget, a set of parameter values should be selected in favor of exploitation. In contrast, if better quality solutions are preferred with relaxed computational constraints, the combination of the parameter values needs to possess more strength on the exploration ability. Building a fundamental understanding of EA's working principles, such as the run-time searching behavior, as opposed to focusing only on the end-of-run performance, is an important future research objective as stated in the position paper by Maier et al. (2014).

As previously outlined, Zheng et al. (2011b) conducted a parametric study on DEA's control parameters (*F* and *CR*), followed by a development of a self-adaptive DE algorithm (Zheng et al., 2013) in order to remove the tedious parameter tuning process. Both studies solely focused on the end-of-run statistics within the given computational budget, ignoring the algorithm's run-time searching properties. Therefore, the question still remains as to why certain algorithms or algorithms with certain parameterization outperformed others for the selected case studies, and how the internal operators and mechanisms alter the DEA's run-time searching behavior that lead up to the end-of-run performance. This paper is such an attempt to address this issue.

To facilitate the search behavior analysis, a range of behavioral measures are developed for the DEA in the current study. The primary run-time statistics of interest concern the population variance, the search quality, the convergence measures, the percentage of the time spent in the feasible and infeasible regions, and the percentage of improved solutions within each generation. The WDS design problem, as one typical type of complex optimization problems in water resources (Fu and Kapelan, 2011), is considered to analyze the search behavior of the DEA with respect to varying parameter values. Three WDS case studies with increased scales and complexity are used in the current study.

Various metrics have been developed to enable the non-dominant set comparison in the multi-objective EA (MOEA) domain. For example, to assess MOEA's final searching performance, Ang et al. (2002) proposed to plot the non-dominated solutions against their distance to the Pareto front and their distance between each other, while Hadka and Reed (2012) utilized a broad range of performance metrics including the hypervolume, the generational distance, the inverse generational distance, the additive epsilon indicator and the spread. These studies have made merit in developing metrics to evaluate MOEA's end-of-run performance, while they significantly differ to the focus of this study that attempts to develop various metrics to measure the DEA's run-time performance in the single-objective space.

Although the impact of different parameter values (*F* and *CR*) on DEA's performance is investigated in this study, it is not intended to derive a quantitive relationship between the case studies (different scales and complexity) and the appropriate parameter values. The aim of the present study is to: (1) develop and demonstrate the utility of metrics for analyzing the DEA's run-time search behavior; and (2) characterize the influence of varying parameter values (*F* and *CR*) on the DEA's searching behavior (exploration and exploitation) through an empirical numerical study, and compare the empirical results with prior theoretical results from Zharie (2002, 2009) using the complex WDS design problems. It is anticipated that such improved knowledge can provide qualitative guidance to design the DEA to possess various exploration and exploitation emphasis according to the problem scales and complexity as well as the available computational budgets.