To be cited as: Wang, Q., Guidolin, M., Savic, D., and Kapelan, Z. (2015). "Two-Objective Design of Benchmark Problems of a Water Distribution System via MOEAs: Towards the Best-Known Approximation of the True Pareto Front." J. Water Resour. Plann. Manage., 141(3), 04014060.

Two-Objective Design of Benchmark Problems of Water Distribution System via MOEAs: Towards the Best-Known Approximation to the True Pareto Front

Qi Wang[1], Michele Guidolin[2], Dragan Savić[3], Zoran Kapelan[4]

Abstract: Various multi-objective evolutionary algorithms (MOEAs) have been applied to solve the optimal design problems of a Water Distribution System (WDS). Such methods are able to find the near-optimal trade-off between cost and performance benefit in a single run. Previously published work used a number of small benchmark networks and/or a few large, real-world networks to test MOEAs on design problems of WDS. A few studies also focused on the comparison of different MOEAs given a limited computational budget. However, no consistent attempt has been made before to investigate and report the best-known approximation of the true Pareto front (PF) for a set of benchmarks problems, and thus there is not a single point of reference. This paper applied five state-of-the-art MOEAs, with minimum time invested in parameterisation (i.e., using the recommended settings), to twelve design problems collected from the literature. Three different population sizes were implemented for each MOEA with respect to the scale of each problem. The true Pareto fronts for small problems and the best-known Pareto fronts for the other problems were obtained. Five MOEAs were complementary to each other on various problems, which implies that no one method was completely superior to the others. The non-dominated sorting genetic algorithm-II (NSGA-II), with minimum parameters tuning, remains a good choice as it showed generally the best achievements across all the problems. In addition, a small population size can be used for small and medium problems (in terms of the number of

[1] Mr. Qi Wang, Centre for Water Systems; qw212@exeter.ac.uk; College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Exeter EX4 4QF, United Kingdom
[2] Dr. Michele Guidolin, Centre for Water Systems; M.Guidolin@exeter.ac.uk; College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Exeter EX4 4QF, United Kingdom
[3] Prof. Dragan Savić, Centre for Water Systems; D.Savic@exeter.ac.uk; College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Exeter EX4 4QF, United Kingdom
[4] Prof. Zoran Kapelan, Centre for Water Systems; Z.Kapelan@exeter.ac.uk; College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Exeter EX4 4QF, United Kingdom

decision variables). However, for intermediate and large problems, different sizes and random seeds are recommended to ensure a wider Pareto front. The publicly available best-known PFs obtained from this work are a good starting point for researchers to test new algorithms and methodologies for WDS analysis.

Subject Headings: Algorithms, Optimization, Water distribution systems, Design, Rehabilitation

Keywords: two-objective design; water distribution system; multi-objective evolutionary algorithm; hybrid algorithm; best-known Pareto front; benchmark problem

Introduction

Multi-objective design of a Water Distribution System (WDS) has received increasing attention during the past two decades. Much effort (Farmani et al., 2006; Fu et al., 2012a; Halhal et al., 1997; Keedwell and Khu, 2003; Prasad and Park, 2004; Prasad and Tanyimboh, 2008) has been made to identify the trade-off, expressed as the Pareto front (PF), between cost and performance type benefit using various indicators. Multi-objective evolutionary algorithms (MOEAs) are widely accepted for addressing this kind of problem as they are capable of approximating the PF effectively and efficiently in a single run (Farmani et al., 2005a). Many benchmark networks and some real networks have been used to demonstrate the strength of MOEAs. For instance, Cheung et al. (2003) applied both strength Pareto evolutionary algorithm (SPEA) and multi-objective genetic algorithm (MOGA) to the rehabilitation problem of a hypothetical network. Minimisation of cost and the total pressure deficit were taken as two objectives. Farmani et al. (2004) contributed a new benchmark network based on a real system and used the non-dominated sorting genetic algorithm-II (NSGA-II) to solve the two-objective rehabilitation of this large network, minimising cost and number of nodes with head deficiency. Besides using pressure deficit and its analogues as the second objective, other formulations were aimed at optimising resilience based indicators (Basupi et al., 2013; Farmani et al., 2005b; Prasad and Park, 2004), flow entropy (Prasad and Tanyimboh, 2008), and other mixed indicators (Raad et al., 2010).

Apart from the aforementioned MOEAs, other methods have been applied to solve benchmark problems, such as SPEA2 (Farmani et al., 2003), Cross Entropy (Perelman et al., 2008), Particle Swarm Optimisation (PSO) (Montalvoa et al., 2010), and Cuckoo search (Wang et al., 2012). Recently, hybrid algorithms (Vrugt and Robinson, 2007; di Pierro et al., 2009; Hadka and Reed, 2013), which combine different schemes and components together in an attempt to further enhance search ability, have demonstrated significant improvement over previous MOEAs, like NSGA-II (Deb et al., 2002) and SPEA2 (Zitzler et al., 2002). The encouraging performance of these newly-proposed MOEAs has gained interest for the design of WDS. Raad et al. (2009) for the first time applied a modified multi-algorithm, genetically adaptive multi-objective (AMALGAM) to the two-objective design of WDS, using cost and network resilience (Prasad and Park, 2004) as objectives. Wang et al. (in press) compared the performance of two distinct hybrid algorithms (including the original AMALGAM) against NSGA-II on a wide range of benchmark problems. Creaco and Franchini (2012; 2013) set up a hybrid procedure where NSGA-II coordinates various subordinate algorithms to perform the multi-objective design under pressure and velocity constraints. Fu et al (2012b) proposed a novel hybrid approach where global sensitivity analysis is used before applying the ε-NSGA-II method to reduce the complexity of the search space size of a multi-objective WDS design problem.

Most comparative studies concerned the ultimate performance of MOEAs. However, as Kollat and Reed (2006) emphasised, it is equally important to assess the dynamic performance of MOEAs. To this end, a reference set of the true PF is generally required to calculate the metrics of convergence and diversity. In practice, a reference set is usually generated by extracting the non-dominated solutions obtained by one or more algorithms through multiple runs. However, none of the aforementioned studies paid attention to generating the best-known PF for benchmark problems. One reason for this lies in the fact that the problem is non-deterministic polynomial-time hard (NP-hard) (Papadimitriou and Steiglitz, 1998), which cannot be enumerated in an acceptable time frame. Another one is probably due to the lack of universally accepted formulation for design problems. To date, a limited effort has been made to provide the suitable reference sets for various benchmark problems in a single location. For this reason, the paper is aimed at finding the best-known approximation to the true PF of each benchmark design problem of WDSs collected from the literature. In addition, for small problems full enumeration is implemented to produce the true PF.

The multi-objective design of benchmark WDSs is formulated to minimise the total cost and to maximise the network resilience (Prasad and Park, 2004). Five predominant MOEAs including two hybrid algorithms are implemented, and then their attainments are aggregated to produce the optimal front for each problem. Also, these MOEAs are compared in terms of the ultimate performance, which assists with identification of the overall best candidate for the task. This paper contributes to the best-known approximations to the true PFs of a wide range of benchmark problems. Hence, these fronts can facilitate rigorous assessment of both ultimate and dynamic performance of newly-developed algorithms. On the other hand, researchers and practitioners alike can decide which algorithm to choose when facing real-world problems, which are complex and inevitably time-consuming.

Multi-objective design of a WDS

The optimal design of a WDS is an intractable problem due to its size, discrete (combinatorial) nature and non-linearity associated with a number of complex constraints. Strictly speaking, for the design problem all the components should be determined, e.g., pipe diameters, pump capacities, valve settings, and tank sizes, to name a few; while for the extended design problem replacement of existing components and/or adding new elements into the system should be carefully decided. Nonetheless, most existing work focuses on the narrow sense of the task, i.e. the pipe sizing problem without considering the other components. Even with this simplification, the problem is still NP-hard and thus a great challenge to tackle especially for large, real-world networks.

Historically, the design problem was treated as single-objective optimisation focusing on economic considerations, but the drawbacks of this formulation have been criticised broadly (Engelhardt et al., 2000; Fu et al., 2012a; Walski, 2001). In transforming the least-cost formulation to the multi-objective, or more precisely, two-objective formulation, many indicators have been proposed as additional objectives. At the very beginning, variants of pressure deficit were used to account for the second objective, such as minimising the total pressure deficit, minimising the maximum of pressure deficiency, and minimising the number of nodes with pressure deficit (Cheung et al., 2003; Farmani et al., 2005a). However, these aforementioned formulations do not necessarily result in looped networks, which are reliable configurations under abnormal conditions (e.g., pipe burst). On the other hand, a resilience index formulation (Todini, 2000) was introduced as a surrogate measure for hydraulic benefits. The index is based on the concept that the total input power into a network consists of the power dissipated in the network and the power delivered at demand nodes. So, less power consumed internally to overcome the friction results in more surplus power at demand nodes and thus being able to counter the failure scenarios. Later on, an improved version of the resilience indicator (Prasad and Park, 2004), called network resilience, was proposed taking the uniformity of pipes around each demand node into account. Network resilience considers the effect of redundancy of a pipe network and maximising this indicator can ensure reliable loops. It is proved that using network resilience as another objective alleviates the shortcomings of the resilience index (e.g., resulting in impracticable loops) and yields the solutions which are robust under pipe failure conditions (Prasad and Park, 2004; Raad et al., 2010). For this reason, network resilience is used as the second objective for the two-objective optimisation of benchmark problems.

In this paper, only the expenditure of pipe components (new pipes and/or existing pipes) is considered for the total cost of a design solution. The unit cost of a specific diameter for each problem is derived from the relevant paper. EPANET 2 software (Rossman, 2000) is taken to run the hydraulic simulation, in which the variables required for the evaluation of network resilience are obtained. The formulation of the objectives is given in Eq. 1 and Eq. 2-3, respectively.

$$\min C = \sum_{i=1}^{np} a \times D_i^b \times L_i \tag{1}$$

Where C=total cost (monetary units problem dependant); np=number of pipes; a and b=constants depending on a specific problem; D_i=diameter of pipe i; L_i=length of pipe i.

$$\max I_n = \frac{\sum_{j=1}^{nn} C_j Q_j (H_j - H_j^{req})}{(\sum_{k=1}^{nr} Q_k H_k + \sum_{i=1}^{npu} \frac{P_i}{\gamma}) - \sum_{j=1}^{nn} Q_j H_j^{req}} \tag{2}$$

$$C_j = \frac{\sum_{i=1}^{npj} D_i}{npj \times \max\{D_i\}} \tag{3}$$

Where I_n=network resilience; nn=number of demand nodes; C_j, Q_j, H_j and H_j^{req}=uniformity, demand, actual head and minimum head of node j; nr=number of reservoirs; Q_k and H_k=discharge and actual head of reservoir k; npu=number of pumps; P_i=power of pump i;

γ=specific weight of water; npj=number of pipes connected to node j; D_i=diameter of pipe i connected to demand node j.

Most papers in the literature solved two or three benchmark problems for explanatory purposes. Some also tried to tackle large, real-world networks. However, normally only a small number of problems were tested, thus making it hard to generalise the conclusions and guide practitioners in dealing with new problems. On the other hand, several benchmark networks already exist, derived from papers and reports, which makes it possible to set up an archive of benchmark problems and to benefit other researchers in this community. In this paper, twelve such networks were collected and categorised into four groups according to the size of search space. Table 1 gives a summary of these benchmark problems including the number of demand loading conditions, water sources, decision variables and pipe diameter options. For small problems, the true Pareto front is obtained via full enumeration which can be completed within a short time using a modern personal computer. For the other three groups, the aim is to approximate the true PFs by taking advantage of five state-of-the-art MOEAs given various computational budgets. A very brief introduction to the various benchmark networks is given below. Readers are referred to the corresponding papers for additional details.

In these problems, the two-loop network (Alperovits and Shamir, 1977) is a hypothetical network, while the others are real networks or simplified networks in the real-world. The New York tunnel network (Schaake and Lai, 1969) and the Exeter network (Farmani et al., 2004) were originally presented as extended design problems. The rest are design problems except the BakRyan network (Lee, 2001) and the Two-Reservoir network (Gessler, 1985) which are a mix of design and extended design. There are both minimum and maximum pressure requirements for demand nodes in the Blacksburg network (Sherali et al., 2001), the Fossolo network (Bragalli et al., 2008), the Pescara network (Bragalli et al., 2008), and the Modena network (Bragalli et al., 2008), whereas the others only have minimum pressure requirements. In addition, there are upper bounds for velocity in the pipes of the four aforementioned networks. The Hanoi network and the GoYang network are taken from (Fujiwara and Khang, 1990) and (Kim et al., 1994), respectively. Unlike a WDS, the water consumption is fixed at 5.55 l/s across all the demand nodes in the Balerma irrigation network (Reca and Martínez, 2006).

[Table 1 goes here]

MOEAs used in the analysis

In this paper, five MOEAs in total are applied to solve the benchmark problems. Two of them are high-level hybrid algorithms (Talbi, 2002), namely AMALGAM (Vrugt and Robinson, 2007) and Borg (Hadka and Reed, 2013). The others are NSGA-II (Deb et al., 2002), ε-MOEA (Deb et al., 2005), and ε-NSGA-II (Kollat and Reed, 2006), which is an enhanced algorithm based on NSGA-II.

The reasons for choosing these MOEAs are as follows. NSGA-II has been widely used as a benchmark MOEA in water engineering (Farmani et al., 2005a; Kollat and Reed, 2006; Raad et al., 2009), and it serves as the prototype of AMALGAM (except in the case of the genetically adaptive multi-operators). Borg was developed based on ε-MOEA as it is a highly efficient steady-state model (Hadka and Reed, 2013). ε-NSGA-II proved to be superior to NSGA-II and ε-MOEA on a four-objective long-term groundwater monitoring design case (Kollat and Reed, 2006). Most recently, Guidolin et al. (2012) further highlighted the strength of ε-NSGA-II by winning the title in the Battle of Water Networks II (BWN-II, Marchi et al 2013), using a master-slave parallel version of this algorithm. Fu et al. (2012a) also applied ε-NSGA-II as well as a tool for visually interactive decision-making (Kollat and Reed, 2007) to the many-objective (up to six) rehabilitation of Anytown network (Walski et al., 1987), revealing the complex tradeoffs that would not be identified in a lower-dimensional formulation. On the other hand, Hybrid algorithms have been developed in an attempt to overcome the "No Free Lunch" theorem (Wolpert and Macready, 1997) by combining the power of different methods. Therefore, it is worth comparing their performance with benchmark MOEAs on various design cases. Note that no other MOEAs were considered in the paper due to the findings in the relevant comparative studies (Raad et al., 2011; Reed et al., 2013). A brief introduction to these algorithms is given below.

Hybrid MOEAs

AMALGAM

AMALGAM is a hybrid optimisation framework which employs simultaneously four sub-algorithms within its structure, including NSGA-II, adaptive metropolis search (AMS) (Haario et al., 2001), particle swarm optimisation (PSO) (Kennedy and Eberhart, 2001) and differential evolution (DE) (Storn and Price, 1997). It is designed to overcome the drawbacks of using an individual algorithm. The strategies of global information sharing and genetically adaptive offspring creation are implemented in the process of population evolution. Each sub-algorithm is allowed to produce a specific number of offspring based on the survival history of its solutions in the previous generation. The pool of current best solutions is shared among sub-algorithms for reproduction. Simulation results on a set of well-known multi-objective benchmark functions suggest that AMALGAM achieves a tenfold improvement over current MOEAs for the more complex, higher dimensional problems (Vrugt and Robinson, 2007). In addition, AMALGAM provides a general template which is flexible and extensible, and can easily accommodate any other population-based algorithm. Raad et al. (2011) subsequently demonstrated that this hybrid framework, with other ingredients tailored for WDS design, convincingly outperformed NSGA-II for a large problem.

Borg

Using ε-MOEA as its predecessors, Borg incorporates more advanced features into a unified framework, including ε-dominance (Laumanns et al., 2002), ε-progress (a measure of convergence speed), randomised restart, and auto-adaptive multi-operator recombination (similar to AMALGAM). The comparative study on 33 instances of three well-known test suites reveals that it is efficient and reliable on various problems with difficult characteristics. Besides its flexibility, another point that should be highlighted is its large regions of high-performing parameterisations (Goldberg, 1989) in terms of so-called sweet spots (Purshouse and Fleming, 2007).

The advantages of Borg are threefold: (1) usage of ε-box dominance archive contributes to maintaining the convergence and diversity concurrently throughout search; (2) the combination of time continuation (Srivastava, 2002), adaptive population sizing, and two types of randomised restart (i.e. ε-progress triggered restart and population-to-archive ratio triggered restart) boosts the algorithm towards global optima; (3) simultaneous employment of multiple recombination operators enhances performance on a wide assortment of problem domains. In addition, the adoption of the steady-state, elitist model of ε-MOEA (Deb et al., 2005) make it easily extendable for use on parallel architectures. Borg has also been successfully used to solve challenging, many-objective, real-world problems in the domain of water resources, a detailed review of which can be found in (Reed et al., 2013).

Referential MOEAs

NSGA-II

NSGA-II (Deb et al., 2002) is arguably the most popular MOEA and is regarded as an "industry standard" algorithm which has been successfully applied to a variety of fields. As a result, it is usually taken as a benchmark MOEA and compared with other algorithms (Deb et al., 2001; Zitzler et al., 2002; Farmani et al., 2005a; Raad et al. 2011). NSGA-II features a fast non-dominated sorting approach, implicit elitist selection method based on Pareto dominance rank and a secondary selection method based on crowding distance, which significantly improve its performance on difficult multi-objective problems. Moreover, it provides a constraint-handling technique to deal with constrained problems efficiently and supports both binary and real coding representations. Since it serves as the outer framework of AMALGAM (Vrugt and Robinson, 2007), it is included in this comparative study.

ε-MOEA

Unlike the NSGA-II, ε-MOEA (Deb et al., 2005) is a steady-state MOEA in which only one solution is generated per iteration. It incorporates the concept of epsilon-dominance (Laumanns et al., 2002), being able to preserve a good representation of Pareto front in terms of convergence and diversity. At the beginning, a population is initialised randomly and the non-dominated solutions are retained in an archive. Next, a solution is created via crossover and mutation using two parents each of which is selected from the population and the archive. Then, this solution is checked for acceptance or rejection by the population and the archive, using Pareto dominance and ε-dominance, respectively. The abovementioned procedure is repeated until a stopping criterion is met. Deb et al. (2005) compared ε-MOEA with four other state-of-the-art MOEAs on many test problems and concluded that it was able to find well-converged and well-distributed solutions in a shorter computational time. Since Borg uses ε-MOEA as the basic framework, it is taken into account for comparative purposes.

ε-NSGA-II

The ε-NSGA-II method (Kollat and Reed, 2006; Tang et al., 2006) goes beyond the common implementation of MOEA by building on NSGA-II (Deb et al., 2002) and three key components, namely ε-dominance archiving (Laumanns et al., 2002), adaptive population sizing with archive injection, and automatic termination. The ε-NSGA-II differs from the

NSGA-II primarily in two aspects: (1) while the population evolves in the same manner as NSGA-II, an offline archive is frequently updated by selecting the ε-non-dominated solutions from the elitist population at current generation; (2) the optimisation is implemented in consecutive epochs, each of which is terminated automatically according to the user-specified improvement criteria. The next epoch is populated by injecting the members in the archive and generating new random solutions. The ε-dominance archiving maintains the convergence and diversity of the archive concurrently. It also allows users to specify the precision of each objective and thus is more flexible in practice. Adaptive population sizing contributes to balance the exploration and exploitation throughout the search, which is achieved by increasing or decreasing the capacity of the population based on the number of members in the archive. Additionally, several connected runs, known as time continuation (Srivastava, 2002), enhance the possibility to explore other regions of search space. The comparative study on ε-NSGA-II as well as three benchmark MOEAs (NSGA-II, ε-MOEA and SPEA2) showed its superiority in terms of efficiency and reliability (Kollat and Reed, 2006). Moreover, the aforementioned key components of ε-NSGA-II remedy the issue of parameterisation commonly found in MOEAs, thus making it easy-to-use for a wide range of applications.

Benchmarking setup

Each benchmark problem was formulated as two-objective design or extended design, taking total cost and network resilience into account. For a design problem, the decision variables were the diameters of individual pipes in the network. While for an extended design problem, the decision variables included the diameters of duplicate pipes as well as the other two options, i.e. leaving alone (do-nothing option) and cleaning of existing pipes. Note that all the MOEAs used real coding but all the decision variables were of integer type. So the real values passed in by MOEAs were rounded down (e.g., 12.9457 becomes 12). For each problem, a solution was considered as infeasible if there were violations of pressure requirements (minimum and maximum if any) and upper bound of flow velocity (if any). Note that no penalty function is used to handle infeasible solutions; instead, infeasibility situations are dealt with by implementing the constrained-domination principle (Deb et al., 2002).

<u>Parameter settings of MOEAs</u>

To ensure as fair comparison as possible, a uniform computational budget in terms of number of objective function evaluations (NFE) has been allowed to solve each benchmark problem across all the MOEAs considered in the paper. In addition, it is worth mentioning that there are various individual parameters in each MOEA, particularly in hybrid algorithms, which can have an impact on the algorithm performance. In this paper, the individual parameters are not fine tuned for three main reasons. Firstly, Borg and ε-NSGA-II both feature adaptive population sizing and "time continuation" strategy (involving several connected runs triggered by automatic restart). In fact, one of the main advantages of these algorithms is to eliminate the need for parameterisation, resulting in highly reliable and efficient MOEAs. Secondly, the primary control parameters in AMALGAM are not fixed by default. Instead, these parameters are randomly sampled from the high-performance ranges recommended in relevant papers (Parsopoulos and Vrahatis, 2002; Hu et al., 2003; Gelman et al., 2003; Iorio and Li, 2005). Hence, it is expected to reduce the issue of parameterisation to some extent. Thirdly, NSGA-II and ε-MOEA are implemented as referential MOEAs and parameterised according to the widely recommended settings from the literature (Deb and Agrawal, 1995; Deb et al., 2002; Kollat and Reed, 2006). Most recently, Reed et al. (2013) conclude that Borg, ε-NSGA-II and AMALGAM represent the top performing MOEAs which demonstrate a satisfactory achievement in terms of effectiveness, efficiency, reliability and scalability. However, it should be noted that these default parameter settings do not necessarily result in the best performance for a variety of benchmark problems. In practice, it is recommended to fine-tune some key parameters of an MOEA via the sensitivity analysis before application. This is usually feasible for solving small problems with a limited number of decision variables. However, it may be extremely computationally expensive to do so for solving large and complex problems. Since this paper is aimed at obtaining the best-known PFs for many benchmark problems given extensive computational budgets, but not at comparing different MOEAs, the default parameter settings for these MOEAs are adopted.

In addition, all the algorithms considered in this paper share similarities in that they use tournament selection, real-valued simulated binary crossover (SBX), polynomial mutation (PM) (Deb et al., 2002). Therefore, unified settings of these factors are kept the same for all the algorithms. More specifically, a tournament size of 2 is applied except for Borg and ε-NSGA-II (tournament size changed due to adaptive population sizing). The probabilities of SBX and PM are 0.9 and the inverse of the number of decision variables respectively, and the distribution index of SBX and PM are 15 and 7 respectively. Note that these values are selected according to the most commonly recommended parameter settings in the literature (Deb et al., 2002; Kollat and Reed, 2006). A smaller distribution index of SBX or PM enables the search operator to create solutions with more spread (Deb and Agrawal, 1995), hence it is effective to avoid premature.

In short, default settings of control parameters in each algorithm are maintained across the experiments except for those related to population size. Table 2 gives a summary of the specific parameter settings of each MOEA used in the paper.

[Table 2 goes here]

<u>Epsilon precision</u>

Of the five MOEAs, three of them (ε -MOEA, ε -NSGA-II, and Borg) require the specification of epsilon precision for both objectives of each problem. Table 3 gives the specifications of epsilon values which were obtained via trial runs with respect to the range of each objective and for each problem.

[Table 3 goes here]

<u>Computational budget</u>

The benchmark problems cover a wide range of complexity, hence the need for investing different computational budgets for different cases. Based on preliminary tests, each type of problem was solved by complying with the budget in terms of NFE specified in Table 4, and each MOEA was run ten times independently using a certain population size (thirty runs in total) to solve each problem. It is worth noting that all the MOEAs took full use of these budgets. In other words, each algorithm was run on each benchmark problem for equal NFE. Thus, any other stopping criteria (or techniques allowing early stopping) were omitted. Additionally, in order to avoid premature convergence of optimisation for large problems by using a small population size, three varied population sizes (referred to as 'group' in the subsequent paragraphs) with respect to the number of decision variables and the size of search space are implemented for each problem. This approach also assists in exploring the impact of the population size on the final achievement of MOEAs.

[Table 4 goes here]

Performance assessment

This paper aims at finding the best-known PF of each benchmark problem, thus facilitating quantitative comparison of the performance of different MOEAs in future research. Meanwhile, the impact of population size on the achievement of MOEAs involved is also investigated.

To achieve the two goals above, a dual-stage procedure for data post-processing is illustrated in Fig. 1. In stage I, raw data reported by each MOEA for each problem via thirty runs are rounded according to the epsilon precision specified in Table 3 and duplicates in the dataset of each group (obtained using a specific population size) are checked and removed. Then, data from different groups are merged together and duplicates are checked and removed once again. Next, the non-dominated sorting procedure (Deb et al., 2002) is applied to the aggregated dataset to produce the best PF obtained by the current MOEA. Finally, the contribution from each group to the best PF is counted. Three kinds of number of solutions were recorded during stage I, namely the number of solutions finally obtained from multiple runs by each MOEA (before being processed, denoted as Sol_{FO}), the number of solutions excluding duplicates (denoted as Sol_{ED}), and the number of solutions contributed from each group to the best PF of each MOEA (denoted as Sol_{CT}), respectively.

In stage II, for each problem, the best PF obtained by each MOEA are firstly aggregated and duplicates in the merged dataset are checked and removed. Next, the non-dominated sorting procedure (Deb et al., 2002) is used to generate the best-known PF of the current problem. Lastly, the contribution from each MOEA to the best-known PF is identified.

[Figure 1 goes here]

As there is no reference set of the true PFs to hand, various performance indicators existing in the literature (Deb et al., 2002; Knowles and Corne, 2002; Zitzler et al., 2003) cannot be applied directly. Therefore, the number of solutions contributed by each MOEA to the best-known PF of each problem is counted. This will demonstrate the general capability (including convergence and diversity) of an MOEA to find the optimal non-dominated solutions to a problem. On the other hand, as mentioned before, for each MOEA three different population sizes are implemented for each benchmark problem (see Table 4). To investigate the impact of population size on the performance of each MOEA, the number of solutions contributed from each population size to the best PF is also counted. It is also worth noting that the computational time spent by each MOEA on each problem are not compared due to the fact that these algorithms were developed in different languages and implemented on various machines with different operating systems. For example, AMALGAM was built in *Matlab* and run on a desktop computer (Windows 7) with 2.66GHz and 3GB RAM. While a parallel version of ε-NSGA-II was written in *C* language and executed on a supercomputer, called "Zen", which consists of diskless compute nodes each with twelve cores and 24GB of RAM. Instead, a rough observation about the runtime spent on algorithm steps and objective function evaluations (including hydraulic simulations) is provided as follows. Generally speaking, all the methods spent a higher proportion of CPU time on the objective function evaluations for large problems. AMALGAM took more CPU time on algorithm steps on average than the *C* language based MOEAs because it was developed and implemented in *Matlab* which is an interpreted language. For the *C* language based MOEAs, Borg and ε-MOEA spent less CPU time on algorithm steps compared with NSGA-II and ε-NSGA-II (non-Parallel version) because they followed the steady-state algorithmic framework, which did not involve time-consuming ranking and sorting as in NSGA-II and ε-NSGA-II.

In addition, since the number of solutions from each MOEA alone cannot demonstrate the distribution of solutions in the best-known PF, a novel projection plot is developed to illustrate the distribution of solutions contributed by a specific MOEA. A clear advantage of using this projection plot is that it can deliver the preferred information of convergence (secondary) and diversity but avoid showing the overlaps between different Pareto fronts, when they are drawn in the same objective space (commonly seen in comparative studies). The procedure of generating such a projection plot is explained as follows. For each problem, firstly, data in the best-known PF are sorted according to the values of either objective in ascending or descending order. Here the ascending order of cost objective was chosen for the purpose of demonstration. Note that the other objective is inevitably ignored by using the projection plot, but this will not affect the interpretation of the results. Then, these solutions (in the space of *Cost* vs. I_n) are evenly projected on to a 1-D axis, from 1 to the length of data set (i.e., the number of solutions in the best-known PF). Next, by comparing the overlaps (duplicates) of solutions from each MOEA with those in the best-known PF, the corresponding positions of solutions contributed from each MOEA on the 1-D axis can be identified. Finally, these solutions from each MOEA are also projected on to the 1-D axis in a stacking fashion (as shown in Fig. 2). It is worth noting that the convergence of each algorithm is implicitly considered by using the novel projection plot, because only the fully converged solutions with respect to the non-dominated solutions in the best-known PF are shown for each MOEA. Therefore, this approach facilitates a direct comparison of the relative distribution of solutions from each MOEA.

Results and Discussion

Table 5 shows the number of solutions found in the best-known PF as well as the percentage of contribution from each MOEA. Note that the impact of ε-precision has been taken into account by rounding off the solutions in each best-known PF and the approximation set obtained by each MOEA according to the settings specified in Table 3. It can be observed that MOEAs using the ε-dominance concept contributed on average less solutions than NSGA-II and AMALGAM, which take the ordinary dominance concept for sorting the solutions. For small problems, as well as FOS, PES and EXN cases, ε-dominance based MOEAs were comparable or even superior to AMALGAM. NSGA-II demonstrated the best overall performance in terms of solutions found in the best-known PF across the whole spectrum of problems. AMALGAM was quite close to NSGA-II for small and medium sized problems, and exceeded NSGA-II's performance for BLA and GOY cases. However, it showed poor results for FOS, PES, BIN and EXN cases as it found less than half of the solutions obtained by NSGA-II. ε-MOEA demonstrated the worst performance in the experiment as it was dominated by the other MOEAs in half of the test problems. It is hard to distinguish Borg and ε-NSGA-II but the latter consistently found the solutions in the best-known PF of all the problems. Borg failed to contribute a single solution to the best-known PF of BIN, while ε-MOEA encountered the same difficulty for EXN problem. Surprisingly, Borg proved to be exceptional powerful for EXN problem by finding more than 60% solutions in the best-known PF followed by ε-NSGA-II. In addition, all the MOEAs failed to discover the entire solutions in the true PFs of small problems, which can be partly attributed to the usage of ε-dominance concept. Nevertheless, NSGA-II and AMALGAM performed satisfactorily for a problem of such size. Here, it is worth noting that the comparison of MOEAs according to the number of solutions contributed to the best-known PFs can be biased, as it did not consider the spread of solutions in the objective space. Due to the lack of reference sets for benchmark problems, it is currently difficult to explain why certain MOEA performed better than others for particular cases. However, it is believed that the best-known

PFs obtained in this paper can facilitate a more comprehensive comparison in a quantitative way, which provides an opportunity to find the reasons.

[Table 5 goes here]

Fig. 2(a) to 2(d) illustrates the relative distribution of solutions contributed by each MOEA in the best-known PF of four selected benchmark problems (i.e. BAK, NYT, PES and EXN cases), selected from each type of problem (small to large). The true PF of BAK problem is also added to Fig. 2(a) for the purpose of comparison. By using the innovative projection plot, it is much easier to compare their performance (both convergence and diversity) in an intuitive sense. Since the solutions in the best-known PF are evenly mapped (except Fig. 2(a) due to the existence of true PF), the gaps appearing in the graph for each MOEA denote the absence of solutions in the corresponding position within the set of best-known PF. Therefore, a longer (in the absolute sense) and more uniform band indicates better achievement of the particular MOEA.

Generally speaking, NSGA-II and AMALGAM were able to provide consistently long and uniform bands of solutions for small and medium sized problems. The other ε-dominance based MOEAs showed acceptably good performance on these problems except that ε-MOEA was unable to cover the high cost region (high network resilience) for NYT problem. Contrastingly, for intermediate and large sized problems, all MOEAs were capable of locating only a portion of solutions in the best-known PFs, which implies that no one method is versatile enough for complex cases. However, except for the EXN problem, NSGA-II always captured the solutions in the region of low to medium cost, while AMALGAM was effective at finding solutions in the region of medium to high cost. Three ε-dominance based MOEAs did not perform well on complex problems as their bands were short and/or discontinuous or even missing (Borg for BIN problem and ε-MOEA for EXN problem). More interestingly, a collaborative effort from different MOEAs was made to solve the EXN problem as there are no overlaps on the bands discovered by each individual MOEA.

[Figure 2 goes here]

As shown in Fig. 1, variations in the number of solutions during data post-processing are recorded for further analysis. The number of solutions in different steps varies due to the existence of duplicates in the non-dominated sets and the stochastic nature of MOEAs. For non-ε-dominance based MOEAs, Sol_{FO} is equal to the size of each group times the number of runs (10 in this paper). While Sol_{FO} of ε-MOEA is the number of solutions in the aggregated archive across multiple runs. This figure is expected to be different from those of NSGA-II and AMALGAM as the size of archive keeps changing during optimisation. Similar to ε-MOEA, the archive sizes of ε-NSGA-II and Borg change over time due to the strategy of adaptive population sizing with time continuation.

Fig. 3(a) to 3(d) demonstrates these variations for each MOEA on four benchmark problems (i.e. BAK, NYT, PES and EXN cases) according to the different population sizes. The light grey bar, medium grey bar and dark grey bar represent Sol_{FO}, Sol_{ED}, and Sol_{CT} in percentage, respectively. Note that Sol_{CT} is the most important value as it indicates the efficiency of solutions found by an algorithm. There is a clear trend of fewer redundancies occurring for large problems, which is to be expected as they have larger search spaces. In other words, for all MOEAs the majority of solutions overlapped with each other for small problems, but the degree of overlap decreased gradually with the increasing complexity of problems. From the viewpoint of efficiency of non-dominated solutions, ε-dominance based MOEAs, especially ε-NSGA-II and Borg, showed consistently more stable convergence than non-ε-dominance based approaches. This is demonstrated by the differences between Sol_{ED} and Sol_{CT}, which are smaller for small and medium problems (i.e. BAK and NYT cases). However, on intermediate and large sized problems, all the MOEAs suffered from inefficiency of solutions, or even failed to discover any solutions in their best PFs. Note that Sol_{CT} of each MOEA on each problem does not represent the amount of solutions appearing in the best-known PF of that problem as these solutions may be dominated by those reported from other MOEAs.

On the other hand, in response to the impact of population size on the final achievement, it seems that a small population size is good enough for small and medium problems. However, as shown in the EXN problem, larger population sizes (group 2 and 3) generally produced more solutions of high quality. Therefore, different population sizes are still recommended, no matter which MOEA is chosen, for solving intermediate and large sized problems. In fact, for each MOEA there are rarely overlapping parts in the set of best PF from different groups for intermediate and large problems.

[Figure 3 goes here]

The best-known PF of each benchmark problem is provided in Appendix A. Note that these best-known PFs are different from the ones used in Results and Discussion, since the data were not processed according to the epsilon precision in Table 3. Thus, a more complete reference set of each benchmark problem is provided. The corresponding data of these best-known PFs can be downloaded from the website of the Centre for Water Systems.

Conclusions

This paper set up the methodology of benchmarking MOEAs for two-objective design of Water Distribution System. Five representative MOEAs were applied to solve twelve benchmark problems. An innovative projection plot was applied to facilitate the comparison of MOEAs in terms of convergence and diversity. The best-known Pareto front of each problem was obtained in the space of cost against network resilience. Note that these Pareto fronts are not necessarily uniformly distributed (as shown in Appendix A) due to the discrete nature of Water Distribution System design. The benchmark problems (including the EPANET input files) written in C code and the associated best-known Pareto fronts (as reference sets) are provided on the website of Centre for Water Systems. This is expected to benefit future research work which formulates the problem in the same manner. In particular, the capability of newly-proposed algorithms can be rigorously tested (both ultimate and dynamic performance) in a much easier way, since various performance indicators are ready for use, requiring only the reference set.

On the other hand, the strength of MOEAs tested in the paper, including two modern hybrid MOEAs and three frequently used MOEAs, was compared in the context of optimal design of Water Distribution System. The results obtained proved that NSGA-II remains one of the best MOEAs, which is suitable for two-objective optimisation of a Water Distribution System. It generally outperformed the other MOEAs in terms of the number of solutions contributed to the best-known PF of each problem. The spread (both extent and uniformity) of its contribution was also comparable, if not better, to those of other MOEAs. AMALGAM is promising for this task as it contributes more than 85% non-dominated solutions in the best-known PFs for small and medium problems consistently. It always discovered solutions in the region of high network resilience, although there is a clear drop in performance on intermediate and large problems. Three ε-dominance based MOEAs failed to demonstrate any clear advantage over NSGA-II and AMALGAM in the experiment. Nevertheless, all of them showed evident advantage in the convergence and efficiency of non-dominated solutions for small and medium problems. Besides, Borg was shown to be exceptionally superior to the other MOEAs for the EXN problem by finding more than 60% solutions in the best-known PF. In short, the MOEAs considered were complementary to each other especially for complex cases, albeit no versatile MOEA was found in this study. Therefore, when facing large sized problems, different MOEAs should be considered to ensure a reliable Pareto front if both time and computational resources are available.

In addition, the impact of the combination of population size and generations on the performance of each MOEA was also investigated given the same computational budget. Small size (e.g. less than 100) seems to work well for small and medium problems. On the other hand, it is advisable to use multiple runs of different population sizes and random seeds as they can cover different parts in the best-known PFs for large problems.

It is worth noting that there is no attempt in this work to fine tune the specific parameters of each MOEA. So the conclusions drawn here should not be generalised especially when a certain MOEA is well adjusted for a particular purpose. However, if resources (time and/or hardware) are limited for fine-tuning the parameters of an optimisation algorithm, NSGA-II is probably a good choice for two-objective optimisation of Water Distribution Systems.

In the future, Sensitivity Analysis can be carried out to investigate the parameterisation issue of MOEAs, especially hybrid algorithms, for the design of a Water Distribution System. Future work is to diagnose the failure mode of MOEAs, like Borg or AMALGAM, for further improvement. Moreover, the many-objective (more than two) formulation should be considered for benchmarking these MOEAs towards a more realistic design perspective.

Acknowledgements

The authors thank Jasper Vrugt (University of California, Irvine, USA) and David Hadka (from the Hydroinformatics Research Group lead by Prof. Patrick Reed) for kindly providing the source code of AMALGAM, ε-NSGA-II and Borg for testing in this paper. Also, we would like to thank Mr. Xiaolei Dong and Mrs Alix Slater for providing some benchmark networks and the proofreading, respectively. The insightful comments and suggestions given by four anonymous reviewers substantially improve the quality of this paper, which are gratefully acknowledged.

References

Alperovits, E., Shamir, U., 1977. Design of optimal water distribution systems. Water Resour. Res. 13(6), 885-900.

Basupi, I., Kapelan, Z., Butler, D., 2013. Reducing Life-cycle Carbon Footprint in the Redesign of Water Distribution Systems. J. Water Climate Change.

Bragalli, C., D'Ambrosio, C., Lee, J., Lodi, A., Toth, P., 2008. Water Network Design by MINLP. Technical report. Report No.: RC24495.

Cheung, P.B., Reis, L.F.R., Formiga, K.T.M., Chaudhry, F.H., Ticona, W.G.C., 2003. Multiobjective evolutionary algorithms applied to the rehabilitation of a water distribution system: a comparative study, Proceedings of the 2nd international conference on Evolutionary multi-criterion optimization. Springer-Verlag: Faro, Portugal.

Creaco E., Franchini M., 2012. Fast network multi-objective design algorithm combined with an a-posteriori procedure for reliability evaluation under various operational scenarios. Urban Water J. 9(6), 385-399.

Creaco E., Franchini M., 2013. Low level hybrid procedure for the multi-objective design of water distribution networks. In: Proceedings of the 12th International Conference on Computing and Control for the Water Industry, CCWI2013, Elsevier.

Deb, K., Agrawal, R. B. 1995. Simulated Binary Crossover for Continuous Search Space. Complex Systems, 9, 115-148.

Deb, K., Mohan, M., Mishra, S., 2005. Evaluating the epsilon-domination based multi-objective evolutionary algorithm for a quick computation of Pareto-optimal solutions. Evolut. Comput. 13(4), 501-525.

Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. IEEE T. Evolut. Comput. 6(2), 182-197.

Deb K, Thiele L, Laumanns M, Zitzler E. 2001. Scalable test problems for evolutionary multi-objective optimization. Tech Rep Kan-GAL 2001001, Indian Institute of Technology Kanpur.

di Pierro, F., Khu, S.-T., Savic, D., Berardi, L., 2009. Efficient multi-objective optimal design of water distribution networks on a budget of simulations using hybrid algorithms. Environ. Modell. Softw. 24(2), 202-213.

Engelhardt, M.O., Skipworth, P.J., Savic, D.A., Saul, A.J., Walters, G.A., 2000. Rehabilitation strategies for water distribution networks: a literature review with a UK perspective. Urban Water 2(2), 153-170.

Farmani, R., Savic, D., Walters, G.A., 2006. A Hybrid Technique For Optimisation Of Branched Urban Water Systems, 7th International Conference on Hydroinformatics: Nice, France.

Farmani, R., Savic, D.A., Walters, G.A., 2003. Multi-objective optimization of water system: A comparative study, In: Jr., C.a.C. (Ed.), Pumps, Electromechanical Devices and Systems Applied to Urban Water Management, pp. 247-256.

Farmani, R., Savic, D.A., Walters, G.A., 2004. "EXNET" Benchmark Problem for Multi-Objective Optimization of Large Water Systems, Modelling and Control for Participatory Planning and Managing Water Systems, IFAC workshop: Venice, Italy.

Farmani, R., Savic, D.A., Walters, G.A., 2005a. Evolutionary multi-objective optimization in water distribution network design. Eng. Optimiz. 37(2), 167-183.

Farmani, R., Walters, G.A., Savic, D.A., 2005b. Trade-Off between Total Cost and Reliability for Anytown Water Distribution Network. J. Water Resour. Plann. Manage. 131(3), 161-171.

Fu, G., Kapelan, Z., Kasprzyk, J., Reed, P., 2012a. Optimal Design of Water Distribution Systems Using Many-Objective Visual Analytics. J. Water Resour. Plann. Manage. doi:10.1061/(ASCE)WR.1943-5452.0000311

Fu, G., Kapelan, Z. and Reed, P.M., 2012b, 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.

Fujiwara, O., Khang, D., 1990. A two-phase decomposition method for optimal design of looped water distribution networks. Water Resour. Res. 26(4), 539-549.

Gessler, J., 1985. Pipe network optimization by enumeration, In Proc., Computer Applications for Water Resources. ASCE: New York, N.Y., pp. 572-581.

Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.

Guidolin, M., Fu, G., Reed, P., 2012. Parallel Evolutionary Multiobjective Optimization of Water Distribution System Design, Water Distribution Systems Analysis 2012: Adelaide, Australia.

Haario, H., Saksman, E., Tamminen, J., 2001. An adaptive Metropolis algorithm. Bernoulli 7(2), 223-242.

Hadka, D., Reed, P., 2013. Borg: An Auto-Adaptive Many-Objective Evolutionary Computing Framework. Evolut. Comput. 21(2), 231-259.

Halhal, D., Walters, G., Ouazar, D., Savic, D., 1997. Water Network Rehabilitation with Structured Messy Genetic Algorithm. J. Water Resour. Plann. Manage. 123(3), 137-146.

Hu, X., Eberhart, R.C., Shi, Y., 2003. Swarm intelligence for permutation optimization: a case study of n-queens problem, Proceedings of the 2003 IEEE Swarm Intelligence Symposium, 243-246.

Iorio, A.W., Li, X., 2005. Solving rotated multi-objective optimization problems using differential evolution, AI 2004: Advances in Artificial Intelligence. Springer, 861-872.

Keedwell, E.C., Khu, S.T., 2003. More choices in water system design through hybrid optimisation, In: Maksimovic, C., Butler, D., Memon, F.A. (Eds.), Computing and Control for the Water Industry 2003. Taylor & Francis: London, UK, pp. 257-264.

Kennedy, J., Eberhart, R.C., 2001. Swarm intelligence. Morgan Kaufmann Publishers Inc., San Francisco.

Kim, J.H., Kim, T.G., Kim, J.H., Yoon, Y.N., 1994. A study on the pipe network system design using non-linear programming. J. Korean Water Resour. Assoc. 27(4), 59-67.

Knowles, J., Corne, D., 2002. On metrics for comparing nondominated sets, Proceedings of the 2002 Congress on Evolutionary Computation, pp. 711-716.

Kollat, J.B., Reed, P., 2007. A framework for Visually Interactive Decision-making and Design using Evolutionary Multi-objective Optimization (VIDEO). Environ. Modell. Softw. 22(12), 1691-1704.

Kollat, J.B., Reed, P.M., 2006. Comparing state-of-the-art evolutionary multi-objective algorithms for long-term groundwater monitoring design. Adv. Water Resour. 29(6), 792-807.

Laumanns, M., Thiele, L., Deb, K., Zitzler, E., 2002. Combining convergence and diversity in evolutionary multiobjective optimization. Evolut. Comput. 10(3), 263-282.

Lee, S.C., Lee, S.I., 2001. Genetic algorithms for optimal augmentation of water distribution networks. J. Korean Water Resour. Assoc. 34(5), 567-575.

Marchi, A., Salomons, E., Ostfeld, A. Kapelan, Z., Simpson, A.R., Zecchin, A.C., Maier, H.R., Wu, Z.Y., Elsayed, S.M., Song, Y., Walski, T., Stokes, C., Wu, W., Dandy, G.C., Alvisi, S., Creaco, E., Franchini, M., Saldarriaga, J., Páez, D., Hernández, D., Bohórquez, J., Bent, R., Coffrin, C., Judi, D., McPherson, T., van Hentenryck, P., Matos, J.P., Monteiro. A.J., Matias, N., Yoo, D.G., Lee, H.M., Kim, J.H., Iglesias-Rey, P.L., Martínez-Solano, F.J., Mora-Meliá, D., Ribelles-Aguilar, J.V., Guidolin, M., Fu, G., Reed, P., Wang, Q., Liu, H., McClymont, K., Johns, M., Keedwell, E., Kandiah, V., Jasper, M.N., Drake, K., Shafiee, E., Barandouzi, M.A., Berglund, A.D., Brill, D., Mahinthakumar, G., Ranjithan, R., Zechman, E.M., Morley, M.S., Tricarico, C., Tolson, B.A., Khedr, A. and Asadzadeh, M., 2013, "The Battle of the Water Networks II (BWN-II)", ASCE Journal of Water Resources Planning and Management, doi: 10.1061/(ASCE)WR.1943-5452.0000378.

Montalvoa, I., Izquierdo, J., Schwarzeb, S., Pérez-García, R., 2010. Multi-objective particle swarm optimization applied to water distribution systems design: An approach with human interaction. Math. Comput. Model. 52(7-8), 1219-1227.

Papadimitriou, C.H., Steiglitz, K., 1998. Combinatorial Optimization: Algorithms and Complexity. Dover Publications, New York.

Parsopoulos, K.E., Vrahatis, M.N., 2002. Particle swarm optimization method in multiobjective problems, Proceedings of the 2002 ACM symposium on Applied computing. ACM: Madrid, Spain.

Perelman, L., Ostfeld, A., Salomons, E., 2008. Cross Entropy multiobjective optimization for water distribution systems design. Water Resour. Res. 44(9), W09413.

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

Prasad, T.D., Tanyimboh, T., T., 2008. Entropy Based Design of "Anytown" Water Distribution Network, In: Van Zyl, J.E., Ilemobade, A.A., Jacobs, H.E. (Eds.), Water Distribution Systems Analysis 2008. ASCE: Kruger National Park, South Africa, pp. 450-461.

Purshouse, R.C., Fleming, P.J., 2007. On the Evolutionary Optimization of Many Conflicting Objectives. IEEE T. Evolut. Comput. 11(6), 770-784.

Raad, D., Sinske, A., van Vuuren, J., 2009. Robust multi-objective optimization for water distribution system design using a meta-metaheuristic. Int. Trans. Oper. Res. 16(5), 595-626.

Raad, D.N., Sinske, A., van Vuuren, J.H., 2011. Water distribution systems design optimisation using metaheuristics and hyperheuristics. ORiON 27(1), 17-44.

Raad, D.N., Sinske, A.N., van Vuuren, J.H., 2010. Comparison of four reliability surrogate measures for water distribution systems design. Water Resour. Res. 46(5), W05524.

Reca, J., Martínez, J., 2006. Genetic algorithms for the design of looped irrigation water distribution networks. Water Resour. Res. 42(5), W05416.

Reed, P.M., Hadka, D., Herman, J.D., Kasprzyk, J.R., Kollat, J.B., 2013. Evolutionary multiobjective optimization in water resources: The past, present, and future. Adv. Water Resour. 51(0), 438-456.

Rossman, L.A., 2000. EPANET 2 USERS MANUAL, U.S. Environment Protection Agency, Cincinnati, Ohio, USA.

Schaake, J., Lai, D., 1969. Linear programming and dynamic programming application to water distribution network design, Dept. of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Sherali, H., Subramanian, S., Loganathan, G.V., 2001. Effective Relaxations and Partitioning Schemes for Solving Water Distribution Network Design Problems to Global Optimality. J. Global Optim. 19(1), 1-26.

Srivastava, R.P., 2002. Time continuation in genetic algorithms. Technical report. Illinois Genetic Algorithm Laboratory.

Storn, R., Price, K., 1997. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. J. Global Optim. 11, 341-359.

Talbi, E.G., 2002. A Taxonomy of Hybrid Metaheuristics. J. Heuristics 8, 541-564.

Tang, Y., Reed, P., Wagener, T., 2006. How effective and efficient are multiobjective evolutionary algorithms at hydrologic model calibration? Hydrol. Earth Syst. Sci. 10(2), 289-307.

Todini, E., 2000. Looped water distribution networks design using a resilience index based heuristic approach. Urban Water 2(2), 115-122.

Vrugt, J.A., Robinson, B.A., 2007. Improved Evolutionary Optimization from Genetically Adaptive Multimethod Search. Proc. Natl. Acad. Sci. U.S.A. 104(3), 708-711.

Walski, T.M., 2001. The Wrong Paradigm, Why Water Distribution Optimization Doesn't Work. J. Water Resour. Plann. Manage. 127(4), 203-205.

Walski, T.M., Brill, J.E.D., Gessler, J., Goulter, I.C., Jeppson, R.M., Lansey, K., Lee, H.-L., Liebman, J.C., Mays, L., Morgan, D.R., Ormsbee, L., 1987. Battle of the Network Models: Epilogue. J. Water Resour. Plann. Manage. 113(2), 191-203.

Wang, Q., Liu, S., Wang, H., Savić, D., 2012. Multi-Objective Cuckoo Search for the Optimal Design of Water Distribution Systems, Civil Engineering and Urban Planning 2012, pp. 402-405.

Wang, Q., Savić, D.A., Kapelan, Z., in press. Hybrid Metaheuristics for Multi-Objective Design of Water Distribution Systems. J. Hydroinform. doi:10.2166/hydro.2013.009.

Wolpert, D.H., Macready, W.G., 1997. No free lunch theorems for optimization. IEEE T. Evolut. Comput. 1(1), 67-82.

Zitzler, E., Laumanns, M., Thiele, L., 2002. SPEA2: Improving the strength pareto evolutionary algorithm for multiobjective optimization, In: Giannakoglou, K., Tsahalis, D., Periaux, J., Papailiou, K., Fogarty, T. (Eds.), Evolutionary Methods for Design, Optimisation and Control. International Center for Numerical Methods in Engineering (CIMNE): Barcelona, Spain, pp. 95-100.

Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C.M., da Fonseca, V.G., 2003. Performance assessment of multiobjective optimizers: an analysis and review. IEEE T. Evolut. Comput. 7(2), 117-132.

To be cited as: Wang, Q., Guidolin, M., Savic, D., and Kapelan, Z. (2015). "Two-Objective Design of Benchmark Problems of a Water Distribution System via MOEAs: Towards the Best-Known Approximation of the True Pareto Front." J. Water Resour. Plann. Manage., 141(3), 04014060.

Appendices

Figures

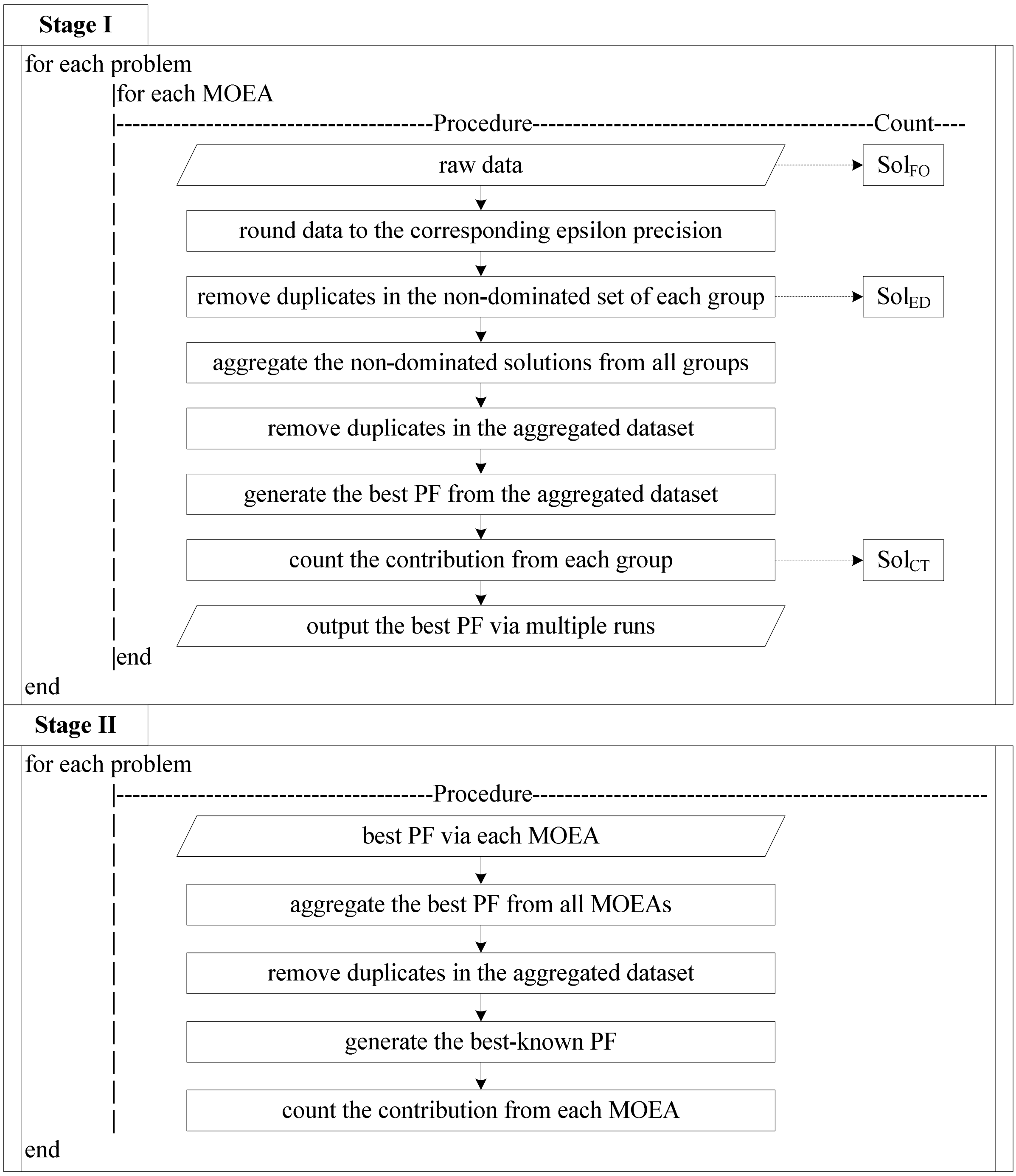
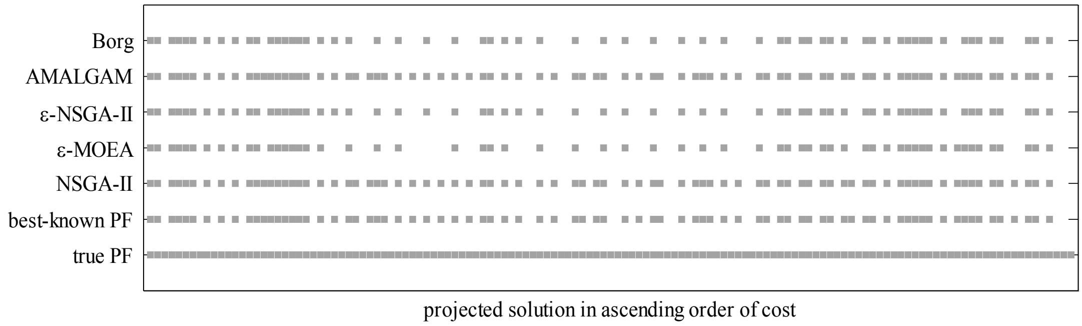
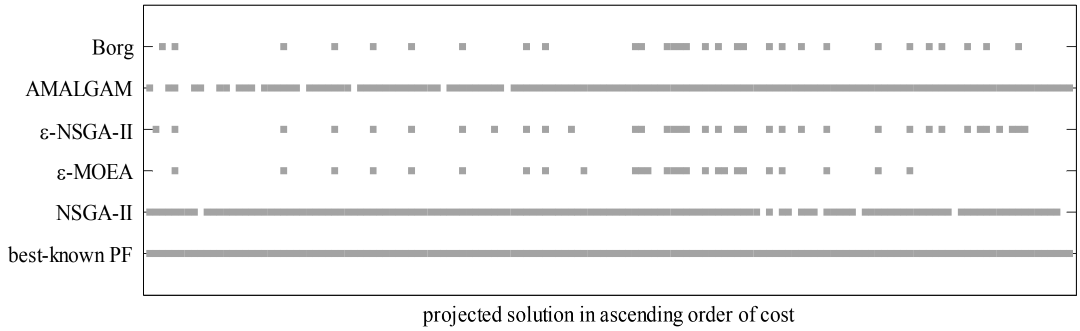


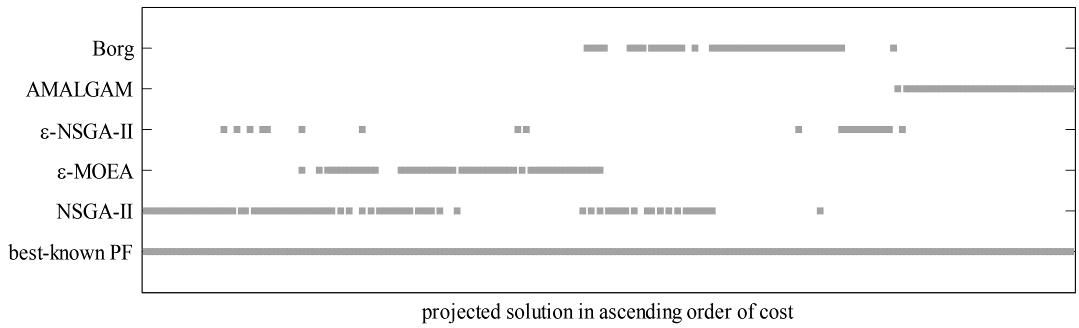
Fig. 1. Flowchart of data post-processing



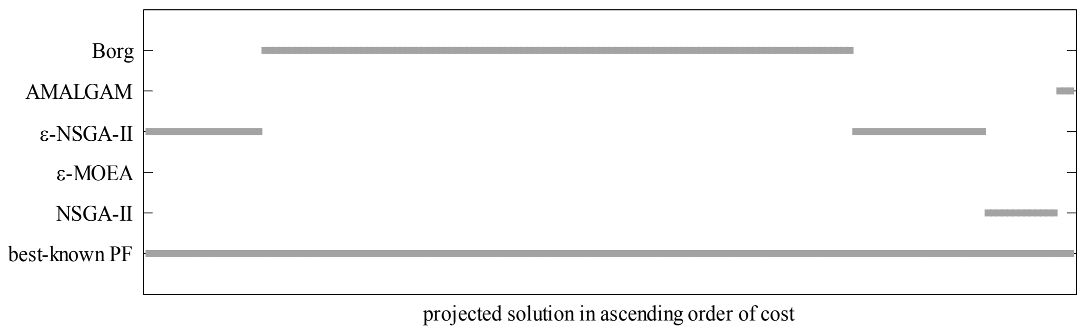
(a) BAK Problem



(b) NYT Problem

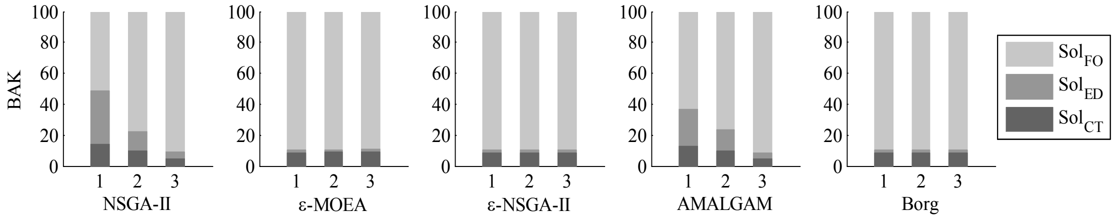


(c) PES Problem

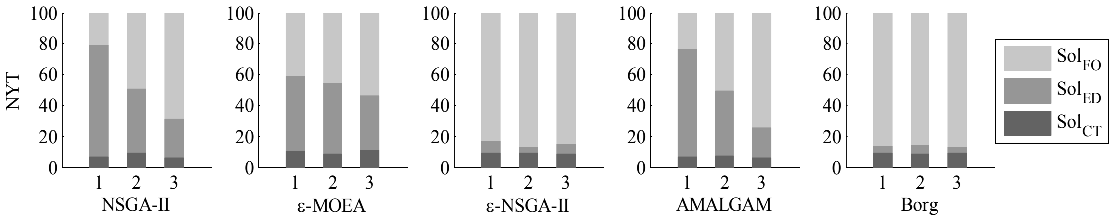


(d) EXN Problem

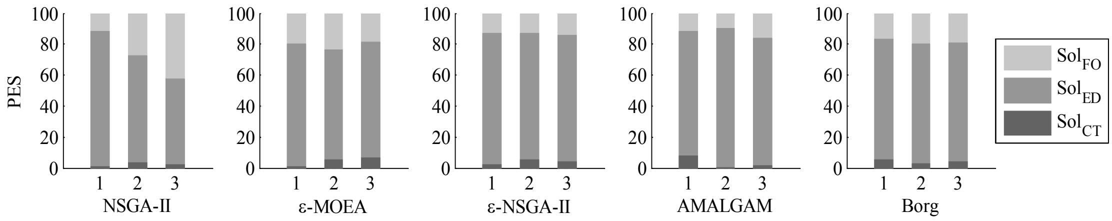
Fig. 2. Distribution of non-dominated solutions from each MOEA in the best-known PF



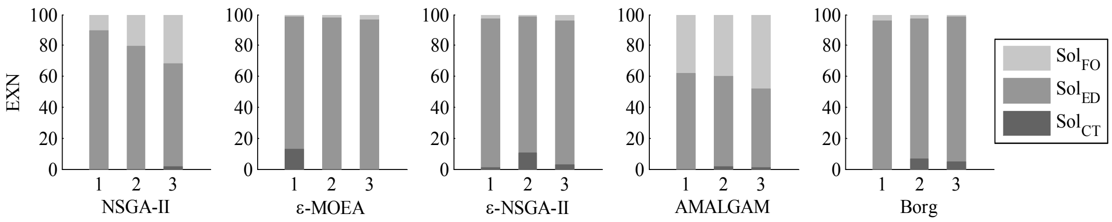
(a) BAK Problem



(b) NYT Problem

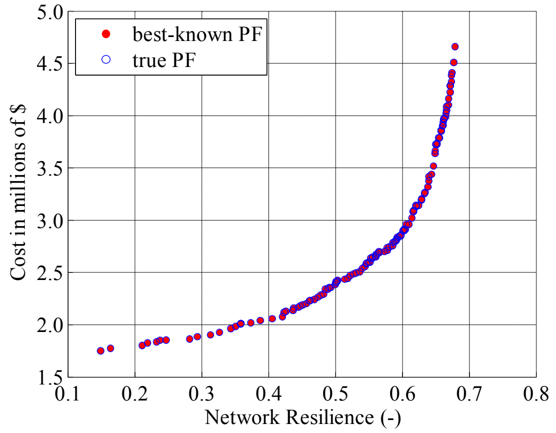


(c) PES Problem

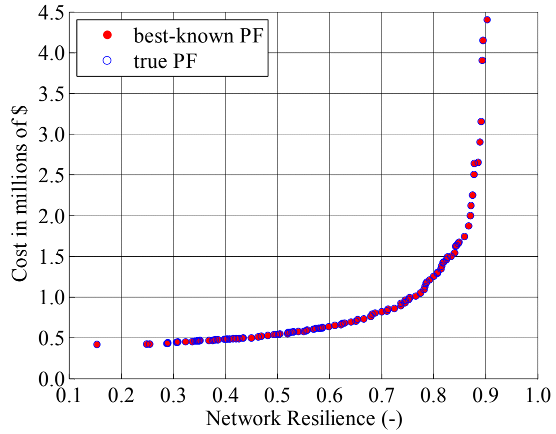


(d) EXN Problem

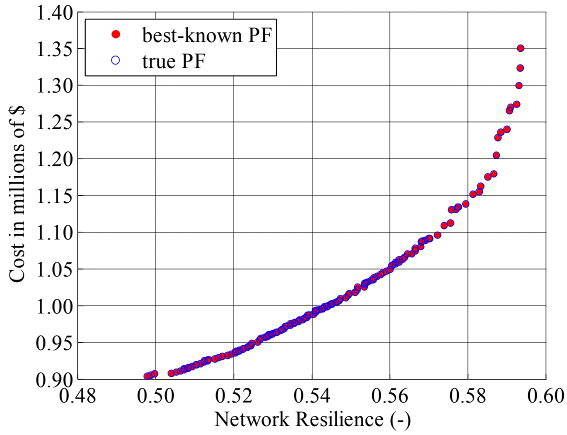
Fig. 3. Variation in the number of solutions in percentage during data post-processing (based on the data given in Table B.1 in Appendix B)



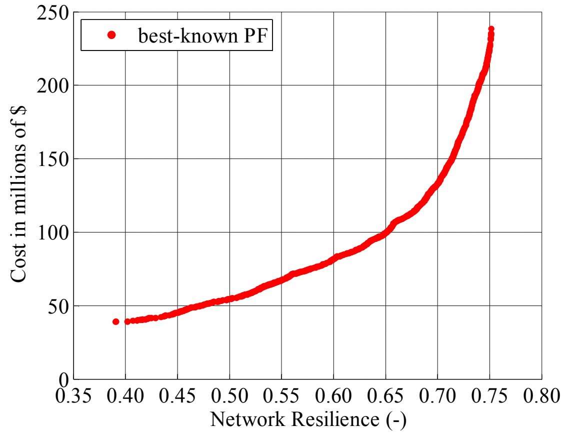
(a) TRN Problem



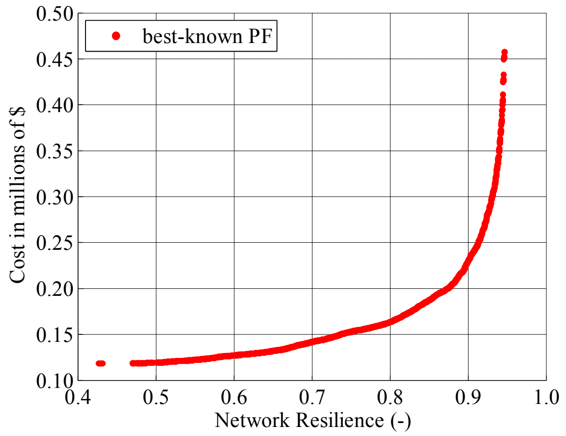
(b) TLN Problem



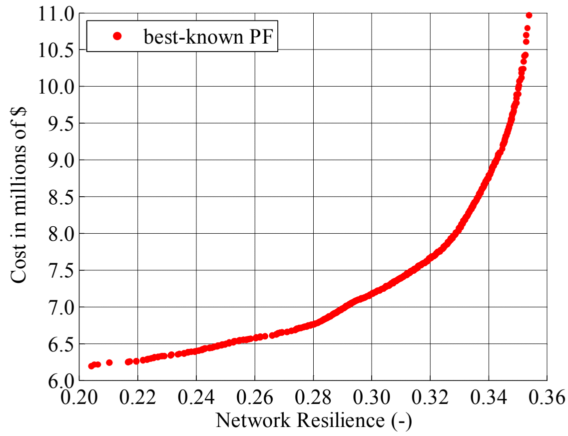
(c) BAK Problem



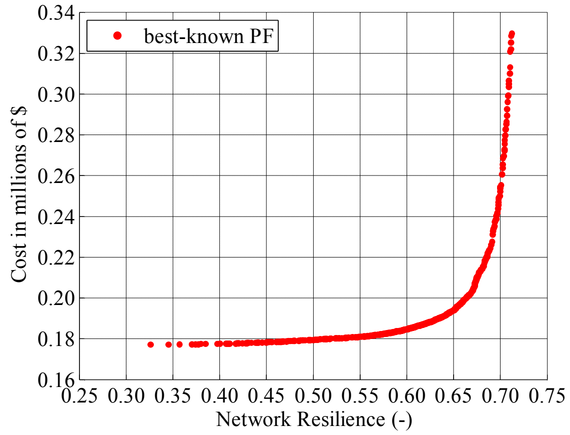
(d) NYT Problem



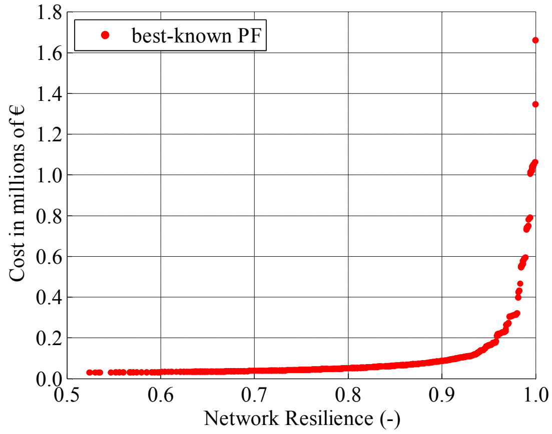
(e) BLA Problem



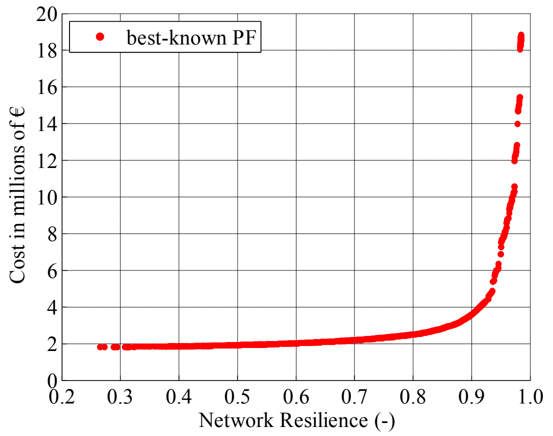
(f) HAN Problem



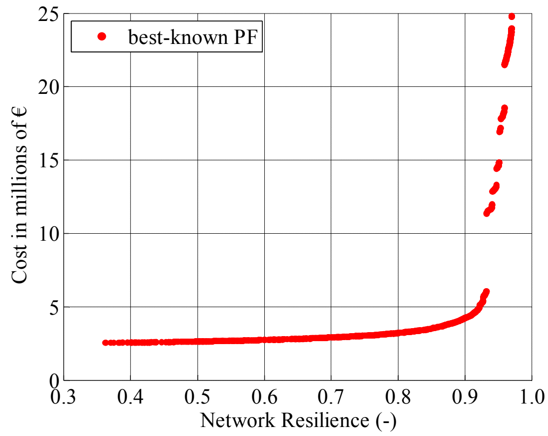
(g) GOY Problem



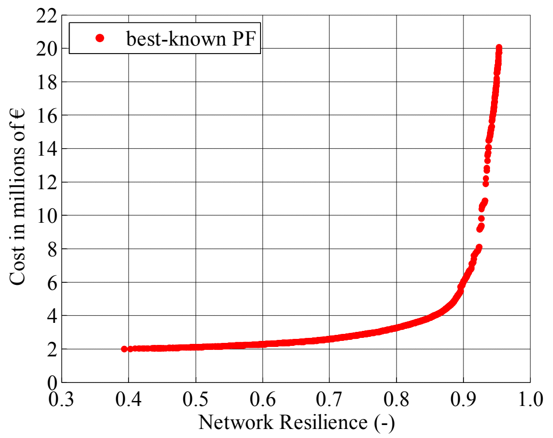
(h) FOS Problem



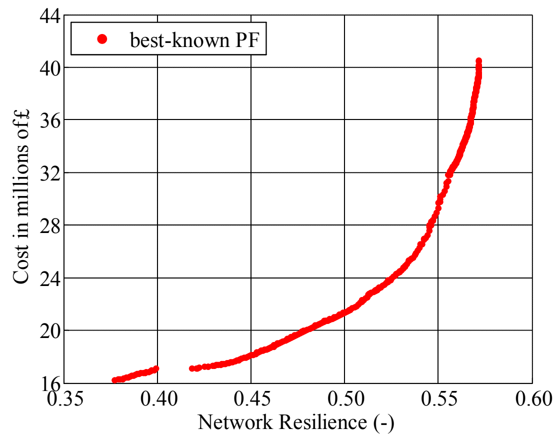
(i) PES Problem



(j) MOD Problem



(k) BIN Problem



(l) EXN Problem

Fig. A.1 Best-known PF of each benchmark problem

To be cited as: Wang, Q., Guidolin, M., Savic, D., and Kapelan, Z. (2015). "Two-Objective Design of Benchmark Problems of a Water Distribution System via MOEAs: Towards the Best-Known Approximation of the True Pareto Front." J. Water Resour. Plann. Manage., 141(3), 04014060.

Tables

Table 1. Benchmark design problems considered in the paper

Type	Problem	Acronym	Number of				Search Space Size
			LC	WS	DV	PD	
SP	Two-Reservoir Network	TRN	3	2	8	8*	3.28×10^{7}
	Two-Loop Network	TLN	1	1	8	14	1.48×10^{9}
	BakRyan Network	BAK	1	1	9	11	2.36×10^{9}
MP	New York Tunnel Network	NYT	1	1	21	16	1.93×10^{25}
	Blacksburg Network	BLA	1	1	23	14	2.30×10^{26}
	Hanoi Network	HAN	1	1	34	6	2.87×10^{26}
	GoYang Network	GOY	1	1	30	8	1.24×10^{27}
IP	Fossolo Network	FOS	1	1	58	22	7.25×10^{77}
	Pescara Network	PES	1	3	99	13	1.91×10^{110}
LP	Modena Network	MOD	1	4	317	13	1.32×10^{353}
	Balerma Irrigation Network	BIN	1	4	454	10	1.00×10^{455}
	Exeter Network	EXN	1	7	567	11	2.95×10^{590}

Note: SP-Small Problems; MP-Medium Problems; IP-Intermediate Problems; LP-Larger Problems; LC-number of loading conditions; WS-number of water sources; DV-number of decision variables; PD-number of pipe diameter options. *For TRN problem, three existing pipes have 8 diameter options for duplication and 2 extra options, i.e. cleaning and leaving alone.