

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.

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

3 Qi Wang¹, Michele Guidolin², Dragan Savić³, Zoran Kapelan⁴

4 Abstract: Various multi-objective evolutionary algorithms (MOEAs) have been applied to
5 solve the optimal design problems of a Water Distribution System (WDS). Such methods are
6 able to find the near-optimal trade-off between cost and performance benefit in a single run.
7 Previously published work used a number of small benchmark networks and/or a few large,
8 real-world networks to test MOEAs on design problems of WDS. A few studies also focused
9 on the comparison of different MOEAs given a limited computational budget. However, no
10 consistent attempt has been made before to investigate and report the best-known
11 approximation of the true Pareto front (PF) for a set of benchmarks problems, and thus there
12 is not a single point of reference. This paper applied five state-of-the-art MOEAs, with
13 minimum time invested in parameterisation (i.e., using the recommended settings), to twelve
14 design problems collected from the literature. Three different population sizes were
15 implemented for each MOEA with respect to the scale of each problem. The true Pareto
16 fronts for small problems and the best-known Pareto fronts for the other problems were
17 obtained. Five MOEAs were complementary to each other on various problems, which
18 implies that no one method was completely superior to the others. The non-dominated sorting
19 genetic algorithm-II (NSGA-II), with minimum parameters tuning, remains a good choice as
20 it showed generally the best achievements across all the problems. In addition, a small
21 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

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.

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

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

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

30

31 Introduction

32 Multi-objective design of a Water Distribution System (WDS) has received increasing
33 attention during the past two decades. Much effort (Farmani et al., 2006; Fu et al., 2012a;
34 Halhal et al., 1997; Keedwell and Khu, 2003; Prasad and Park, 2004; Prasad and Tanyimboh,
35 2008) has been made to identify the trade-off, expressed as the Pareto front (PF), between
36 cost and performance type benefit using various indicators. Multi-objective evolutionary
37 algorithms (MOEAs) are widely accepted for addressing this kind of problem as they are
38 capable of approximating the PF effectively and efficiently in a single run (Farmani et al.,
39 2005a). Many benchmark networks and some real networks have been used to demonstrate
40 the strength of MOEAs. For instance, Cheung et al. (2003) applied both strength Pareto
41 evolutionary algorithm (SPEA) and multi-objective genetic algorithm (MOGA) to the
42 rehabilitation problem of a hypothetical network. Minimisation of cost and the total pressure
43 deficit were taken as two objectives. Farmani et al. (2004) contributed a new benchmark
44 network based on a real system and used the non-dominated sorting genetic algorithm-II
45 (NSGA-II) to solve the two-objective rehabilitation of this large network, minimising cost
46 and number of nodes with head deficiency. Besides using pressure deficit and its analogues
47 as the second objective, other formulations were aimed at optimising resilience based
48 indicators (Basupi et al., 2013; Farmani et al., 2005b; Prasad and Park, 2004), flow entropy
49 (Prasad and Tanyimboh, 2008), and other mixed indicators (Raad et al., 2010).

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.

50 Apart from the aforementioned MOEAs, other methods have been applied to solve
51 benchmark problems, such as SPEA2 (Farmani et al., 2003), Cross Entropy (Perelman et al.,
52 2008), Particle Swarm Optimisation (PSO) (Montalvoa et al., 2010), and Cuckoo search
53 (Wang et al., 2012). Recently, hybrid algorithms (Vrugt and Robinson, 2007; di Pierro et al.,
54 2009; Hadka and Reed, 2013), which combine different schemes and components together in
55 an attempt to further enhance search ability, have demonstrated significant improvement over
56 previous MOEAs, like NSGA-II (Deb et al., 2002) and SPEA2 (Zitzler et al., 2002). The
57 encouraging performance of these newly-proposed MOEAs has gained interest for the design
58 of WDS. Raad et al. (2009) for the first time applied a modified multi-algorithm, genetically
59 adaptive multi-objective (AMALGAM) to the two-objective design of WDS, using cost and
60 network resilience (Prasad and Park, 2004) as objectives. Wang et al. (in press) compared the
61 performance of two distinct hybrid algorithms (including the original AMALGAM) against
62 NSGA-II on a wide range of benchmark problems. Creaco and Franchini (2012; 2013) set up
63 a hybrid procedure where NSGA-II coordinates various subordinate algorithms to perform
64 the multi-objective design under pressure and velocity constraints. Fu et al (2012b) proposed
65 a novel hybrid approach where global sensitivity analysis is used before applying the ϵ -
66 NSGA-II method to reduce the complexity of the search space size of a multi-objective WDS
67 design problem.

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

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.

81 literature. In addition, for small problems full enumeration is implemented to produce the true
82 PF.

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

93 Multi-objective design of a WDS

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

103 Historically, the design problem was treated as single-objective optimisation focusing on
104 economic considerations, but the drawbacks of this formulation have been criticised broadly
105 (Engelhardt et al., 2000; Fu et al., 2012a; Walski, 2001). In transforming the least-cost
106 formulation to the multi-objective, or more precisely, two-objective formulation, many
107 indicators have been proposed as additional objectives. At the very beginning, variants of
108 pressure deficit were used to account for the second objective, such as minimising the total
109 pressure deficit, minimising the maximum of pressure deficiency, and minimising the number
110 of nodes with pressure deficit (Cheung et al., 2003; Farmani et al., 2005a). However, these
111 aforementioned formulations do not necessarily result in looped networks, which are reliable

112 configurations under abnormal conditions (e.g., pipe burst). On the other hand, a resilience
 113 index formulation (Todini, 2000) was introduced as a surrogate measure for hydraulic
 114 benefits. The index is based on the concept that the total input power into a network consists
 115 of the power dissipated in the network and the power delivered at demand nodes. So, less
 116 power consumed internally to overcome the friction results in more surplus power at demand
 117 nodes and thus being able to counter the failure scenarios. Later on, an improved version of
 118 the resilience indicator (Prasad and Park, 2004), called network resilience, was proposed
 119 taking the uniformity of pipes around each demand node into account. Network resilience
 120 considers the effect of redundancy of a pipe network and maximising this indicator can
 121 ensure reliable loops. It is proved that using network resilience as another objective alleviates
 122 the shortcomings of the resilience index (e.g., resulting in impracticable loops) and yields the
 123 solutions which are robust under pipe failure conditions (Prasad and Park, 2004; Raad et al.,
 124 2010). For this reason, network resilience is used as the second objective for the two-
 125 objective optimisation of benchmark problems.

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

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

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

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

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

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

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.

137 γ =specific weight of water; n_{pj} =number of pipes connected to node j ; D_i =diameter of pipe i
138 connected to demand node j .

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

154 In these problems, the two-loop network (Alperovits and Shamir, 1977) is a hypothetical
155 network, while the others are real networks or simplified networks in the real-world. The
156 New York tunnel network (Schaake and Lai, 1969) and the Exeter network (Farmani et al.,
157 2004) were originally presented as **extended design** problems. The rest are design problems
158 except the BakRyan network (Lee, 2001) and the Two-Reservoir network (Gessler, 1985)
159 which are a mix of design and **extended design**. There are both minimum and maximum
160 pressure requirements for demand nodes in the Blacksburg network (Sherali et al., 2001), the
161 Fossolo network (Bragalli et al., 2008), the Pescara network (Bragalli et al., 2008), and the
162 Modena network (Bragalli et al., 2008), whereas the others only have minimum pressure
163 requirements. In addition, there are upper bounds for velocity in the pipes of the four
164 aforementioned networks. The Hanoi network and the GoYang network are taken from
165 (Fujiwara and Khang, 1990) and (Kim et al., 1994), respectively. Unlike a WDS, the water
166 consumption is fixed at 5.55 l/s across all the demand nodes in the Balerma irrigation
167 network (Reca and Martínez, 2006).

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.

168 [Table 1 goes here]

169 MOEAs used in the analysis

170 In this paper, five MOEAs in total are applied to solve the benchmark problems. Two of them
171 are high-level hybrid algorithms (Talbi, 2002), namely AMALGAM (Vrugt and Robinson,
172 2007) and Borg (Hadka and Reed, 2013). The others are NSGA-II (Deb et al., 2002), ϵ -
173 MOEA (Deb et al., 2005), and ϵ -NSGA-II (Kollat and Reed, 2006), which is an enhanced
174 algorithm based on NSGA-II.

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

193 Hybrid MOEAs

194 *AMALGAM*

195 AMALGAM is a hybrid optimisation framework which employs simultaneously four sub-
196 algorithms within its structure, including NSGA-II, adaptive metropolis search (AMS)
197 (Haario et al., 2001), particle swarm optimisation (PSO) (Kennedy and Eberhart, 2001) and

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.

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

210 *Borg*

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

219 The advantages of Borg are threefold: (1) usage of ε -box dominance archive contributes to
220 maintaining the convergence and diversity concurrently throughout search; (2) the
221 combination of time continuation (Srivastava, 2002), adaptive population sizing, and two
222 types of randomised restart (i.e. ε -progress triggered restart and population-to-archive ratio
223 triggered restart) boosts the algorithm towards global optima; (3) simultaneous employment
224 of multiple recombination operators enhances performance on a wide assortment of problem
225 domains. In addition, the adoption of the steady-state, elitist model of ε -MOEA (Deb et al.,
226 2005) make it easily extendable for use on parallel architectures. Borg has also been
227 successfully used to solve challenging, many-objective, real-world problems in the domain of
228 water resources, a detailed review of which can be found in (Reed et al., 2013).

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.

229 Referential MOEAs

230 *NSGA-II*

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

241 ε -MOEA

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

254 ε -NSGA-II

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

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.

259 NSGA-II primarily in two aspects: (1) while the population evolves in the same manner as
260 NSGA-II, an offline archive is frequently updated by selecting the ε -non-dominated
261 solutions from the elitist population at current generation; (2) the optimisation is implemented
262 in consecutive epochs, each of which is terminated automatically according to the user-
263 specified improvement criteria. The next epoch is populated by injecting the members in the
264 archive and generating new random solutions. The ε -dominance archiving maintains the
265 convergence and diversity of the archive concurrently. It also allows users to specify the
266 precision of each objective and thus is more flexible in practice. Adaptive population sizing
267 contributes to balance the exploration and exploitation throughout the search, which is
268 achieved by increasing or decreasing the capacity of the population based on the number of
269 members in the archive. Additionally, several connected runs, known as time continuation
270 (Srivastava, 2002), enhance the possibility to explore other regions of search space. The
271 comparative study on ε -NSGA-II as well as three benchmark MOEAs (NSGA-II, ε -MOEA
272 and SPEA2) showed its superiority in terms of efficiency and reliability (Kollat and Reed,
273 2006). Moreover, the aforementioned key components of ε -NSGA-II remedy the issue of
274 parameterisation commonly found in MOEAs, thus making it easy-to-use for a wide range of
275 applications.

276 Benchmarking setup

277 Each benchmark problem was formulated as two-objective design or **extended design**, taking
278 total cost and network resilience into account. For a design problem, the decision variables
279 were the diameters of individual pipes in the network. While for **an extended design** problem,
280 the decision variables included the diameters of duplicate pipes as well as the other two
281 options, i.e. leaving alone (do-nothing option) and cleaning of existing pipes. Note that all the
282 MOEAs used real coding but all the decision variables were of integer type. So the real
283 values passed in by MOEAs were rounded down (e.g., 12.9457 becomes 12). For each
284 problem, a solution was considered as infeasible if there were violations of pressure
285 requirements (minimum and maximum if any) and upper bound of flow velocity (if any).
286 **Note that no penalty function is used to handle infeasible solutions; instead, infeasibility**
287 **situations are dealt with by implementing the constrained-domination principle (Deb et al.,**
288 **2002).**

289 Parameter settings of MOEAs

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

316 In addition, all the algorithms considered in this paper share similarities in that they use
317 tournament selection, real-valued simulated binary crossover (SBX), polynomial mutation
318 (PM) (Deb et al., 2002). Therefore, unified settings of these factors are kept the same for all
319 the algorithms. More specifically, a tournament size of 2 is applied except for Borg and ε -
320 NSGA-II (tournament size changed due to adaptive population sizing). The probabilities of
321 SBX and PM are 0.9 and the inverse of the number of decision variables respectively, and the

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.

322 distribution index of SBX and PM are 15 and 7 respectively. Note that these values are
323 selected according to the most commonly recommended parameter settings in the literature
324 (Deb et al., 2002; Kollat and Reed, 2006). A smaller distribution index of SBX or PM enables
325 the search operator to create solutions with more spread (Deb and Agrawal, 1995), hence it is
326 effective to avoid premature.

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

330 [Table 2 goes here]

331 Epsilon precision

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

336 [Table 3 goes here]

337 Computational budget

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

350 [Table 4 goes here]

351 Performance assessment

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

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

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

372 [Figure 1 goes here]

373 As there is no reference set of the true PFs to hand, various performance indicators existing in
374 the literature (Deb et al., 2002; Knowles and Corne, 2002; Zitzler et al., 2003) cannot be
375 applied directly. Therefore, the number of solutions contributed by each MOEA to the best-
376 known PF of each problem is counted. This will demonstrate the general capability (including
377 convergence and diversity) of an MOEA to find the optimal non-dominated solutions to a
378 problem. On the other hand, as mentioned before, for each MOEA three different population
379 sizes are implemented for each benchmark problem (see Table 4). To investigate the impact
380 of population size on the performance of each MOEA, the number of solutions contributed

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.

381 from each population size to the best PF is also counted. It is also worth noting that the
382 computational time spent by each MOEA on each problem are not compared due to the fact
383 that these algorithms were developed in different languages and implemented on various
384 machines with different operating systems. For example, AMALGAM was built in *Matlab*
385 and run on a desktop computer (Windows 7) with 2.66GHz and 3GB RAM. While a parallel
386 version of ϵ -NSGA-II was written in *C* language and executed on a supercomputer, called
387 "Zen", which consists of diskless compute nodes each with twelve cores and 24GB of RAM.
388 Instead, a rough observation about the runtime spent on algorithm steps and objective
389 function evaluations (including hydraulic simulations) is provided as follows. Generally
390 speaking, all the methods spent a higher proportion of CPU time on the objective function
391 evaluations for large problems. AMALGAM took more CPU time on algorithm steps on
392 average than the *C* language based MOEAs because it was developed and implemented in
393 *Matlab* which is an interpreted language. For the *C* language based MOEAs, Borg and ϵ -
394 MOEA spent less CPU time on algorithm steps compared with NSGA-II and ϵ -NSGA-II
395 (non-Parallel version) because they followed the steady-state algorithmic framework, which
396 did not involve time-consuming ranking and sorting as in NSGA-II and ϵ -NSGA-II.

397 In addition, since the number of solutions from each MOEA alone cannot demonstrate the
398 distribution of solutions in the best-known PF, a novel projection plot is developed to
399 illustrate the distribution of solutions contributed by a specific MOEA. A clear advantage of
400 using this projection plot is that it can deliver the preferred information of convergence
401 (secondary) and diversity but avoid showing the overlaps between different Pareto fronts,
402 when they are drawn in the same objective space (commonly seen in comparative studies).
403 The procedure of generating such a projection plot is explained as follows. For each problem,
404 firstly, data in the best-known PF are sorted according to the values of either objective in
405 ascending or descending order. Here the ascending order of cost objective was chosen for the
406 purpose of demonstration. Note that the other objective is inevitably ignored by using the
407 projection plot, but this will not affect the interpretation of the results. Then, these solutions
408 (in the space of *Cost* vs. I_n) are evenly projected on to a 1-D axis, from 1 to the length of data
409 set (i.e., the number of solutions in the best-known PF). Next, by comparing the overlaps
410 (duplicates) of solutions from each MOEA with those in the best-known PF, the
411 corresponding positions of solutions contributed from each MOEA on the 1-D axis can be
412 identified. Finally, these solutions from each MOEA are also projected on to the 1-D axis in a

413 stacking fashion (as shown in Fig. 2). It is worth noting that the convergence of each
414 algorithm is implicitly considered by using the novel projection plot, because only the fully
415 converged solutions with respect to the non-dominated solutions in the best-known PF are
416 shown for each MOEA. Therefore, this approach **facilitates** a direct comparison of the
417 relative distribution of solutions from each MOEA.

418 Results and Discussion

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

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

447 [Table 5 goes here]

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

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

471 [Figure 2 goes here]

472 As shown in Fig. 1, variations in the number of solutions during data post-processing are
473 recorded for further analysis. The number of solutions in different steps varies due to the
474 existence of duplicates in the non-dominated sets and the stochastic nature of MOEAs. For

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.

475 non- ε -dominance based MOEAs, Sol_{FO} is equal to the size of each group times the number
476 of runs (10 in this paper). While Sol_{FO} of ε -MOEA is the number of solutions in the
477 aggregated archive across multiple runs. This figure is expected to be different from those of
478 NSGA-II and AMALGAM as the size of archive keeps changing during optimisation. Similar
479 to ε -MOEA, the archive sizes of ε -NSGA-II and Borg change over time due to the strategy
480 of adaptive population sizing with time continuation.

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

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

504 [Figure 3 goes here]

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.

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

510 Conclusions

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

524 On the other hand, the strength of MOEAs tested in the paper, including two modern hybrid
525 MOEAs and three frequently used MOEAs, was compared in the context of optimal design of
526 Water Distribution System. The results obtained proved that NSGA-II remains one of the best
527 MOEAs, which is suitable for two-objective optimisation of a Water Distribution System. It
528 generally outperformed the other MOEAs in terms of the number of solutions contributed to
529 the best-known PF of each problem. The spread (both extent and uniformity) of its
530 contribution was also comparable, if not better, to those of other MOEAs. AMALGAM is
531 promising for this task as it contributes more than 85% non-dominated solutions in the best-
532 known PFs for small and medium problems consistently. It always discovered solutions in the
533 region of high network resilience, although there is a clear drop in performance on
534 intermediate and large problems. Three ε -dominance based MOEAs failed to demonstrate
535 any clear advantage over NSGA-II and AMALGAM in the experiment. Nevertheless, all of

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.

536 them showed evident advantage in the convergence and efficiency of non-dominated
537 solutions for small and medium problems. Besides, Borg was shown to be exceptionally
538 superior to the other MOEAs for the EXN problem by finding more than 60% solutions in the
539 best-known PF. In short, the MOEAs considered were complementary to each other
540 especially for complex cases, albeit no versatile MOEA was found in this study. Therefore,
541 when facing large sized problems, different MOEAs should be considered to ensure a reliable
542 Pareto front if both time and computational resources are available.

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

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

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

558 Acknowledgements

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

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.

566

567 References

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

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

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

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

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

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

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

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

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

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

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

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.

- 598 Engelhardt, M.O., Skipworth, P.J., Savic, D.A., Saul, A.J., Walters, G.A., 2000.
599 Rehabilitation strategies for water distribution networks: a literature review with a UK
600 perspective. *Urban Water* 2(2), 153-170.
- 601 Farmani, R., Savic, D., Walters, G.A., 2006. A Hybrid Technique For Optimisation Of
602 Branched Urban Water Systems, 7th International Conference on Hydroinformatics:
603 Nice, France.
- 604 Farmani, R., Savic, D.A., Walters, G.A., 2003. Multi-objective optimization of water system:
605 A comparative study, In: Jr., C.a.C. (Ed.), *Pumps, Electromechanical Devices and*
606 *Systems Applied to Urban Water Management*, pp. 247-256.
- 607 Farmani, R., Savic, D.A., Walters, G.A., 2004. "EXNET" Benchmark Problem for Multi-
608 Objective Optimization of Large Water Systems, *Modelling and Control for*
609 *Participatory Planning and Managing Water Systems*, IFAC workshop: Venice, Italy.
- 610 Farmani, R., Savic, D.A., Walters, G.A., 2005a. Evolutionary multi-objective optimization in
611 water distribution network design. *Eng. Optimiz.* 37(2), 167-183.
- 612 Farmani, R., Walters, G.A., Savic, D.A., 2005b. Trade-Off between Total Cost and
613 Reliability for Anytown Water Distribution Network. *J. Water Resour. Plann. Manage.*
614 131(3), 161-171.
- 615 Fu, G., Kapelan, Z., Kasprzyk, J., Reed, P., 2012a. Optimal Design of Water Distribution
616 Systems Using Many-Objective Visual Analytics. *J. Water Resour. Plann. Manage.*
617 doi:10.1061/(ASCE)WR.1943-5452.0000311
- 618 Fu, G., Kapelan, Z. and Reed, P.M., 2012b, Reducing the complexity of multi-objective
619 water distribution system optimization through global sensitivity analysis, *Journal of*
620 *Water Resources Planning and Management*, 138(3), 196–207.
- 621 Fujiwara, O., Khang, D., 1990. A two-phase decomposition method for optimal design of
622 looped water distribution networks. *Water Resour. Res.* 26(4), 539-549.
- 623 Gessler, J., 1985. Pipe network optimization by enumeration, In *Proc., Computer*
624 *Applications for Water Resources*. ASCE: New York, N.Y., pp. 572-581.
- 625 Goldberg, D.E., 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning*.
626 Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.
- 627 Guidolin, M., Fu, G., Reed, P., 2012. Parallel Evolutionary Multiobjective Optimization of
628 Water Distribution System Design, *Water Distribution Systems Analysis 2012:*
629 *Adelaide, Australia*.

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.

- 630 Haario, H., Saksman, E., Tamminen, J., 2001. An adaptive Metropolis algorithm. Bernoulli
631 7(2), 223-242.
- 632 Hadka, D., Reed, P., 2013. Borg: An Auto-Adaptive Many-Objective Evolutionary
633 Computing Framework. Evolut. Comput. 21(2), 231-259.
- 634 Halhal, D., Walters, G., Ouazar, D., Savic, D., 1997. Water Network Rehabilitation with
635 Structured Messy Genetic Algorithm. J. Water Resour. Plann. Manage. 123(3), 137-
636 146.
- 637 Hu, X., Eberhart, R.C., Shi, Y., 2003. Swarm intelligence for permutation optimization: a
638 case study of n-queens problem, Proceedings of the 2003 IEEE Swarm Intelligence
639 Symposium, 243-246.
- 640 Iorio, A.W., Li, X., 2005. Solving rotated multi-objective optimization problems using
641 differential evolution, AI 2004: Advances in Artificial Intelligence. Springer, 861-872.
- 642 Keedwell, E.C., Khu, S.T., 2003. More choices in water system design through hybrid
643 optimisation, In: Maksimovic, C., Butler, D., Memon, F.A. (Eds.), Computing and
644 Control for the Water Industry 2003. Taylor & Francis: London, UK, pp. 257-264.
- 645 Kennedy, J., Eberhart, R.C., 2001. Swarm intelligence. Morgan Kaufmann Publishers Inc.,
646 San Francisco.
- 647 Kim, J.H., Kim, T.G., Kim, J.H., Yoon, Y.N., 1994. A study on the pipe network system
648 design using non-linear programming. J. Korean Water Resour. Assoc. 27(4), 59-67.
- 649 Knowles, J., Corne, D., 2002. On metrics for comparing nondominated sets, Proceedings of
650 the 2002 Congress on Evolutionary Computation, pp. 711-716.
- 651 Kollat, J.B., Reed, P., 2007. A framework for Visually Interactive Decision-making and
652 Design using Evolutionary Multi-objective Optimization (VIDEO). Environ. Modell.
653 Softw. 22(12), 1691-1704.
- 654 Kollat, J.B., Reed, P.M., 2006. Comparing state-of-the-art evolutionary multi-objective
655 algorithms for long-term groundwater monitoring design. Adv. Water Resour. 29(6),
656 792-807.
- 657 Laumanns, M., Thiele, L., Deb, K., Zitzler, E., 2002. Combining convergence and diversity
658 in evolutionary multiobjective optimization. Evolut. Comput. 10(3), 263-282.
- 659 Lee, S.C., Lee, S.I., 2001. Genetic algorithms for optimal augmentation of water distribution
660 networks. J. Korean Water Resour. Assoc. 34(5), 567-575.
- 661 Marchi, A., Salomons, E., Ostfeld, A., Kapelan, Z., Simpson, A.R., Zecchin, A.C., Maier,
662 H.R., Wu, Z.Y., Elsayed, S.M., Song, Y., Walski, T., Stokes, C., Wu, W., Dandy,

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.

- 663 G.C., Alvisi, S., Creaco, E., Franchini, M., Saldarriaga, J., Páez, D., Hernández, D.,
664 Bohórquez, J., Bent, R., Coffrin, C., Judi, D., McPherson, T., van Hentenryck, P.,
665 Matos, J.P., Monteiro, A.J., Matias, N., Yoo, D.G., Lee, H.M., Kim, J.H., Iglesias-
666 Rey, P.L., Martínez-Solano, F.J., Mora-Meliá, D., Ribelles-Aguilar, J.V., Guidolin,
667 M., Fu, G., Reed, P., Wang, Q., Liu, H., McClymont, K., Johns, M., Keedwell, E.,
668 Kandiah, V., Jasper, M.N., Drake, K., Shafiee, E., Barandouzi, M.A., Berglund, A.D.,
669 Brill, D., Mahinthakumar, G., Ranjithan, R., Zechman, E.M., Morley, M.S., Tricarico,
670 C., Tolson, B.A., Khedr, A. and Asadzadeh, M., 2013, "The Battle of the Water
671 Networks II (BWN-II)", *ASCE Journal of Water Resources Planning and*
672 *Management*, doi: 10.1061/(ASCE)WR.1943-5452.0000378.
- 673 Montalvoa, I., Izquierdo, J., Schwarzeb, S., Pérez-García, R., 2010. Multi-objective particle
674 swarm optimization applied to water distribution systems design: An approach with
675 human interaction. *Math. Comput. Model.* 52(7-8), 1219-1227.
- 676 Papadimitriou, C.H., Steiglitz, K., 1998. *Combinatorial Optimization: Algorithms and*
677 *Complexity*. Dover Publications, New York.
- 678 Parsopoulos, K.E., Vrahatis, M.N., 2002. Particle swarm optimization method in
679 multiobjective problems, *Proceedings of the 2002 ACM symposium on Applied*
680 *computing*. ACM: Madrid, Spain.
- 681 Perelman, L., Ostfeld, A., Salomons, E., 2008. Cross Entropy multiobjective optimization for
682 water distribution systems design. *Water Resour. Res.* 44(9), W09413.
- 683 Prasad, T.D., Park, N.-S., 2004. Multiobjective Genetic Algorithms for Design of Water
684 Distribution Networks. *J. Water Resour. Plann. Manage.* 130(1), 73-82.
- 685 Prasad, T.D., Tanyimboh, T., T., 2008. Entropy Based Design of "Anytown" Water
686 Distribution Network, In: Van Zyl, J.E., Ilemobade, A.A., Jacobs, H.E. (Eds.), *Water*
687 *Distribution Systems Analysis 2008*. ASCE: Kruger National Park, South Africa, pp.
688 450-461.
- 689 Purshouse, R.C., Fleming, P.J., 2007. On the Evolutionary Optimization of Many Conflicting
690 Objectives. *IEEE T. Evolut. Comput.* 11(6), 770-784.
- 691 Raad, D., Sinske, A., van Vuuren, J., 2009. Robust multi-objective optimization for water
692 distribution system design using a meta-metaheuristic. *Int. Trans. Oper. Res.* 16(5),
693 595-626.
- 694 Raad, D.N., Sinske, A., van Vuuren, J.H., 2011. Water distribution systems design
695 optimisation using metaheuristics and hyperheuristics. *ORiON* 27(1), 17-44.

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.

- 696 Raad, D.N., Sinske, A.N., van Vuuren, J.H., 2010. Comparison of four reliability surrogate
697 measures for water distribution systems design. *Water Resour. Res.* 46(5), W05524.
- 698 Reca, J., Martínez, J., 2006. Genetic algorithms for the design of looped irrigation water
699 distribution networks. *Water Resour. Res.* 42(5), W05416.
- 700 Reed, P.M., Hadka, D., Herman, J.D., Kasprzyk, J.R., Kollat, J.B., 2013. Evolutionary
701 multiobjective optimization in water resources: The past, present, and future. *Adv.*
702 *Water Resour.* 51(0), 438-456.
- 703 Rossman, L.A., 2000. EPANET 2 USERS MANUAL, U.S. Environment Protection Agency,
704 Cincinnati, Ohio, USA.
- 705 Schaake, J., Lai, D., 1969. Linear programming and dynamic programming application to
706 water distribution network design, Dept. of Civil Engineering, Massachusetts Institute
707 of Technology, Cambridge, Massachusetts.
- 708 Sherali, H., Subramanian, S., Loganathan, G.V., 2001. Effective Relaxations and Partitioning
709 Schemes for Solving Water Distribution Network Design Problems to Global
710 Optimality. *J. Global Optim.* 19(1), 1-26.
- 711 Srivastava, R.P., 2002. Time continuation in genetic algorithms. Technical report. Illinois
712 Genetic Algorithm Laboratory.
- 713 Storn, R., Price, K., 1997. Differential evolution - a simple and efficient heuristic for global
714 optimization over continuous spaces. *J. Global Optim.* 11, 341-359.
- 715 Talbi, E.G., 2002. A Taxonomy of Hybrid Metaheuristics. *J. Heuristics* 8, 541-564.
- 716 Tang, Y., Reed, P., Wagener, T., 2006. How effective and efficient are multiobjective
717 evolutionary algorithms at hydrologic model calibration? *Hydrol. Earth Syst. Sci.*
718 10(2), 289-307.
- 719 Todini, E., 2000. Looped water distribution networks design using a resilience index based
720 heuristic approach. *Urban Water* 2(2), 115-122.
- 721 Vrugt, J.A., Robinson, B.A., 2007. Improved Evolutionary Optimization from Genetically
722 Adaptive Multimethod Search. *Proc. Natl. Acad. Sci. U.S.A.* 104(3), 708-711.
- 723 Walski, T.M., 2001. The Wrong Paradigm, Why Water Distribution Optimization Doesn't
724 Work. *J. Water Resour. Plann. Manage.* 127(4), 203-205.
- 725 Walski, T.M., Brill, J.E.D., Gessler, J., Goulter, I.C., Jeppson, R.M., Lansey, K., Lee, H.-L.,
726 Liebman, J.C., Mays, L., Morgan, D.R., Ormsbee, L., 1987. Battle of the Network
727 Models: Epilogue. *J. Water Resour. Plann. Manage.* 113(2), 191-203.

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.

- 728 Wang, Q., Liu, S., Wang, H., Savić, D., 2012. Multi-Objective Cuckoo Search for the
729 Optimal Design of Water Distribution Systems, Civil Engineering and Urban
730 Planning 2012, pp. 402-405.
- 731 Wang, Q., Savić, D.A., Kapelan, Z., in press. Hybrid Metaheuristics for Multi-Objective
732 Design of Water Distribution Systems. J. Hydroinform. doi:10.2166/hydro.2013.009.
- 733 Wolpert, D.H., Macready, W.G., 1997. No free lunch theorems for optimization. IEEE T.
734 Evolut. Comput. 1(1), 67-82.
- 735 Zitzler, E., Laumanns, M., Thiele, L., 2002. SPEA2: Improving the strength pareto
736 evolutionary algorithm for multiobjective optimization, In: Giannakoglou, K.,
737 Tsahalis, D., Periaux, J., Papailiou, K., Fogarty, T. (Eds.), Evolutionary Methods for
738 Design, Optimisation and Control. International Center for Numerical Methods in
739 Engineering (CIMNE): Barcelona, Spain, pp. 95-100.
- 740 Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C.M., da Fonseca, V.G., 2003. Performance
741 assessment of multiobjective optimizers: an analysis and review. IEEE T. Evolut.
742 Comput. 7(2), 117-132.

Appendices

Figures

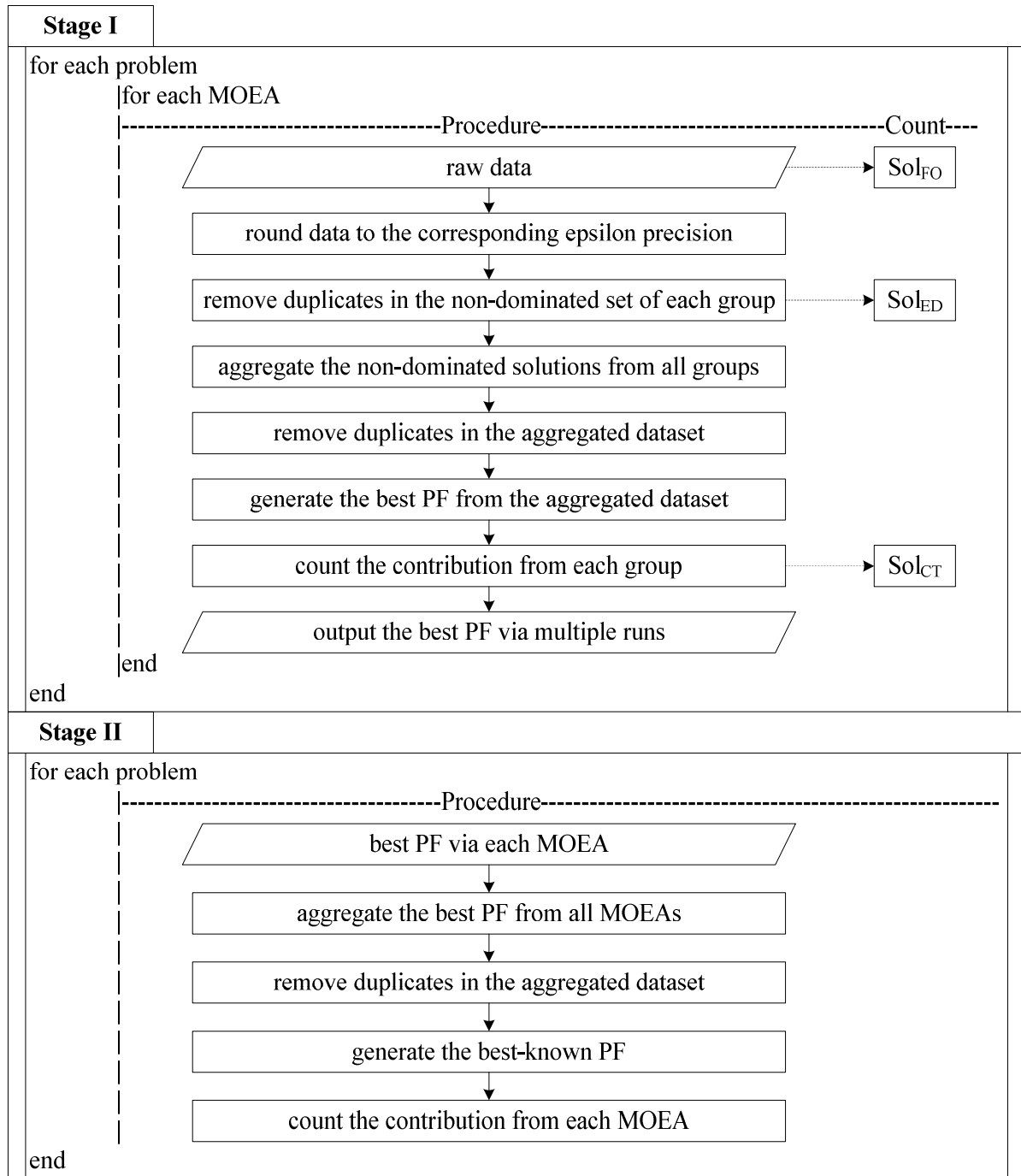
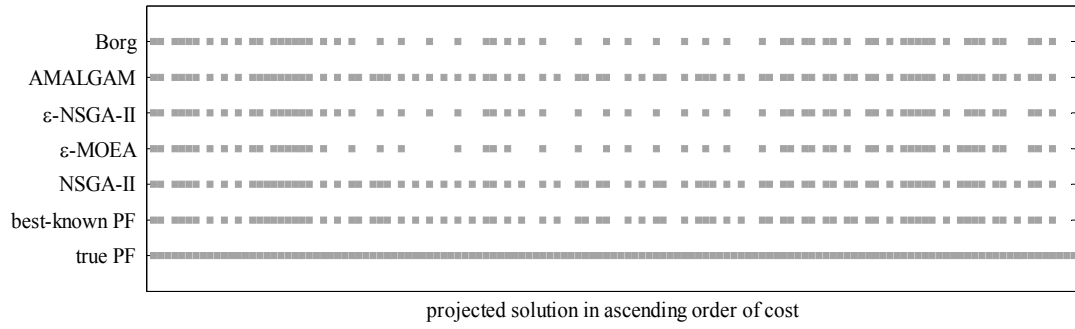
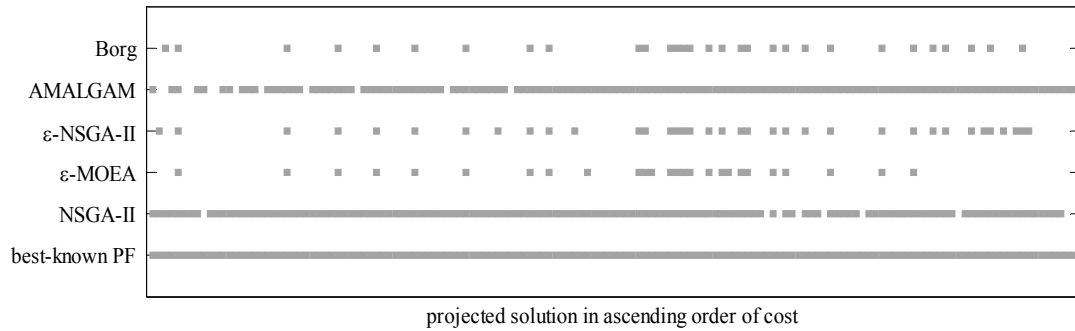


Fig. 1. Flowchart of data post-processing

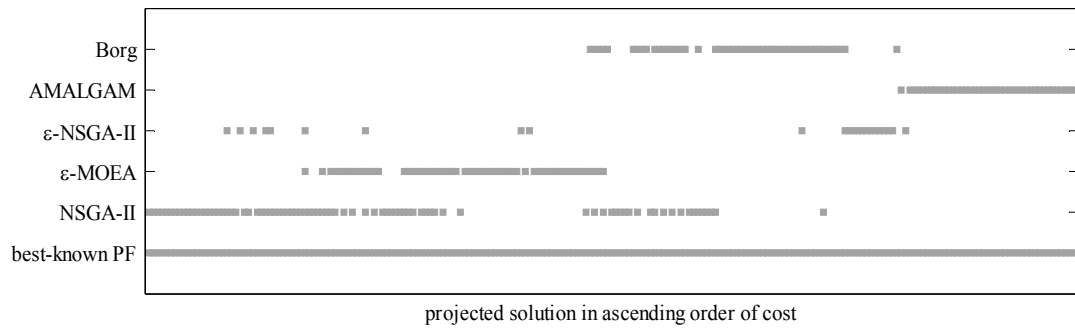
To be cited as: Wang, Q., Guidolin, M., Savic, D., and Kapelan, Z. (2015). "Two-Objective Design of Benchmark Problems of a Water Distribution System via MOEAs: Towards the Best-Known Approximation of the True Pareto Front." *J. Water Resour. Plann. Manage.*, 141(3), 04014060.



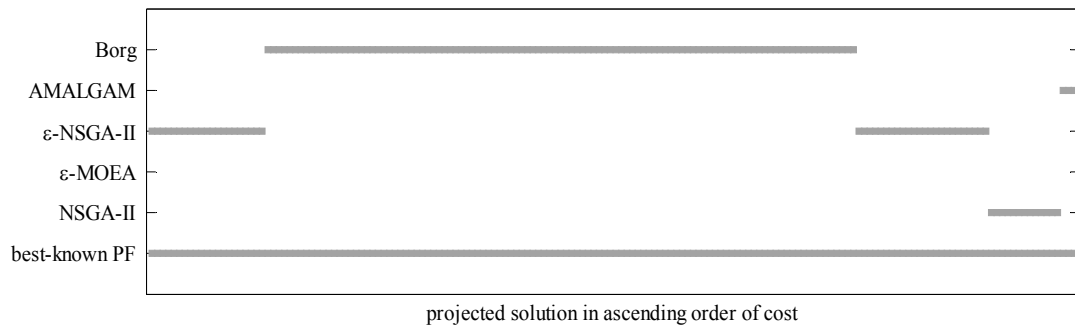
(a) BAK Problem



(b) NYT Problem



(c) PES Problem



(d) EXN Problem

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

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.

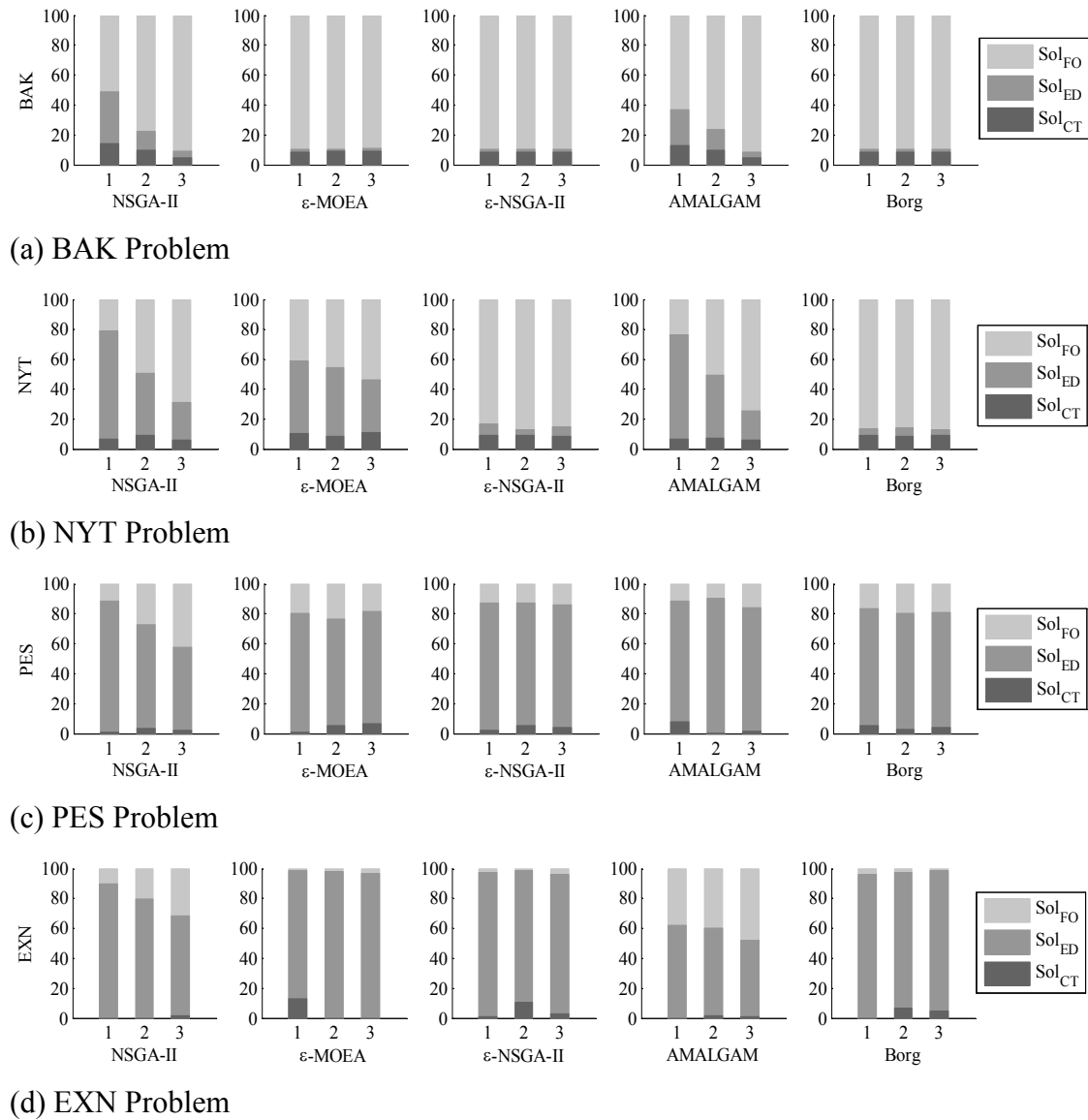
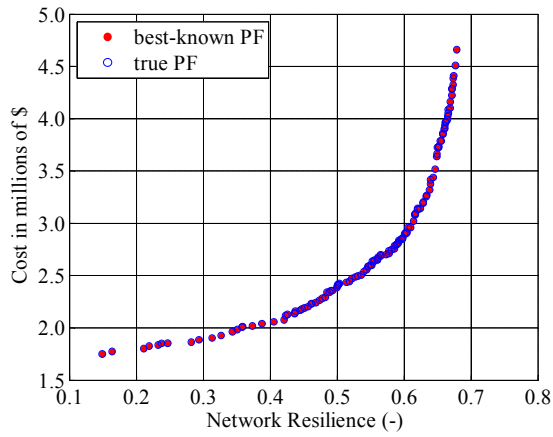
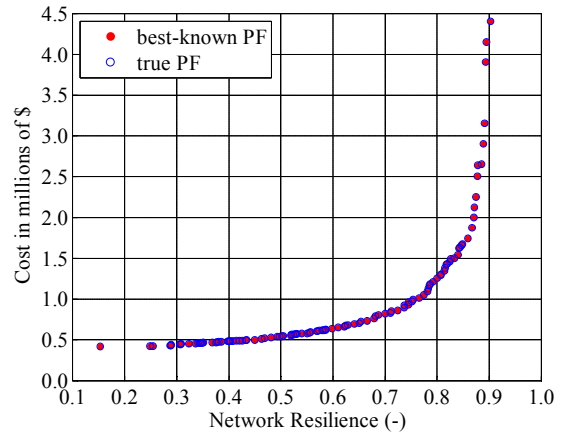


Fig. 3. Variation in the number of solutions in percentage during data post-processing (based on the data given in Table B.1 in Appendix B)

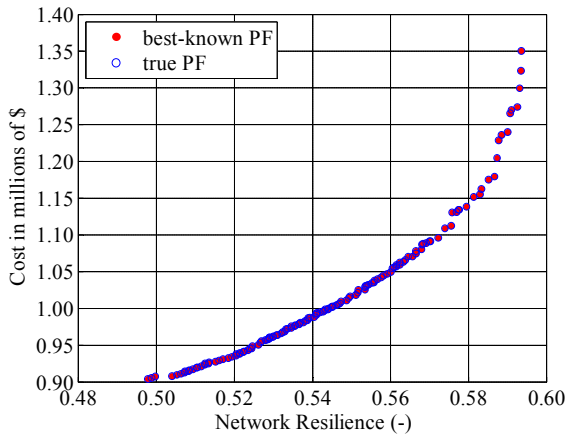
To be cited as: Wang, Q., Guidolin, M., Savic, D., and Kapelan, Z. (2015). "Two-Objective Design of Benchmark Problems of a Water Distribution System via MOEAs: Towards the Best-Known Approximation of the True Pareto Front." *J. Water Resour. Plann. Manage.*, 141(3), 04014060.



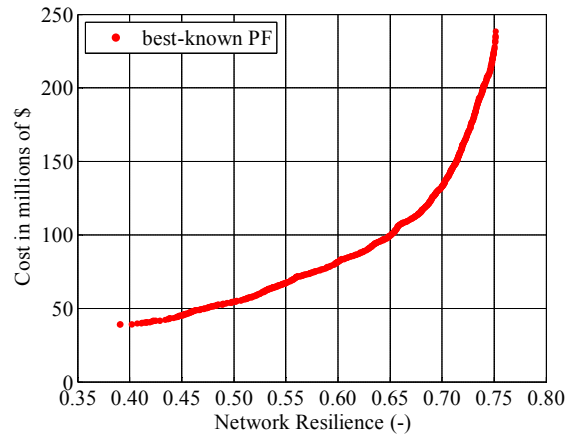
(a) TRN Problem



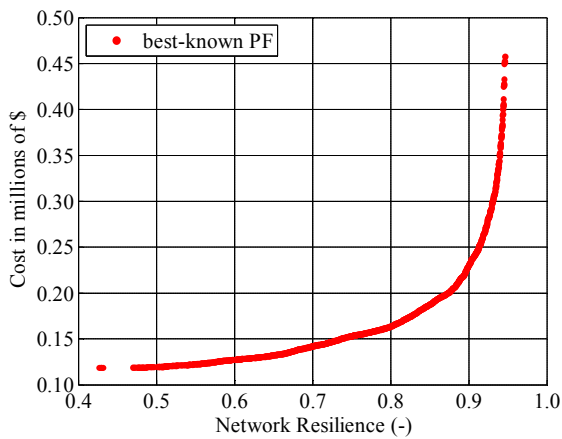
(b) TLN Problem



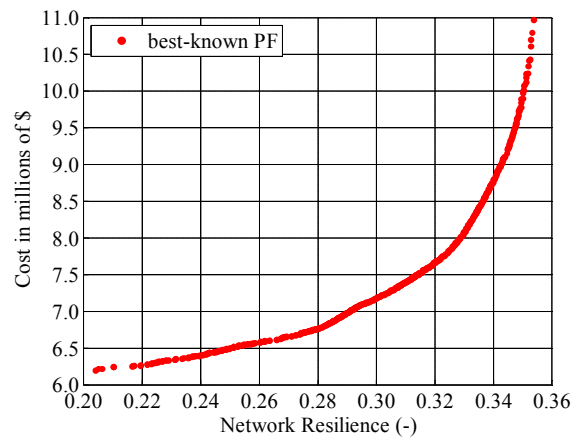
(c) BAK Problem



(d) NYT Problem

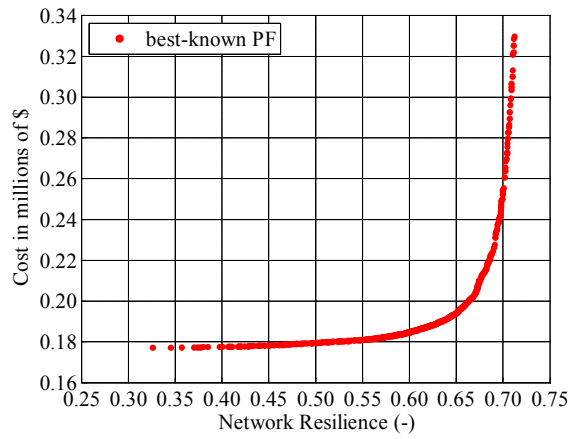


(e) BLA Problem

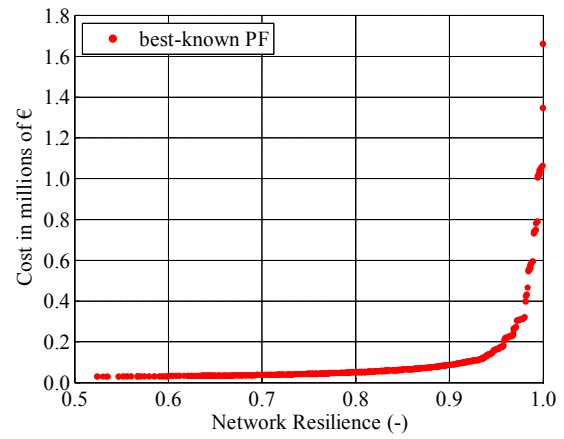


(f) HAN Problem

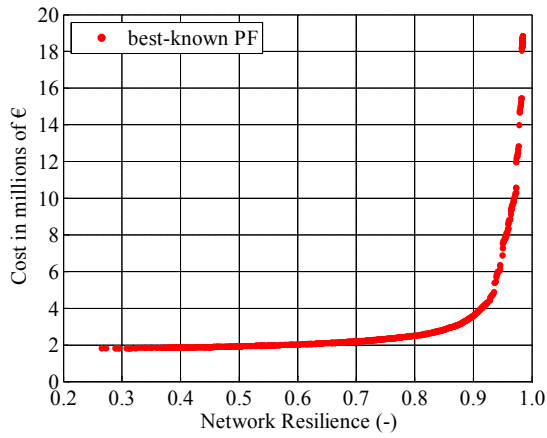
To be cited as: Wang, Q., Guidolin, M., Savic, D., and Kapelan, Z. (2015). "Two-Objective Design of Benchmark Problems of a Water Distribution System via MOEAs: Towards the Best-Known Approximation of the True Pareto Front." *J. Water Resour. Plann. Manage.*, 141(3), 04014060.



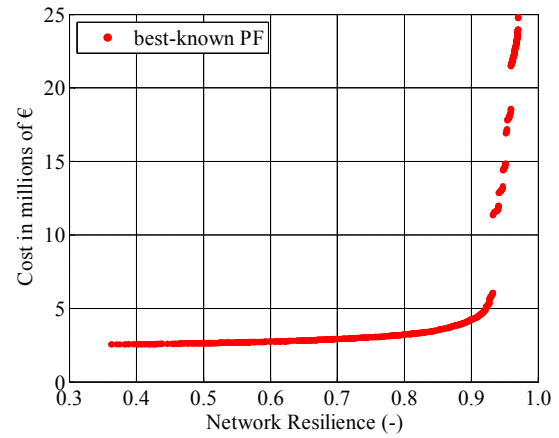
(g) GOY Problem



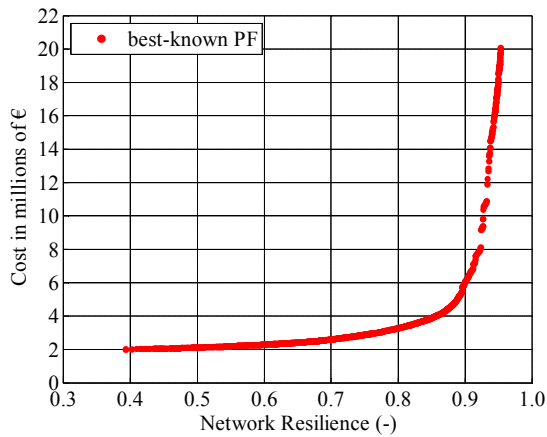
(h) FOS Problem



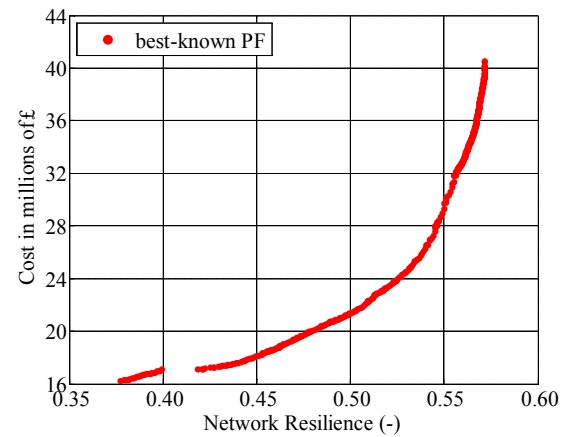
(i) PES Problem



(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.