Multiobjective Valve Management Optimization Formulations for Water Quality Enhancement in Water Distribution Networks

Claudia Quintiliani[1]; Oscar Marquez-Calvo[2]; Leonardo Alfonso[3]; Cristiana Di Cristo[4]; Angelo Leopardi[5]; Dimitri P. Solomatine[6]; and Giovanni de Marinis[7]

Abstract: Water distribution networks (WDNs) need to guarantee that water is delivered with adequate quality. This paper compares the performance of 12 multiobjective procedures to limit water quality deterioration in a WDN through the optimal operation of valves. The first objective (*ObF1*) is to minimize the water age, chosen as a surrogate parameter of quality deterioration, and the second objective (*ObF2*) is to minimize the number of valve closures. The 12 procedures are derived from the combination of 4 different optimization algorithms and 3 formulations of *ObF1*, namely, to minimize the maximum, the arithmetic mean, and the demand-weighted mean water age. The optimization algorithms considered are random search (RS), Loop for Optimal Valve Status Configuration (LOC), and a combination of each of these two with the Archive-based Micro Genetic Algorithm. The procedures are tested on two networks of different complexity. Results show how LOC is able to find near-optimal solutions using a fraction of the computational time required by a brute force search. Furthermore, among the *ObF1* formulations, the use of the averages (either arithmetic or demand-weighted) gives better results in terms of impact on the population served by a WDN. **DOI: 10.1061/(ASCE)WR.1943-5452.0001133.**

Author keywords: Water distribution network; Multiobjective optimization; Valve operation; Water age.

Introduction

Water distribution networks (WDNs) are commonly designed to meet future situations, such as population growth and industrial development, or to handle extraordinary events, such as urban fire. Therefore, utilities often have to manage oversized-pipe systems characterized by reduced velocities and high water age, defined as the time required for a drop of water to travel from the main delivery point to a consumer. An increment of water residence time can negatively impact the microbiological quality of the potable water (USEPA 2002). In particular, a high age value implies deteriorated water quality in terms of chlorine residual concentration reduction and of disinfection byproduct (DBP) formation, which may have carcinogenic effects on human health.

This study proposes a methodology to optimally manage the operational status of valves to modify a network configuration solving a multiobjective optimization (MOO) problem in order to reduce water quality deterioration expressed in terms of age.

Different techniques have been widely used for optimizing WDN design and operation (Mala-Jetmarova et al. 2018). In WDN design, optimization problems have been mainly formulated considering the minimization of construction and operational costs and the maximization of resilience or head pressure. For example, Cembrano et al. (2000) adopted a generalized reduced gradient to minimize WDN operational costs, while Giustolisi et al. (2012) addressed the same problem considering leaks and using evolutionary optimization algorithms. Creaco et al. (2015) used a multiobjective approach to optimize design and operation considering installation and operational costs as objective functions. For the efficient operation of a WDN, optimization problems have been formulated mainly considering operating cost minimization (e.g., Jamieson et al. 2007) and pump scheduling optimization (e.g., Castro Gama et al. 2015).

Some works suggest optimizing WDN operation using valve management with different solvers and for different purposes, including pressure control, backflow prevention, and sectorization for demand control (e.g., Di Nardo et al. 2014). For instance, Jowitt and Germanopoulos (1992) proposed optimal scheduling of pumps and valves to minimize energy consumption using linear programming, while Carpentier and Cohen (1993) used discrete

[1]Scientific Researcher, Water Infrastructure Team, KWR Water Research Institute, Groningenhaven 7, 3433 PE Nieuwegein, Netherlands; formerly, Research Fellow, Dept. Civil Engineering and Architecture, Univ. of Pavia, Via Ferrata 3, 27100 Pavia, Italy; formerly, Ph.D. Fellow, Dept. of Civil and Mechanical Engineering, Univ. of Cassino and Southern Lazio, Via di Biasio 43, 03043 Cassino, Italy (corresponding author). Email: claudia.quintiliani@kwrwater.nl

[2]Ph.D. Fellow, Hydroinformatics Chair Group, IHE-Delft, Institute for Water Education, P.O. Box 3015, NL-2601 DA Delft, Netherlands. ORCID: https://orcid.org/0000-0002-4585-7513. Email: o.marquezcalvo@un-ihe.org

[3]Senior Lecturer, Hydroinformatics Chair Group, IHE-Delft, Institute for Water Education, P.O. Box 3015, NL-2601 DA Delft, Netherlands. ORCID: https://orcid.org/0000-0002-8471-5876. Email: l.alfonso@un-ihe.org

[4]Assistant Professor, Dept. of Civil, Architectural and Environmental Engineering, Univ. of Naples Federico II, Via Claudio 21, 80100 Naples, Italy. Email: cristiana.dicristo@unina.it

[5]Assistant Professor, Dept. of Civil and Mechanical Engineering, Univ. of Cassino and Southern Lazio, Via G. Di Biasio 43, 03043 Cassino, Italy. ORCID: https://orcid.org/0000-0001-6314-0279. Email: a.leopardi@unicas.it

[6]Professor, Hydroinformatics Chair Group, IHE-Delft, Institute for Water Education, P.O. Box 3015, NL-2601 DA Delft, Netherlands; Water Resources Section, Delft Univ. of Technology, Postbus 5, 2600 AA Delft, Netherlands; Water Problems Institute of RAS, Gubkina St., 3, 119333 Moscow, Russia. Email: d.solomatine@un-ihe.org

[7]Professor, Dept. of Civil and Mechanical Engineering, Univ. of Cassino and Southern Lazio, Via G. Di Biasio 43, 03043 Cassino, Italy. Email: demarinis@unicas.it

Note. This manuscript was submitted on October 25, 2018; approved on April 17, 2019; published online on October 3, 2019. Discussion period open until March 3, 2020; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Water Resources Planning and Management***, © ASCE, ISSN 0733-9496.

dynamic programming. Minimization of operational costs by valve scheduling was solved by Ulanicki and Kennedy (1994) using an augmented Lagrangian method. The same problem was also addressed solving one part using a projected gradient method and the other part by a complex method (Cohen et al. 2000a, b). While water quality has been taken into account only recently in the design of WDNs, it has been often considered in the optimization of WDN operation, for example, through effective booster disinfection (e.g., Boccelli et al. 1998) or considering the minimization of rechlorination costs (e.g., Ostfeld and Salomons 2006; Li et al. 2015). In optimization problems, water quality has been considered either as objective (Fu et al. 2013; Shokoohi et al. 2017) or constrained (Bi and Dandy 2014; Kanta et al. 2011; Andrade et al. 2016), in terms of either chlorine residual concentration or water age.

Owing to the uncertainty related to the adoption of existing formulations and to the relative reaction coefficients used to model water quality parameters (for example, to predict DBP formation or chlorine decay), it is preferable to use a more general and less uncertain parameter such as age, as has been done in other studies (Fu et al. 2013; Shokoohi et al. 2017). Instead of using chlorine (Bi and Dandy 2014; Kanta et al. 2011; Andrade et al. 2016) or DBP concentrations (Quintiliani et al. 2018), in this study water age is chosen as the parameter since many aspects of water quality deterioration depend on it (Machell and Boxall 2014). Moreover, defining and evaluating water age is not a trivial task. In this paper, water age is computed following the common approach of estimating it as the flow-weighted average age value of merged flow at a node, even if such an approach has some limitations. Other enhanced approaches could be adopted (Machell et al. 2009; Zhao et al. 2018) as alternatives to the presented methodology.

Depending on the flow velocities in the system, water age can be modified by varying the fluxes through tank- level regulation, changing the network configuration using valves, or opening hydrants to increase discharges. As in Prasad and Walters (2006), the methodology presented in this paper minimizes water age by means of valve management. In fact, this option makes it possible to intervene without losing a precious resource, and the valves can be reopened during critical scenarios. Since reopening may cause the release of accumulated material, in the proposed procedure their movements are intended as a long-term operation for the reconfiguration of the fluxes in the network, and not necessarily as a real-time management procedure.

In Prasad and Walters (2006), the optimization of pipe closures to minimize residence time was formulated as a single-objective problem solved using genetic algorithms. The novelty of the presented contribution consists of three main aspects: first, the adoption of a multiobjective optimization problem formulation, introducing a second objective function; second, the evaluation of different optimization algorithms, from the simplest random search (RS) to the advanced evolutionary algorithm Archive-based Micro Genetic Algorithm (AMGA2) (Tiwari et al. 2011); third, the application of a new algorithm suitable for this specific problem, namely, Loop for Optimal valve status Configuration (LOC). The same three objective functions proposed by Prasad and Walters (2006) are evaluated, and their effectiveness is investigated. Considering 4 different optimization algorithms (with the third and fourth ones being a combination of AMGA2 with RS and LOC) and the 3 objective functions, 12 different procedures are obtained and compared. They are applied to two distribution networks of different complexity: the example network used by Prasad and Walters (2006) and a real network system in Kentucky (Jolly et al. 2012).

The paper is structured as follows. First, the formulation of the optimization problem is presented and then the general methodology is described. Next, the two considered networks are introduced, followed by the analysis of results and discussion. Finally, conclusions are presented and future works discussed.

Definition of Optimization Problem

Objective Functions

Two objective functions are considered in the optimization problem formulation. The first one (*ObF1*) aims to minimize water age at demand nodes, and the following three formulations are explored one at a time (Prasad and Walters 2006):

Maximum Water Age, *MaWA*, represents the maximum age that occurs during the simulation period across all demand nodes:

$$ObF1 = \min\{MaWA\} = \min\{\max\{WA_{i,t}\} \quad \forall\ i = 1 \ldots T_n,\ t = 0 \ldots TST\} \quad (1)$$

Mean Water Age, *MeWA*, represents the arithmetic average of the ages at all nodes:

$$ObF1 = \min\{MeWA\} = \min\left\{\frac{1}{T_n * T_{\text{step}}}\sum_{i=1}^{T_n}\sum_{t=0}^{\text{TST}} WA_{i,t}\right\} \quad (2)$$

- Demand-weighted Mean Water Age, *DeMeWA*, represents the average of the ages calculated assigning at each node a weight equals the demand requested at each time step:

$$ObF1 = \min\{DeMeWA\} = \min\left\{\frac{\sum_{i=1}^{T_n}\sum_{t=0}^{\text{TST}} WA_{i,t} * q_{i,t}}{\sum_{i=1}^{T_n}\sum_{t=0}^{\text{TST}} q_{i,t}}\right\} \quad (3)$$

where $WA_{i,t}$ = water age at ith node at time step t; T_n = number of demand nodes of network; T_{step} = number of time steps into which total simulation time (TST) is divided; and $q_{i,t}$ = demand requested at ith node at time step t. The three proposed formulations of Eqs. (1)–(3) represent different ways to approach water quality evaluation. For example, with reference to DBP formation, the use of Eq. (1) implies that more attention is given to the maximum concentration at those nodes far from the disinfection points. The minimization of the mean water age [Eq. (2)] considers the behavior of the network in average, without controlling the extreme values. Finally, Eq. (3) is based not only on the DBP concentrations but also takes into account the quantity of users exposed to higher values. To provide recommendations on the selection of the most suitable formulation, a comparison of performances of the three *ObF1* formulations is presented.

The second objective function, *ObF2*, minimizes the number of valve closures (NoC):

$$ObF2 = \min\{NoC\} \quad (4)$$

NoC is defined as the number of valves to be closed to reroute the flow in the network. The aim of *ObF2* is to contain interventions in the network to reduce investment costs for placing new valves and to limit their movement. In fact, if only the *ObF1* objective is considered, solutions with a huge number of valve operations may be generated, implying an unacceptable effort by the water utility. Moreover, the valves could be successively reopened if required for a change in system functioning. However, this may cause the release of accumulated material behind the closed section, an aspect that is addressed by minimizing the number of closures.

Decision Variables and Constraints

It is assumed that every pipe in the network has a potential shut-off valve. The decision variables in the optimization problem are the valves' status, represented at that stage by binary values (open or close) (Alfonso et al. 2010). Further investigations will consider the effects of percentages/degrees of valve closures or openings (Kang and Lansey 2009; Ostfeld and Salomons 2006).

The constraints are fixed considering that the operational status of the valves needs to guarantee the required service also in terms of pressure. Hence, the considered constraints are as follows: (1) any valve configuration status must guarantee the supply of water to all nodes, i.e., nodes cannot be disconnected; (2) the pressure $P_{i,t}$ at each ith node at each time t should be within a fixed range:

$$P_{\min} < P_{i,t} < P_{\max} \tag{5}$$

Methodology

Procedures

Twelve different procedures combining different optimization algorithms and formulations are compared (Table 1). The four algorithms used, described in detail in the following sections, are RS, LOC, and a combination of each of these two with AMGA2, a multiobjective evolutionary algorithm based on genetic algorithms. The first objective function is *MaWA* [Eq. (1)], *MeWA* [Eq. (2)], or *DeMeWA* [Eq. (3)], while the second objective function is always *NoC* [Eq. (4)]. The results are provided as Pareto fronts and maps to compare the different procedures.

Simulation Setup

EPANET (Rossman 2000) is used as a WDN model for hydraulic and quality simulation (water age evaluation). Since the aim of this paper is to present a new and general methodology to reduce water age, at the present stage some simplifying hypotheses are considered:

- Even if in real WDN users are placed along pipes, demands are assumed to be concentrated in nodes. For the mean pipe length of the presented networks the corresponding approximation of water age is on the order of less than 1 s. Further investigations will consider demands distributed along pipes as in Farina et al. (2014) and Menapace et al. (2018).

Table 1. Optimization procedures combining *ObF1* formulations and optimization algorithms

Procedure	*ObF1*	Optimizer
P1	*MaWA*	RS
P2	*MaWA*	LOC
P3	*MaWA*	RS-AMGA2
P4	*MaWA*	LOC-AMGA2
P5	*MeWA*	RS
P6	*MeWA*	LOC
P7	*MeWA*	RS-AMGA2
P8	MeWA	LOC-AMGA2
P9	*DeMeWA*	RS
P10	*DeMeWA*	LOC
P11	*DeMeWA*	RS-AMGA2
P12	*DeMeWA*	LOC-AMGA2

- The pressure-driven approach is not used because the minimum pressure value in the constraint [Eq. (5)] is fixed in order to guarantee demand-driven functioning.
- Leakages are neglected even if they represent a component of demands. Their effect will be analyzed in future research.
- To verify the existence of disconnected nodes, a procedure implemented in EPANET is used. However, other methods could be adopted (e.g., Creaco et al. 2012).
- For water age evaluation complete mixing at nodes is assumed and dispersion is neglected. Although this assumption is questionable (Machell et al. 2009), its correction requires more complex computations, and for this reason they are still adopted in the majority of simulation tools and applications (Boccelli et al. 1998; Di Cristo and Leopardi 2008; Seyoum and Tanyimboh 2017).
- Input data uncertainty (Di Cristo et al. 2015) is not considered herein, but the same authors presented a robust optimization with respect to demand uncertainty in Marquez-Calvo et al. (2018).

A standard model-based optimization framework, commonly used in the literature (e.g., Alfonso et al. 2010; Quintiliani et al. 2017), is adopted. An application compiled in C++ using the library of functions of the EPANET Programmer's Toolkit (Rossman 1999) was developed to set up the valve configurations in the input file and to run the hydraulic and water quality engines. The outputs of the application used by the optimization algorithm are *ObF1* and *ObF2* values.

All objective functions are evaluated with respect to the original status of the network, i.e., with all valves open, corresponding to $ObF2 = 0$. This means that the "do-nothing" solution is always included in the Pareto front. In this way, a comparison is made on how much *ObF1* improves for different configurations with respect to the original status.

Optimization Algorithms

To describe the RS and LOC algorithms, the Class P network is defined as a network that has P pipes that can be closed through valve operation.

Random Search

Given a maximum number N of objective function evaluations and a maximum number P of valves to close, $M = N/P$ network configurations belonging to the same class are considered. The RS algorithm generates M random network configurations for each class and selects the one with the lowest *ObF1*. The procedure stops when all P classes have been analyzed.

Loop for Optimal Valve Status Configuration

LOC is an algorithm specifically designed to solve the stated problem, which is based on procedures that find the best possible solution incrementally at each step, similarly to greedy algorithms (e.g., Alfonso et al. 2013; Banik et al. 2017a, b). As in the previous case, LOC is used to find P configurations of a network.

Starting from Class 0, corresponding to an initial condition where all valves of the network are open, LOC investigates all possible configurations and selects the valve that produces the highest *ObF1* reduction in the entire network when it is closed. Then it is removed from the set of "Remaining Valves" and added to the set of "Best Configurations." To set the second valve to close, the algorithm considers the configurations with the valves previously closed, selecting within the "Remaining Valves" set the valve that offers the *ObF1* highest reduction. This valve is added to the "Best Configuration" set. The procedure stops when the P class has been reached.

LOC uses a predetermined, limited number of function evaluations to find a (suboptimal) Pareto front. This number of evaluations is given by the expression

$$Ne = \sum_{i=NP-P+1}^{NP} i \tag{6}$$

where Ne = number of function evaluations; NP = total number of pipes of network; and P = maximum number of valves to close.

AMGA2

The AMGA2 by Tiwari et al. (2011) is a multiobjective evolutionary algorithm to find optimal solutions. It is considered a steady-state genetic algorithm because its main Pareto front has a small number of solutions, although other good solutions are stored in an archive. To produce the next generation of populations, it uses all solutions in the main Pareto front mated with some of the solutions in the archive. To decide which solutions to include in the new Pareto front, two criteria are used: the degree of dominance of the solution and the diversity of the solution. In this way two goals are reached, namely, a small number of function evaluations and the advantage of the diversity of solutions in the archive. The good solutions that are not selected for the new Pareto front are included in the archive. To maintain the archive, the solutions crowding a specific region of the solution space are eliminated using the nearest-neighbor search strategy.

Some experiments, not reported in this paper, demonstrated that AMGA2 alone was not able to find a satisfactory number of solutions because most of the generated networks were characterized by disconnected nodes. To deal with this problem, Prasad and Walters (2006) modified their algorithm to avoid the generation of networks with disconnections. In contrast, in this work the search space is reduced to minimize the generation of networks with disconnected nodes by combining AMGA2 with either RS or LOC (named RS-AMGA2 and LOC-AMGA2, respectively). In this way, two objectives are met. First, some sets of candidate valves to be used as decision variables by AMGA2 are generated, drastically reducing the search space. Second, a reference initial population is given to AMGA2, improving its efficiency.

Performance Indicators

To measure the improvement of RS and LOC algorithms by combining them with AMGA2, the following index of improvement (IoI) is used:

$$IoI(\boldsymbol{F}_k, \boldsymbol{F}_j) = \frac{1}{|\boldsymbol{C}(\boldsymbol{F}_k, \boldsymbol{F}_j)|} \sum \boldsymbol{C}(\boldsymbol{F}_k, \boldsymbol{F}_j) \frac{\boldsymbol{f}_{j,m}^{(1)}}{\boldsymbol{f}_{k,h}^{(1)}} \tag{7}$$

where $\boldsymbol{F}_k$ and $\boldsymbol{F}_j$ = solution of Pareto fronts of AMGA2 (subscript k) and of each of its counterpart LOC or RS (subscript j), respectively, for a fixed value of *ObF2* (*NoC*); $\boldsymbol{C}$ = set containing all couples ($\boldsymbol{F}_k$, $\boldsymbol{F}_j$) and $|\boldsymbol{C}(\boldsymbol{F}_k, \mathbf{F}_j)|$ as its cardinality; $\boldsymbol{f}_{k,h}^{(1)}$ = value of *ObF1* of hth tuple in Pareto front k; and $\boldsymbol{f}_{j,m}^{(1)}$ = value of *ObF1* of mth tuple in Pareto front j.

In other words, considering a solution with the same number of operations *NoC* (*ObF2*), Eq. (7) estimates the ratio of the *ObF1* value of the solution in the counterpart to the *ObF1* value of the solution with AMGA2. The summation of all these ratios is divided by the number of solutions with the same *ObF2* to consider a global value representing the efficiency of the procedures, regardless of the *ObF1* formulation used. Then, the weighted average of the *IoI* (*WAIoI*) is evaluated:

$$WAIoI(\boldsymbol{F}_k, \boldsymbol{F}_j) = \frac{1}{\sum_{ObF1} |\boldsymbol{C}(\boldsymbol{F}_{k(ObF1)}, \boldsymbol{F}_{j(ObF1)})|} \times \sum_{ObF1} [|\boldsymbol{C}(\boldsymbol{F}_{k(ObF1)}, \boldsymbol{F}_{j(ObF1)})| * IoI(\boldsymbol{F}_{k(ObF1)}, \boldsymbol{F}_{j(ObF1)})] \tag{8}$$

where $\sum_{ObF1}$ = summation of sets $\boldsymbol{C}$ for all *ObF1* formulations.

To compare the performances of different *ObF1* formulations, the differences between the initial condition values and the optimized ones of the following parameters are computed in each node:

$$MaWA_i = \max\{WA_t, \forall\ t = 0 \ldots TST\}_i \quad \text{for MaWA as } ObF1 \tag{9}$$

$$MeWA_i = \left(\frac{1}{T_{\text{step}}} \sum_{t=0}^{\text{TST}} WA_t\right)_i \quad \text{for } MeWA \text{ as } ObF1 \tag{10}$$

$$DeMeWA_i = \left(\frac{\sum_{t=0}^{\text{TST}} WA_t \cdot q_t}{\sum_{t=0}^{\text{TST}} q_t}\right)_i \quad \text{for } DeMeWA \text{ as } ObF1 \tag{11}$$

In particular, $MaWA_i$, $MeWA_i$, and $DeMeWA_i$ = maximum, arithmetic mean, and demand-weighted mean of ages observed at ith node during TST, respectively. A negative value of the differences between the initial condition values and the optimized ones, indicated as $\Delta MaWA_i$, $\Delta MeWa_i$, and $\Delta DeMeWA_i$, means a reduction of the age formulation value at the ith node.

To evaluate the quality of the solutions, the average (μ) and standard deviation (σ) of the variations $\Delta MaWA_i$, $\Delta MeWA_i$, and $\Delta DeMeWA_i$ observed in all nodes of the network are computed. Negative values of μ indicate an average reduction of the age in the network. A higher negative average indicates a better performance; a lower standard deviation indicates good homogeneity in the variation age in the network.

Case Studies

Two distribution networks with different characteristics are selected to explore the performance of the proposed procedures: Network PW06 by Prasad and Walters (2006) and Network J14 from the database developed by the Kentucky Infrastructure Authority (Jolly et al. 2012).

The PW06 network [Fig. 1(a)] has 47 pipes and 33 demand nodes, with elevations that vary between 10 and 30 m, and it is supplied from a single source (reservoir). The demands assigned in the nodes are the same as those in the original paper.

Network J14 [Fig. 1(b)] has the following characteristics: 377 demand nodes with elevations between 200 and 274 m, 3 tanks, 473 pipes with a total length of about 104 km, and 5 pump stations. The system is supplied from four sources, one at a head of 274 m and the others at around 200 m. In the schematization [Fig. 1(b)], while two sources are visible, the others are indicated as INLET 1 and INLET 2, located respectively at 12 and 62 km from the WDN. In all nodes, the same demand pattern is assigned, characterized by a 1-h time step multiplier with two picks of request around 10:00 a.m. and 9:00 p.m.

In both cases, the simulations were run long enough to guarantee stability of the hydraulic conditions. The latter was achieved after 72 h of simulation for Network PW06 and 168 h of simulation for Network J14.

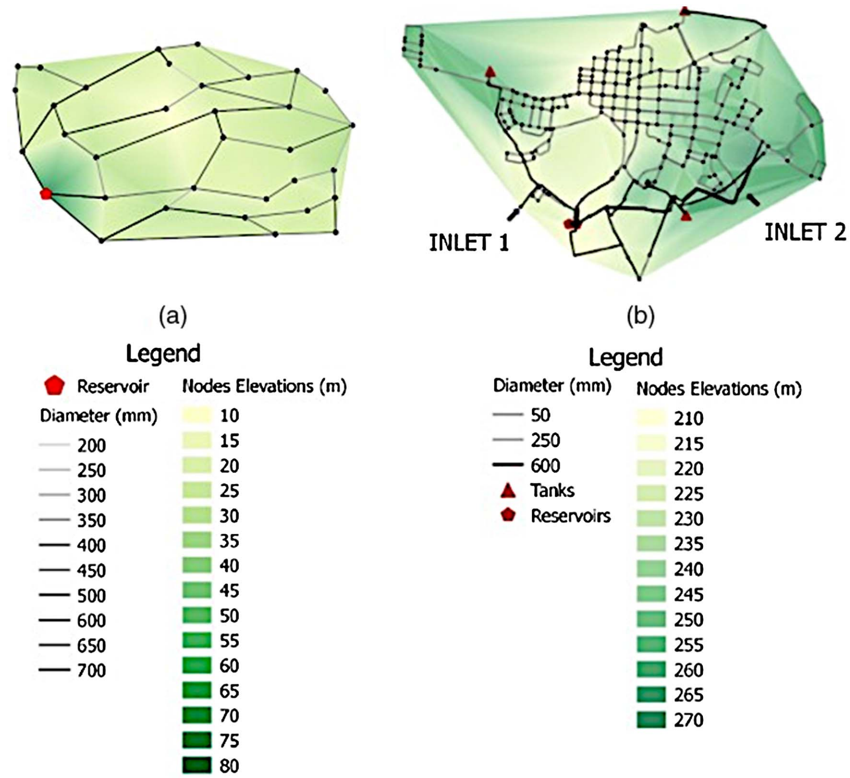

Fig. 1. Distribution network schemes: (a) PW06 (data from Prasad and Walters 2006); and (b) J14 (data from Jolly et al. 2012).

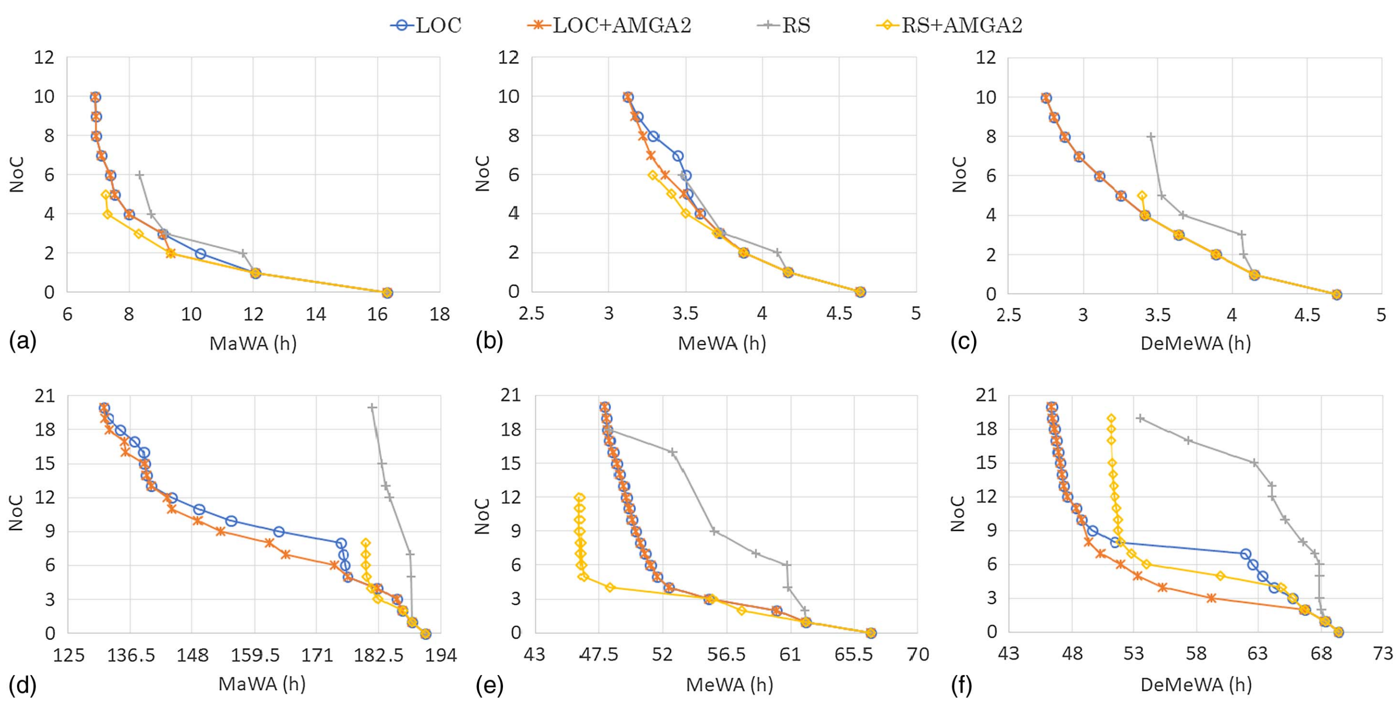

Fig. 2. Results in terms of Pareto fronts for (a–c) PW06 (data from Prasad and Walters 2006); and (d–f) J14 (data from Jolly et al. 2012): (a and d) Procedures P1–P4; (b and e) Procedures P5–P8; and (c and f) Procedures P9–P12.

Analysis of Results and Discussion

The LOC algorithm requires a predefined number of evaluations, Ne [Eq. (6)]. In contrast, the other algorithms do not use a predetermined Ne, which means that their performance depends directly on the required function evaluations. The analysis of the

Table 2. Values of *WAIoI* for both case studies

Performance indicator	J14	PW06
$WAIoI\ (F_{LOC\text{-}AMGA2}, F_{LOC})$	1.021	1.007
$WAIoI\ (F_{RS\text{-}AMGA2}, F_{RS})$	1.134	1.060
$WAIoI\ (F_{RS\text{-}AMGA2}, F_{LOC})$	1.010	1.022

performance is done considering the fixed Ne of LOC as the baseline.

As described in more detail in the following paragraphs, Fig. 2 shows the results of the procedures listed in Table 1 in terms of Pareto fronts for both case studies, while Table 2 reports the values of the indicator *WAIoI* [Eq. (8)] used to evaluate the performances of the optimization algorithms.

PW06 Network

In PW06 the required number of function evaluations is $Ne = 425$ [Eq. (6)] to obtain a 10-point Pareto front. The values used as pressure thresholds in the constraint of Eq. (5), expressed in terms of piezometric height, are $P_{\max} = 100$ m and $P_{\min} = 10$ m.

For PW06, the solutions reported in terms of Pareto fronts in Figs. 2(a–c) show that for all considered *ObF1* formulations, LOC generates a better front than that from RS. Moreover, RS and RS-AMGA2 algorithms are able to find a limited number of solutions with respect to LOC and LOC-AMGA2.

AMGA2 barely improves the Pareto front found by LOC. However, its improvement over RS is significant. In fact, the use of AMGA2 in combination with RS makes it possible to reach the same *ObF1* values of RS by operating fewer valves. Moreover, this combination is also slightly better than LOC and LOC-AMGA2 solutions. This is confirmed by the *WAIoI* values reported in Table 2, which suggest that the addition of AMGA2 produces an improvement of 6.0% and 0.7% with respect to the solutions of RS and LOC, respectively, while the Pareto front of RS-AMGA2 is about 2% better than the one from LOC.

Fig. 3 represents for all procedures the heat maps showing the frequency of the valves included in the solutions of the Pareto front; a darker dot indicates that the valve is more often considered. A RS algorithm (P1-P5-P9) is characterized by the use of a large number of valves in the network, which is not convenient in the operational context. The application of AMGA2 after RS (P3-P7-P11) improves the solutions, focusing on only five or six valves to operate. LOC algorithm has better behavior also without having to apply AMGA2 afterwards. Moreover, LOC

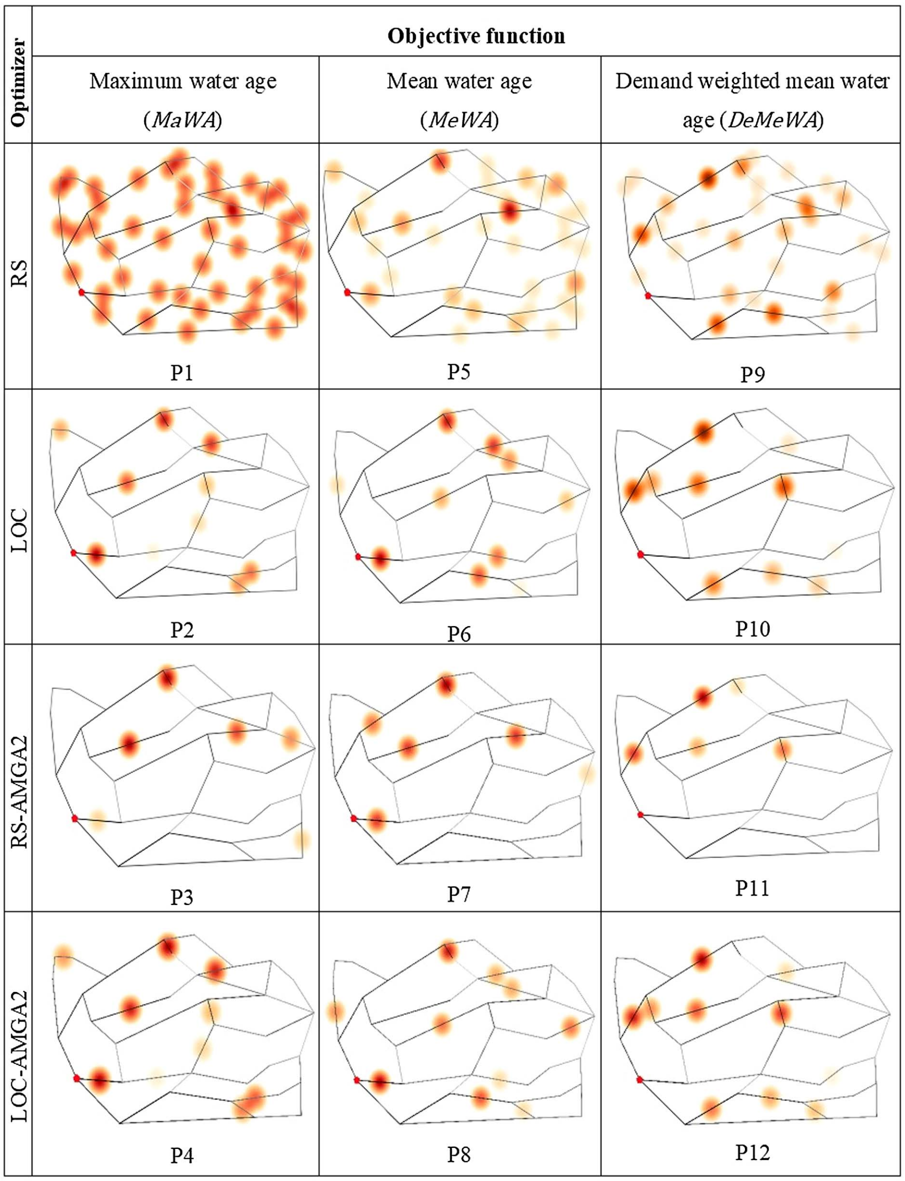

Fig. 3. Heat maps showing frequency of valve closure from solutions of Procedures P1 to P12 for Network PW06 (data from Prasad and Walters 2006).

and LOC-AMGA2 consider almost the same valves, mainly placed on the largest diameters.

To compare the performances of different *ObF1* formulations, the average (μ) and standard deviation (σ) of the variation $\Delta MaWa_i$, $\Delta MeWa_i$, and $\Delta DeMeWa_i$ for the optimized solutions obtained with LOC and LOC-AMGA2 for $NoC = 5$ are computed. This *NoC* number was selected considering that additional closures reduce *ObF1* only marginally. For all cases, the obtained μ values are negative, showing for all formulations a reduction in the average age with respect to the original condition. Insignificant differences have been observed among considered age formulations and between LOC and LOC-AMGA2 results.

The performance of each *ObF1* is also estimated extracting the optimal network configurations and evaluating how well they performed for the remaining *ObF1* formulations. It is observed that the use of each of the *ObF1* formulations implies, on average, a reduction in the values of the other objective functions, when compared with the do-nothing option, almost reaching the values obtained when they are used as the optimization target.

J14 Network

For the J14 network, assuming that a maximum of 20 valves can be operated, the number of function evaluations, Ne, is 9270. The values used as pressure thresholds in the constraint of Eq. (5), expressed in terms of piezometric height, are $P_{\max} = 100$ m and $P_{\min} = 10$ m.

The Pareto fronts obtained for the J14 network are presented in Figs. 2(d–f), where the comparison among the different algorithms shows a similar tendency of what is obtained for the PW06 case. In particular, LOC generates a better Pareto front than RS; AMGA2 improves slightly the solutions of LOC, while those of RS are improved significantly. The *WAIoI* values (Table 2) indicate that by adding AMGA2, LOC is improved by approximately 2% and RS by approximately 13%. Finally, RS-AMGA2 produces an improvement of about 1% with respect to LOC.

In summary, the results suggest that the LOC algorithm produces a better Pareto front than RS. Also, although the combination RS-AMGA2 works better than LOC, it requires more function evaluations. The improvement that AMGA2 offers over LOC is negligible, whereas for RS it is more significant.

Fig. 4 shows the heat maps to provide a spatial indication of where and how frequently the pipes were selected by different procedures (Table 1). As expected, the solutions using the RS algorithm (P1, P5, and P9) do not focus on specific sectors of the network because the closures are randomly spread over the whole system. Independently of the selected *ObF1*, around 33% of the

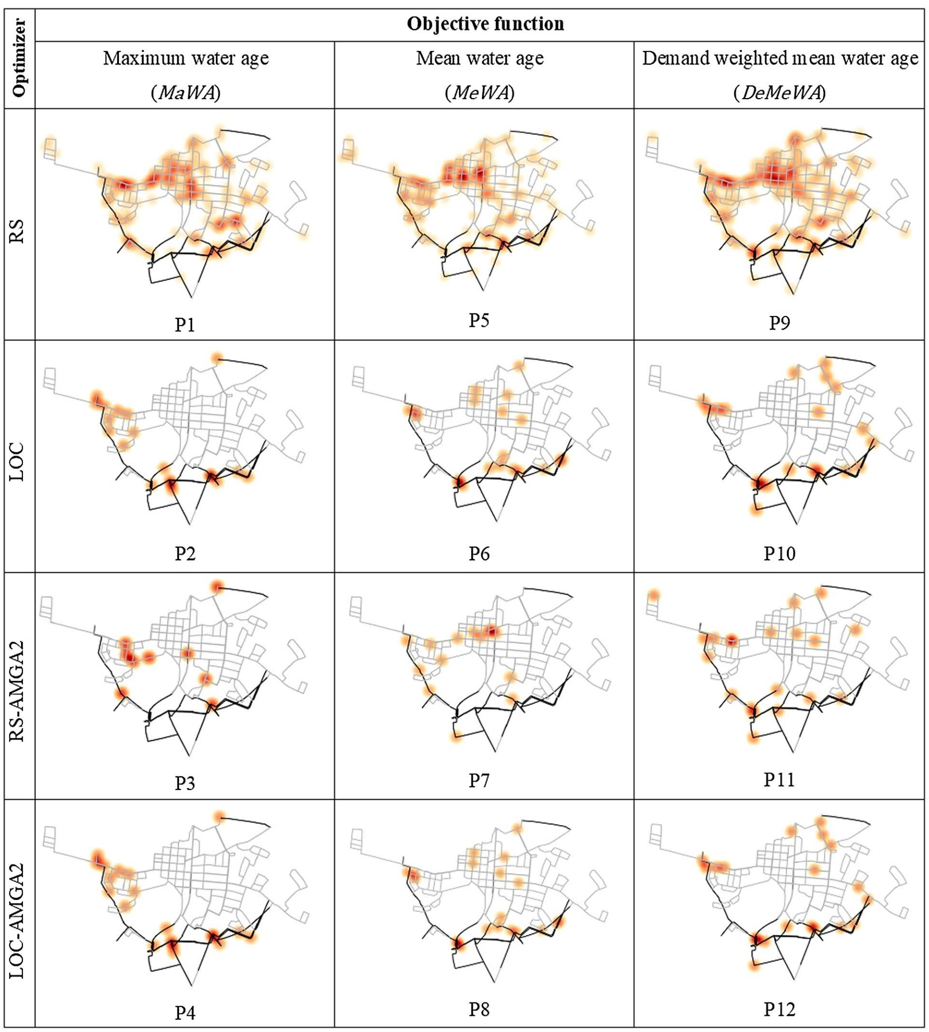

Fig. 4. Heat maps showing frequency of valve closure from solutions of Procedures P1 to P12 for Network J14 (data from Jolly et al. 2012).

valves are included in at least one solution, meaning RS requires a large number of valves to be operated.

The solutions obtained with the RS-AMGA2, LOC, and LOC-AMGA2 algorithms are characterized by a reduced selection of valves to close, varying from 3% to 4.2% among all the possible decision variables. This confirms again that AMGA2 performs significantly better than RS. A closer look at the valves selected in each experiment reveals that RS-AMGA2 individuates different areas with respect to LOC and LOC-AMGA2. For the latter algorithms the considered valves are concentrated in specific areas of the network involving mainly the larger diameters located in the southern part of the system.

The average (μ) and standard deviation (σ) of the variation $\Delta MaWA_i$, $\Delta MeWA_i$, and $\Delta DeMeWA_i$ calculated between the initial values and those for the solutions of LOC and LOC-AMGA2 with $NoC = 10$ are reported in Table 3. The NoC number has been again selected considering that additional closures reduce *ObF1* only marginally. $\Delta MaWA_i$ has a positive μ, indicating an average increase of $MaWA_i$ in the network, suggesting a bad performance of $MaWA_i$ as *ObF1*. Both $\Delta MeWA_i$ and $\Delta DeMeWA_i$ have negative μ values and lower σ with respect to $\Delta MaWA_i$. $\Delta DeMeWA_i$ is characterized by the highest negative average and the lowest standard deviation, which indicate its better performance as *ObF1*. No differences are observed between the LOC and LOC-AMGA2 results.

Regarding the performance of the *ObF1* formulations, extracting the optimal network configurations and evaluating how well they performed for the remaining set of *ObF1* not selected, the results show mixed behaviors. Considering the configuration valve sets obtained using *MaWA* as *ObF1*, this leads to almost no improvements for the other formulations with respect to the case of $NoC = 0$. This has serious consequences for the majority of users, because minimizing *MaWA* does not imply a diminution of the residence time for a large part of the WDN. The solutions obtained with *MeWA* do not modify the values of *MaWA* but improve those of *DeMeWA*. This means that the majority of users would have a partial improvement, but not those with high water residence time. Similarly, for the solution with *DeMeWA*, *MaWA* remains, on average, near the zero-closure values regardless of the number of closures, while *MeWA* is reduced to optimal levels. This means that most users would have access to water with a reduced age.

Table 3. Average (μ) and standard deviation (σ) of variations of $MaWA_i$, $MeWA_i$, and $DeMeWA_i$ in the J14 network ($NoC = 10$)

	LOC		LOC-AMGA2	
Formulation	μ	σ	μ	σ
$\Delta MaWA_i$	12.96	40.52	10.03	38.14
$\Delta MeWA_i$	−16.83	33.86	−16.83	33.86
$\Delta DeMeWA_i$	−25.97	29.94	−25.97	29.94

Performance of LOC Algorithm

To evaluate the performance of the LOC algorithm, its results are compared with the method proposed by Prasad and Walters (2006) and the brute-force search (BFS) procedure. Those tests were executed considering the PW06 network and fixing a constraint of 15 m as the minimum head in the network in accordance with the value used by Prasad and Walters (2006).

A comparison of the results obtained by Prasad and Walters (2006) with those of LOC is shown in Fig. 5. For the *MaWA* function, LOC finds several solutions that achieve a similar reduction in water age with fewer pipe closures. Using the objective function *MeWA* [Fig. 5(b)], the LOC solution with 9 closures is as good as the solution of Prasad and Walters (2006) with 11 closures. For *DeMeWA* [Fig. 5(c)], LOC with 10 operations marginally dominates the solution by Prasad and Walters (2006). Unfortunately, Prasad and Walters (2006) do not make any reference to the number of evaluations required to obtain their results so the efficiency of the algorithms cannot be compared.

A further experiment was designed to prove that the LOC method is suitable for finding a close-to-optimal solution. An exhaustive search of all solutions was carried out with a BFS in the smallest network, PW06, taking into account *DeMeWA* as *ObF1*. To reduce the execution time, an array of 28 CPU cores was used to perform the simulations in parallel. Both BFS and LOC were run for eight pipe closures to achieve the *DeMeWA* maximum reduction.

The solution found by BFS reduced the water age down to 2.8735 h, and it was available after 16.6 days of computational effort. Remarkably, the solution found by LOC reduced the water age down to 2.8736 h, requiring only 3 s. This demonstrates the efficiency of the proposed LOC algorithm.

To ensure the reliability of this comparison, the experiment was repeated considering different pipe closures, from one to seven. The results are reported in Table 4. In all cases LOC performed as well as BFS, with an advantage of several orders of magnitude in terms of computational time. Unfortunately, it was not feasible to run BFS for $NoC = 9$, 10, and 11. Indeed, these would take 55, 145, and 299 days, respectively, because the required number of simulations are 5.44×10^8, 1.44×10^9, and 2.97×10^9, respectively. Moreover, when LOC runs for X closures, the solutions for $X - 1, X - 2, \ldots, 1$ are immediately available, contrasting with

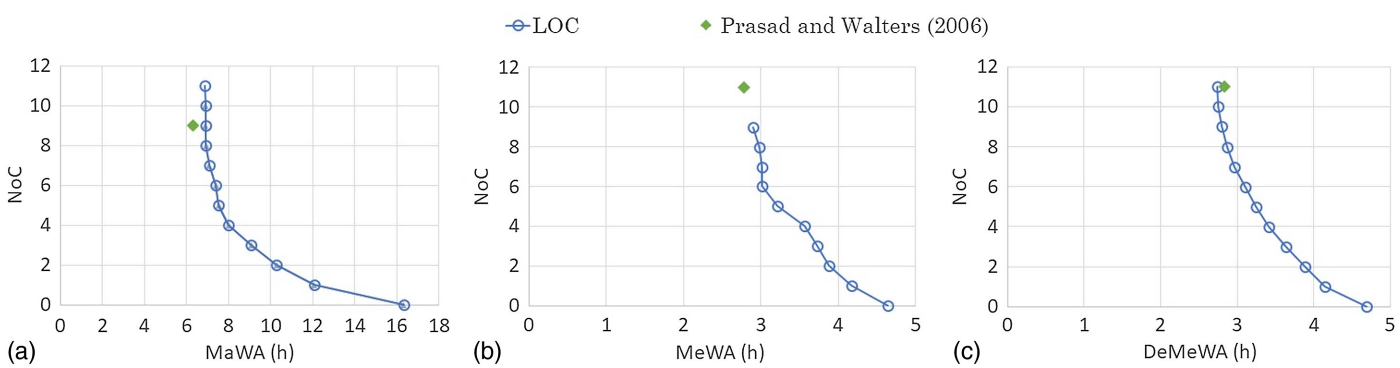

Fig. 5. Comparison of LOC and data from Prasad and Walters (2006) solutions using: (a) MaWA; (b) MeWA; and (c) DeMeWA.

Table 4. Comparison between BFS and LOC solutions for PW06 network

Number of closures (*NoC*)	*DeMeWA* (h) found		Number of simulations		Computational time (days) required	
	BFS	LOC	BSF	LOC	BFS	LOC
1	4.1482	4.1482	4.70×10^1	4.70×10^1	4.73×10^{-6}	4.73×10^{-6}
2	3.8869	3.8869	1.07×10^3	9.30×10^1	1.07×10^{-4}	9.36×10^{-6}
3	3.6402	3.6402	1.55×10^4	1.38×10^2	1.56×10^{-3}	1.39×10^{-5}
4	3.3797	3.4119	1.61×10^5	1.82×10^2	1.62×10^{-2}	1.83×10^{-5}
5	3.1795	3.2528	1.28×10^6	2.25×10^2	1.29×10^{-1}	2.27×10^{-5}
6	3.0672	3.1072	8.02×10^6	2.67×10^2	8.07×10^{-1}	2.69×10^{-5}
7	2.9670	2.9670	4.03×10^7	3.08×10^2	4.06×10^0	3.10×10^{-5}
8	2.8735	2.8736	1.65×10^8	3.48×10^2	1.66×10^1	3.50×10^{-5}

BFS, which requires a separate experiment for each number of closures.

Conclusions

The present paper compares the performances of 12 multiobjective optimization procedures to optimize valve management in WDNs for improving water quality, evaluated in terms of water age. The procedures derive from the combination of four different algorithms (RS, LOC, RS-AMGA2, and LOC-AMGA2) and of three water quality objective function formulations (*MaWA*, *MeWA* and *DeMeWA*). Two distribution networks of different complexity are considered.

The results show that the proposed LOC algorithm always produces better solutions with respect to RS, obtaining lower age values with the same number of closures. Moreover, heat maps show that LOC considers candidate valves concentrated in specific areas of the network, which is an advantage for operators. Its codification is very simple, and it produces a good compromise between the quality of the Pareto front and the required number of function evaluations.

The alternatives LOC-AMGA2 and RS-AMGA2 offer only a marginal improvement with respect to the solutions found by LOC, at the expense of having double function evaluations. This implies that, for this particular optimization problem, the LOC algorithm is the most convenient. The heat maps obtained with LOC show also that the operation on the larger pipes are more efficient for the reduction of water age. The comparison of LOC with BFS demonstrates that, despite its simplicity, LOC achieves near-optimal results with very small computational effort, which justifies its use in large networks.

Regarding the comparison among the *ObF1* formulations, the analysis of the average and standard deviation of the variations $\Delta MaWA_i$, $\Delta MeWA_i$, and $\Delta DeMeWA_i$ observed in all nodes indicates similar performances for the smaller Network PW06. For the more complex J14, the results suggest better performances of *MeWA* and *DeMeWA*, indicating that the latter is the best one. The evaluation of the different *ObF1* shows that the minimization of *MaWA* does not improve *MeWA* and *DeMeWA*, meaning most water consumers would be affected at the expense of improving the water quality of a few. In conclusion, the use of averages, in particular the demand-weighted average, is recommended, because it would bring better water quality to most users.

Data Availability Statement

The data of the models of WDNs analyzed in the paper, the complete results of the simulations, and an executable file of the code generated to solve the optimization problem are available from the corresponding author by request.

Acknowledgments

C. Quintiliani was financially supported by the Ph.D. program of the University of Cassino and Southern Lazio, sponsored by the Italian Ministry for Education, University and Research. O. Marquez thanks the support of the Mexican government through CONACYT to fund this research. This research was in part supported by the Russian Science Foundation (Grant No. 17-77-30006) and by IHE Delft Hydroinformatics Research Fund.

References

Alfonso, L., L. He, A. Lobbrecht, and R. Price. 2013. "Information theory applied to evaluate the discharge monitoring network of the Magdalena River." *J. Hydroinf.* 15 (1): 211–228. https://doi.org/10.2166/hydro.2012.066.

Alfonso, L., A. Jonoski, and S. Dimitri. 2010. "Multiobjective optimization of operational responses for contaminant flushing in water distribution networks." *J. Water Resour. Plann. Manage.* 136 (1): 48–58. https://doi.org/10.1061/(ASCE)0733-9496(2010)136:1(48).

Andrade, M. A., C. Y. Choi, K. Lansey, and D. Jung. 2016. "Enhanced artificial neural networks estimating water quality constraints for the optimal water distribution systems design." *J. Water Resour. Plann. Manage.* 142 (9): 04016024. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000663.

Banik, B. K., L. Alfonso, C. Di Cristo, and A. Leopardi. 2017a. "Greedy algorithms for sensor location in sewer systems." *Water* 9 (11): 856. https://doi.org/10.3390/w9110856.

Banik, B. K., L. Alfonso, C. Di Cristo, A. Leopardi, and A. Mynett. 2017b. "Evaluation of different formulations to optimally locate sensors in sewer systems." *J. Water Resour. Plann. Manage.* 143 (7): 04017026. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000778.

Bi, W., and G. Dandy. 2014. "Optimization of water distribution systems using online retrained metamodels." *J. Water Resour. Plann. Manage.* 140 (11): 04014032. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000419.

Boccelli, D. L., M. E. Tryby, J. G. Uber, L. A. Rossman, M. L. Zierolf, and M. M. Polycarpou. 1998. "Optimal scheduling of booster disinfection in water distribution systems." *J. Water Resour. Plann. Manage.* 124 (2): 99–111. https://doi.org/10.1061/(ASCE)0733-9496(1998)124:2(99).

Carpentier, P., and G. Cohen. 1993. "Applied mathematics in water supply network management." *Automatica* 29 (5): 1215–1250. https://doi.org/10.1016/0005-1098(93)90048-X.

Castro Gama, M. E., Q. Pan, S. Salman, and A. Jonoski. 2015. "Multivariate optimization to decrease total energy consumption in the water supply of abbiategrasso (Milan, Italy)." *Environ. Eng. Manage. J.* 14 (9): 2019–2029.

Cembrano, G., G. Wells, J. Quevedo, R. P. Pérez, and R. Argelaguet. 2000. "Optimal control of a water distribution network in a supervisory control system." *Control Eng. Pract.* 8 (10): 1177–1188. https://doi.org/10.1016/S0967-0661(00)00058-7.

Cohen, D., U. Shamir, and G. Sinai. 2000a. "Optimal operation of multi-quality water supply systems. I: Introduction and the QC model." *Eng. Optim. A35* 32 (5): 549–584. https://doi.org/10.1080/03052150008941313.

Cohen, D., U. Shamir, and G. Sinai. 2000b. "Optimal operation of multi-quality water supply systems. II: The QH model." *Eng. Optim. A35* 32 (6): 687–719. https://doi.org/10.1080/03052150008941318.

Creaco, E., S. Alvisi, and M. Franchini. 2015. "Multistep approach for optimizing design and operation of the C-town pipe network model." *J. Water Resour. Plann. Manage.* 142 (5): C4015005. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000585.

Creaco, E., M. Franchini, and S. Alvisi. 2012. "Evaluating water demand shortfalls in segment analysis." *Water Resour. Manage.* 26 (8): 2301–2321. https://doi.org/10.1007/s11269-012-0018-0.

Di Cristo, C., and A. Leopardi. 2008. "Pollution source identification of accidental contamination in water distribution networks." *J. Water Resour. Plann. Manage.* 134 (2): 197–202. https://doi.org/10.1061/(ASCE)0733-9496(2008)134:2(197).

Di Cristo, C., A. Leopardi, and G. de Marinis. 2015. "Assessing measurements uncertainty on trihalomethanes prediction through kinetic models in water supply systems." *J. Water Supply Res. Technol. AQUA* 64 (5): 516–528. https://doi.org/10.2166/aqua.2014.036.

Di Nardo, A., M. Di Natale, and G. F. Santonastaso. 2014. "A comparison between different techniques for water network sectorization." *Water Sci. Technol. Water Supply* 14 (6): 961–970. https://doi.org/10.2166/ws.2014.046.

Farina, G., E. Creaco, and M. Franchini. 2014. "Using EPANET for modelling water distribution systems with users along the pipes." *Civ. Eng. Environ. Syst.* 31 (1): 36–50. https://doi.org/10.1080/10286608.2013.820279.

Fu, G., Z. Kapelan, J. Kasprzyk, and P. Reed. 2013. "Optimal design of water distribution systems using many-objective visual analytics." *J. Water Resour. Plann. Manage.* 139 (6): 624–633. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000311.

Giustolisi, O., D. Laucelli, and L. Berardi. 2012. "Operational optimization: Water losses versus energy costs." *J. Hydraul. Eng.* 139 (4): 410–423. https://doi.org/10.1061/(ASCE)HY.1943-7900.0000681.

Jamieson, D. G., U. Shamir, F. Martinez, and M. Franchini. 2007. "Conceptual design of a generic, real-time, near-optimal control system for water-distribution networks." *J. Hydroinf.* 9 (1): 3–14. https://doi.org/10.2166/hydro.2006.013.

Jolly, M. D., A. D. Lothes, L. S. Bryson, and L. Ormsbee. 2012. "Research database of water distribution system models." *J. Water Resour. Plann. Manage.* 140 (4): 410–416. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000352.

Jowitt, P. W., and G. Germanopoulos. 1992. "Optimal pump scheduling in water-supply networks." *J. Water Resour. Plann. Manage.* 118 (4): 406–422. https://doi.org/10.1061/(ASCE)0733-9496(1992)118:4(406).

Kang, D., and K. Lansey. 2009. "Real-time optimal valve operation and booster disinfection for water quality in water distribution systems." *J. Water Resour. Plann. Manage.* 136 (4): 463–473. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000056.

Kanta, L., E. Zechman, and K. Brumbelow. 2011. "Multiobjective evolutionary computation approach for redesigning water distribution systems to provide fire flows." *J. Water Resour. Plann. Manage.* 138 (2): 144–152. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000156.

Li, C., J. Z. Yu, T. Q. Zhang, X. W. Mao, and Y. J. Hu. 2015. "Multiobjective optimization of water quality and rechlorination cost in water distribution systems." *Urban Water J.* 12 (8): 646–652. https://doi.org/10.1080/1573062X.2014.939093.

Machell, J., and J. Boxall. 2014. "Modeling and field work to investigate the relationship between age and quality of tap water." *J. Water Resour. Plann. Manage.* 140 (9): 04014020. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000383.

Machell, J., J. Boxall, A. Saul, and D. Bramley. 2009. "Improved representation of water age in distribution networks to inform water quality." *J. Water Resour. Plann. Manage.* 135 (5): 382–391. https://doi.org/10.1061/(ASCE)0733-9496(2009)135:5(382).

Mala-Jetmarova, H., N. Sultanova, and D. Savic. 2018. "Lost in optimisation of water distribution systems? A literature review of system design." *Water (Switzerland)* 10 (3): 307. https://doi.org/10.3390/w10030307.

Marquez-Calvo, O., C. Quintiliani, L. Alfonso, C. Di Cristo, A. Leopardi, D. Solomatine, and G. de Marinis. 2018. "Robust optimization of valve management to improve water quality in WDNs under demand uncertainty." *Urban Water J.* 15 (10): 943–952. https://doi.org/10.1080/1573062X.2019.1595673.

Menapace, A., D. Avesani, M. Righetti, A. Bellin, and G. Pisaturo. 2018. "Uniformly distributed demand EPANET extension." *Water Resour. Manage.* 32 (6): 2165–2180. https://doi.org/10.1007/s11269-018-1924-6.

Ostfeld, A., and E. Salomons. 2006. "Conjunctive optimal scheduling of pumping and booster chlorine injections in water distribution systems." *Eng. Optim.* 38 (03): 337–352. https://doi.org/10.1080/03052150500478007.

Prasad, T. D., and G. A. Walters. 2006. "Minimizing residence times by rerouting flows to improve water quality in distribution networks." *Eng. Optim.* 38 (8): 923–939. https://doi.org/10.1080/03052150600833036.

Quintiliani, C., L. Alfonso, C. Di Cristo, A. Leopardi, and G. de Marinis. 2017. "Exploring the use of operational interventions in water distribution systems to reduce the formation of TTHMs." *Procedia Eng.* 186: 475–482. https://doi.org/10.1016/j.proeng.2017.03.258.

Quintiliani, C., C. Di Cristo, and A. Leopardi. 2018. "Vulnerability assessment to trihalomethane exposure in water distribution systems." *Water* 10 (7): 912. https://doi.org/10.3390/w10070912.

Rossman, L. A. 1999. "The EPANET programmer's toolkit for analysis of water distribution systems." In *Proc., Annual Water Resources Planning and Management Conf.*, 1–10. Reston, VA: ASCE.

Rossman, L. A. 2000. *EPANET 2: User's manual*. EPA/600/R-00/057. Cincinnati: National Risk Management Research Laboratory, USEPA.

Seyoum, A. G., and T. T. Tanyimboh. 2017. "Integration of hydraulic and water quality modelling in distribution networks: EPANET-PMX." *Water Resour. Manage.* 31 (14): 4485–4503. https://doi.org/10.1007/s11269-017-1760-0.

Shokoohi, M., M. Tabesh, S. Nazif, and M. Dini. 2017. "Water quality based multi-objective optimal design of water distribution systems." *Water Resour. Manage.* 31 (1): 93–108. https://doi.org/10.1007/s11269-016-1512-6.

Tiwari, S., G. Fadel, and K. Deb. 2011. "AMGA2: Improving the performance of the archive-based micro-genetic algorithm for multi-objective optimization." *Eng. Optim.* 43 (4): 377–401. https://doi.org/10.1080/0305215X.2010.491549.

Ulanicki, B., and P. R. Kennedy. 1994. "An optimization technique for water network operations and design." In *Proc., 33rd Conf. on Decision and Control*, 4114–4115. Piscataway, NJ: IEEE.

USEPA. 2002. *Effects of water age on distribution system water quality.* Washington, DC: USEPA.

Zhao, Y., Y. J. Yang, Y. Shao, Y. Lee, and T. Zhang. 2018. "Demand-driven spatiotemporal variations of flow hydraulics and water age by comparative modeling analysis of distribution network." *J. Water Resour. Plann. Manage.* 144 (12): 04018074. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000995.