Optimization of Potable Water Distribution and Wastewater Collection Networks: A Systematic Review and Future Research Directions

Wanqing Zhao, *Member, IEEE,* Thomas H. Beach, and Yacine Rezgui

***Abstract*—Potable water distribution networks (WDNs) and wastewater collection networks (WWCNs) are the two fundamental constituents of the complex urban water infrastructure. Such water networks require adapted design interventions as part of retrofitting, extension and maintenance activities. Consequently, proper optimization methodologies are required to reduce the associated capital cost while also meeting the demands of acquiring clean water and releasing wastewater by consumers. In this paper, a systematic review of the optimization of both WDNs and WWCNs, from the preliminary stages of development through to the state-of-the-art, is jointly presented. Firstly, both WDNs and WWCNs are conceptually and functionally described along with illustrative benchmarks. The optimization of water networks across both clean and waste domains is then systematically reviewed and organized, covering all levels of complexity from the formulation of cost functions and constraints, through to traditional and advanced optimization methodologies. The rationales behind employing these methodologies as well as their advantages and disadvantages are investigated. The paper then critically discusses current trends and identifies directions for future research by comparing the existing optimization paradigms within WDNs and WWCNs and proposing common research directions for optimizing water networks. Optimization of urban water networks is a multidisciplinary domain, within which this paper is anticipated to be of great benefit to researchers and practitioners.**

***Index Terms*—Artificial intelligence, hydraulics, network optimization, wastewater collection networks, water distribution networks.**

I. Introduction

POTABLE water distribution networks (WDNs) [1]–[3] and wastewater collection networks (WWCNs) [4]–[6] are essential components of the urban water value chain that generally encompasses abstracting and treating raw water, distributing and consuming potable water, collecting and treating wastewater and discharging or reusing the final effluent or sludge. Here, WDNs are defined as the networks employed to deliver the potable water from treatment works to various residential and business consumers, while WWCNs are the networks used to collect wastewater (residential and industrial sewage, stormwater, etc.) and convey it to wastewater treatment plants (WWTPs). Typical components that make up a WDN include pipes, valves, reservoirs/tanks and clean water pumping stations; while the main components that constitute a WWCN include sewers, manholes and sewage pumping stations. Due to the distinct functions of WDNs and WWCNs, their components and functionalities within the water networks are different, as are their interior hydraulic behaviors. The final complexity is that, although both WDNs and WWCNs are part of the essential infrastructure for an urban environment, they are often operated by different water utilities or local authorities.

Within the water value chain, WDNs and WWCNs respectively represent the upstream water infrastructure dealing with clean water and the downstream water infrastructure dealing with wastewater. Due to the process of urbanization, changing legislative standards and the ageing of existing infrastructure, the design of new water networks or the rehabilitation of existing networks is one of the most pressing issues faced by public service providers. The capital cost of these networks usually incurs huge water infrastructure investment. For instance, the Thames Tideway Tunnel project [7], currently being explored by the UK government, aims to tackle the flushing of raw sewage overflows from central London's aging Victorian drainage system directly into the river Thames. The proposed new tunnel will be roughly 25km long and the estimated capital cost of the project is currently £4.2bn. It is hoped that the outcomes will bring the UK into compliance with the EU Urban Wastewater Treatment Directive.

For WDNs, networks are designed to enable the delivery of clean water to consumers, while meeting requirements such as achieving sufficient tap pressure. However, oversized WDNs can also lead to large capital and operational costs and poor water quality issues [8], [9]. On the other hand, WWCNs are designed to be capable of conveying wastewater to WWTPs with little or no overflow in order to avoid/alleviate urban pollution problems, such as the adverse impact on public health and the environment. The capacity of WWCNs should be large enough to carry the peak sewage and/or stormwater flows to WWTPs. If not, combined sewer overflows (CSOs) or sanitary sewer overflows (SSOs), or even sewer flooding [10] can occur, releasing partially treated or untreated wastewater into the environment. Hence, the design of both WDNs and WWCNs is considered as an optimization task that searches for optimal (near-optimal) solutions for the undetermined decision variables in the network of interest. Considerable savings are thus anticipated from using optimization methods to appropriately dimension the water network components (e.g., the pipes/sewers) while typically following a given network layout

This work is supported by the EU Seventh Framework Programme (FP7) under Grant 619795 involving eleven partners, including Universities and Institutes, local authorities, water utilities and ICT companies.

The authors are with Cardiff School of Engineering, Cardiff University, Cardiff CF24 3AA, U.K. (E-mail: ZhaoW9@cardiff.ac.uk; BeachTH@cardiff.ac.uk; RezguiY@cardiff.ac.uk).

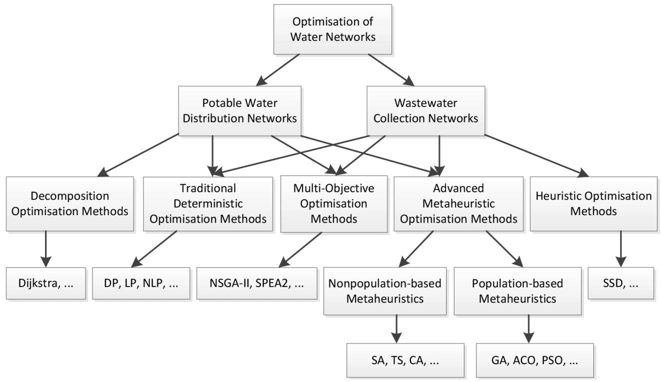


Fig. 1. Taxonomy of the optimization methodologies of water networks.

and satisfying a number of hydraulic and physical constraints such as the hydraulic water and wastewater behaviors.

The key task of water network optimization entails selecting between alternative solutions, which leads to finding the best or locally optimal solution(s) for dimensioning components involved in WDNs and WWCNs. It is a complex multidisciplinary task covering water and environmental management, artificial intelligence and ICT (Information Communication Technology) fields. A substantial number of optimization methodologies have been developed in this domain over the past few decades, which can be roughly categorized as shown in Fig. 1 though overlaps might exist between them. Firstly, two key categories of optimization methods employed in both WDNs and WWCNs are traditional deterministic and advanced metaheuristic methods. Deterministic optimization methods usually rely heavily on rigorous mathematical calculations. For example, the gradient of an objective function is required in a nonlinear programming approach. Depending on the number of solutions being dealt with simultaneously in the optimization process, the metaheuristic approach can be further divided into nonpopulation-based metaheuristics (usually only with a single solution being processed per iteration/generation) and population-based metaheuristics (with multiple solutions being processed per iteration/generation). Then, multi-objective optimization methods considering more than one (usually conflicting) objective are also a promising research area for the optimization of both WDNs and WWCNs. According to the distinct characteristics between WDNs and WWCNs, the decomposition and the intuitive heuristic approaches are therefore popularly studied. Furthermore, it is noted that the hybridization of various types of optimization methods is also seen in the domain.

In this paper, the contribution first lies in a systematic and comprehensive review of the optimization of both WDNs and WWCNs from the preliminary stages of development through to the state-of-the-art. Subsequently, current trends and future research directions are identified and discussed. A large spectrum of optimization research into the design of water networks is accessibly brought together in a multidisciplinary setting. It is worth noting that the two types of water networks across both clean and waste domains are jointly and critically analyzed, as opposed to traditional approaches which usually consider these in isolation. The motivation for jointly addressing the optimization of WDNs and WWCNs comes from the fact that *a)* they are essentially both an integral part of the complex water network but deal with different types of water (i.e., potable water and wastewater) and *b)* it is useful to learn from the differences and research gaps in designing such water networks from the optimization point of view, to mutually benefit from research advances and experiences, and to identify common future research directions in a united way. In that respect:

- the water networks, objective functions and associated constraints are clearly presented providing an introduction to the underpinning theory for the optimization task in order to broaden the reach of the paper, from the original field of water and environmental management to other related fields, especially artificial intelligence;
- an in-depth review of various traditional and advanced optimization methodologies for the optimization of both WDNs and WWCNs are systematically presented. The paper clearly presents and focuses on the intrinsic optimization problem of both types of water networks and the underpinning rationales of the optimization methods being used in the domain. The strengths and limitations behind using these methods are also thereby discussed. It has been found that research into optimization techniques within WWCNs is far less developed than within WDNs, especially in the use of advanced optimization approaches such as metaheuristics and multi-objective optimization;
- critical discussion of current trends and identification of potential future research directions are then presented by comparing the existing optimization paradigms in WDNs and WWCNs and by proposing common research directions for optimizing both types of water networks. These include a variety of aspects such as network characteristics, optimization methodologies, hydraulic simulator adoption, constraints handling, unified objective functions, central repository construction, fair comparisons of optimization methods, network modeling, etc. High level guidance and suggestions for the development of new efficient and effective metaheuristic optimization paradigms for the domain are also described in the paper.

It is envisioned that this paper will enable more researchers, especially those from the field of artificial intelligence, ICT and water and environmental management, to focus on the key optimization problems in the domain, i.e., how to effectively (in terms of finding optimal solutions) and efficiently (in terms of developing computational inexpensive algorithms) optimize water networks with a view to minimizing network capital cost while addressing various physical and hydraulic constraints.

The paper is organized as follows. Section II presents the background description for WDNs and WWCNs. The mathematical formulation of the optimization problems and the systematic review of various optimization methodologies for both WDNs and WWCNs respectively, are given in Sections III and IV. Section V discusses current trends and potential opportunities for future research directions. Finally, Section VI concludes the paper.

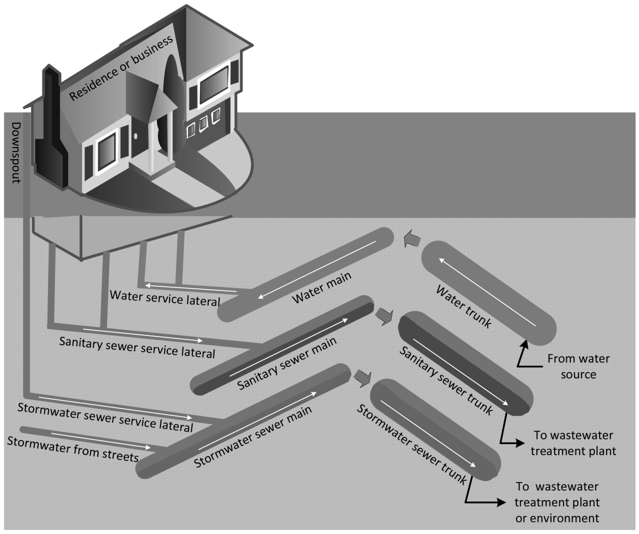


Fig. 2. Schematic of potable water distribution and wastewater collection services for a property.

II. Water Network Description

As basic constituents of the urban water infrastructure used to deliver clean water to consumers and to collect wastewater from consumers and surface runoffs (commonly shown as in Fig. 2) [11], water distribution networks (WDNs) and wastewater collection networks (WWCNs) are fundamental to our daily lives. The roles and characteristics of WDNs and WWCNs are described in this section to elicit the network optimization tasks. Well-known benchmarks are also depicted to provide an intuitive understanding of water networks.

A. Potable Water Distribution Networks

Generally, potable water distribution networks (WDNs) [1]–[3] refer to the water infrastructure used to distribute potable water obtained from treatment works to consumers within a local region. A typical WDN is comprised of pumps, reservoirs/tanks, pipes and valves. Normally, multiple treatment works (indicating the availability of multiple sources of water) exist to provide a sufficient amount of potable water for the region and also to increase the network reliability. In reality, the adoption of loops in the design of WDNs further helps improve the network reliability as alternative water flow paths are created to address pipe failures or maintenance work [12]. Once the raw water is treated, potable water is obtained and leaves the water treatment works through pipes normally driven by pumping, gravity or a combination of both. The potable water is then pressurized in the network and delivered to different consumer nodes. Additionally, within a WDN, potable water is often stored in reservoirs/tanks for the short term, to aid with pressurizing the network, satisfying peak water demand throughout the day and emergency situations. Fig. 2 shows the point at which the WDN arrives at the consumer's property, indicating how a property is typically connected to the water utility or local authority owned water main.

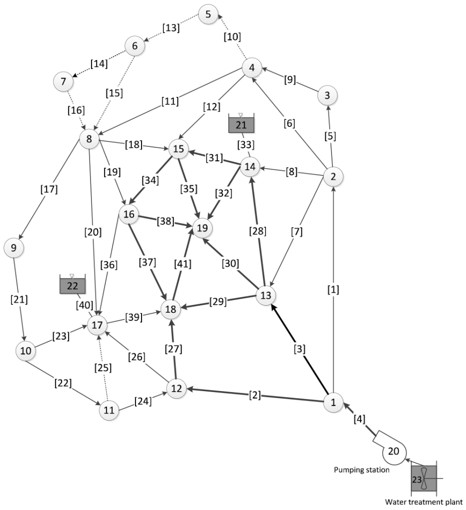
To obtain an initial impression of WDNs and a better understanding of the optimization tasks to be investigated, Fig. 3a depicts the commonly used Anytown benchmark (presented in the battle of the network models (BNM) in 1985), aimed to provide a relatively realistic network having typical features of real systems [13]. Here, the arcs denote the pipes and the nodes denote the treatment plants, reservoirs/tanks, pumps, junctions and demand points where necessary. This benchmark includes a water treatment works, a pumping station, two tanks, 19 consumer nodes and 41 pipelines. Three identical pumps in the pumping station are employed to take potable water from the clear well at the treatment works and pump it into the WDN. The pipes denoted as solid lines are the existing ones, of which the thick and the thin lines are in the central city and residential area, respectively. The dashed lines represent new pipelines. There are also some other frequently used benchmarks, such as the two-loop network [14], Hanoi network [15], NYCT (New York City Tunnel) [16] and the two-reservoir network [17]. The length of each pipeline, the demand, elevation and minimum required head of each node, the available types of pipes and their cost, etc., can all be found in the corresponding references.

There are two types of hydraulic mechanisms that have to be followed within a WDN, i.e., the conservation law of mass and the conservation law of energy [1]. The conservation law of mass is used to describe the continuity of flow at every node in the network, while the energy conservation law states that the pressure head losses accumulated along a closed loop should be zero. Given the nodal demands and the network properties for a specified network layout, the water flows and pressures in a WDN can be determined by utilizing these hydraulic mechanisms. The famous simulation model EPANET [19] is widely used to simulate the functions of the hydraulic mechanisms and is commonly used in the domain of WDN optimization. There are also some other simulation models that have been developed in the last forty years, e.g., InfoWorks WS [20], Cross [21], HYDROFLO [22], KYPipe [23]. It is also worthwhile mentioning that a WDN can also be conveniently managed by dividing the network into a set of district metering areas (DMAs) using isolation valves [24]. These DMAs are hydraulically isolated from one another and the amount of incoming and outgoing water to a DMA can thus be easily measured and controlled, where a DMA manager can be involved for various operations. These characteristics allow burst pipes to be easily located, and this was perhaps the initial motivation behind the adoption of DMAs within WDNs.

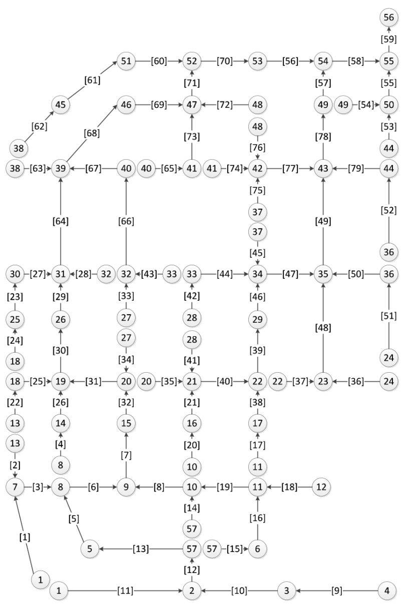
B. Wastewater Collection Networks

Wastewater collection networks (WWCNs), on the other hand, refer to the water infrastructure used to collect and convey wastewater from consumers and/or surface runoffs to wastewater treatment plants (WWTPs). As a result, wastewater can be treated in the WWTPs rather than being directly discharged in order to protect the environment and to prevent outbreaks of disease. Fig. 2 shows how the WWCN connects to both properties and street drainage by adopting a separate sewer system.

According to the different types of wastewater (sewage and/or stormwater) collected in WWCNs, there are generally



(a) Anytown network [13]



(b) Li and Matthew network [18]

Fig. 3. Benchmark of potable water distribution and wastewater collection networks.

two types of WWCNs, i.e., combined sewer systems and separate sewer systems. The combined sewer systems [25] (aka combined sewage or sewerage systems) which are common in old urban wastewater systems are adopted to transport both stormwater (e.g., rainwater, snow, hail, etc.) and sewage in the same sewers. As the stormwater volumes are hard to predict and can reach large quantities in a short time period, the phenomenon of combined sewer overflows (CSOs) is reasonably common, resulting in serious environment contamination. Modern wastewater collection systems are often comprised of both sanitary sewer systems and stormwater systems as shown in Fig. 2, which transport sewage separate from stormwater. On the one hand, sanitary sewer systems [26] (aka sewage or sewerage systems) are only used to transport sewage from domestic properties and wastewater from businesses to WWTPs. It should be noted that businesses are required to have appropriate approvals from local authorities for discharging their wastewater. On the other hand, stormwater systems [27] are used to collect stormwater that flows across surfaces, such as downspouts, streets and footpaths. The stormwater is then transported to waterways with little or no treatment (Some legislation may require a certain level of stormwater treatment).

Apart from the different types of wastewater collected in WWCNs, there are also varying ways in which wastewater is conveyed, i.e., gravity sewer systems, pressure sewer systems, septic tank effluent drainage (STED, aka effluent sewer system) and vacuum sewer systems [28]. Of these, gravity sewer systems are the most common, which convey wastewater mainly by gravity to a WWTP in order to save energy. Therefore, the optimization of gravity sewer systems constitutes the main topic to be investigated for WWCNs in this paper. In a gravity sewer system, the sewers have to be buried at proper depths with sufficient gradients in order to keep the wastewater moving through the network. In addition, manholes are used to connect different sewers to meet the changes in flow directions and sewer slopes, and also for the convenience of sewer cleaning and flushing. Depending on the geography, excavation cost and other constraints, lift pumps are occasionally required to raise wastewater from a lower sewer line to a higher one.

As in the WDN description, Fig. 3b also depicts a relatively complex gravity sewer network proposed by Li and Matthew [18], in which the arcs denote the sewers and the nodes denote the manholes. It was designed for the residential area of Shengli Oilfield in China with a drainage area and population of 2.6 km^2 and 10,000, respectively, where the wastewater is delivered to outlet 56 for treatment. There are a total of 80 manholes (some have the same locations) and 79 sewers. The maximum and minimum velocities, the minimum slope, the maximum flow depth-to-diameter ratio, the minimum cover and the network characteristics were also defined therein. Another commonly seen benchmark in the domain is the hypothetical Mays and Wenzel network [29]. Unlike WDNs, WWCNs normally have a tree-like structure, and the hydraulic behavior lies in each sewer within the network and exhibits complex nonlinear relationships between wastewater flows, sewer types and slopes and flow depths, which can be simulated by using SWMM (storm water management model) [30].

III. Water Distribution Network Optimization

Assuming that the nodal demands are known a priori, the goal of WDN optimization is to choose the type of each network component (e.g., pipes, reservoirs, pumps) at minimum capital cost while satisfying all hydraulic and physical constraints. This task together with that of WWCN optimization (to be presented in the next section) appears when designing a completely new network for a new urban area or rehabilitating the existing network to cope with aging network components, increased urbanization or upgraded standards. It is usually assumed that the network layout is predefined together with specified nodal demands [31]. In this context, most interest has focused on selecting and dimensioning the necessary types of pipes for WDNs.

A. Cost Function and Constraints

Suppose there are a total of N_n network nodes, N_p pipes to be installed and N_t different types of pipes available in the market, the basic mathematical objective (cost) function to be minimized together with the associated constraints [1], [2] can be formulated as

$$J(\boldsymbol{\theta}) = \sum_{i=1}^{N_p} L_i u(\theta_i) \ \text{ or } \ J(\boldsymbol{\theta}) = \sum_{i=1}^{N_p}\sum_{j=1}^{N_t} \vartheta_{i,j} L_i u(\zeta_j), \tag{1}$$

$$\text{s.t.} \begin{cases} \begin{cases} \theta_i \in \{\zeta_1, \ldots, \zeta_{N_t}\},\ i = 1, \ldots, N_p, & (2) \\ \sum_{j=1}^{N_t} \vartheta_{i,j} = 1,\ i = 1, \ldots, N_p, & (3) \\ \vartheta_{i,j} \in \{0, 1\},\ i = 1, \ldots, N_p,\ j = 1, \ldots, N_t, & (4) \end{cases} \\ Q_i^{\text{ext})} + \sum_{j=1}^{N_n} Q_{j,i}^{\text{in})} = Q_i^{n)} + \sum_{k=1}^{N_n} Q_{i,k}^{\text{out})},\ i = 1, \ldots, N_n, & (5) \\ \begin{cases} \sum_{i \in \mathbf{P}} \Delta H_i = H_k - H_l,\ \mathbf{P} \in \mathbf{NP}, & (6) \\ \Delta H_i = H_{i,1} - H_{i,2}, i = 1, \ldots, N_p, & (7) \end{cases} \\ H_{i,\text{min}} \le H_i \le H_{i,\text{max}},\ i = 1, \ldots, N_n, & (8) \end{cases}$$

where L_i (m) is the length of the ith pipe and $u(\theta_i)$ is the cost per unit length of the pipe having type θ_i, $\{\zeta_1, \ldots, \zeta_{N_t}\}$ is a series of commercially available pipe types, $Q_i^{\text{ext})}$ (m^3/s) is the quantity of external inflow of water (e.g., purificated rainwater) for the ith node per second, $Q_i^{n)}$ (m^3/s) is the quantity of water consumed at the ith node per second, $Q_{j,i}^{\text{in})}$ (m^3/s) and $Q_{i,k}^{\text{out})}$ (m^3/s) are respectively the quantity of incoming water and outgoing water from the jth node and to the kth node per second with respect to the ith node, ΔH_i (m) is the head loss in the ith pipe, $\mathbf{P}$ is one of the path from the path set $\mathbf{NP}$ (consisted of a series of successively connected pipes) involved in the network, $H_{i,1}$ (m) and $H_{i,2}$ (m) are respectively the head at each end of the ith pipe, H_i (m), $H_{i,\text{min}}$ (m) and $H_{i,\text{max}}$ (m) are respectively the ith nodal head together with its minimum and maximum head allowances. Specifically, H_k and H_l are respectively the pressure head at each end of the path $\mathbf{P}$.

The aim is to find an optimum set of WDN decision variables ($\hat{\boldsymbol{\theta}} = [\hat{\theta}_1, \ldots, \hat{\theta}_{N_p}]^{\text{T}}$) that minimize the total cost $J(\boldsymbol{\theta})$ in the left-hand side of (1) based on the constrained decision variables defined in (2) where they have to be chosen from a series of commercially available pipe types. Alternatively, if the unit cost of the jth available type of pipes is denoted by $u(\zeta_j)$ and the decision variable is denoted by $\vartheta_{i,j}$ determining whether the jth type has been chosen for the ith pipe (1 for yes and 0 for no), the objective function can be reformulated as in the right-hand double summation manner of (1). In this manner, the decision variables are required to be subject to the constraints that they are binary and that only one type of pipe can be chosen for a particular pipeline, as defined in (3) and (4). If distinct available types of pipes are used for different pipelines, a specific unit cost function instead of a united one can be adopted for each pipeline, i.e., $u_i(\cdot)$, $i = 1, \ldots, N_p$. Other costs in WDNs design other than those generated from the pipes, such as those associated with the sizing and location of pumps and reservoirs, can also be formulated in a similar way with expanded design decision vector $\boldsymbol{\theta}$.

Apart from the decision variable constraints described in (2), or alternatively in (3) and (4), the formulas (5)-(8) list the basic hydraulic constraints that have to be satisfied for the distribution network optimization, i.e., the conservation laws of mass and energy and the minimum and maximum head requirements in the demand nodes. The mass conservation law (5) indicates the continuity of flow where the total incoming amount of water per second for a node in the network is equal to the amount consumed in the node plus the outgoing amount of water, which applies on every node ($i = 1, \ldots, N_n$) in a WDN. It should be noted that $Q_{j,i}^{\text{in})} = 0$ and $Q_{i,k}^{\text{out})} = 0$ hold if there are no direct pipelines connected between the corresponding two nodes (j and i and, i and k). Regarding (6) and (7), they are used to describe the energy conservation law, where the head losses (mainly caused by friction in pipes) accumulated along a path between two nodes (the kth and lth) should be equal to the difference of their nodal heads. It should be noted that if the path is a closed loop, then $H_k = H_l$ as $k = l$. The head loss ΔH_i (m) in the ith pipe refers to the head difference between $H_{i,1}$ (m) and $H_{i,2}$ (m) at each end of the pipe and this can roughly be approximated by using Hazen-Williams formula [19]:

$$\Delta H_i = \alpha \frac{(Q_i)^\beta L_i}{(C_i^{\text{HW})})^\beta (D_i)^\gamma},\ i = 1, \ldots, N_p, \tag{9}$$

where α, β and γ are the associated parameters (they are set as 10.667, 1.852 and 4.871, respectively, in EPANET 2.0) and, Q_i (m^3/s), L_i (m), D_i (m) and $C_i^{\text{HW})}$ are the water flow rate, length, diameter and Hazen-Williams roughness coefficient (depending on the material of pipes) for the ith pipe. Here, the roughness coefficient $C_i^{\text{HW})}$ and diameter D_i are related to the types of pipes. Another widely used approximation formula is the Darcy-Weisbach [32] which is more accurate in certain circumstances but more time-consuming to solve. Finally, equation (8) simply indicates that there exists a minimum and maximum head constraint, i.e., $H_{i,\text{min}}$ (m) and $H_{i,\text{max}}$ (m), for the actual head H_i (m) on every node i ($i = 1, \ldots, N_n$) in order to meet the compliance requirements. Furthermore, some other constraints/goals may also be studied, such as the bound of flow velocity [2], epistemic and aleatory uncertainties [33], water quality [34], reliability in case of pipe breaking or pump station failure [35]. While many studies focus on the optimization of potable water networks for single loading where the demand pattern is fixed for every node in the network (can be taken as the estimated peak loading), multiple loadings where several different demand patterns are involved, are also seen in some designs [14], [36] in order to satisfy varying nodal demands in different time periods. For instance, if only the peak loading is considered, a reservoir may not be filled during periods of low demand (usually in the night) by consumers which would in turn affect the continuous provision of potable water to consumers during peak demands.

B. Traditional Deterministic Optimization Methods

Given the objective function and constraints listed in (1)-(8), the distribution network determination is an NP-hard (where it can be reduced from other known NP-hard problems, e.g., 0-1 Knapsack problem, in polynomial time) and combinatorial optimization problem [37], where the decision variables have to be chosen from a series of commercially available types of pipes. It is also a highly nonlinear problem mainly due to the set of nonlinear head loss equations (depending on the number of loops in the network) (6), (7) and (9) that need to be solved for an ordinary looped network. Due to their ability to provide unequivocal results, deterministic optimization methods were first introduced to optimize distribution networks. Here, the deterministic optimization is defined as the optimization methods that usually comply with rigorous mathematics [38]. Essentially, only an exhaustive search method such as enumeration can guarantee a globally optimal solution of WDN optimization [37]. Unfortunately, its computational demand increases exponentially with the number of pipes in the network (where a total of $(N_t)^{N_p}$ solutions space needs to be examined). Gessler also proposed a selective enumeration method to reduce the search space although the global optimum could be eliminated during the pruning process [17]. It is worth pointing out that there are also some heuristic methods proposed for the optimization of WDNs, but they are not as popular as in the optimization of WWCNs (to be presented in the next section). For example, recently, a computationally efficient heuristic pipe sizing procedure was proposed in [39], where two stages were involved. Initially, all the pipes were set as their minimum size from the commercially available types. Then, in the first stage, the pipes with the maximum flow velocity were successively selected to increase in size until the nodal pressure requirement was satisfied for every demand node. In the second stage, at each step a pipe was selected according to one of the six alternative selection indices suggested by the author, for possible size reduction to the next commercially available size. This continuous process stops when a node violates the pressure head requirement. These kinds of pure heuristic design processes in general are limited by the lack of solution optimality although potentially very efficient.

Relying on exact mathematical derivation, linear programming (LP) and its variants were amongst the traditional mathematical attempts to solve the problem with low computational complexity (computational burden in terms of obtaining the final solution). However, their disadvantages are deemed in the local optimality (the obtained solutions are locally optimal with respect to the overall objective) and in the resulted split pipe solutions (a pipeline could thereby consist of subpipes of different types, which is impractical from an engineering perspective) [14], [15], [40]. Dynamic programming (DP) was also applied to optimize the distribution network stage by stage rather than making simultaneous decisions for the whole network [41], [42], although it would be very complex to deal with looped systems in stages. Due to the nonlinear nature of the problem of interest, nonlinear programming (NLP) was thus applied usually through the use of the generalized reduced gradient method [43], [44], in which the conservation laws of mass and energy were implicitly solved by a hydraulic simulator, while the nodal head constraints were considered in an augmented Lagrangian manner. However, given the nature of NLP, it is commonly recognized that only continuous types of pipes are normally dealt with by the optimization (rounding solutions is thus required) and local optima still occur (highly dependent on the initial solution).

To handle the discrete types of pipes, integer linear programming (ILP) [2] was proposed wherein the decision variables were denoted as a series of zero-unity variables as in the right-hand side objective formulation defined in (1). An iterative searching process running between the hydraulic simulation (finding the pipe flows for the looped networks) and an ILP solver (finding the intermediate solution of the network), was carried out based on an initial solution, until the whole optimization process converged to give the final solution. However, the optimality and the convergence of this iterative approach, especially when used with large scale networks have been queried [45], [46]. Particularly, the optimal solution also depends on the selected paths used to impose nodal head constraints, which in turn means that all the pipes included in the network are not globally optimized together.

C. Advanced Metaheuristic Optimization Methods

Metaheuristic optimization algorithms have recently been attracting substantial interest in the domain as they can easily handle various constraints such as the hydraulics and the discrete solution space, as well as being able to search amongst large solution space and locate near-optimal solutions. The terminology "metaheuristic" was derived from the composition of two Greek words, i.e., heuriskein (meaning "to find or to discover") and meta (meaning "beyond or higher level") [47]. Metaheuristics can thereby be defined as the high level strategies devised to efficiently and effectively explore and exploit search spaces of the problem of interest, in order to find the optimal solutions. Metaheuristics themselves are usually independent of the problem of interest, and are thus somewhat distinct from the heuristics which are often problem-specific and aim to make use of the peculiarities of the problem which can be perceived by human cognition such as the engineer's empirical knowledge. The constraints in metaheuristic optimization are most commonly handled by employing penalty functions which are easily aggregated in the cost function. A binary coding scheme can be simply adopted to choose the pipe type from the set of available types. However, for the metaheuristic algorithms that are good at dealing with real numbers, it is natural to round these numbers to the nearby commercially available pipe types after the various evolving operations performed at each generation. The nonpopulation-based and population-based metaheuristics for WDN optimization are now discussed as follows.

1) Nonpopulation-based Metaheuristics: As one branch of metaheuristics, the nonpopulation-based metaheuristic optimization methods refer to the ones in which only a single solution (no population) is dealt with per iteration/generation during the optimization. To avoid local optima, simulated annealing (SA) inspired by the physical annealing process operates on the single solution basis and iterates between neighboring solutions. Besides the usual cost reduction at an iteration, it also allows cost increment (in case of minimization problems), according to a probability function associated with the current solution, the new generated solution and a so-called controllable "temperature" parameter. It was adopted into the looped WDN optimization where the Newton-Raphson method was used to solve the hydraulic constraints during the optimization [48]. Tabu search (TS) is another well-known nonpopulation-based local search method which iteratively generates and examines neighboring solutions [49], controlled by memory structures (defining a tabu list, used to manage the visited solutions and/or some user defined rules which can vary with time). It was used for WDN optimization in [1] by incorporating a hydraulic simulator at each iteration, where a diversification procedure for establishing the rules for the generation of new solutions was involved when local optima occurred.

As mathematical models of complex systems with many simple identical components, cellular automata (CAs) consist of a number of interconnected cells, each accompanied by some cell states [50]. The states of all cells are able to evolve synchronously over time-steps according to the local transition rules used to define the interactions between the cells and their neighbors. The key characteristics of a CA embody parallelism, localist representation and homogeneity [51]. CA was applied to WDN optimization by Keedwell and Khu, in which CANDA (cellular automaton network design algorithm) was designed [51]. The network nodes were considered as the cells with the diameter of their inflow pipes as the cell states. Simple heuristic local rules obtained from engineers' knowledge were used on each of these cells to update the cell states. Unlike ordinary optimization methods, the best solution was found from the entire optimization process between the starting point and the repeating stable state (where oscillations started appearing). The advantages of this method lie in the small number of network evaluations and the fact that it does not need to bear the global objective in mind (but is concerned with local changes of cells), although the optimality of the final solution and the change of flow directions are possible weaknesses. Furthermore, other nonpopulation-based metaheuristic optimization methods such as iterated local search (ILS) [52] and variable neighborhood search (VNS) [53] have also recently arisen in the domain of WDN optimization.

2) Population-based Metaheuristics: As the name implies, population-based metaheuristic optimization methods deal with a set of solutions (populations do thus exist) at each iteration/generation during the optimization in order to avoid local optima. Genetic algorithms (GAs) are the most representative algorithms that fall into this category. They are based on the principle of natural selection usually comprising operations of reproduction, crossover and mutation for the evolution of populations. A set of chromosomes (solutions) are successively evaluated as fitness values according to the objective of the problem during a number of generations. GAs and their variants are amongst the most popular alternatives to the deterministic methods used in WDN optimization [54]–[56], where binary or integer coding can usually be adopted to deal with the discrete pipe diameters. It is noted that Nicklow et al. presented a dedicated synthesis on the use of various GAs and the associated operators for the general field of water resources planning and management (e.g., WDN optimization, groundwater monitoring and remediation) [57]. Ant colony optimization (ACO) belonging to the swarm intelligence family was inspired by the foraging behavior of ant colonies and it works in an iterative manner [58]. In ACO, each ant incrementally finds the elements to construct a trial solution at each iteration according to pheromone intensities and local information (environment). This environmental information is then also updated based on the cost of the constructed trial solutions at each iteration. To map the WDN optimization problem onto a graph that can be handled by ACO, Maier et al. [59] considered each pipe as a decision point together with the available pipe diameters as the choices at each decision point. The simulation results demonstrated that ACO produced slightly better solutions than GA in terms of lower capital costs and less computational times when given complex networks. A comparative study on the application of various variant ACOs to WDN optimization was also reported in [60].

To simulate the social behavior in bird flocking or fish schooling, particle swarm optimization (PSO) [61] was devised and has now been widely adopted in the general field of computational intelligence for various multidisciplinary applications. This algorithm deploys a number of particles (potential solutions) in a swarm. Each particle has a position vector employed to represent the current solution and a velocity vector used to show its search direction which is related to the best position locally tracked by it and the best position globally found by all the particles. Its application in WDN optimization has been widely reported [62], [63]. Differential evolution (DE) proposed by Storn and Price, as a continuous searching metaheuristic, often outperforms many other metaheuristic methods such as GAs [64]. As in GAs, DE operates on the basis of populations. The population at each generation successively undergoes the processes of mutation, crossover and selection, from which the contained individuals are gradually evolved to give better performance. Several researchers have recently successfully applied DE in the optimization of WDNs [65], [66]. Harmony search (HS) is a relatively recent metaheuristic inspired by the music improvisation process [67]. HS is initialized by constructing a harmony memory (HM) filled with a number of harmonies (solutions). In each iteration, only one new harmony is improvised according to the processes of memory consideration, pitch adjustment and random selection. The performance of this new harmony is then compared to the worst performing harmony stored in the HM to determine whether the latter needs to be replaced by the former. Interestingly, the original HS was also verified by the authors in the WDN optimization problem using the Hanoi network [67].

To avoid the tedious parameter settings for HMCR and PAR in the HS, Geem and Cho further proposed a parameter-setting-free method and used it to optimally design WDNs [68]. In this method, an extra OTM (operation type memory) was defined to record the types of operations (i.e., random selection with a rate of 1-HMCR, memory consideration with a rate of HMCR×(1-PAR), pitch adjustment with a rate of HMCR×PAR) used to produce each decision variable in the HM and was continuously updated during the HS evolving process. The early stage of the algorithm performed the same as in the conventional HS by using some central values for HMCR and PAR (where both parameters were set as 0.5). The following decision vectors in the remaining iterations were generated according to the dynamic values of HMCR and PAR which were computed from the frequency of appearance of their correspondingly incurred types of operations in the OTM obtained at the beginning of each iteration. Based on this mechanism, each decision variable has its own values of HMCR and PAR to be used at each iteration. A noise inserting scheme was also devised by adding noise into the obtained values of HMCR and PAR, in order to alleviate the potential problem where the optimal settings could be very close to one or zero. Favorable results using the method were shown on two benchmark WDNs, i.e., two-loop and Hanoi networks.

There are also a few population-based metaheuristics that have recently appeared in the domain. The shuffled frog leaping algorithm (SFLA), being a representative of the memetic algorithms (MAs) based on memetic evolution, was proposed according to the cooperative evolution memes of frogs [69]. The SFLA was linked to EPANET and its toolkit to develop the so-called SFLANET in [70] for general WDN optimization problems. The authors claim that SFLANET had found the optimal solutions faster than GA and SA on the two-loop, Hanoi and New York networks. Baños et al. [71] also presented a new memetic algorithm and compared its performance on the two-loop, Hanoi and Balerma irrigation networks with several metaheuristic and deterministic optimization methods. Moreover, the scatter search (SS) (aiming at maintaining diverse and high quality solutions) [72], the immune algorithm (IA) (motivated by immunology in protecting the host organism from invaders) [73] and the honey-bee mating method (motivated by the biological behavior of honey bees) [74] were also applied to WDN optimization. A comparative study between ACO, IA and SS based on the NYCT network indicated that the ACO always found the global optimum in 20 runs [75]. Furthermore, the cross entropy (CE) method originated from rare event simulation [76] where the generation of random sample vectors and the update of some random mechanism take inputs from and iterate between each other until convergence is reached. This method was also applied in the optimization of WDNs [77] and with uncertain nodal demands [78].

D. Decomposition Optimization Approaches

Since the size of the real distribution networks is usually substantially larger with hundreds to thousands of nodes and pipelines, decomposition approaches were therefore proposed to partition large networks into smaller sub-networks. The results from the decomposition can help monitor, manage, and understand the various components within the network and their interactions [79]. District metering areas (DMAs) can be obtained by enabling pipes to remain open with the flows metered or be closed off using isolation valves [24], thus allowing for the convenience of management and operation of WDNs based on DMAs. One of the well-known primary thrusts of using DMAs is driven by leakage management concerns, where it is easier to find burst pipes and repair them based on manageable smaller DMAs. Leakage teams can take the inflow and outflow meter readings at night when the consumer demands are at their minimum level, thus more accurately estimating the leakage locations and severities.

From a network optimization perspective, the decomposed small sub-networks can also be more readily handled. It is worth noting that the deduction of the number of commercially available types of pipes for each pipeline could also help reduce the solution space. In [80], the original looped network was first converted into a distribution tree such that each demand node has only one path (determined according to the shortest route) connecting the node to a source. Considering the minimum pressure head required by the end nodes, the heads of intermediate demand nodes were then determined by iteratively constructing the so-called critical path according to the least average friction slope in the obtained distribution tree. The flows in the pipes not included in the paths can therefore be approximated using (9) with the minimum pipe diameters as the initial diameters for these pipes, and the flows in the pipes included in the paths can then be obtained according to (5). As a result, the initial diameters for the path pipes can be calculated by again using (9). A GA method with a self-organizing penalty was then employed to optimize the original network according to the reduced sizes of discrete diameters for each pipe, which were clipped according to those obtained initial diameters.

In [81], the shortest-distance tree consisting of a similar path as above for every node in the looped network was found by Dijkstra's algorithm and the diameters of the pipes within the tree were optimized by NLP. The approximately optimal solution was thus given by the NLP solution for the pipes involved in the shortest-distance tree, together with the minimum allowable pipe diameters for the pipes involved in the chords. The differential evolution (DE) was then used to optimize the original network by using an initial population generated based on the approximated solution in order to accelerate the learning. Other types of tree decomposition related multi-stage optimization methods, such as forest and core, have also been well researched [82], [83]. Apart from partitioning a WDN into trees, partitioning into subnetworks was also pursued in [84]. The obtained subnetworks were then optimized separately using DE. The whole network was finally determined by again employing DE given the initial solution gathered from each of the optimized subnetworks and the optimal source partitioning cut-set. Another subnetwork decomposition based optimization approach with the aid of DE was also recently reported in [85].

E. Multi-Objective Optimization Approaches

It is worth noting that the reckless pursuit of lower capital cost by reducing the sizes of the network components can result in low network reliability in terms of failing to provide required standard of services when uncertainties appear. Apart from aggregating all goals and constraints into a single cost function for single-objective optimization, an alternative is to use multi-objective optimization approaches, where the goals and constraints for the network of interest can all or partly be considered as the objectives to be optimized. The trade-off between different conflicting objectives (where this is usually true) can be automatically determined and thus be more easily handled by the network designers. According to the roles of the decision maker (DM) in the solutions searching and determination processes, the multi-objective optimization can generally be divided into *no-preference* methods (without DM's preferences articulation, e.g., the global criterion method), *a posteriori* methods (with DM's preferences articulated after optimization, e.g., mathematical programming and multi-objective evolutionary algorithms (MOEAs)), *a priori* methods (with DM's preferences articulated before optimization, e.g., the weighted sum method) and *interactive* methods (with DM's preferences articulated progressively during optimization, e.g., step method) [86], [87], though a method from one class with some modifications may turn into another class. For the *no-preference* methods, usually only a single solution is generated without requiring preferences information, leaving the DM no options but to accept it no matter if it is satisfied or not. As the DM's preferences have to be given before optimization, the *a priori* methods have the problems that the DM may not be entirely sure about the impact of his/her preferences on the final solution and thus may miss useful solutions. In contrast, *a posteriori* methods are able to first produce a set of solutions from which the DM can then choose one that meets his/her requirements, although the computational demand is usually quite high. Regarding the *interactive* methods, a DM would have the chance to be involved in the solutions searching process and progressively provide his/her preferences while gaining more information from the system through observing intermediate solutions. However, there is still an argument that the final solutions can be highly dependent on a particular DM and the involvement of different DMs may result in significantly different results [88].

The most important aim of performing multi-objective optimization is to find the Pareto optimal solutions, where each Pareto optimal solution can define a relative importance/preference of different objectives, such that no other solutions can be found to decrease any objective(s) defined by this solution without increasing the other objectives, in the case of minimization problems. While, in practice, algorithms able to produce weakly Pareto optimal solutions in which no other solutions exist to decrease all the objectives at the same time, are also acceptable. Since the rapid growth of computing power, multi-objective evolutionary algorithms (MOEAs) [89], [90] are now amongst the most popular multi-objective optimization techniques due to a number of merits such as the ability to produce several Pareto optimal solutions in a single run normally without the need of preference information in order to approximate the entire Pareto front and being robust to the shape (e.g., concave) or continuity (e.g., discontinuous) of the Pareto front [91]. Among various MOEAs, the vector evaluated genetic algorithm (VEGA) proposed by Schaffer was the first attempt to cope with multi-objective optimization problems [92]. This is a population-based multi-objective optimization approach implemented by using a modified sub-population-based selection mechanism in GAs; however, the main drawback is widely considered to be that the (potential) Pareto optimal solutions could be destroyed during the solution selection procedure. The real sense of Pareto-based multi-objective optimization approaches for MOEAs was thereafter developed to incorporate the idea of nondominated ranking and selection suggested by Goldberg [93]. These can generally be categorized as non-elitist methods (e.g., multi-objective genetic algorithm (MOGA) [94], niched-Pareto genetic algorithm (NPGA) [95] and non-dominated sorting genetic algorithm (NSGA) [96]) and elitist methods (e.g., strength Pareto evolutionary algorithm (SPEA) [97], SPEA2 [98], NSGA-II [99], Pareto envelope-based selection algorithm (PESA) [100], PESA-II [101], Pareto archived evolution strategy (PAES) [102] and NPGA2 [103]).

The NSGA assigns the fitness values of solutions by means of layers, each being associated with a dummy value of fitness which decreases over layers. However, the iterative manner of obtaining layers of solutions to acquire the nondominated ranking is time-consuming. More efficiently, the NPGA uses the idea of tournament selection to save computational time, where a mating pool is successively constructed by each time only comparing against two random solutions using the Pareto dominance concept and fitness sharing, based on a randomly selected set of solutions from the population. Regarding MOGA, the fitnesses are calculated according to the solutions' rank in the population which equals the number of solutions by which they are dominated. Although the performance of an algorithm varies with applications, the three non-elitist methods are generally ranked in the following sequence: MOGA, NPGA and NSGA [91]. It has been found that in the non-elitist methods, the determination of the ranking of Pareto dominance is only based on the underlying generation of population. In contrast, in the elitist methods, the situation is extended to a wide temporal scope normally by combing another set of population containing the nondominated solutions that have been visited by the algorithm, thus leading to the succeeder or the new generation of MOEAs. For example, the SPEA separately computes a strength value for each solution in this extra population and a fitness for each solution in the main population before carrying out the selection procedure based on both populations. Readers having particular interest regarding this aspect are referred to [91], [104].

Within WDN optimization, NSGA was applied to simultaneously minimize the network capital cost and to maximize the network reliability in [105]. NSGA-II [99] is an improved version of NSGA and has been popularly applied to WDN optimization [106], [107] with low computational complexity and ability to find a good set of diverse solutions. The minimum nodal head across the whole network and the capital cost were considered as the two objectives in [108] and the optimization was undertaken by the proposal of a computationally efficient decomposition and dual-stage multi-objective optimization (DDMO) method. In DDMO, the original network was first broken down into a number of small sub-networks by using a graph decomposition algorithm. The optimal front for each of the sub-networks was then determined separately and quickly using NSGA-II, followed by the recombination of these fronts based on their hydraulic compatibility and the final generation of Pareto front for the original network using another NSGA-II. In [109], the construction phases rather than the static design were considered, where the network expanded and nodal requirements varied with time. The decision variables are the diameters of the pipes to be installed in the new areas and the pipes to be laid in parallel to the existing pipes for every construction phase. A modified NSGA-II to allow integer number encoding was used as the multi-objective optimizer with the minimization of network capital cost and the maximization of pressure head surplus over the whole construction period as the two conflicting objectives.

A comparative study between the non-elitist MOGA and elitist SPEA for capital cost minimization and pressure deficit minimization was carried out for the network rehabilitation problem, to demonstrate the superiority of SPEA in terms of Pareto fronts and computational time [110]. An improved version SPEA2 [98] generally having better performance than SPEA, PESA and NSGA-II was also applied in WDN optimization in [111]. Furthermore, Kapelan et al. [33] proposed an RNSGA-II (Robust NSGA-II) with less computational demand for the minimization of capital cost and maximization of network robustness regarding uncertainties in water consumption and pipe roughness coefficients. Recently, some new population-based approaches have been devised in the WDN domain for multi-objective optimization, such as the multi-objective PSO for the minimization of pipe cost and nodal pressure deficit and/or the maximization of network reliability [112]. The CE algorithm was also extended for multi-objective optimization with some features derived from MOGA, where its performance on WDN optimization was compared to NSGA-II [113].

There are also some researchers focusing on optimizing more conflicting objectives including those arising from network designs and operations. Farmani et al. [8], [36] tried to solve the optimization and operation of WDNs together by using the Anytown's benchmark as an example, devoting their approach to the minimization of capital costs from pipes and tanks, and energy cost during a specified operational period. Multi-objective optimization method NSGA-II was adopted where the minimization of the total cost served as one of the objectives, while other objectives such as resilience index (representing reliability of the network), minimum surplus head and residence time (representing water quality) were gradually considered therein. Kurek and Ostfeld [34] utilised SPEA2 with the aid of EPANET to develop a multi-objective model for the optimization of water quality (disinfectant residuals concentrations and water age), pumping cost and tank sizing in WDNs. The continuity of flow and pressure, the tank water levels and the storage-reliability were thereby considered as the constraints embedded in the optimization process. In [114], the authors discussed the differences between the reliability index and robustness index for system design, where they claimed that although the system robustness can be enhanced by increased system reliability, robustness should also include the variation of system performance. A measure of system robustness was then defined to reflect the variation of nodal pressure under system uncertainties. The NSGA-II was then used to minimize the capital and operational costs and maximize the system robustness. A total of six objectives with interests from different stakeholders, including capital and operational costs, hydraulic failure (due to low nodal pressure and/or tank water level), leakage, water age and fire-fighting capacity had also been investigated in [115] for Anytown network by using epsilon NSGA-II (ϵ-NSGA-II) with the decision variables taken from pipe and tank sizing, tank siting (regarding the locations and elevations of tanks in a network), and pump scheduling. The trade-offs between these conflicting objectives were analyzed by interactive visual analytics.

F. Hybrid Optimization Approaches

Since the downside of metaheuristic optimization methods mainly lies in their low convergence rate, there is a tendency to incorporate deterministic methods into metaheuristic optimization in order to reduce the search space imposed on metaheuristics and thus to accelerate learning. GAs were combined with ILP [116] to remedy the path dependent optimality issue in ILP [2], where the ignored pipes from the selected paths for iterative ILP optimization were determined by GAs. The ILP was only used to optimize the pipes involved in the paths, under which the types of ignored pipes were fixed through GA designation. It can be found that the iterative running of the hybridization of hydraulic simulator and ILP (for every individual involved in a population for a number of generations in GAs), still exhibited high computational demands.

Interestingly, it can be more effective and efficient to combine the ILP method with metaheuristics using decomposition techniques. For example, in [83], the authors decomposed the original network into trees and core. Then, the pipes involved in the trees were optimally determined by the ILP (where a series of optimal solutions were obtained for each tree based on different root nodal heads) since the flow rates in the trees can be linearly solved. The pipes involved in the core were optimized by DE with the aid of the sets of optimal solutions for the trees. A penalty cost was used to measure the infeasible solutions and was added to the pipe cost. Compared to [116], there was no need to iteratively run ILP to obtain one solution for part of the network, and in every generation of the heuristic learning process only a table lookup was executed for each individual instead of performing iterative ILP. In addition, the initial solutions of the metaheuristic methods could also be improved by employing more efficient algorithms rather than having them randomly generated [117]. The integration of such algorithms can usually reduce the number of generations required by metaheuristics.

To reduce computational burden, besides concentrating on the optimization algorithms, the saving of the hydraulic simulation time for the network of interest can also be looked at. For example, a combinatorial model of DE-ANN was designed in [118], where ANNs (artificial neural networks) rather than the hydraulic simulators were used to capture the hydraulic and water quality behaviors while using DE for optimizing WDNs. The objective function considered both the pipe cost and the chlorine cost while the decision variables were taken as pipe sizes and chlorine dosage rates at water treatment plants. It is noted that the determination of critical nodes amongst the networks was performed to reduce the ANN training time and a local search heuristic was additionally devised to polish the solutions obtained from DE-ANN. The integration of such modeling techniques then also raises the question as to how accurate are the replicated hydraulic models and how to (dynamically) get the artificial models trained, which further brings up research topics from the field of system identification.

Before finishing this section, it is worth noting that making an exact comparison between different optimization methodologies currently reported is somewhat improper and unfair, as different factors were involved in conducting the experiments/simulations, such as different constraint values (or sometimes the reported solutions may just simply violate some constraints), different hydraulic models, different computation software and platforms, different benchmark networks, different calibrated optimization methods, etc. All of these factors would undoubtedly affect the solution optimality and/or the computational time spent to achieve the final solution(s). One of the extreme cases is that if the computational time is no longer a concern, a global optimal solution could always be found by some algorithm such as enumeration (where the entire solution space for the objective functions and the associated constraints of a nonconvex problem is explored). This is also evident from a recent competition in the domain: the battle of the water networks II (BWN-II) held in Adelaide 2012 [119]. A total of fourteen participants/research groups from the domain presented their approaches in the BWN-II competition specifically to design a D-Town network consisting of five DMAs that need to be upgraded and one additional new zone that needs to be constructed [119]. It is noted that even though excellent and dedicated work has been done on the same benchmark, the best approach in terms of both finding optimal solutions and consuming less computational times is still hard to determine. In more detail, heuristic (e.g., engineering knowledge), metaheuristic (e.g., GAs), single objective or multi-objective (e.g., NSGA-II) approaches were applied with the decision variables taken from design variables and/or operational variables, where two types of scenarios were considered, i.e., normal loading and emergency scenarios. The total annualized cost (including capital and operational costs), the estimated greenhouse gas emissions (as a result of the energy incurred by operation of pumps and by manufacturing, transportation and installation of new pipes) and the water age (as a water quality indicator) were generally required as the system performance criteria although only part of them were finally adopted as objective(s) to be optimized by some approaches. The detailed analysis of the various solutions obtained from the competition can be found in the summarization paper [119]. It is also worth pointing out that although similar optimization algorithms were used by different participants (such as the GAs used by Matos et al. and Kandiah et al., or the NSGA-II by Stokes et al. and Wang et al.), the obtained results can be significantly distinct from each other, partially due to the diverse engineering heuristics involved and/or the algorithms' settings. Since there is always a balance between the algorithms' computational demand and solution optimality which can be reflected as the trade-off between the exploitation/intensification and exploration/diversification abilities, it is unfortunate that the computational times of different approaches were not reported for comparison.

IV. Wastewater Collection Network Optimization

Wastewater collection networks (WWCNs) (gravity sewer systems specifically in this paper) should be properly designed to drive flows of wastewater towards WWTPs without causing surcharging or pressurizing issues. WWCNs differentiate from WDNs mainly in aspects of gravity driven flows and tree-like network structures. To facilitate the optimization of WWCNs, design peak flows, ground elevations, manhole locations, and, usually, the system layout and flow directions (if they are not part of the optimization), are known a priori. Research has thus been devoted to the network optimization such that the sewer types, excavation depths, manhole depths and the existence of lift pumping stations are determined to be able to convey peak flows at minimum capital cost while also satisfying all the hydraulic and physical constraints. The peak flow of a sewer is practically assumed by a peaking factor of 2.5-3.5 times larger than the average daily flow (which can be estimated according to the water consumption in a service area) plus the infiltration and inflow (I/I) allowance [5].

A. Cost Function and Constraints

Differing from potable water in WDNs, wastewater in the ith sewer is normally flowing partially full and the resultant central angle ϕ_i (radian) to the water surface can simply be computed as

$$\phi_i = 2\cos^{-1}(1 - 2h_i/D_i), \quad i = 1, \dots, N_s, \tag{10}$$

according to the cross sectional area of a circular sewer [6] shown in Fig. 4a, where h_i denotes the flow depth (m) in the ith sewer, D_i is the diameter (m) of the ith sewer, and N_s is the total number of sewers in WWCNs. For accessibility of the paper, the notations of variables with similar physical meanings involved in WDNs and WWCNs are no longer differentiated, such as the pipe/sewer diameters D_i. Then, the associated wetted perimeter P_i (m), flow area A_i (m^2) and hydraulic radius R_i (m) [4] can be easily calculated as

$$P_i = \frac{D_i\phi_i}{2}, \quad i = 1, \dots, N_s, \tag{11}$$

$$A_i = \frac{(D_i)^2}{8}(\phi_i - \sin\phi_i), \quad i = 1, \dots, N_s, \tag{12}$$

$$R_i = \frac{D_i}{4}\left(1 - \frac{\sin\phi_i}{\phi_i}\right), \quad i = 1, \dots, N_s. \tag{13}$$

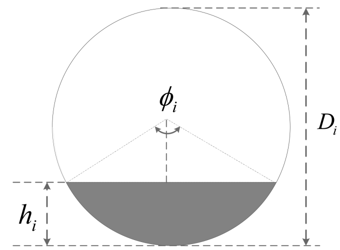
Fig. 4b presents a typical geometry profile of a WWCN [6], [120]. In comparison with the optimization of WDNs, apart from the sewer types/sizes (related to the sewer expenses), the optimization of WWCNs has also to consider sewer slopes (excavation depths, related to the excavation expenses) as decision variables in order to fulfill wastewater hydraulic restrictions. The basic objective function for the optimization of WWCNs together with the associated constraints [4], [6], [120] can be formulated as

$$J(\boldsymbol{\theta}) = \sum_{i=1}^{N_s} \mu_i(D_i, UD_i, DD_i, L_i) \tag{14}$$

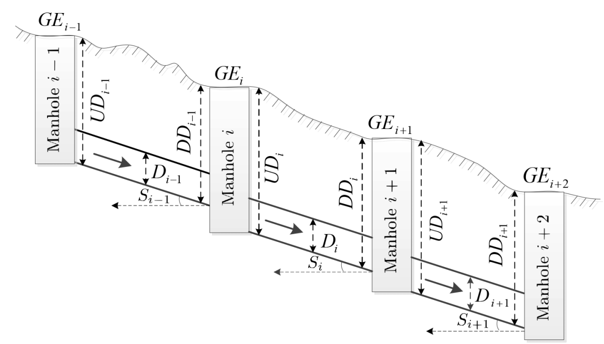
$$\text{s.t.} \begin{cases} D_i \in \{\zeta_1, \ldots, \zeta_{N_t}\},\ i = 1, \ldots, N_s & (15) \\ \begin{cases} Q_i = f_i(D_i, \phi_i, S_i),\ i = 1, \ldots, N_s, & (16) \\ V_i = Q_i/A_i,\ i = 1, \ldots, N_s, & (17) \end{cases} \\ V_{\text{min}} \le V_i \le V_{\text{max}},\ i = 1, \ldots, N_s, & (18) \\ \begin{cases} UD_i \le AD_{\text{max}},\ i = 1, \ldots, N_s, & (19) \\ DD_i \le AD_{\text{max}},\ i = 1, \ldots, N_s, & (20) \end{cases} \\ \begin{cases} UD_i - D_i \ge AC_{\text{min}},\ i = 1, \ldots, N_s, & (21) \\ DD_i - D_i \ge AC_{\text{min}},\ i = 1, \ldots, N_s, & (22) \end{cases} \\ D_i \ge D_{i-1},\ i = 1, \ldots, N_s, & (23) \\ \phi_i \le \phi_{\text{max}},\ i = 1, \ldots, N_s, & (24) \\ \begin{cases} GE_i - UD_i \le GE_i - DD_{i-1}, \\ \qquad i = 1, \ldots, N_s, & (25) \\ GE_i - UD_i + D_i \le GE_i - DD_{i-1} + D_{i-1}, \\ \qquad i = 1, \ldots, N_s, & (26) \end{cases} \\ (GE_i - UD_i - (GE_{i+1} - DD_i))/L_i \ge S_{\text{min}}, \\ \qquad i = 1, \ldots, N_s, & (27) \end{cases}$$

where $\mu_i(\cdot)$ and L_i denotes the construction cost (£) and length (m) of sewer i, UD_i and DD_i are the upstream and downstream excavation depths (m) of sewer i, N_s is the total number of sewers in a WWCN, $\{\zeta_1, \ldots, \zeta_{N_t}\}$ is a series of commercially available sewer diameters (m), Q_i and $f_i(\cdot)$ represent the flow rate (m^3/s) and the nonlinear hydraulic behaviors in the ith sewer, S_i is the slope of the ith sewer, V_i is the flow velocity (m/s) in the ith sewer, V_{min} and V_{max} are the minimum and maximum allowable velocity, AD_{max} is the maximum allowable excavation depth, AC_{min} is the minimum allowable cover, ϕ_{max} is the maximum allowable central angle, GE_i is the ground elevation (m) at manhole i, and S_{min} is the minimum allowable slope. It should be noted that the values of these hydraulic process variables (Q_i, V_i, ϕ_i and A_i) incurred here should be considered in the case of design peak flows as required by the nature of WWCN problems.

As shown in (14), the sewer construction cost including the sewer cost and the excavation cost is generally a function of sewer diameter, excavation depths (or average excavation depth) and sewer length. It is difficult to accurately compute this sewer construction cost as it also depends on specific conditions (e.g., geological conditions and construction methods) and different measures [121]–[124] including cost models and look-up tables have been used for approximating it. On this aspect, the multiple regression analysis [125] is often used to estimate the unknown associated parameters in the approximated cost model. It should be noted that the slope S_i of the ith sewer can be computed from the excavation depths (UD_i and DD_i) by using the elevations (GE_i and GE_{i+1}) and sewer length L_i. Other capital expenditure to construct a WWCN could also include those incurred from manhole construction and pumping station construction. The manhole cost could be a function of the manhole depth (the deepest excavation depth of the sewers connecting to the manhole) [121] and also the maximum sewer diameter connecting to the manhole [6]. The construction cost of pumping stations can be related to the design flow rate [18]. Some studies have also annualized the capital expenditure based on the rate of interest and, the life, salvage factor and maintenance cost of various physical components [4].



(a) Cross sectional area of the ith sewer



(b) Geometry profile of a sewer system

Fig. 4. Illustration of wastewater collection networks.

A number of hydraulic and physical, equality and inequality constraints have been adopted in the domain [4], [6], [120] which are listed in (15)-(27). In (15), the type (material and diameter) of every sewer i must be chosen from a set of discrete commercially available types. In contrast to the potable water behaviors described in Subsection III-A, the fundamental hydraulic behaviors of the wastewater flowing through the ith sewer can be deemed as a nonlinear equality constraint in (16), while (17) states the flow continuity. A highly nonlinear relationship can be observed between flow rate Q_i, central angle ϕ_i, diameter D_i and slope S_i in every sewer i. The velocity constraints (18) (e.g., V_{min} = 0.6 m/s and V_{max} = 3.0 m/s) at peak flows are the primary requirements of a WWCN, which are employed to enable sewers self-cleansing (to prevent clogging) and also to reduce sewers scouring (to extend sewer lifetime). Due to the underground structures, as stated in (19) and (20), the excavation depth for both ends of the ith sewer is usually confined to a certain level (e.g., AD_{max} = 8.0 m) from economical and practical concerns. Constraints (21) and (22) are adopted to guarantee that both ends of sewer i are buried at enough depth (e.g., AC_{min} = 1.0 m) in order to prevent any damages from ground surface loading and to facilitate collecting wastewater from properties/drains.

To ensure that the downstream sewers are not surcharging, sometimes the diameter progression constraint (23) is also placed on the successive sewers, as the volume of downstream wastewater generally increases. In addition, the wastewater depth-to-diameter ratio at peak flow in the ith sewer is sometimes constrained as in (24). Since wastewater is generally conveyed by gravity via the sewer system, the usage of pumping stations should be kept at minimum. Therefore, in (25) and (26), it states that the upstream invert (crown) elevation of sewer i should be less than or equal to the downstream invert (crown) elevation of its upstream sewer $i-1$, in the absence of a pumping station at the ith manhole. Finally, constraint (27) describes the minimum allowable slope for sewer i, where sewers with very flat slopes (e.g., flatter than $S_{\min}$ = 0.08%) are not suitable for laying. It should be mentioned that within a specific research into the optimization of WWCNs, only part of these described constraints may be considered.

The most common expression for describing the wastewater hydraulic behaviors in WWCN optimization problems is to use Manning's equation [122],

$$V_i = \frac{1}{C_i^{\mathrm{MN)}}}(R_i)^{\frac{2}{3}} S_i^{\frac{1}{2}}, \ \ i = 1, \ldots, N_s, \tag{28}$$

where $C_i^{\mathrm{MN)}}$ is the Manning's roughness coefficient (typical value for concrete sewers is 0.013, and 0.010 for Polyvinyl chloride (PVC) sewers). To be more accurate, other equations such as the modified Hazen-Williams equation and the Darcy-Weisbach equation can also be employed in order to encompass the impact of flow velocity, pipe diameter and fluid viscosity on the roughness coefficient [4], [6]. Obeying the flow continuity, the wastewater flow Q_i can be obtained [122] using $Q_i = V_i A_i$, for $i = 1, \ldots, N_s$, such that

$$Q_i = \frac{1}{20C_i^{\mathrm{MN)}}}(\phi_i)^{-\frac{2}{3}}(\phi_i - \sin\phi_i)^{\frac{5}{3}}(D_i)^{\frac{8}{3}} S_i^{\frac{1}{2}}. \tag{29}$$

It is worth mentioning that (28) and (29) (originally proposed for flows in open channels) are used to describe the wastewater behaviors under partially filled flows as defined in (16) and (17), while in the design of WWCNs (with the assumption of partially filled peak flows), pressurized flows or surcharging issues should be avoided by using the constraint defined in (24), i.e., the central angle ϕ_i or depth-to-diameter ratio h_i/D_i at peak flow is upper bounded.

B. Traditional Deterministic Optimization Methods

Amongst the traditional deterministic methods, dynamic programming (DP) techniques (especially the discrete differential dynamic programming (DDDP) approach) had dominated the optimization of WWCNs for many years [4], [18], [29], [121]. These algorithms cope with the problem by first decomposing it into multiple stages corresponding to sewers/manholes with pertinent design variables (e.g., sewer downstream/upstream invert/crown elevations) as the states (discrete alternatives) and then dealing with the problem stage by stage based on recursive equations. The minimum required sewer diameters can be computed based on the wastewater hydraulic equations given the states (determining slopes), assumed central angle and design flows. At every stage, DP only considers and records the least-cost solution associated with each state within the stage instead of visiting all possible solutions up to the stage as in an enumeration technique, thus saving considerable computational time. The entire programme may need to rerun several times using more closely spaced states based on the previously obtained results in order to improve performance. However, the downsides of DP are widely regarded as the local performance caused by noncontinuous states, limitations related to large size networks and the assumption of a central angle in computing the sewer sizes together with the resultant continuous sewer sizes. Other than DP techniques, the indirect use of linear programming (LP) by piecewise linearizing the nonlinear objective function and constraints [123], as well as the direct use of nonlinear programming (NLP) [126], [127] has also been tried in the optimization of WWCNs. Nevertheless, the local optima still occur in both methods and the obtained diameters are required to round to the commercially available ones which may cause further reduction of performance.

C. Traditional Heuristic Optimization Methods

Heuristic optimization methods were also amongst the earliest attempts at optimizing the design of WWCNs. To satisfy the minimum velocity constraint defined in (18) at peak flows, it can be realized by selecting appropriate sewer diameters and slopes based on estimated flows or depth-to-diameter ratios (see (13), (28) and (29)). According to (28), heuristically, one can adopt a depth-to-diameter ratio of a half for the peak flows to find the minimum sewer slopes based on every commercially available pipe diameter [122]. It was found thereby that the required minimum sewer slopes generally decreased when using higher sewer diameters. The actual flow depths and velocities in a sewer would then depend on the real-time quantity of flows, the diameter, the slope and the roughness coefficient. Therefore, the disadvantage noted by the authors is that if the actual peak flow depth is lower than half the diameter, the corresponding velocity would be smaller than the required minimum value.

Desher and Davis [122] proposed the sanitary sewer design (SSD) method to perform the least costly design by considering the sewer construction cost. Under the assumption of the sewer diameters already suggested by experienced engineers, the minimum sewer slopes were determined according to the ground slopes and the required minimum velocities (during peak flows), based on a half depth-to-diameter ratio ($\phi_i = \pi$) by using the Manning's equation. The resulted flow depth and velocity in every sewer from the determined sewer slope, the sewer diameter and an estimated sewer flow were assessed to avoid breaking the minimum and maximum velocity constraints and surcharging issues by further adjusting the sewer slopes and diameters. Interestingly, the relaxation of the sewer size progression constraint with careful consideration of peak flows was found to possibly lead to substantial cost savings.

Charalambous and Elimam [128] also presented a heuristic design approach using either the Manning or the modified

Hazen-Williams equation with the convenience of introducing lift pumping stations, although its downsides lie in non-optimal solutions and continuous diameters (where an extra standardization step was needed). Moreover, the spreadsheet method [120] in which hydraulic calculations can be conveniently tabulated, was developed to tentatively evaluate the effect of different sewer sizes and slopes on the system cost, given designed flow while also satisfying minimum cover and minimum and maximum velocities constraints. Compared to DP, it had been demonstrated that better solutions in terms of constraints handling and cost saving can be found using the same network and system cost model. The design procedure is obviously transparent but tedious to engineers although still with restricted performance in terms of optimality due to a limited number of trial solutions being examined.

D. Advanced Metaheuristic Optimization Methods

Due to the discrete-continuous characteristics of the WWCN optimization problem together with the high nonlinearity inherent within it, metaheuristics have been naturally employed to provide better performance. Similar to their usage in the WDN optimization, the adoption of metaheuristics for optimizing WWCNs includes those from nonpopulation-based and population-based categories.

1) Nonpopulation-based Metaheuristics: As in WDN optimization, the simulated annealing (SA), tabu search (TS), and cellular automata (CA) have also been applied in the framework of WWCN optimization. Yeh et al. [129] applied SA to design a local WWCN in central Taiwan with significantly varied elevations, where the sewer construction cost was considered as the objective function. The average slope of sewers from SA was found to be larger than the original official design. Also, compared to the original design, all the sewers designed by SA were able to satisfy the minimum velocity constraint though the resulted construction cost was a little higher. Moreover, Karovic and Mays [130] recently applied SA in Microsoft Excel for the sewer system design to make it more convenient from the engineers' perspective. In [131], the authors proposed an integrated approach to combine the determination of network layout and network components since the two problems are naturally related to each other. Overall, TS was used as the optimizer with the associated parameters from both problems as the decision variables, while the network construction cost was taken as the objective and the constraints were systematically satisfied to improve algorithm's efficiency. A comparison between SA and TS was also studied in [132] for WWCN optimization, in which SA outperformed TS in terms of both efficiency and robustness in a number of runs.

Moreover, CA has also been actively employed in WWCN optimization [133], [134]. In these, the sewer junction nodes were considered as the cells with the corresponding elevations regarded as cell states (decision variables) in [134]. The local transition rule for each cell was then mathematically derived (gradient-based local optimization) to minimize the local construction cost formed in its neighborhood (which was defined by the sewers connecting to it). Sewer diameters were determined heuristically based on the slopes given by CA and the maximum flow depth constraint. Significant computational efficiency had been demonstrated in the Mays and Wenzel benchmark, although inferior quality of solutions was obtained compared to the ACO methods [135], [136] to be discussed later. Furthermore, an iterative two-stage method was also proposed in [137], where, in each stage, CA was used to search for either the cover depths or the diameters respectively whilst keeping the other fixed. Transition rules could be obtained mathematically for updating both cover depths and diameters, while ad hoc engineering rules were additionally designed for updating the diameters considering their discrete nature.

2) Population-based Metaheuristics: With respect to population-based methods, genetic algorithms (GAs) [124], [138] are again popularly employed for the optimization of WWCNs. Both sewer diameters and slopes were coded as binary strings in [138] for the application of GAs in WWCN optimization. The sewer construction cost together with the penalty cost regarding the diameter progression constraint was used as the fitness function during the optimization. The impact of different population sizes, crossover rates and mutation rates on algorithm convergence was also demonstrated. In [124], the authors used a GA to define the diameter for each sewer with binary coding at every generation, while the sewer slope was computed from the design flow and an initial assigned velocity by satisfying the maximum flow depth and minimum slope constraints. The resultant cost from the solution of all the sewers can thus guide the search direction of GA. The elitist adaptive genetic algorithm (EAGA) developed by integrating the elitist genetic algorithm (EGA) with the adaptive genetic algorithm (AGA) was employed in this study and its improved performance was compared to the EGA and AGA.

Some researchers also proposed the use of GAs in combination with a wastewater hydraulic simulator (such as TRANSPORT module of SWMM, now at version 5.0 [30]) to minimize the capital cost of a WWCN. Two optimization schemes were devised in [139], i.e., GA-TRANS1 and GA-TRANS2, the first using GA to find both diameters and elevations of sewers and using TRANSPORT to perform hydraulic analysis, while the second using GA to find only the sewer elevations with the diameters and hydraulics analyzed by TRANSPORT (employing trial-and-error to determine the diameters). The solutions that conflict with various constraints considered for the problem were penalized in the objective function during the optimization. Simulated results have been produced on the Mays and Wenzel benchmark to demonstrate cost savings in comparison with traditional methods, such as DP and spreadsheet methods. Due to the large number of constraints involved in the optimization process which may often cause infeasible solutions of GA, Haghighi and Bakhshipour proposed an adaptive sequential constraint handling strategy being used in the decoding stage of GA [140]. The decoded solutions can thus be maintained in the feasible regions of the solution space without requiring a penalty function, resulting in better optimization performance in terms of efficiency and effectiveness as compared to DDDP and GA-QP on the Li and Matthew benchmark. The diameter and slope of sewers and the presence/absence of pumping stations at each manhole were all considered as the decision variables, while taking into account the construction costs of sewers, manholes and pumping stations.

Ant colony optimization (ACO) methods originally designed to solve the combinatorial optimization problems with discrete decision variables have also been used to solve WWCN optimization problems. However, due to the continuous variables (the sewer slopes or elevations) involved in WWCN problems, discretization is usually required. Afshar [135] discussed the effect of the discretization size on algorithm convergence and solution quality, where either too large or too small discretization size has adverse effects. An adaptive refinement procedure was thereby designed to progressively define an appropriate set of discrete variables with reduced range for each decision variable, based on the locally optimal solutions found during the optimization process. The author also devised two partially constrained ant colony optimization methods, i.e., PCACOA1 and PCACOA2 [136], and compared their performance with the unconstrained ACO. The sewer connection nodes were used as the decision points with the sewer elevations as the decision variables where discretization of continuous variables into discrete ones is required for ACO. The diameters can be heuristically selected from a series of commercially available ones given the slopes computed from the ACO decision variables and the design flows by satisfying some hydraulic constraints. The PCACOA1 method employed the minimum slope constraints to refine a tabu list for each decision variable as the ACO incrementally constructed its solution components, while PCACOA2 also considered the maximum flow depth-to-diameter ratio when updating this tabu list. The remaining constraints such as the maximum and minimum velocities were penalized in the objective function as usual. The comparison results on the Mays and Wenzel benchmark showed that significant improvement regarding algorithm convergence and sewer construction cost can be achieved compared to the unconstrained ACO.

Moreover, Moeini and Afshar [141] used the sewer diameters as decision variables with slopes computed by assuming a maximum allowable depth-to-diameter ratio, while also considering the layout determination by combining ACO with a tree growing algorithm. Furthermore, particle swarm optimization (PSO) was recently adopted by Izquierdo et al. [142] in WWCN optimization by considering the sewer diameters and slopes as decision variables with special treatment for the discrete diameters. The simulator SWMM was used to perform hydraulic analysis and the constraints were penalized explicitly in the objective function.

E. Multi-Objective Optimization Approaches

Regarding the multi-objective optimization approaches, they have just begun to appear in the domain of WWCN optimization. The hybridization of celluar automata and multi-objective genetic algorithms for sewer network optimization [143]–[145] was devised to unite the strengths and remedy the downsides that come from GA (global searching ability but limited to high computational demand) and CA (low computational demand but confined with local searching ability). The CASiNO (celluar automata for sewers in network optimization) was thereby quickly performed to supply initial solutions for NSGA-II to accelerate multi-objective global optimization, where the minimization of flooding and capital cost were the two objectives and SWMM as a hydraulic simulator was used to evaluate all the generated solutions. However, only the sewer diameters were considered as the variables to be optimized; based on these variables the network cost was computed.

F. Hybrid Optimization Approaches

Generally speaking, metaheuristic approaches face difficulties when the number of decision variables becomes larger as the search space would inevitably get too huge to be visited efficiently. Moreover, the inclusion of substantial constraints in WWCNs can also affect the efficiency of the optimization algorithms in terms of generating feasible solutions. An alternative is to combine metaheuristic methods with traditional deterministic methods as in WDN optimization. Due to the fact that piecewise linearization of the original problem such as in LP could lead to errors, Pan and Kao recently proposed the use of a GA-QP combination to solve the problem [146]. GA was utilized to search for the diameters of sewers and the locations of pumping stations, under which the original problem was then transformed into quadratic forms with the slopes and downstream excavation depths of sewers as decision variables which can be solved by quadratic programming (QP). The computational efficiency of GA was also improved by controlling the generation of feasible solutions using a proper constraints handling strategy.

V. Discussion and Future Research Directions

Reflecting on the detailed and systematic investigation in the previous sections, including WDNs and WWCNs, their associated optimization problems and the development through to the state-of-the-art of optimization methodologies, generally, several important aspects emerge and are worth discussing. Firstly, the scale of the benchmark networks commonly being examined is somewhat limited, requiring the proposal of optimization algorithms that are able to solve large scale water networks. Secondly, more components and/or constraints that exist in real water networks are expected to be included in the network design process. Thirdly, the development of more efficient and effective optimization methodologies is required in order to cope with the high complexity of network optimization problems. Fourthly, providing test benchmarks factoring in different levels of complexities of water networks for pure algorithmic developers would be beneficial. Fifthly, the limitations of various (categorical) optimization methods (e.g., the deterministic and metaheuristic) must be further researched. Finally, more effective many-objective optimization methodologies are needed in the domain to assist the decision making process especially when the DMs have little knowledge about the network. This section will provide a reflective and critical discussion of current trends and potential future research directions in the domain. These are organized into two aspects: *a)* comparison of existing optimization paradigms between WDNs and WWCNs and, *b)* common opportunities for both types of water network optimization. The critical reflections provided in this section are aimed to promote advances in the domain, especially effective integration of the domain with the field of artificial intelligence.

A. Discussion and Research Directions Observed from the Comparison between WDN and WWCN Optimization

Generally speaking, the optimization of WDNs is currently receiving more attention than the optimization of WWCNs. This may be attributed to deviated research interests (as the formulation of WDN optimization is simpler and better established) and the fact that the construction of WDNs to meet citizen's basic survival needs of clean water is somewhat more imperious than WWCNs. As classical methodologies for real-world optimization problems, deterministic optimization methods such as exhaustive search, LP and NLP can be directly or indirectly applied to both WDN and WWCN optimization after some domain specific modifications and derivations. The major difference between WDNs and WWCNs is that the clean water in WDNs is pressurized for distribution to various consumer nodes whereas the collection of wastewater is mainly driven by gravity provided by inclined sewers although lift pumping stations may be involved to transfer wastewater from one sewer to another. This difference leads to different hydraulic behaviors involved in distribution and collection networks, as well as different network structures (looped distribution networks and tree-like collection networks). Due to the tree-like structure of WWCNs, the optimization can be straightforwardly realized stage by stage (where in each stage one or several hydraulic equations are solved) deterministically or heuristically, which explains why DP and heuristic related optimization methods have been widely accepted and used in this area. In contrast, the looped nature of WDNs attracts particular attention to the utilization of decomposition related methodologies which are employed to divide the entire network under consideration into trees or small subnetworks to be easily handled by other optimization methods. From this, additional deterministic and heuristic methods are required, in combination with decomposition techniques, for WDN optimization. At the very least, the results of such combinations will be able to provide good initial solutions as starting points for other methods such as metaheuristics to further examine the problem.

The metaheuristic optimization methods are a relatively recent development in the water network optimization domain, although their adoption in the optimization of WWCNs has been considerably less (where most work is still dominated by DP, LP and heuristic methods). As discussed in Subsection III-C, the hot topics of the nonpopulation-based and population-based metaheuristics applied in distribution network optimization include SA, TS, CA, GA, ACO, PSO, DE, HS, MA, CE and IA approaches, whilst the metaheuristics applied in collection network optimization are relatively new and have currently mainly focused on SA, TS, CA, GA, ACO and PSO approaches, as presented in Subsection IV-D. Some examples of the metaheuristics that had been employed in the optimization of distribution networks are therefore expected to be examined in the domain of WWCN optimization, such as DE, HS, MA and CE approaches. With respect to multi-objective optimization methods that have been used in WDN optimization, popular trials have mainly focused on MOEAs, e.g., VEGA, MOGA, NSGA, SPEA, SPEA2, NSGA-II, PESA, PESA-II, PAES, as described in Subsection III-E; whilst examples of multi-objective optimization of WWCNs have been rare and, as presented in Subsection IV-E, the only real trial to date is the NSGA-II. The adoption of MOGA, SPEA2, PESA-II, PAES, etc., for the optimization of WWCNs can therefore be investigated and compared between one another.

It has also been found that the popularity of using a hydraulic simulator in dealing with the hydraulics involved in WDNs and WWCNs is distinct. On the one hand, the adoption of hydraulic simulators (e.g., EPANET) for simulating the water behaviors in WDNs is pervasive, especially for the most recent advanced optimization methodologies applied in the domain. On the other hand, the adoption of hydraulic simulators (e.g., SWMM) for determining the wastewater behaviors in WWCNs is less common. This is mainly because there are too many hydraulic and physical constraints involved in collection networks and that the direct evaluation of them using SWMM under some given trial solution(s) to all the decision variables would inevitably decrease the optimization efficiency (even making it very difficult to locate feasible solutions for complex networks) as many constraints can be violated during the trial solution generation process. The weighting parameters used in the common penalizing approaches for dealing with constraints are also hard to control. As a result, implicitly satisfying part of these constraints heuristically or mathematically when generating the trial solutions is often pursued in the optimization of WWCNs. Although hydraulic and physical constraints are fewer in WDNs, it will be interesting to develop constraint handling strategies to directly generate feasible solutions rather than penalizing infeasible solutions during the WDN optimization process. This kind of implicit constraint handling may also be related to choosing the appropriate decision variables for water networks (for example, diameters, slopes or a combination of both in collection networks), while the rest of the unknown variables can be determined by mathematics or some heuristic rules considering part of the constraints. This allows the remaining constraints, possibly together with the search space to be significantly reduced, resulting in better optimization efficiency and/or effectiveness.

Furthermore, by comparing the objective functions between the optimization of WDNs and WWCNs, it can be seen that the basic objective function for WDNs is simpler, although it becomes more complicated when other components like pumping stations in the network are optimized. Most studies are therefore based on finding the least-cost design defined in (1) for the distribution networks. Due to the inclined sewers and the network structure, the basic objective function for WWCN optimization is more nonlinear and difficult to estimate, where usually the excavation cost is also included. As shown in (14), the computation of capital cost is not fixed for various research studies carried out in the domain and a variety of measures have been employed to estimate it. This will obviously cause difficulties when comparing the costs obtained from different optimization methods. Therefore, a unified and accurate representation of the capital cost for WWCNs is required. It is also worth mentioning that the full satisfaction of the constraints involved in water networks is needed if comparisons are going to be made between different optimization approaches. Last but not least, according to the multiple loadings that can be used in the optimization of WDNs, the use of multiple design flows rather than single design flows for all the sewers can also be tried in WWCN optimization in order to improve the utility of the constructed networks.

B. Discussion and Common Research Directions Observed for both WDN and WWCN Optimization

Although a few relatively simple benchmarks have appeared in the domain of water network optimization (more for WDNs than WWCNs), more artificial and real-world benchmarks for larger scale systems, featuring the full range of network components suitable for WDNs and WWCNs are desired. A notable phenomenon is that comparisons made on unified benchmarks (same network, hydraulic equations/simulators, constraints, setting values, etc.) by using various optimization methods (running on the same computation platform in order to compare the execution times) are currently quite limited. Although some artificial generators [147], [148] exist to construct virtual water networks in a certain level of complexity, this does not remove the necessity for constructing central repositories of WDNs and WWCNs for the convenience and fairness of comparisons between diverse optimization methodologies. Similar phenomena can also be seen in the field of machine learning where several central repositories exist, for example, the well-known UCI (University of California, Irvine) machine learning repository [149] despite the fact that various artificial sampling data generators exist such as the time-series prediction data generator.

A number of key factors have to be considered in the construction of repositories. Firstly, the repository should capture a significant number and variety of networks appearing within the domain, while providing detailed descriptions for each of these networks. Secondly, the repository should be dynamic and allow developers/researchers to update/upload existing/new benchmark networks. Thirdly, the pertinent hydraulic and physical constraints and the specific objective(s) should be explicitly listed together with the possible setting values of these constraints and/or objective(s), for each benchmark to enable unification. Fourthly, the number and type of hydraulic equations involved are described and the corresponding hydraulic simulators are suggested. Fifthly, results and performance data from various optimization methodologies can be uploaded into the repository thus providing a platform for competition. Depending on the purpose of the repository, different themes may also be included, such as single-objective optimization and multi-objective optimization. Through the construction of this repository, substantial benchmarks for both WDNs and WWCNs are required to be gathered, unified and eventually provided to researchers from different fields working in this multidisciplinary domain. Depending on the complexity (for example, network scale and number and types of components involved) of different benchmarks, each optimization method can then be examined and compared by using a range of benchmarks (from simple to complex). It is also suggested that the same basis (platform, programming languages, etc.) is required for fair comparisons on computational time, etc. In this way, the advantages and disadvantages of each optimization method and their suitability for dealing with simple or complex water networks can be fully analyzed.

To allow fair comparisons, Marchi et al. also suggested a general approach to compare evolutionary algorithms applied in the domain consisting of five steps [150] which is beneficial to enable comparisons to be made between different optimization methodologies within the domain. These steps were: *a)* selecting particular algorithms for comparison, *b)* selecting test benchmarks, *c)* calibrating the selected algorithms, *d)* conducting the simulation/experiment and *e)* analyzing the results obtained. In particular, the calibration of optimization algorithms to find the best algorithm parameters, so as to achieve the best performance on the test benchmarks of interest, was considered essential and is usually a complicated process. Moreover, the tested range of parameter settings were suggested to cover their classical values and combinations, and other values around these classical values. Normally, the same number of executions for each algorithm was performed for statistical analysis purposes (such as computing the average solution, the best solution and the standard deviation across all executions). It is also foreseen that fair comparisons of different optimization algorithms based on large scale and real-world water networks are required in the domain.

In reality, the evaluation of hydraulic behaviors in both WDNs and WWCNs is time-consuming regardless of whether this is done by solving hydraulic equations explicitly or using hydraulic simulators implicitly. In addition, this evaluation is inevitably required to be repeated many times for most advanced optimization methods. Simple numerical models have been emerging to simulate the hydraulic behaviors, however, from the system modeling perspective, more applications of advanced models, such as the fuzzy systems [151] with good model interpretabilities, Gaussian processes [152] with good probabilistic characteristics and support vector machines [153] with good generalization abilities, are expected to be integrated into the network optimization process to speed it up.

From our investigation of optimization methods applied in the domain, there is still scope to apply a large variety of existing optimization methods that have emerged in other fields, to the optimization of water networks, especially metaheuristic, multi-objective and hybrid optimization methods. For the nonpopulation-based metaheuristics category of methods, for example, guided local search (GLS), fast local search (FLS) [154] and greedy randomized adaptive search procedure (GRASP) [155] are applicable and of interest. Regarding population-based metaheuristics, for example, more evolutionary algorithms (e.g., evolution strategies (ES) [156] and estimation of distribution algorithms (EDAs) [157]) and swarm intelligence (e.g., artificial bee colony (ABC) algorithm [158], the bees algorithm [159] and bat algorithm [160]) are expected to be further investigated in the domain. Alternatively, most metaheuristic optimization methods applied in this domain are more or less based on their primitive prototypes, while their extensive modifications and recent revisions, such as those from the general evolutionary computation subject, require further research. On this aspect, it is also worth mentioning that a general review for EAs used in the broad field of water resources systems (such as model calibration, water distribution systems, groundwater management, river-basin planning and management) was recently performed to address some key issues (e.g., optimization problem understanding and formulation) and common challenges being faced in the field (e.g., algorithm performance and actual decision-making process in complex and uncertain contexts) [161].

Moreover, more optimization methods developed by hybridizing the previously mentioned categories of optimization methods (methods between several categories and/or within one category) will certainly be developed in the future to unite the advantages and alleviate the disadvantages of existing methods when applied in water network optimization, while also considering the respective network characteristics for WDNs and WWCNs. It is very important to analyze and consider the specific characteristics of the problem of interest, as overall there is a tradeoff between the generality and individuality for a given optimization method being applied to solve the problem. This tradeoff is considered to have corresponding influence on the resultant performance. The more generically (directly) one applies an optimization method to the problem, the worse it performs. In other words, the performance of an optimization method will undoubtedly be improved by considering the specifics of the problem. This is even true for a generic optimization method such as GAs, where the performance of applying it directly to water network optimization can be worse than also considering the specific network characteristics (an example is to consider part of network constraints for the generation of high-quality trial solutions during the optimization process, rather than penalizing infeasible solutions as discussed before). Furthermore, as the current multi-objective optimization methods applied in the domain are mainly derived from evolutionary algorithms, of which the NSGA-II and SPEA2 are the two well-known representatives. However, others such as those derived from mathematical programming [162] and swarm intelligence (e.g., multi-objective ACO [163]) can also be investigated and/or developed for the multi-objective optimization of water networks.

There is also a challenging and active research area about many-objective optimization problems (where more than three objectives are involved which can be common for the optimization of water networks). The popular Pareto-based approaches have been revealed to have significant shortcomings, such as the deterioration of solution diversities and poor convergence rate due to the reduced Pareto-based selection pressure as a result of excessive non-dominated solutions in each population [164], [165]. Researchers have therefore proposed a number of techniques for tackling these issues, such as incorporating *a priori* or *interactive* ideas to narrow down the Pareto-optimal solutions around the preferred solutions [166]. Recently, Giagkiozis and Fleming [165] also provided theoretical results on using decomposition-based approaches over Pareto-based (where both employ *a posteriori* approaches) for many-objective optimization problems from the probabilistic perspective. The application and/or development of these sorts of new techniques are expected to be further studied within both WDN and WWCN optimization, given that more objectives/constraints are involved in domain.

Although extensive academic research has been carried out in the domain, available software modules, such as the optimal design module within KYPipe [23], for designing water networks by employing advanced optimization techniques are limited. Part of the reason can be attributed to the existing gap between academic interest and industry requirements. The research undertaken is sometimes more focused on relatively small and simplified benchmark networks, while actual water networks are usually large, dynamic/evolving and complex to address. Other factors such as the socio-organizational and political issues (e.g., willingness of using advanced software), stability, simplicity and user friendly interface of the software, compatibility with the existing systems of water utilities and lack of financial support could also affect developing exclusive software in the domain. Due to the rising awareness and importance of the domain to modern urban life, it is anticipated that more commercial software will emerge.

Due to the complex characteristics of water networks, it is anticipated that the network optimization will be dominated by metaheuristic related methods. However, few metaheuristics like HS were originally proposed for water network optimization. It is also noted that, recently, critical comments regarding the usefulness of some recently developed metaheuristics have been seen in the field, where the key argument is about contribution (sometimes being considered as slightly varied from earlier developed metaheuristics or proposed concepts) to the metaheuristic research community [167], [168]. Specifically, for the criticism regarding the HS algorithm, it is claimed equivalent to evolution strategies (ES) by Weyland [168]. Geem as the original author of HS also made a rebuttal, arguing the differences between HS and ES (e.g., algorithms structure and mechanism, characteristics of the problems being targeted, similarity and uniqueness of general metaheuristic algorithms, applicability of a method rather than novelty) [169]. Moreover, besides its wide application in water resource management, HS is becoming popular in other research fields such as steel, electronics, mechanics, telecommunication, medicine, control, power and energy [170], [171]. In addition, HS itself as a global metaheuristic optimizer has also attracted a lot of interest in recent years [172], [173]. Although some levels of equivalencies might (potentially) exist between different metaheuristic methods, it would be useful to discover and understand the underlying ideas being adopted to deal with the well-known exploitation/intensification and exploration/diversification abilities, in order to balance the convergence and the solution quality of an optimization algorithm.

It is therefore expected that more emerging computational paradigms will originate from this domain with improved computational efficiency and solution optimality, meanwhile strengthening the subject of computational intelligence. Generally, the most important stage for developing a metaheuristic algorithm is to observe the interesting behavior/phenomenon in the universe (such as those from biology), and then imitate the internal mechanism that underpins this phenomenon in order to artificially simulate such a phenomenon. This behavior/phenomenon can then be examined to ensure that, after some transformations, it corresponds to the objective of the optimization algorithm (normally formulating the objective as to the minimization and maximization of some indexes). It is therefore vital to capture the key factors that form the internal mechanism determining the phenomenon. These "key factors" can be defined as a minimum set of factors, in which the neglect of any one factor can significantly affect the performance of the observed phenomenon. For instance, the phenomenon of human evolution is explained by the mechanism of natural selection in Darwinism which is mimicked by using several evolving operators as the key factors in GAs. In HS, the phenomenon of the improvisation process of musicians with the associated mechanism is mimicked by a set of pertinent operations on harmonies. Therefore, expert knowledge from the field is important in discovering these key factors. To this end, another research topic is to improve the imperfect/unreasonable aspects of the mechanism that governs a metaheuristic optimization method or to add new mechanisms into the existing mechanism, in order to obtain better algorithm performance. This is reasonable as the biologies and artificial processes continue evolving or upgrading and their present mechanism being imitated is obviously not the best from the temporal perspective. Finally, no matter what type of metaheuristics that will be developed, a special focus should always be placed on the abilities of exploitation/intensification and exploration/diversification.

VI. Conclusion

This paper has investigated network optimization of potable water distribution networks (WDNs) and wastewater collection networks (WWCNs). Both of these types of water networks were conceptually and functionally introduced in this paper together with illustrative benchmarks. The cost function and constraints for each type of network were then separately formulated and explained, in a way that helps clarify the understanding of the problem and facilitates the optimization tasks that are to be considered. The optimization methodologies together with their rationales for use within both types of water networks were then systematically described and analyzed from traditional approaches through to the state-of-the-art. Engagement with the optimization of WWCNs has been found to be somewhat lacking in the domain compared to the more common WDNs, especially in the use of advanced optimization approaches such as metaheuristics and multi-objective optimization. More discussions of current trends and potential future research directions were then given. On the one hand, several critical aspects were identified and discussed for WDNs and WWCNs respectively, for instance, the popularity of using common and distinct optimization methods between WDNs and WWCNs due to their different network characteristics, the suggestions for using deterministic methods and/or heuristic methods in combination with decomposition approaches for WDN optimization, the different complexities of objective functions for WDNs and WWCNs and the corresponding suggestions, etc. On the other hand, common opportunities for the optimization of both WDNs and WWCNs were also identified, for instance, the construction of central repositories considering a number of critical points for fair comparison and research purposes, the investigation of more advanced system modeling techniques for the replication of hydraulics to save network optimization times, the high level guidance and suggestions for the development of new effective and efficient metaheuristic optimization methodologies, etc.

The paper is not intended to provide a completely exhaustive review of optimization methods, but to reflect on current developments and the state-of-the-art technologies being examined in the water network optimization domain, as well as gaps related to common optimization methodologies. As a matter of fact, this is an application area currently not fully considered in the artificial intelligence field, especially within the computational intelligence subject. Apart from researchers from the field of water and environmental management, more researchers from the artificial intelligence field are expected to be working on water network optimization in the future, as important challenges begin to emerge related to the aging/degradation of water networks (some of which date back to the Victorian period), the expansion of urban areas and cities, and the environmental and healthy considerations (including strengthening of the water regulatory framework).

Acknowledgment

The authors would like to thank the Editors and the reviewers for their most constructive comments and suggestions to improve the quality of the paper.

References

[1] M. Cunha and L. Ribeiro, "Tabu search algorithms for water network optimization," *Eur. J. Oper. Res.*, vol. 157, no. 3, pp. 746–758, Sep. 2004.

[2] H. M. V. Samani and A. Mottaghi, "Optimization of water distribution networks using integer linear programming," *J. Hydraul. Eng.*, vol. 132, no. 5, pp. 501–509, May 2006.

[3] S. Atkinson, R. Farmani, F. A. Memon, and D. Butler, "Reliability indicators for water distribution system design: Comparison," *J. Water Resour. Plann. Manage.*, vol. 140, no. 2, pp. 160–168, Feb. 2014.

[4] A. Gupta, S. L. Mehndiratta, and P. Khanna, "Gravity wastewater collection systems optimization," *J. Environ. Eng.*, vol. 109, no. 5, pp. 1195–1209, Oct. 1983.

[5] *New Mexico Wastewater Systems Operator Certification Study Manual*. Santa Fe, New Mexico, USA: Utility Operators Certification Program, Surface Water Quality Bureau, New Mexico Environment Department (NMED), Nov. 2007.

[6] P. K. Swamee and A. K. Sharma, "Optimal design of a sewer line using linear programming," *Appl. Math. Modell.*, vol. 37, no. 6, pp. 4430–4439, Mar. 2013.

[7] Thames Tideway Tunnels Ltd, "Thames Tideway Tunnel-Creating a cleaner, healthier River Thames," accessed July 20, 2015. [Online]. Available: http://www.thamestidewaytunnel.co.uk/

[8] R. Farmani, G. Walters, and D. Savic, "Evolutionary multi-objective optimization of the design and operation of water distribution network: total cost vs. reliability vs. water quality," *J. Hydroinf.*, vol. 8, no. 3, pp. 165–179, 2006.

[9] B. Coelho and A. Andrade-Campos, "Efficiency achievement in water supply systems–a review," *Renewable Sustainable Energy Rev.*, vol. 30, pp. 59–84, Feb. 2014.

[10] R. Field and D. Sullivan, *Wet-weather flow in the urban watershed: technology and management*. CRC Press, 2010.

[11] City of Henderson, NV, "Utility services: Water and sewer laterals," accessed July 20, 2015. [Online]. Available: http://www.cityofhenderson.com/utility-services/customer-care-center/water-and-sewer-laterals

[12] E. Todini, "Looped water distribution networks design using a resilience index based heuristic approach," *Urban Water*, vol. 2, no. 2, pp. 115–122, Jun. 2000.

[13] T. M. Walski, E. Downey Brill, Jr., J. Gessler, I. C. Goulter, R. M. Jeppson, K. Lansey, H.-L. Lee, J. C. Liebman, L. Mays, D. R. Morgan, and L. Ormsbee, "Battle of the network models: Epilogue," *J. Water Resour. Plann. Manage.*, vol. 113, no. 2, pp. 191–203, Mar. 1987.

[14] E. Alperovits and U. Shamir, "Design of optimal water distribution systems," *Water Resour. Res.*, vol. 13, no. 6, pp. 885–900, Dec. 1977.

[15] O. Fujiwara and D. B. Khang, "A two-phase decomposition method for optimal design of looped water distribution networks," *Water Resour. Res.*, vol. 26, no. 4, pp. 539–549, Apr. 1990.

[16] J. C. Schaake and D. Lai, *Linear programming and dynamic programming applications to water distribution network design*, ser. Report 116. MIT Hydrodynamics Laboratory, 1969.

[17] J. Gessler, "Pipe network optimization by enumeration," in *Proc. Comput. Appl. Water Resour.* New York, USA: ASCE, 1985, pp. 572–581.

[18] G. Li and R. G. S. Matthew, "New approach for optimization of urban drainage systems," *J. Environ. Eng.*, vol. 116, no. 5, pp. 927–944, Sep. 1990.

[19] L. A. Rossman, *EPANET 2: users manual*. Cincinnati, OH: National Risk Management Research Laboratory, Office of Research and Development, US Environmental Protection Agency, 2000.

[20] Innovyze, "Water distribution modeling and management–InfoWorks WS," accessed July 20, 2015. [Online]. Available: http://www.innovyze.com/products/infoworks_ws/

[21] Rehm Software, "Program CROSS," accessed July 20, 2015. [Online]. Available: http://www.rehm.de/produkte/WaterPac/Cross/Default.aspx

[22] Tahoe Design Software, "HYDROFLO$^{\text{TM}}$ 2.1 – piping system design software for windows$^{\text{TM}}$," accessed July 20, 2015. [Online]. Available: http://www.tahoesoft.com/html/hydroflo.htm

[23] KYPipe LLC, "Pipe2012: Kypipe overview," accessed July 20, 2015. [Online]. Available: http://kypipe.com/kypipe

[24] S. Alvisi and M. Franchini, "A procedure for the design of district metered areas in water distribution systems," *Procedia Eng.*, vol. 70, pp. 41–50, 2014.

[25] Dwr Cymru Welsh Water, "What is a Combined Sewer Overflow?" accessed July 20, 2015. [Online]. Available: https://www.dwrcymru.co.uk/_library/leaflets_publications_english /combined_sewer_overflow.pdf

[26] D. Strifling, "Sanitary sewer overflows: Past, present, and future regulation," *Marquette Law Rev.*, vol. 87, pp. 225–252, 2003.

[27] L. W. Mays, *Stormwater collection systems design handbook*. McGraw-Hill New York, 2001.

[28] Water Environment Research Foundation (WERF), "Fact sheet C1–Gravity sewer systems," accessed July 20, 2015. [Online]. Available: www.werf.org/c/DecentralizedCost/C1_Gravity_Sewers.aspx

[29] L. W. Mays and H. G. Wenzel, "Optimal design of multilevel branching sewer systems," *Water Resour. Res.*, vol. 12, no. 5, pp. 913–917, Oct. 1976.

[30] L. A. Rossman, *Storm water management model user's manual, version 5.0*. Cincinnati, OH: National Risk Management Research Laboratory, Office of Research and Development, US Environmental Protection Agency, 2010.

[31] A. D. Corte and K. Sörensen, "Optimisation of gravity-fed water distribution network design: A critical review," *Eur. J. Oper. Res.*, vol. 228, no. 1, pp. 1–10, Jul. 2013.

[32] T. Haktanır and M. Ardıçlıoğlu, "Numerical modeling of Darcy-Weisbach friction factor and branching pipes problem," *Adv. Eng. Software*, vol. 35, no. 12, pp. 773–779, Dec. 2004.

[33] Z. S. Kapelan, D. A. Savic, and G. A. Walters, "Multiobjective design of water distribution systems under uncertainty," *Water Resour. Res.*, vol. 41, no. 11, p. W11407, Nov. 2005.

[34] W. Kurek and A. Ostfeld, "Multi-objective optimization of water quality, pumps operation, and storage sizing of water distribution systems," *J. Environ. Manage.*, vol. 115, pp. 189–197, Jan. 2013.

[35] R. Guercio and Z. Xu, "Linearized optimization model for reliability-based design of water systems," *J. Hydraul. Eng.*, vol. 123, no. 11, pp. 1020–1026, Nov. 1997.

[36] R. Farmani, G. A. Walters, and D. A. Savic, "Trade-off between total cost and reliability for Anytown water distribution network," *J. Water Resour. Plann. Manage.*, vol. 131, no. 3, pp. 161–171, May 2005.

[37] D. F. Yates, A. B. Templeman, and T. B. Boffey, "The computational complexity of the problem of determining least capital cost designs for water supply networks," *Eng. Optim.*, vol. 7, no. 2, pp. 143–155, 1984.

[38] M. Cavazzuti, "Deterministic optimization," in *Optimization Methods*. Springer Berlin Heidelberg, 2013, pp. 77–102.

[39] C. R. Suribabu, "Heuristic-based pipe dimensioning model for water distribution networks," *J. Pipeline Syst. Eng. Pract.*, vol. 3, no. 4, pp. 115–124, Nov. 2012.

[40] V. V. Sonak and P. R. Bhave, "Global optimum tree solution for single-source looped water distribution networks subjected to a single loading pattern," *Water Resour. Res.*, vol. 29, no. 7, pp. 2437–2443, Jul. 1993.

[41] T. Liang, "Design conduit system by dynamic programming," *J. Hydraul. Div.*, vol. 97, no. 3, pp. 383–393, Mar. 1971.

[42] T. Liang, K.-P. Yang, and I.-P. Wu, "WRRCTR No. 83 Dynamic programming optimization: A water distribution system design," Water Resour. Res. Center, University of Hawaii at Manoa, Tech. Rep., 1974.

[43] K. E. Lansey and L. W. Mays, "Optimization model for water distribution system design," *J. Hydraul. Eng.*, vol. 115, no. 10, pp. 1401–1418, Oct. 1989.

[44] N. Duan, L. W. Mays, and K. E. Lansey, "Optimal reliability-based design of pumping and distribution systems," *J. Hydraul. Eng.*, vol. 116, no. 2, pp. 249–268, Feb. 1990.

[45] J. B. Martínez, "Discussion of "optimization of water distribution networks using integer linear programming" by Hossein M. V. Samani and Alireza Mottaghi," *J. Hydraul. Eng.*, vol. 134, no. 7, pp. 1023–1024, Jul. 2008.

[46] D. Savic and M. da Conceição Cunha, "Discussion of "optimization of water distribution networks using integer and linear programming" by Hossein M. V. Samani and Alireza Mottaghi," *J. Hydraul. Eng.*, vol. 134, no. 7, pp. 1024–1025, Jul. 2008.

[47] C. Blum and A. Roli, "Metaheuristics in combinatorial optimization: Overview and conceptual comparison," *ACM Comput. Surv.*, vol. 35, no. 3, pp. 268–308, Sep. 2003.

[48] M. d. C. Cunha and J. Sousa, "Water distribution network design optimization: simulated annealing approach," *J. Water Resour. Plann. Manage.*, vol. 125, no. 4, pp. 215–221, Jul. 1999.

[49] F. Glover and M. Laguna, "Tabu search," in *Handbook of Combinatorial Optimization*, D.-Z. Du and P. M. Pardalos, Eds. Springer US, 1999, pp. 2093–2229.

[50] H. Beigy and M. R. Meybodi, "Cellular learning automata with multiple learning automata in each cell and its applications," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 40, no. 1, pp. 54–65, Feb. 2010.

[51] E. Keedwell and S.-T. Khu, "Novel cellular automata approach to optimal water distribution network design," *J. Comput. Civil Eng.*, vol. 20, no. 1, pp. 49–56, Jan. 2006.

[52] A. De Corte and K. Sörensen, "An iterated local search algorithm for water distribution network design optimisation," University of Antwerp, Faculty of Applied Economics, Working Papers, 2014.

[53] Q. Nguyen, "Variable neighborhood search for water distribution network design problem," Master's thesis, the Software Engineering Department, University of Science, Vietnam National University - Ho Chi Minh, September 2014.

[54] A. R. Simpson, G. C. Dandy, and L. J. Murphy, "Genetic algorithms compared to other techniques for pipe optimization," *J. Water Resour. Plann. Manage.*, vol. 120, no. 4, pp. 423–443, Jul. 1994.

[55] F. Zheng, A. C. Zecchin, A. R. Simpson, and M. F. Lambert, "Non-crossover dither creeping mutation-based genetic algorithm for pipe network optimization," *J. Water Resour. Plann. Manage.*, vol. 140, no. 4, pp. 553–557, Apr. 2014.

[56] W. Bi, G. C. Dandy, and H. R. Maier, "Improved genetic algorithm optimization of water distribution system design by incorporating domain knowledge," *Environ. Modell. Softw.*, to be published, 2014.

[57] J. W. Nicklow, P. M. Reed, D. Savic, T. Dessalegne, L. Harrell, A. Chan-Hilton, M. Karamouz, B. Minsker, A. Ostfeld, A. Singh, and E. Zechman, "State of the art for genetic algorithms and beyond in water resources planning and management," *J. Water Resour. Plann. Manage.*, vol. 136, no. 4, pp. 412–432, Jul. 2010.

[58] M. Dorigo, M. Birattari, and T. Stützle, "Ant colony optimization," *IEEE Comput. Intell. Mag.*, vol. 1, no. 4, pp. 28–39, Nov. 2006.

[59] H. R. Maier, A. R. Simpson, A. C. Zecchin, W. K. Foong, K. Y. Phang, H. Y. Seah, and C. L. Tan, "Ant colony optimization for design of water distribution systems," *J. Water Resour. Plann. Manage.*, vol. 129, no. 3, pp. 200–209, May 2003.

[60] A. C. Zecchin, H. R. Maier, A. R. Simpson, M. Leonard, and J. B. Nixon, "Ant colony optimization applied to water distribution system design: Comparative study of five algorithms," *J. Water Resour. Plann. Manage.*, vol. 133, no. 1, pp. 87–92, Jan. 2007.

[61] Y. Valle, G. K. Venayagamoorthy, S. Mohagheghi, J.-C. Hernandez, and R. G. Harley, "Particle swarm optimization: Basic concepts, variants and applications in power systems," *IEEE Trans. Evol. Comput.*, vol. 12, no. 2, pp. 171–195, Apr. 2008.

[62] C. R. Suribabu and T. R. Neelakantan, "Design of water distribution networks using particle swarm optimization," *Urban Water J.*, vol. 3, no. 2, pp. 111–120, 2006.

[63] J. C. Bansal and K. Deep, "Optimal design of water distribution networks via particle swarm optimization," in *Proc. IEEE Int. Adv. Comput. Conf.*, Mar. 2009, pp. 1314–1316.

[64] R. Storn and K. Price, "Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optim.*, vol. 11, no. 4, pp. 341–359, Dec. 1997.

[65] C. R. Suribabu, "Differential evolution algorithm for optimal design of water distribution networks," *J. Hydroinf.*, vol. 12, no. 1, pp. 66–82, 2010.

[66] F. Zheng, A. C. Zecchin, and A. R. Simpson, "Self-adaptive differential evolution algorithm applied to water distribution system optimization," *J. Comput. Civil Eng.*, vol. 27, no. 2, pp. 148–158, Mar. 2013.

[67] Z. W. Geem, J. H. Kim, and G. V. Loganathan, "A new heuristic optimization algorithm: Harmony search," *Simul.*, vol. 76, no. 2, pp. 60–68, Feb. 2001.

[68] Z. Geem and Y. Cho, "Optimal design of water distribution networks using parameter-setting-free harmony search for two major parameters," *J. Water Resour. Plann. Manage.*, vol. 137, no. 4, pp. 377–380, Jul. 2011.

[69] M. Eusuff, K. Lansey, and F. Pasha, "Shuffled frog-leaping algorithm: a memetic meta-heuristic for discrete optimization," *Eng. Optim.*, vol. 38, no. 2, pp. 129–154, Mar. 2006.

[70] M. M. Eusuff and K. E. Lansey, "Optimization of water distribution network design using the shuffled frog leaping algorithm," *J. Water Resour. Plann. Manage.*, vol. 129, no. 3, pp. 210–225, May 2003.

[71] R. Baños, C. Gil, J. Reca, and F. G. Montoya, "A memetic algorithm applied to the design of water distribution networks," *Appl. Soft Comput.*, vol. 10, no. 1, pp. 261–266, Jan. 2010.

[72] M.-D. Lin, Y.-H. Liu, G.-F. Liu, and C.-W. Chu, "Scatter search heuristic for least-cost design of water distribution networks," *Eng. Optim.*, vol. 39, no. 7, pp. 857–876, 2007.

[73] C.-W. Chu, M.-D. Lin, G.-F. Liu, and Y.-H. Sung, "Application of immune algorithms on solving minimum-cost problem of water distribution network," *Math. Comput. Modell.*, vol. 48, no. 11-12, pp. 1888–1900, Dec. 2008.

[74] S. Mohan and K. S. Jinesh Babu, "Optimal water distribution network design with honey-bee mating optimization," *J. Comput. Civil Eng.*, vol. 24, no. 1, pp. 117–126, Jan. 2010.

[75] M.-D. Lin, C.-W. Chu, G.-F. Liu, Z.-H. Wu, and S.-F. Yeh, "Application of evolutionary computing on least-cost water distribution network designs," in *Proc. 2009 IEEE Int. Conf. Ind. Eng. Eng. Manage.*, Dec. 2009, pp. 375–379.

[76] R. Rubinstein, "The cross-entropy method for combinatorial and continuous optimization," *Method. Comput. Appl. Probab.*, vol. 1, no. 2, pp. 127–190, Sep. 1999.

[77] A. Shibu and M. J. Reddy, "Least cost design of water distribution network by cross entropy optimization," in *Proc. IEEE 2011 World Congr. Inf. Commun. Technol.*, Dec. 2011, pp. 302–306.

[78] A. Shibu and M. Janga Reddy, "Reliability-based optimal design of water distribution networks under uncertain demands using cross-entropy method," *ISH J. Hydraul. Eng.*, vol. 18, no. 3, pp. 258–268, 2012.

[79] L. Perelman and A. Ostfeld, "Topological clustering for water distribution systems analysis," *Environ. Modell. Softw.*, vol. 26, no. 7, pp. 969–972, Jul. 2011.

[80] M. S. Kadu, R. Gupta, and P. R. Bhave, "Optimal design of water networks using a modified genetic algorithm with reduction in search space," *J. Water Resour. Plann. Manage.*, vol. 134, no. 2, pp. 147–160, Mar. 2008.

[81] F. Zheng, A. R. Simpson, and A. C. Zecchin, "A combined NLP-differential evolution algorithm approach for the optimization of looped water distribution systems," *Water Resour. Res.*, vol. 47, no. 8, p. W08531, Aug. 2011.

[82] J. W. Deuerlein, "Decomposition model of a general water supply network graph," *J. Hydraul. Eng.*, vol. 134, no. 6, pp. 822–832, Jun. 2008.

[83] F. Zheng, A. R. Simpson, and A. C. Zecchin, "Coupled binary linear programming–differential evolution algorithm approach for water distribution system optimization," *J. Water Resour. Plann. Manage.*, vol. 140, no. 5, pp. 585–597, May 2014.

[84] ——, "A decomposition and multistage optimization approach applied to the optimization of water distribution systems with multiple supply sources," *Water Resour. Res.*, vol. 49, no. 1, pp. 380–399, Jan. 2013.

[85] F. Zheng, A. R. Simpson, A. C. Zecchin, and J. W. Deuerlein, "A graph decomposition-based approach for water distribution network optimization," *Water Resour. Res.*, vol. 49, no. 4, pp. 2093–2109, Apr. 2013.

[86] K. Miettinen, *Nonlinear multiobjective optimization*. Springer US, 1999, vol. 12.

[87] R. T. Marler and J. S. Arora, "Survey of multi-objective optimization methods for engineering," *Struct. Multidiscip. Optim.*, vol. 26, no. 6, pp. 369–395, Apr. 2004.

[88] J. Andersson, "A survey of multiobjective optimization in engineering design," Department of Mechanical Engineering, Linköping University, Sweden, Tech. Rep., 2000.

[89] C. M. Fonseca and P. J. Fleming, "Multiobjective optimization and multiple constraint handling with evolutionary algorithms–Part I: A unified formulation," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 28, no. 1, pp. 26–37, Jan. 1998.

[90] C. Brizuela, "Applications of multi-objective evolutionary algorithms - [book review]," *IEEE Comput. Intell. Mag.*, vol. 1, no. 1, pp. 43–44, Feb. 2006.

[91] C. A. Coello Coello, "Evolutionary multi-objective optimization: a historical view of the field," *IEEE Comput. Intell. Mag.*, vol. 1, no. 1, pp. 28–36, Feb. 2006.

[92] J. D. Schaffer, "Multiple objective optimization with vector evaluated genetic algorithms," in *Proc. First Int. Conf. Genet. Algorithms*. Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 1985, pp. 93–100.

[93] D. E. Goldberg, *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley Publishing Company, Inc., 1989.

[94] C. M. Fonseca and P. J. Fleming, "Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization," in *Proc. 5th Int. Conf. Genet. Algorithms*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1993, pp. 416–423.

[95] J. Horn, N. Nafpliotis, and D. E. Goldberg, "A niched pareto genetic algorithm for multiobjective optimization," in *Proc. First IEEE Conf. Evol. Comput., IEEE World Congr. Comput. Intell.*, Jun. 1994, pp. 82–87, vol.1.

[96] N. Srinivas and K. Deb, "Multiobjective optimization using nondominated sorting in genetic algorithms," *Evol. Comput.*, vol. 2, no. 3, pp. 221–248, 1994.

[97] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach," *IEEE Trans. Evol. Comput.*, vol. 3, no. 4, pp. 257–271, Nov. 1999.

[98] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the strength Pareto evolutionary algorithm for multiobjective optimization," in *Evolutionary Methods for Design, Optimisation, and Control*. CIMNE, Barcelona, Spain, 2002, pp. 95–100.

[99] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.

[100] D. W. Corne, J. D. Knowles, and M. J. Oates, "The Pareto envelope-based selection algorithm for multiobjective optimization," in *Parallel Problem Solving from Nature PPSN VI*, ser. Lect. Notes Comput. Sci. Springer Berlin, 2000, vol. 1917, pp. 839–848.

[101] D. W. Corne, N. R. Jerram, J. D. Knowles, and M. J. Oates, "PESA-II: Region-based selection in evolutionary multiobjective optimization," in *Proc. Genet. Evol. Comput. Conf.* Morgan Kaufmann Publishers, 2001, pp. 283–290.

[102] J. D. Knowles and D. W. Corne, "Approximating the nondominated front using the Pareto archived evolution strategy," *Evol. Comput.*, vol. 8, no. 2, pp. 149–172, Jun. 2000.

[103] M. Erickson, A. Mayer, and J. Horn, "The niched Pareto genetic algorithm 2 applied to the design of groundwater remediation systems," in *Evol. Multi-Criterion Optim.*, ser. Lect. Notes Comput. Sci. Springer Berlin Heidelberg, 2001, vol. 1993, pp. 681–695.

[104] E. Zitzler, "Evolutionary algorithms for multiobjective optimization: Methods and applications," Ph.D. dissertation, ETH Zürich, Switzerland, 1999.

[105] T. D. Prasad and N.-S. Park, "Multiobjective genetic algorithms for design of water distribution networks," *J. Water Resour. Plann. Manage.*, vol. 130, no. 1, pp. 73–82, Jan. 2004.

[106] R. Farmani, D. A. Savic, and G. A. Walters, "Multi-objective optimization of water systems: a comparative study," *Cabrera et al.(eds.) Pumps, Electromechanical Devices and Systems Applied to Urban Water Management*, vol. 1, pp. 247–256, 2003.

[107] M. Atiquzzaman, S.-Y. Liong, and X. Yu, "Alternative decision making in water distribution network with NSGA-II," *J. Water Resour. Plann. Manage.*, vol. 132, no. 2, pp. 122–126, Mar. 2006.

[108] F. Zheng and A. Zecchin, "An efficient decomposition and dual-stage multi-objective optimization method for water distribution systems with multiple supply sources," *Environ. Modell. Softw.*, vol. 55, pp. 143–155, May 2014.

[109] E. Creaco, M. Franchini, and T. M. Walski, "Accounting for phasing of construction within the design of water distribution networks," *J. Water Resour. Plann. Manage.*, vol. 140, no. 5, pp. 598–606, May 2014.

[110] P. B. Cheung, L. F. R. Reis, K. T. M. Formiga, F. H. Chaudhry, and W. G. C. Ticona, "Multiobjective evolutionary algorithms applied to the rehabilitation of a water distribution system: A comparative study," in *Evol. Multi-Criterion Optim.*, ser. Lect. Notes Comput. Sci. Springer Berlin Heidelberg, 2003, vol. 2632, pp. 662–676.

[111] R. Farmani, D. A. Savic, and G. A. Walters, "Evolutionary multi-objective optimization in water distribution network design," *Eng. Optim.*, vol. 37, no. 2, pp. 167–183, 2005.

[112] I. Montalvo, J. Izquierdo, S. Schwarze, and R. Pérez-García, "Multi-objective particle swarm optimization applied to water distribution systems design: An approach with human interaction," *Math. Comput. Modell.*, vol. 52, no. 7-8, pp. 1219–1227, Oct. 2010.

[113] L. Perelman, A. Ostfeld, and E. Salomons, "Cross entropy multiobjective optimization for water distribution systems design," *Water Resour. Res.*, vol. 44, no. 9, p. W09413, Sep. 2008.

[114] D. Jung, D. Kang, J. H. Kim, and K. Lansey, "Robustness-based design of water distribution systems," *J. Water Resour. Plann. Manage.*, vol. 140, no. 11, p. 04014033, Nov. 2014.

[115] G. Fu, Z. Kapelan, J. R. Kasprzyk, and P. Reed, "Optimal design of water distribution systems using many-objective visual analytics," *J. Water Resour. Plann. Manage.*, vol. 139, no. 6, pp. 624–633, Nov. 2013.

[116] A. Haghighi, H. M. V. Samani, and Z. M. V. Samani, "GA-ILP method for optimization of water distribution networks," *Water Resour. Manage.*, vol. 25, no. 7, pp. 1791–1808, May. 2011.

[117] "A hybrid genetic algorithm for the design of water distribution networks," *Eng. Appl. Artif. Intell.*, vol. 18, no. 4, pp. 461–472, 2005.

[118] W. Bi and G. C. Dandy, "Optimization of water distribution systems using online retrained metamodels," *J. Water Resour. Plann. Manage.*, vol. 140, no. 11, p. 04014032, Nov. 2014.

[119] A. Marchi, et al., "Battle of the water networks II," *J. Water Resour. Plann. and Manage.*, vol. 140, no. 7, p. 04014009, Jul. 2014.

[120] S. W. Miles and J. P. Heaney, "Better than "optimal" method for designing drainage systems," *J. Water Resour. Plann. Manage.*, vol. 114, no. 5, pp. 477–499, Sep. 1988.

[121] L. W. Mays and B. C. Yen, "Optimal cost design of branched sewer systems," *Water Resour. Res.*, vol. 11, no. 1, pp. 37–47, Feb. 1975.

[122] D. P. Desher and P. K. Davis, "Designing sanitary sewers with microcomputers," *J. Environ. Eng.*, vol. 112, no. 6, pp. 993–1007, Dec. 1986.

[123] A. A. Elimam, C. Charalambous, and F. H. Ghobrial, "Optimum design of large sewer networks," *J. Environ. Eng.*, vol. 115, no. 6, pp. 1171–1190, Dec. 1989.

[124] W.-X. Peng and R. Jia, "Improved genetic algorithms for optimal design of drainage systems," in *Proc. 8th Int. Conf. Control, Autom., Rob. Vision*, vol. 1, Dec. 2004, pp. 227–231.

[125] V. Marchionni, N. Lopes, L. Mamouros, and D. Covas, "Modelling sewer systems costs with multiple linear regression," *Water Resour. Manage.*, pp. 4415–4431, Oct. 2014.

[126] J. M. Gupta, P. Khanna, and S. K. Agarwal, "Optimal design of wastewater collection systems," *J. Environ. Eng. Div.*, vol. 102, no. 5, pp. 1029–1041, Sep./Oct. 1976.

[127] P. K. Swamee, "Design of sewer line," *J. Environ. Eng.*, vol. 127, no. 9, pp. 776–781, Sep. 2001.

[128] C. Charalambous and A. A. Elimam, "Heuristic design of sewer networks," *J. Environ. Eng.*, vol. 116, no. 6, pp. 1181–1199, Oct. 1990.

[129] S.-F. Yeh, Y.-J. Chang, and M.-D. Lin, "Application of simulated annealing on least-cost design of sewer network," in *Proc. 2010 IEEE Int. Conf. Ind. Eng. Eng. Manage.*, Dec. 2010, pp. 390–394.

[130] O. Karovic and L. W. Mays, "Sewer system design using simulated annealing in Excel," *Water Resour. Manage.*, pp. 1–15, 2014 (to be published).

[131] A. Haghighi and A. E. Bakhshipour, "Deterministic integrated optimization model for sewage collection networks using tabu search," *J. Water Resour. Plann. Manage.*, vol. 141, no. 1, p. 04014045, Jan. 2015.

[132] S.-F. Yeh, C.-W. Chu, Y.-J. Chang, and M.-D. Lin, "Applying tabu search and simulated annealing to the optimal design of sewer networks," *Eng. Optim.*, vol. 43, no. 2, pp. 159–174, 2011.

[133] Y. Guo, G. A. Walters, S. T. Khu, and E. Keedwell, "A novel cellular automata based approach to storm sewer design," *Eng. Optim.*, vol. 39, no. 3, pp. 345–364, 2007.

[134] M. H. Afshar, M. Shahidi, M. Rohani, and M. Sargolzaei, "Application of cellular automata to sewer network optimization problems," *Sci. Iran.*, vol. 18, no. 3, pp. 304–312, Jun. 2011.

[135] M. H. Afshar, "Improving the efficiency of ant algorithms using adaptive refinement: Application to storm water network design," *Adv. Water Resour.*, vol. 29, no. 9, pp. 1371–1382, Sep. 2006.

[136] ——, "Partially constrained ant colony optimization algorithm for the solution of constrained optimization problems: Application to storm water network design," *Adv. Water Resour.*, vol. 30, no. 4, pp. 954–965, Apr. 2007.

[137] M. H. Afshar and M. Rohani, "Optimal design of sewer networks using cellular automata-based hybrid methods: Discrete and continuous approaches," *Eng. Optim.*, vol. 44, no. 1, pp. 1–22, 2012.

[138] L. Y. Liang, R. G. Thompson, and D. M. Young, "Designing wastewater collection systems using genetic algorithms," in *PRICAI 2000 Topics in Artif. Intell.*, ser. Lect. Notes Comput. Sci. Springer Berlin Heidelberg, 2000, vol. 1886, pp. 415–423.

[139] M. H. Afshar, A. Afshar, M. A. Mariño, and A. A. S. Darbandi, "Hydrograph-based storm sewer design optimization by genetic algorithm," *Can. J. Civ. Eng.*, vol. 33, no. 3, pp. 319–325, 2006.

[140] A. Haghighi and A. E. Bakhshipour, "Optimization of sewer networks using an adaptive genetic algorithm," *Water Resour. Manage.*, vol. 26, no. 12, pp. 3441–3456, Sep. 2012.

[141] R. Moeini and M. H. Afshar, "Layout and size optimization of sanitary sewer network using intelligent ants," *Adv. Eng. Software*, vol. 51, pp. 49–62, Sep. 2012.

[142] J. Izquierdo, I. Montalvo, R. Pérez, and V. S. Fuertes, "Design optimization of wastewater collection networks by PSO," *Comput. Math. Appl.*, vol. 56, no. 3, pp. 777–784, Aug. 2008.

[143] Y. Guo, G. Walters, S.-T. Khu, and E. Keedwell, "Optimal design of sewer networks using hybrid cellular automata and genetic algorithm," in *Proc. Fifth IWA World Water Congr.*, 2006.

[144] Y. Guo, E. Keedwell, G. Walters, and S.-T. Khu, "Hybridizing cellular automata principles and NSGAII for multi-objective design of urban water networks," in *Evol. Multi-Criterion Optim.*, ser. Lect. Notes Comput. Sci. Springer Berlin, 2007, vol. 4403, pp. 546–559.

[145] Y. Guo, G. Walters, S.-T. Khu, and E. Keedwell, "Efficient multiobjective storm sewer design using cellular automata and genetic algorithm hybrid," *J. Water Resour. Plann. Manage.*, vol. 134, no. 6, pp. 511–515, Nov. 2008.

[146] T.-C. Pan and J.-J. Kao, "GA-QP model to optimize sewer system design," *J. Environ. Eng.*, vol. 135, no. 1, pp. 17–24, Jan. 2009.

[147] M. Möderl, R. Sitzenfrei, T. Fetz, E. Fleischhacker, and W. Rauch, "Systematic generation of virtual networks for water supply," *Water Resour. Res.*, vol. 47, no. 2, p. W02502, Feb. 2011.

[148] A. De Corte and K. Sörensen, "Hydrogen: an artificial water distribution network generator," *Water Resour. Manage.*, vol. 28, no. 2, pp. 333–350, Jan. 2014.

[149] K. Bache and M. Lichman, "UCI machine learning repository," University of California, Irvine, School of Information and Computer Sciences, accessed July 20, 2015. [Online]. Available: http://archive.ics.uci.edu/ml

[150] A. Marchi, G. Dandy, A. Wilkins, and H. Rohrlach, "Methodology for comparing evolutionary algorithms for optimization of water distribution systems," *J. Water Resour. Plann. Manage.*, vol. 140, no. 1, pp. 22–31, Jan. 2014.

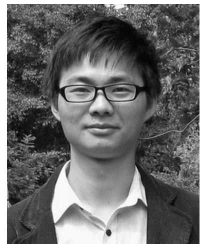
[151] W. Zhao, Q. Niu, K. Li, and G. W. Irwin, "A hybrid learning method for constructing compact rule-based fuzzy models," *IEEE Trans. Cybern.*, vol. 43, no. 6, pp. 1807–1821, Dec. 2013.

[152] M. Jadaliha, Y. Xu, J. Choi, N. S. Johnson, and W. Li, "Gaussian process regression for sensor networks under localization uncertainty," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 223–237, Jan. 2013.

[153] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 513–529, Apr. 2012.

[154] C. Voudouris and E. P. K. Tsang, "Guided local search," in *Handbook of Metaheuristics*, ser. International Series in Operations Research & Management Science. Springer US, 2003, vol. 57, pp. 185–218.

[155] T. A. Feo and M. G. C. Resende, "Greedy randomized adaptive search procedures," *J. Global Optim.*, vol. 6, no. 2, pp. 109–133, Mar. 1995.
[156] Q. Zhou and Y. Li, "Directed variation in evolution strategies," *IEEE Trans. Evol. Comput.*, vol. 7, no. 4, pp. 356–366, Aug. 2003.
[157] T. Chen, K. Tang, G. Chen, and X. Yao, "Analysis of computational time of simple estimation of distribution algorithms," *IEEE Trans. Evol. Comput.*, vol. 14, no. 1, pp. 1–22, Feb. 2010.
[158] D. Karaboga, B. Gorkemli, C. Ozturk, and N. Karaboga, "A comprehensive survey: artificial bee colony (ABC) algorithm and applications," *Artif. Intell. Rev.*, vol. 42, no. 1, pp. 21–57, Jun. 2014.
[159] D. T. Pham, A. Ghanbarzadeh, E. Koç, S. Otri, S. Rahim, and M. Zaidi, "The bees algorithm-a novel tool for complex optimisation problems," in *Proc. 2nd Virtual Int. Conf. Intell. Prod. Mach. Syst.*, 2006, pp. 454–459.
[160] A. H. Gandomi, X.-S. Yang, A. H. Alavi, and S. Talatahari, "Bat algorithm for constrained optimization tasks," *Neural Comput. Appl.*, vol. 22, no. 6, pp. 1239–1255, May 2013.
[161] H. R. Maier, Z. Kapelan, J. Kasprzyk, J. Kollat, L. S. Matott, M. C. Cunha, G. C. Dandy, M. S. Gibbs, E. Keedwell, A. Marchi, A. Ostfeld, D. Savic, D. P. Solomatine, J. A. Vrugt, A. C. Zecchin, B. S. Minsker, E. J. Barbour, G. Kuczera, F. Pasha, A. Castelletti, M. Giuliani, and P. M. Reed, "Evolutionary algorithms and other metaheuristics in water resources: Current status, research challenges and future directions," *Environ. Modell. Softw.*, vol. 62, pp. 271–299, Dec. 2014.
[162] G. Mavrotas, "Effective implementation of the ε-constraint method in multi-objective mathematical programming problems," *Appl. Math. Comput.*, vol. 213, no. 2, pp. 455–465, Jul. 2009.
[163] D. Angus and C. Woodward, "Multiple objective ant colony optimisation," *Swarm Intell.*, vol. 3, no. 1, pp. 69–85, Mar. 2009.
[164] H. Ishibuchi, N. Tsukamoto, and Y. Nojima, "Evolutionary many-objective optimization: A short review," in *Proc. IEEE 2008 Congr. Evol. Comput.*, Jun. 2008, pp. 2419–2426.
[165] I. Giagkiozis and P. J. Fleming, "Methods for multi-objective optimization: An analysis," *Inf. Sci.*, vol. 293, pp. 338–350, Feb. 2015.
[166] K. Deb, A. Sinha, P. J. Korhonen, and J. Wallenius, "An interactive evolutionary multiobjective optimization method based on progressively approximated value functions," *IEEE Trans. Evol. Comput.*, vol. 14, no. 5, pp. 723–739, Oct. 2010.
[167] K. Sörensen, "Metaheuristics–the metaphor exposed," *Int. Trans. Oper. Res.*, vol. 22, no. 1, pp. 3–18, Jan. 2015.
[168] D. Weyland, "A rigorous analysis of the harmony search algorithm: How the research community can be misled by a "novel" methodology," *Int. J. Appl. Metaheuristic Comput.*, vol. 1, no. 2, pp. 50–60, 2010.
[169] Z. W. Geem, "Research commentary: Survival of the fittest algorithm or the novelest algorithm?" *Int. J. Appl. Metaheuristic Comput.*, vol. 1, no. 4, pp. 75–79, 2010.
[170] D. Manjarres, I. Landa-Torres, S. Gil-Lopez, J. Del Ser, M. N. Bilbao, S. Salcedo-Sanz, and Z. W. Geem, "A survey on applications of the harmony search algorithm," *Eng. Appl. Artif. Intell.*, vol. 26, no. 8, pp. 1818–1831, Sep. 2013.
[171] R. Srinivasa Rao, S. V. L. Narasimham, M. Ramalinga Raju, and A. Srinivasa Rao, "Optimal network reconfiguration of large-scale distribution system using harmony search algorithm," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1080–1088, Aug. 2011.
[172] W.-L. Xiang, M.-Q. An, Y.-Z. Li, R.-C. He, and J.-F. Zhang, "An improved global-best harmony search algorithm for faster optimization," *Expert Syst. Appl.*, vol. 41, no. 13, pp. 5788–5803, Oct. 2014.
[173] E. Valian, S. Tavakoli, and S. Mohanna, "An intelligent global harmony search approach to continuous optimization problems," *Appl. Math. Comput.*, vol. 232, pp. 670–684, Apr. 2014.



Wanqing Zhao (M'13) received the B.Eng. degree in automation from Anhui Polytechnic University, Anhui, China, in 2006, the M.Eng. degree in control theory and control engineering from Shanghai University, Shanghai, China, in 2009, and the Ph.D. degree from Intelligent Systems and Control group, Queen's University Belfast, Belfast, U.K, in 2012.

He is currently a Research Fellow with the School of Engineering, Cardiff University, Cardiff, U.K. He was a Research Associate with the Department of Computer Science, Loughborough University, Loughborough, U.K. His current research interests include system identification, fuzzy regression, neural networks, machine learning, heuristic optimization methods, autonomous systems, and water resource management.

Thomas Beach received the Ph.D. degree in computer science, Cardiff University, Cardiff, U.K.

He is currently a lecturer in Construction Informatics, Cardiff University, Cardiff, U.K. He specializes in the application of high performance and distributed computing technologies to the construction industry. His research is currently focusing on the application of cloud computing technologies for the storage, management and processing of BIM data and is currently involved in EU(FP7) and TSB projects in this area.

Yacine Rezgui received the M.Sc. degree in architecture from University Paris 6, Paris, France, and the Ph.D. degree in architecture from Ecole Nationale des Ponts et chaussees, Paris, France.

He is a Professor in Building Systems and Informatics, Cardiff University, Cardiff, U.K., and the founding director of the BRE Centre in Sustainable Engineering, sponsored by the Building Research Establishment (BRE) in the U.K. In 1995, he joined Salford University in the U.K. as a Research Fellow, then Academic: Lecturer (1996), Senior Lecturer (1998), and Professor (2001). He was the founding Director (2003-2007) of the 5* (RAE 2001) rated Informatics Research Institute at Salford University. He conducts research in informatics, including ontology engineering and artificial intelligence applied to the built environment.

Prof. Rezgui has over 150 refereed publications in the above areas, which appeared in international journals such as IEEE Transactions on Services Computing, Information Sciences, Data and Knowledge Engineering, and Computer-Aided Design. He has successfully completed over 40 research and development projects at a national (UK EPSRC and TSB) and international (European Framework Programs 5, 6, 7) levels.