



# Algorithms for optimization of branching gravity-driven water networks

Ian Dardani and Gerard F. Jones

College of Engineering, Villanova University, Villanova, PA 19085, USA

**Correspondence:** Ian Dardani (ian.dardani@villanova.edu)

Received: 31 January 2017 – Discussion started: 24 February 2017

Revised: 3 February 2018 – Accepted: 7 March 2018 – Published: 15 May 2018

**Abstract.** The design of a water network involves the selection of pipe diameters that satisfy pressure and flow requirements while considering cost. A variety of design approaches can be used to optimize for hydraulic performance or reduce costs. To help designers select an appropriate approach in the context of gravity-driven water networks (GDWNs), this work assesses three cost-minimization algorithms on six moderate-scale GDWN test cases. Two algorithms, a backtracking algorithm and a genetic algorithm, use a set of discrete pipe diameters, while a new calculus-based algorithm produces a continuous-diameter solution which is mapped onto a discrete-diameter set. The backtracking algorithm finds the global optimum for all but the largest of cases tested, for which its long runtime makes it an infeasible option. The calculus-based algorithm's discrete-diameter solution produced slightly higher-cost results but was more scalable to larger network cases. Furthermore, the new calculus-based algorithm's continuous-diameter and mapped solutions provided lower and upper bounds, respectively, on the discrete-diameter global optimum cost, where the mapped solutions were typically within one diameter size of the global optimum. The genetic algorithm produced solutions even closer to the global optimum with consistently short run times, although slightly higher solution costs were seen for the larger network cases tested. The results of this study highlight the advantages and weaknesses of each GDWN design method including closeness to the global optimum, the ability to prune the solution space of infeasible and suboptimal candidates without missing the global optimum, and algorithm run time. We also extend an existing closed-form model of Jones (2011) to include minor losses and a more comprehensive two-part cost model, which realistically applies to pipe sizes that span a broad range typical of GDWNs of interest in this work, and for smooth and commercial steel roughness values.

## 1 Introduction

A gravity-driven water network (GDWN) is commonly used to deliver potable water from a source at a high elevation, such as a natural spring or reservoir, to households or public tap stands (Fig. 1). When feasible, gravity-driven water networks are attractive in comparison to pumped networks because of their simplicity and lower capital, operational, and maintenance costs. In addition, in many locations where GDWN are considered, there may be little or no access to reliable grid-based electrical power for pumps. To improve reliability, networks may be designed with loops or multiple water sources, although often material cost considerations restrict attention to single-source branched networks.

Water networks are modeled as a collection of nodes, each representing a point of water demand or supply, which are connected with links representing pipes. The geometrical layout of the site is known and fixed, including water source and demand locations and elevations of all nodes. For the present work, design flow rates are determined from community survey data, which are extrapolated for future population growth. Networks in this category are referred to as “demand-driven” designs. Bhave (1978, 1983) refers to these as “Q-specified” designs. Thus, to design a network of this type, pipe diameters for each link must be chosen such that acceptable but arbitrary minimum (positive) pressure heads are maintained at each node given the design flow rate at

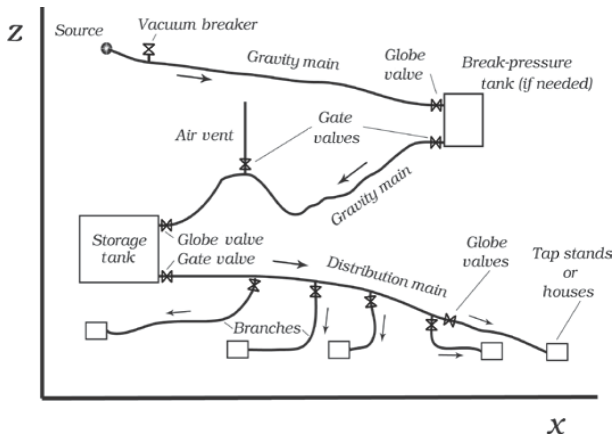


Figure 1. Element schematic of a GDWN.

the node. Furthermore, application of the energy equation to each link in the network demonstrates that the design problem is nonunique; i.e., choosing different pressure heads at the nodes will result in a different pipe diameter solution for the network, and thus different network costs. Minimizing network cost will produce a unique solution to the design problem, i.e., unique link diameters and nodal pressure heads.

In practice, gravity-driven water networks are commonly designed by a marching method, where diameters for each link of the network are chosen sequentially. After selecting a reasonable diameter for each link, the designer calculates the pressure head at the link outlet, and proceeds to the next link if this result is acceptable. In this way, the designer marches through the network until all pipe diameters have been selected. This method produces a feasible solution, but not a cost optimized one. As noted by Bhave (2003), cost savings of 20–30% can result from the use of optimization techniques. In developing regions, the cost of a water network can be prohibitive, adding to the importance of optimizing network design.

Within the provided framework, the global optimum can be found through an exhaustive search of the solution space, known as complete enumeration, although this is infeasible when considering networks with many links and diameter choices (Kadu et al., 2008; González-Cebollada et al., 2011). To reduce the computational time required by enumeration, authors have proposed various partial enumeration methods which prune the search space (Kadu et al., 2008), although some of these techniques may remove the global optimum (Simpson et al., 1994). The most common types of algorithms that have been applied to optimize water network design include traditional deterministic methods, heuristic methods, metaheuristic methods, multi-objective methods, and decomposition methods (Zhao et al., 2016).

Deterministic methods include linear programming (LP), dynamic programming, and nonlinear programming (NLP),

and typically involve rigorous mathematical approaches (Zhao et al., 2016). A brief overview and comparison of these algorithms is given in Kansal et al. (1996), who use a single-part cost correlation for metric pipe diameters between 100 and 350 mm. Linear programming techniques have relatively low computational complexity and allow each link to be composed of two diameters, called a split-pipe solution, although these may not always be practical to implement (Bhave, 1983; Kessler and Shamir, 1989; Swamee and Sharma, 2000; Samani and Mottaghi, 2006). LP can also get stuck in a local optimum (Zhao et al., 2016), although combining LP with metaheuristic techniques can help with the problem's non-smoothness properties (Krapivka and Ostfeld, 2009). Dynamic programming has been used by Yang et al. (1975) and Martin (1980) to optimize networks in stages. This approach begins at the discharge nodes, proceeding to select feasible diameters and joints for upstream stages and storing these partial candidates in memory until the source node is reached. At this point, the algorithm reviews the feasible segment design options and selects a combination of stage solutions producing the lowest cost overall solution. This method, however, requires the designer to allow a relatively narrow range for the design pressure of each node, or otherwise store a large set of feasible candidate solutions in memory and also allow adjoining branches to arrive at different heads at the same node.

Nonlinear programming, a calculus-based method, deals with each link's diameter as a continuous variable. Using Lagrange multipliers and a one-part, pipe-cost model with minor-lossless flow, Swamee and Sharma (2000) developed systems of equations for both continuous and discrete pipe diameters for branch networks, assuming a constant friction factor. When solved, the solution gives diameter values that minimize distribution main cost, not network cost. In carrying out the solution, iteration is required to update the value of the friction factor. For the discrete diameter case, large computational times were noted by Swamee and Sharma because of the stiffness of the mathematical system. Cases where one or more nodal pressure heads are not acceptable need to be treated manually by the designer in various ways as discussed by the authors. For branching networks, Jones (2011) showed that by restricting the focus to smooth-turbulent (turbulent flow in a smooth pipe) minor-lossless flow, and the use of a one-part, pipe-cost model, a simple nonlinear algebraic equation for each internal node in the distribution main could be developed. The development of this algorithm, as well as solution methodology, differs from that of Bhave (1978), which assumes constancy in several terms and thus requires iteration to solve. The Jones algorithm has been extended in the present work to include minor losses and rough pipe. When solved simultaneously with the energy equation for each link, a unique solution for all link diameters and nodal pressure head values is obtained which produces minimum network cost, as opposed to the distribution main cost as in Swamee and Sharma (2000). The

method of Jones also applies to serial and loop networks because of its generality.

Heuristic methods follow specific rules to incrementally build better solutions, although the rules are not strictly formulated to trend towards local or global optima. An approach by Monbaliu et al. (1990) sets all network pipes to their minimum size, where the pipe that has a maximum head loss gradient is incremented to its next-highest size until all nodal head requirements are satisfied. Similarly, an algorithm by Keedwell and Khu (2006) selects an initial solution and iteratively responds to nodal head deficits and surpluses by incrementing or decrementing pipe sizes accordingly until a feasible solution is found. Suribabu (2012) proposed a heuristic that identifies pipes to increment or decrement in size based on flow velocity and alternative metrics such as proximity to the source node, achieving acceptable cost results with computational efficiency. While these algorithms are typically computationally efficient, they do not guarantee a global optimum.

Metaheuristic optimization methods allow for a set of solutions to evolve through random processes that are guided with an objective function which rewards low network costs and penalizes hydraulic insufficiencies. Examples include evolutionary algorithms, which are most commonly genetic algorithms (Krapivka and Ostfeld, 2009; Simpson et al., 1994; Kadu et al., 2008; Prasad and Park, 2004), simulated annealing (Vasan and Simonovic, 2010; Tospornsampan et al., 2007), ant colony optimization (Maier et al., 2003), and differential evolution (Vasan and Simonovic, 2010). As reviewed by Nicklow et al. (2010), evolutionary algorithms are an emerging popular alternative to the deterministic methods, and they offer the opportunity to accommodate unique constraints and multiple design objectives. The main challenges for evolutionary algorithms are the difficulty of incorporating constraints into objective functions, the optimum selection of parameters, and a relatively large amount of computational effort. In addition to optimizing for cost, multi-objective methods, often based on evolutionary algorithms, allow the designer to choose from a Pareto optimal front of objectives, such as cost and reliability (Prasad and Park, 2004). In addition to water network design, metaheuristic algorithms have been used for a range of problems in water resources engineering, such as rainfall and runoff modeling (Taormina and Chau, 2015).

Decomposition methods involve the partitioning of networks into smaller sub-networks which are each optimized using one of many types of techniques and then combined into an overall solution. In some cases, the loops in the sub-networks are removed, producing branching trees which are then optimized individually. Techniques used to optimize the sub-networks can involve multiple methods, including linear programming (Saldarriaga et al., 2013) and differential evolution (Zheng et al., 2013), with a later stage optimizing the network as a whole using the sub-network solutions as inputs. Note that another distinct use of the term “decompo-

sition” refers to the approach of iteratively solving “inner” and “outer” mathematical problem formulations, and has been used in the literature by Krapivka and Ostfeld (2009) who traces its use in this context back to Alperovits and Shamir (1977).

In the present study, we present three algorithms, each from one of three major categories of methods applied to the cost optimization of water distribution networks, and compare their performance on five cases adapted from real GDWNs. These algorithms include (1) the calculus-based (CB) optimization model of Jones (2011), an NLP method; (2) backtracking (BT), a partial enumeration method; and (3) a genetic algorithm (GA), a metaheuristic method. Major distinguishing features of these algorithms include their working use of continuous diameters (CB) versus discrete diameters (BT and GA), their deterministic (CB and BT) versus stochastic (GA) search process, and their relative scalability as better (CB, GA) and worse (BT) for larger networks. In terms of their ability to find a global optimum solution for the problem formulation, CB finds a global optimum for continuous diameters but cannot guarantee a discrete diameter global optimum in its mapped solution, BT can guarantee a discrete global optimum, and GA cannot guarantee an optimum. For a direct comparison of techniques, the pipe costs used for all algorithms are found by interpolating a two-part cost formula based on a curve fit of real cost data for available diameter values. The three algorithms are tested against networks adapted from field data on five actual GDWNs installed in Panama, Nicaragua, and the Philippines.

Within the broader context of water network problem formulations, this paper is concerned with finding cost-optimal single-diameter solutions to branching water distribution networks with steady-state demand flows and pre-specified pipe locations. By implication of being gravity-driven, the problem does not involve the use of pumping stations. This problem formulation is directly applicable to typical gravity-driven water networks, and is also useful for multi-objective algorithms, the consideration of sub-networks in a decomposition technique, pumped networks, and looped system optimization, which can involve reformulating the problem into a branching configuration.

The results of this study highlight the advantages and weaknesses of each GDWN design method including closeness to the global optimum, the ability to prune the solution space of infeasible and suboptimal candidates without missing the global optimum, and also computational time. We present two pre-processors which discrete-diameter search methods can use to reduce the search space without pruning the global optimum. To the authors’ knowledge, this is the first implementation of “pre-processor 1” in enumeration methods and the first implementation of “pre-processor 2” in any water network design method. We also extend the Jones closed-form model to include minor losses, a more comprehensive two-part cost model, which realistically applies to pipe sizes that span a broad range typical of GDWNs of inter-

est in this work, and for smooth and commercial steel roughness values.

## 2 Problem formulation

Branching networks are considered (Fig. 1), where all branches connect a distribution main node with a delivery node, shown as tap stands or houses. For each link in a network of  $N_L$  links, pipe length ( $L$ ) and the net elevation change ( $\Delta z$ ) are considered known and fixed. Steady-state flow rates ( $Q$ ) are prescribed for each link based on the demand flow data at delivery nodes. As noted above, demand flows are determined by community surveys and extrapolated in time to quantitatively account for population growth. Minor losses are accounted for through a minor loss coefficient  $K$  or a dimensionless equivalent pipe length, ( $L_e/D$ , or in symbol form,  $L_{\text{ebyD}}$ ), where  $L_e$  is the pipe length of diameter  $D$  whose frictional loss results in the corresponding minor loss. An optimal solution is obtained by selecting pipe diameters ( $D$ ) from a set of commercially available diameters such that the network's material cost is minimized. With  $N_D$  choices of diameters for  $N_L$  links, the problem has  $N_D^{N_L}$  candidate solutions.

For all nodes, pressure head,  $h$ , is greater than or equal to a chosen minimum,  $h_{\text{min}}$ . The value for  $h_{\text{min}}$  is selected to eliminate possible leakage of contaminated ground water into the network should the operating conditions change in an unanticipated way. The change in pressure head,  $\Delta h$ , across each link is calculated with the energy equation for pipe flow as follows:

$$\Delta h = -\Delta z + \left( \alpha + K + f \left( \frac{L}{D} + L_{\text{ebyD}} \right) \right) \frac{8Q^2}{\pi^2 g D^4}, \quad (1)$$

where for each link,  $\alpha$  is the kinetic energy correction factor and  $f$  is the Darcy friction factor, calculated with the Colebrook–White equation (Colebrook and White, 1937) or Churchill correlation (Churchill, 1977), and  $g$  is acceleration of gravity. The kinetic energy correction factor,  $\alpha$ , is considered only in the first link, where acceleration from a zero-velocity source is sometimes non-negligible for the smallest of GDWNs that have been encountered. Thus,

$$\alpha = \begin{cases} 2 & Re \leq 2300 \\ 1.05 & Re > 2300 \end{cases}$$

where  $Re$  is the Reynolds number for pipe flow,  $4Q/\pi\nu D$ , and  $\nu$  is the kinematic viscosity of water. The possibility of laminar flow ( $Re \leq 2300$ ) is permitted since branches from the smallest GDWN observed in practice have been in this regime.

The pressure upper bound is not incorporated into the optimization process. Worst-case pressure conditions occur under hydrostatic conditions, which are directly related to the maximum elevation change in the network and where no flow occurs. Therefore, before the optimization process is undertaken, the selections of appropriate pressure ratings for the

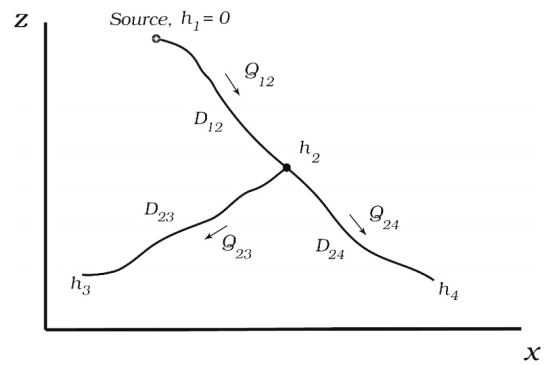


Figure 2. Three-pipe branch network.

pipe and, if needed, break-pressure tanks are left to the correct judgment of the designer under no-flow conditions. In addition, precautions against water hammer are left to the designer.

## 3 New calculus-based algorithm

In this section we develop a new calculus-based algorithm for pipe diameters that minimize overall pipe cost for the network. First appearing in Jones (2011), this algorithm is solved simultaneously with the energy equation for each link to produce unique solutions for  $D$  and nodal pressure head values that minimize network pipe cost, as opposed to only the distribution main cost as in Swamee and Sharma (2000). The method also applies to serial and loop networks but the focus for the present work is on branching networks.

We assume continuous pipe diameters in this section; values that result from the solution of the energy equation. Mapping between continuous diameters and the discrete nominal sizes, required to complete the design, will be addressed below.

Consider the three-pipe network shown in Fig. 2. Pipes 1–2, 2–3, and 2–4 meet where head  $h_2$  is unknown. Each pipe has a prescribed volume flow rate and length and unknown diameter  $D$  as shown. The change in elevation between the top and bottom of each pipe is  $\Delta z$ , and  $\Delta h$  is the change in pressure head. There is a prescribed head at each outlet for pipes 2–3 and 2–4.

To facilitate insight, we at first assume turbulent flow (which can be verified post-calculation if necessary) in smooth pipe and that minor losses are negligible. Two sources for the friction factor for smooth-turbulent flow are considered, namely the classical Blasius equation (reported in Streeter et al., 1998),  $f = 0.316Re^{-1/4}$ , and the Swamee–Jain correlation (Swamee and Jain, 1976),  $f = 0.175Re^{-0.1923}$  (though not explicitly appearing in this reference,  $f$  from the Swamee–Jain correlation is obtained by writing it for smooth pipe and comparing this with the energy equation, where  $f$  is assumed to be in the form  $aRe^n$ ).

The Blasius equation has higher accuracy (2% for low  $Re$  and 3% for high  $Re$ ) in the range  $10^4 < Re < 10^5$ , over which most of the GDWNs in this work operate, compared with the Swamee–Jain correlation of +8%/–3%, thus the Blasius equation is chosen for this work. A combination of the Blasius equation with the energy equation gives explicit formulas for  $D$  for the three links in Fig. 2. For simplicity, and to reduce the number of free parameters, the conditions for pipes 2–3 and 2–4 are assumed to be identical without loss of generality. We therefore obtain

$$D_{12} = 0.741 \left( \frac{\Delta z_{12} + \Delta h_{12}}{L_1} \right)^{-4/19} \left( \frac{Q_{12} v^{1/7}}{g^{4/7}} \right)^{7/19} \quad (2)$$

$$D_{23} = D_{24} = 0.741 \left( \frac{\Delta z_{23} + \Delta h_{23}}{L_2} \right)^{-4/19} \left( \frac{Q_{23} v^{1/7}}{g^{4/7}} \right)^{7/19}.$$

With our assumptions and inspection of Fig. 2,  $\Delta h_{12} = -h_2$  and  $\Delta h_{23} = \Delta h_{24} = h_2 - h_3 = h_2 - h_4$ , we furthermore obtain

$$D_{12} = 0.741 \left( \frac{\Delta z_{12} - h_2}{L_1} \right)^{-4/19} \left( \frac{Q_{12} v^{1/7}}{g^{4/7}} \right)^{7/19} \quad (3)$$

$$D_{23} = D_{24} = 0.741 \left( \frac{\Delta z_{23} - h_3 + h_2}{L_2} \right)^{-4/19} \left( \frac{Q_{23} v^{1/7}}{g^{4/7}} \right)^{7/19}.$$

The single-part pipe-cost model can be assumed to follow a power-law relationship (Swamee and Sharma, 2008)

$$C = a \left( \frac{D}{D_u} \right)^b, \quad (4)$$

where  $C$  is cost per unit length of pipe,  $a$  is a constant coefficient,  $b$  is a constant exponent, and  $D_u$  an assumed unit diameter. A more robust, two-part model, valid for a greater range of pipe sizes than that of Swamee and Sharma (2008), will be used below. The use of pipe material cost as the objective function was assumed because of relevance. In most GDWNs of interest in this work, installation labor comes from the local community and has no well-defined associated cost, such that the material cost for the network is of prime importance. For a more general case, the economics of a GDWN may be more encompassing and include materials, labor, operation and maintenance, depreciation, taxes, and salvage, among others. The time value of money may also need to be considered, which includes interest rates and estimation of the network lifetime.

With Eq. (4) the general expression for the total cost for the pipe material,  $C_T$ , is obtained by summing over all links  $ij$ ,

$$C_T = a \sum_{ij} L_{ij} \left( \frac{D_{ij}}{D_u} \right)^b, \quad (5)$$

which, for the present problem, becomes

$$C_T = a \left[ L_{12} \left( \frac{D_{12}}{D_u} \right)^b + L_{23} \left( \frac{D_{23}}{D_u} \right)^b + L_{24} \left( \frac{D_{24}}{D_u} \right)^b \right]$$

$$= a \left[ L_{12} \left( \frac{D_{12}}{D_u} \right)^b + 2L_{23} \left( \frac{D_{23}}{D_u} \right)^b \right]. \quad (6)$$

The mathematical basis for a unique solution for  $h_2$  with cost minimization is now presented. In addition to the fixed pipe lengths, the total cost depends on the diameters for all pipes in the network. For the case of Fig. 2, where we now allow pipe 2–3 and pipe 2–4 to be different, we get

$$C_T = C_T(D_{12}(h_2), D_{23}(h_2), D_{24}(h_2)). \quad (7)$$

Using the chain rule from the calculus, the total differential of Eq. (7) is

$$dC_T = \frac{\partial C_T}{\partial D_{12}} \frac{\partial D_{12}}{\partial h_2} dh_2 + \frac{\partial C_T}{\partial D_{23}} \frac{\partial D_{23}}{\partial h_2} dh_2 + \frac{\partial C_T}{\partial D_{24}} \frac{\partial D_{24}}{\partial h_2} dh_2. \quad (8)$$

The minimum value of  $C_T$  is found once  $dC_T = 0$  (and once it is verified that the second derivative of  $C_T$  is positive thus indicating that  $C_T$  is indeed a minimum). Requiring this, we obtain

$$0 = \frac{\partial C_T}{\partial D_{12}} \frac{\partial D_{12}}{\partial h_2} + \frac{\partial C_T}{\partial D_{23}} \frac{\partial D_{23}}{\partial h_2} + \frac{\partial C_T}{\partial D_{24}} \frac{\partial D_{24}}{\partial h_2}. \quad (9)$$

The cost  $C_T$  is from Eq. (5), so the derivatives like  $\partial C_T / \partial D_{12}$  in Eq. (9) are written in general as

$$\frac{\partial C_T}{\partial D_{ij}} = ab \frac{D_{ij}^{b-1}}{D_u^b} L_{ij} \quad (10)$$

for any link  $ij$ .

The derivatives like  $\partial D_{12} / \partial h_2$  in Eq. (9) are obtained by taking the partial derivative of the pipe diameter with respect to the relevant pressure head in the appropriate energy equation. For the full energy equation, where  $D$  appears in a non-linear way in more than one location, this would be done using numerical methods. However, if we assume minor-lossless, smooth-turbulent flow as noted above, we can use the energy equations like Eq. (3). We therefore obtain the following for pipe 1–2:

$$\frac{\partial D_{12}}{\partial h_2} = 0.156 \left( \frac{\Delta z_{12} - h_2}{L_{12}} \right)^{-\frac{23}{19}} \left( \frac{v^{1/7} Q_{12}}{g^{4/7} L_{12}^{19/7}} \right)^{\frac{7}{19}}; \quad (11)$$

for pipe 2–3, we get

$$\frac{\partial D_{23}}{\partial h_2} = -0.156 \left( \frac{\Delta z_{23} + h_2 - h_3}{L_{23}} \right)^{-\frac{23}{19}} \left( \frac{v^{1/7} Q_{23}}{g^{4/7} L_{23}^{19/7}} \right)^{\frac{7}{19}}; \quad (12)$$

and for pipe 2–4,

$$\frac{\partial D_{24}}{\partial h_2} = -0.156 \left( \frac{\Delta z_{24} + h_2 - h_4}{L_{24}} \right)^{-\frac{23}{19}} \left( \frac{v^{1/7} Q_{24}}{g^{4/7} L_{24}^{19/7}} \right)^{\frac{7}{19}}. \quad (13)$$

Equations (10)–(13) are combined with Eq. (9) to produce a single algebraic equation that depends on  $h_2$ , as well as  $D_{12}$ ,  $D_{23}$ , and  $D_{24}$ . Introducing  $D_{12}$ ,  $D_{23}$ , and  $D_{24}$  from Eq. (3) into this algebraic equation, we get

$$0 = Q_{12}^{7b/19} \left( \frac{\Delta z_{12} - h_2}{L_{12}} \right)^{-(1+4b/19)} - Q_{23}^{7b/19} \left( \frac{\Delta z_{23} + h_2 - h_3}{L_{23}} \right)^{-(1+4b/19)} - Q_{24}^{7b/19} \left( \frac{\Delta z_{24} + h_2 - h_4}{L_{24}} \right)^{-(1+4b/19)}. \quad (14)$$

The general form of Eq. (14), written at any internal node is

$$0 = \sum_{ij,in} Q_{ij}^{7b/19} S_{ij}^{-(1+4b/19)} - \sum_{ij,out} Q_{ij}^{7b/19} S_{ij}^{-(1+4b/19)}, \quad (15)$$

where the hydraulic gradient,  $S_{ij}$ , is

$$S_{ij} = \frac{\Delta z_{ij} + \Delta h_{ij}}{L_{ij}}. \quad (16)$$

In Eq. (15) the indices  $ij,in$  and  $ij,out$  on the summations refer to inflows and outflows at the node (e.g., in Fig. 2,  $ij,in = 12$  and  $ij,out = 23$  and 24). Equation (15), the new CB algorithm proposed in this work, is written for each internal node in the network and solved simultaneously with the energy equation for each link to obtain unique and optimal values of  $D_{ij}$  for all links and  $h_j$  for all internal nodes. It is understood that the nodal pressure heads determined from the solution of this system must be greater than or equal to the  $h_{min}$  prescribed for the network. For nodes that do not satisfy this condition, the pressure head is set equal to  $h_{min}$ , as part of the CB algorithm. Thus,  $h_j \geq h_{min}$ .

Minor losses using the equivalent-length method can be included in the above developments by artificially extending the length of the link by  $L_e$  in which minor loss occurs, thus contributing a non-zero  $L_{eByD}$  term in Eq. (1). We also extend the cost model of Eq. (5) from Swamee and Sharma (2008) to encompass two different ranges of pipe diameters having two different coefficients  $a$  and exponents  $b$ . The link between the two ranges starts at discrete pipe size  $D_{co}$ , at and below which the cost model for the small (subscript  $s$ ) pipe sizes applies, and discrete pipe size  $D_{co+1}$ , at and above which the cost model for the large (subscript  $l$ ) pipe sizes applies. The cutoff diameter,  $D_{co}$  is chosen by the designer based on inspection of cost vs. diameter data. Thus,

$$C_{ij} = L_{ij} \begin{cases} a_s \left( \frac{D_{ij}}{D_u} \right)^{b_s}, & D_{ij} \leq D_{co} \\ c_1 + c_2 \frac{D_{ij}}{D_u} + c_3 \left( \frac{D_{ij}}{D_u} \right)^2 + c_4 \left( \frac{D_{ij}}{D_u} \right)^3, & D_{co} < D_{ij} < D_{co+1} \\ a_l \left( \frac{D_{ij}}{D_u} \right)^{b_l}, & D_{ij} \geq D_{co+1}. \end{cases} \quad (17)$$

In Eq. (17),  $a_s$  and  $a_l$  are the coefficients for the small (designated by subscript  $s$ ) and large (subscript  $l$ ) pipe size regions, respectively, and  $b_s$  and  $b_l$  are the exponents for the small and large pipe size regions, respectively. A cubic spline is fit between pipe sizes  $D_{co}$  and  $D_{co+1}$  to complete the transition between small and large pipe sizes. The coefficients of this polynomial are  $c_1$ ,  $c_2$ ,  $c_3$ , and  $c_4$  as seen in Eq. (17). These coefficients are evaluated by matching the cubic polynomial and pipe data at  $D_{co}$  and  $D_{co+1}$  and the first derivative of the polynomial with respect to  $D_{ij}/D_u$  to  $a_s b_s \left( \frac{D_{co}}{D_u} \right)^{b_s-1}$

at  $D_{ij} = D_{co}$  and to  $a_l b_l \left( \frac{D_{co+1}}{D_u} \right)^{b_l-1}$  at  $D_{ij} = D_{co+1}$ . An example of data for polyvinyl chloride (PVC) pipe and the curve fit is shown in Fig. 3. The results of the curve fit are as follows:  $D_{co} = 2.067$  in.,  $D_{co+1} = 2.469$  in.,  $a_s = \text{USD } 1.349 \text{ m}^{-1}$ ,  $b_s = 1.157$ ,  $a_l = \text{USD } 1.381 \text{ m}^{-1}$ ,  $b_l = 1.344$ ,  $c_1 = \text{USD } 237.516 \text{ m}^{-1}$ ,  $c_2 = \text{USD } 316.125 \text{ m}^{-1}$ ,  $c_3 = \text{USD } 140.450 \text{ m}^{-1}$ ,  $c_4 = -\text{USD } 20.499 \text{ m}^{-1}$ . It is clear from inspection of Fig. 3 that a one-part cost model would not have produced an acceptable curve-fit to pipe-cost data.

With the inclusion of the two-part cost model and minor loss term, Eq. (15) becomes

$$0 = \sum_{ij,in} \frac{C'_{ij} A_{ij}^{\frac{4}{19}} (1 + \epsilon_{ij})^{\frac{4}{19}} S_{ij}^{-\frac{23}{19}} \left( \frac{Q_{ij}^7 v}{g^4 D_u^{19}} \right)^{\frac{1}{19}}}{1 - B A_{ij}^{\frac{4}{19}} \epsilon'_{ij} (1 + \epsilon_{ij})^{-\frac{15}{19}} S_{ij}^{-\frac{4}{19}} \left( \frac{Q_{ij}^7 v}{g^4 D_u^{19}} \right)^{\frac{1}{19}}} - \sum_{ij,out} \frac{C'_{ij} A_{ij}^{\frac{4}{19}} (1 + \epsilon_{ij})^{\frac{4}{19}} S_{ij}^{-\frac{23}{19}} \left( \frac{Q_{ij}^7 v}{g^4 D_u^{19}} \right)^{\frac{1}{19}}}{1 - B A_{ij}^{\frac{4}{19}} \epsilon'_{ij} (1 + \epsilon_{ij})^{-\frac{15}{19}} S_{ij}^{-\frac{4}{19}} \left( \frac{Q_{ij}^7 v}{g^4 D_u^{19}} \right)^{\frac{1}{19}}}, \quad (18)$$

where  $B = 0.1989$  and

$$\epsilon_{ij} = \sum_k \left( \frac{L_e}{D} \right)_{k,ij} \frac{D_{ij}}{L_{ij}} \quad (19)$$

$$\epsilon'_{ij} = \sum_k \left( \frac{L_e}{D} \right)_{k,ij} \frac{D_u}{L_{ij}}$$

$$A_{ij} = \begin{cases} 0.318, & \text{smooth pipe} \\ 0.420, & \text{steel pipe} \end{cases} \quad (20)$$

and  $A$  accounts for the effect of pipe roughness (smooth and commercial steel). The term  $C'_{ij}$  is the derivative of the cost function per unit length with respect to  $D/D_u$ . For the two-part cost model from above, we obtain

$$C'_{ij} = \begin{cases} a_s b_s \left( \frac{D_{ij}}{D_u} \right)^{b_s-1}, & D_{ij} \leq D_{co} \\ c_2 + 2c_3 \left( \frac{D_{ij}}{D_u} \right) + 3c_4 \left( \frac{D_{ij}}{D_u} \right)^2, & D_{co} < D_{ij} < D_{co+1} \\ a_l b_l \left( \frac{D_{ij}}{D_u} \right)^{b_l-1}, & D_{ij} \geq D_{co+1}. \end{cases} \quad (21)$$

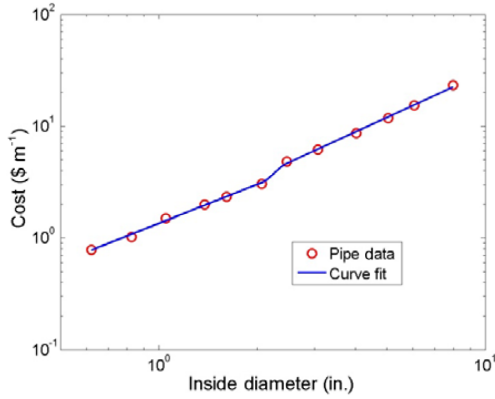


Figure 3. PVC pipe cost from 2011 data.

Equation (18), and its simpler form Eq. (15) for minor-lossless flow and a single-part pipe-cost model (it is easy to show that Eq. (18) regresses to Eq. (15) for these conditions), is the root of the calculus-based optimization in this work and is applied at all internal nodes to uniquely determine  $h_j$ . Equation (18) is valid over the range of  $\sim 4000 < Re < \sim 300\,000$ . Algorithms to solve a general set of independent, nonlinear algebraic equations using, for example, the Levenberg–Marquardt, quasi-Newton, Newton–Raphson, or conjugate gradient methods are available in most commercial math packages including Matlab (1 Apple Hill Drive, Natick, MA USA 01760) and Mathcad (<http://www.ptc.com> 31 January 2018). We used the package Mathcad in the present work. Thus, compared with an iterative solution procedure, a solution flowchart is not relevant here.

Bhave (1978) first proposed an algorithm like Eq. (15) using slightly different notation than here. For clarity, we represent Eq. (15) using Bhave’s notation as

$$0 = \sum Q_{ij}^{7b/19} S_{ij}^{-(1+4b/19)} - \sum Q_{jk}^{7b/19} S_{jk}^{-(1+4b/19)},$$

where the  $ij$  and  $jk$  notation are shown in Fig. 4. Index  $j$  spans all internal nodes along the distribution main. A quantitative comparison between Eq. (18) and the method of Bhave is presented below.

#### 4 Backtracking algorithm and genetic algorithm

Backtracking (BT) and genetic algorithm (GA) assess candidate solutions composed of discrete diameters from a commercially available set. These candidates are represented by a vector of  $N_L$  elements where each element corresponds to a commercially available diameter of a network link. To reduce the computational time associated with these evaluations, the constraints imposed by the energy equation and cost minimization may be more efficiently evaluated through lookup tables. With fixed  $L$ ,  $\Delta z$ ,  $K$ ,  $L_{\text{ebyD}}$ , and  $\alpha$ , the change in pressure head  $\Delta h$  is evaluated for all  $N_D \times N_L$  combinations

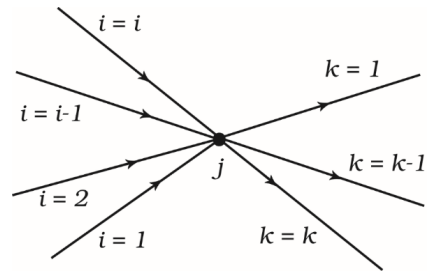


Figure 4. Bhave (1978) index notation at an internal node,  $j$ .

of pipe diameter and link index:

$$\Delta \mathbf{h} = \begin{bmatrix} \Delta h_{11} & \cdots & \Delta h_{1N_L} \\ \vdots & \ddots & \vdots \\ \Delta h_{N_D 1} & \cdots & \Delta h_{N_D N_L} \end{bmatrix}. \quad (22)$$

While an algorithm evaluates a candidate solution, the pressure head at each node is sequentially calculated by “marching” through the network. Starting with the fixed source pressure head, the algorithm finds the pressure head  $h_i$  for a given node by adding the head at the upstream node,  $h_{i-1}$  to the change in head for that link  $i_L$  and the diameter  $i_D$  under consideration. Thus,

$$h_i = h_{i-1} + \Delta \mathbf{h}(i_D, i_L). \quad (23)$$

Along with the hydraulic evaluation of a candidate solution, the cost of the partial candidate is found through the use of a lookup table  $\mathbf{C}$ :

$$\mathbf{C} = \begin{bmatrix} C_{11} & \cdots & C_{1N_L} \\ \vdots & \ddots & \vdots \\ C_{N_D 1} & \cdots & C_{N_D N_L} \end{bmatrix}, \quad (24)$$

where  $\mathbf{C}(i_D i_L)$  returns the additional cost of assigning a diameter with index  $i_D$  to link  $i_L$ . In this way, the candidate solution’s hydraulic performance and cost are incorporated into the genetic algorithm and backtracking approaches. In contrast to GA, the backtracking algorithm evaluates pressure head and cost upon consideration of each partial candidate, where GA calculates these values on full candidates as part of the objective function.

##### 4.1 BT and GA pre-processor 1: maximum available diameter

To increase the efficiency of BT and GA, it is advantageous to limit the number of pipe diameters in the available set, especially those outside of the range of the optimal solution. For the BT algorithm in particular, larger diameters can require considerable computational effort, since they tend not to violate static head requirements and require multiple-link partial candidates for the algorithm to reject them once

their cost exceeds that of an already-found viable candidate. Therefore, a pre-processor is used to provide a maximum diameter ( $D_{\max}$ ) that should be considered during the optimization process. This procedure, which produces a conservative estimate, finds the smallest diameter at which a network with a single pipe diameter choice produces no nodes with a pressure head below  $h_{\min}$ , similar to the technique used by Mohan and Jinesh Babu (2009). After this diameter is found, the next larger diameter in the set is selected as  $D_{\max}$  to allow the algorithm to select a larger than necessary diameter if this is able to save cost elsewhere. It worth noting that Kadu et al. (2008) presents another method to further prune the search space with the critical path concept, where Dongre and Gupta (2011) noted the computational advantages of having just four diameter choices per link. This method, however, may prune the global optimum and may not produce feasible head values at intermediate nodes, as in the case of networks with a local high point.

#### 4.2 BT and GA pre-processor 2: adjusted minimum pressure head

A second pre-processor adjusts the minimum pressure head requirement for each internal node by considering the total head required at downstream nodes. It can be recognized that, without the use of a pump, the total head cannot increase at nodes downstream of a given node  $i$ . Furthermore, the total head must decline at a minimum grade that is determined by the demand volume flow rate and the largest pipe diameter available ( $D_{\max}$ ) for selection. This energy constraint is utilized to reduce the number of candidates to be considered by increasing the minimum pressure head at nodes where these rules produce a higher minimum head than the original  $h_{\min}$ . For example, nodes upstream of a local network high point can have their minimum pressure head increased beyond the normal minimum, since the pressure head must be great enough to ensure adequate flow to the higher elevation downstream node. To begin this process, each node  $i$  is initialized with a baseline minimum total head:

$$th_{\min,i} = z_i + h_{\min}. \quad (25)$$

$th_{\min,i}$  is thus initialized by considering only the node's hydraulic requirements in isolation, i.e., without acknowledging the neighboring downstream nodes. The pre-processor then considers updating  $th_{\min,i}$  by checking the following condition, which is false when the minimum pressure head at downstream nodes produces further constraints on an upstream node  $i$ . Thus, for all nodes  $i$  which are upstream of some node  $j$ , the following inequality can be evaluated:

$$th_{\min,i} - th_{\min,j} \geq \left( \alpha_{i-j} + K_{i-j} + f_{i-j} \left( \frac{L_{i-j}}{D_{i-j}} + L_{\text{eby}D_{i-j}} \right) \right) \frac{8Q_{i-j}^2}{\pi^2 g D_{\max}^4}. \quad (26)$$

Also, consider that when flow rate  $Q_{i-j}$  is small and  $D_{\max}$  is large, the right-hand side of Eq. (26) approaches zero, representing the simple statement that upstream total head must always be greater than downstream total head. When the condition in Eq. (26) is false, the minimum total head can be updated in node  $i$  such that the maximum diameter size in link  $i-j$  is able to meet the downstream node's minimum total head, or

$$th_{\min,i} = th_{\min,j} + \left( \alpha_{i-j} + K_{i-j} + f_{i-j} \left( \frac{L_{i-j}}{D_{i-j}} + L_{\text{eby}D_{i-j}} \right) \right) \frac{8Q_{i-j}^2}{\pi^2 g D_{\max}^4}. \quad (27)$$

In this way,  $th_{\min,i}$  may be updated for each node until the condition in Eq. (26) is true for all nodes  $i$  with a downstream node  $j$  connected by a single link.

After the values for  $th_{\min,i}$  are updated, they are converted back into minimum pressure head values by subtracting the elevation  $z_i$  from  $th_{\min,i}$ . This pre-processor serves to narrow the search for viable candidate solutions by potentially increasing the minimum pressure head. Since backtracking and GA consider network links in the downstream direction, these algorithms are otherwise blind to future downstream pressure head requirements. This limitation is alleviated by the pre-processor, which allows these algorithms some implicit information about what local diameter choices will be viable for the full network solution. Note that both pre-processors discussed will not prune the global optimum from the solution.

#### 4.3 Backtracking algorithm (BT)

The backtracking algorithm is a partial enumeration method that employs a systematic search of candidate solutions to find a global optimum. The algorithm works recursively to incrementally build candidate solutions while checking the candidates for hydraulic and cost acceptability. The strength of the BT is that, upon discovery of an infeasible partial candidate, all extensions of that candidate can be eliminated from consideration. In this way, many solutions can be pruned from the solution tree to achieve greater computational efficiency.

Two backtracking methods in the literature are those by Gessler (1985) and González-Cebollada et al. (2011). The algorithm proposed by Gessler proposes a pipe-grouping strategy which speeds up the algorithm but risks pruning the global optimum. Additionally, pipe grouping represents its own optimization problem (Raad, 2011). The González-Cebollada algorithm does not include such pipe-grouping criteria, although it does halt its search after finding the first feasible solution, thus it does not guarantee a global optimum. The present study's BT algorithm, once run to completion, does guarantee a global optimum. It operates similarly to the method presented by González-Cebollada et al. (2011), with the major differences being that the algorithm contin-



ues searching once it has found its first feasible solution and uses pre-processors 1 and 2 to further reduce the search space. This implementation of BT, however, scales poorly with larger network sizes and would not be appropriate for use on large urban networks. Its appropriateness is shown here for many of the GDWNs encountered in practice, as evidenced by its use on real-world GDWN test cases in this paper. Moreover, it serves as a benchmark against which other algorithms can be compared.

BT uses two rejection criteria to discard candidate solutions from further consideration. The first rejection criterion is that when a candidate violates pressure head constraints, all candidates with equal or lesser diameter sizes can be discarded. This condition is leveraged even more effectively with pre-processor 2 above, which can increase pressure heads at individual nodes by anticipating the head requirements at surrounding nodes. The second rejection criterion is that once a feasible candidate has been found, all other partial candidates with a higher cost can also be discarded. The BT algorithm further extends this second criterion by considering that the links yet to be considered in a partial candidate, an “extension” to the partial candidate, will cost at a minimum that of the entire extension being composed of the smallest available diameter.

The backtracking algorithm begins its search of the solution tree by considering the partial candidate with the smallest diameter size assigned to the first network link. The pressure head and the partial candidate cost at the outlet node are calculated with the  $\mathbf{A}h$  and  $\mathbf{C}$  lookup tables. If this partial candidate meets pressure head and cost requirements, the algorithm extends this partial candidate by assigning the smallest diameter to the downstream link. If a partial candidate produces a node that is rejected on the basis of pressure head, the next largest larger diameter is chosen for the link upstream of the node. If no diameter satisfies the pressure head condition, the algorithm backtracks to the upstream link and assigns a larger diameter to the link. In this way, the algorithm continues to extend and reject candidate solutions until a full candidate satisfies the pressure head requirements. Once a working solution has been found, candidate solutions may be rejected based on cost. For each new candidate, cost is calculated by adding the cost of diameters that have already been assigned to the cost of assigning all downstream links with the smallest diameter available. If this cost exceeds the cost of the running optimum, the partial candidate is rejected. While the minimum pressure head criterion tends to prune candidates with diameters that are too small, the cost-based criterion tends to prune candidates of diameters that are too large.

#### 4.4 Modified backtracking algorithm (BT-NoUp)

A modification to the BT algorithm was made to further improve its computational speed, although at the risk of pruning the global optimum from the search. This modified algorithm

(BT-NoUp) rejects all candidates that feature a smaller diameter that is upstream of a larger diameter when an equal or smaller flow rate is present in the downstream link. Typically, a network designer would not consider such designs, and in cases where a single source feeds into a network with constant-length links, it is advantageous (or equivalent) to place larger diameters upstream of smaller diameters. However, due to the discrete nature of diameter choices and link lengths, an optimization problem may, in fact, have an optimal candidate with a larger diameter downstream from smaller ones. For this reason, the BT-NoUp algorithm, unlike the BT algorithm, may miss the global optimum at the expense of its greater computational efficiency.

#### 4.5 Genetic algorithm (GA)

Genetic algorithms are stochastic optimization techniques that mimic the process of natural selection, and numerous variations of GAs have demonstrated acceptable performance on WDN design (Nicklow et al., 2010). Given their popularity, the GA included in this study is meant to provide a point of comparison to the BT and CB algorithms when applied to GDWNs.

When implemented in water network design, each candidate solution represents a selection of pipe diameters. The algorithm is initialized with a population of candidates of size  $N_c$  that repeatedly undergoes the processes of mutation, crossover, and selection

$$c_i = [D_{1,i} \quad D_{2,i} \quad \dots \quad D_{N_L,i}], \quad (28)$$

where each candidate in the population  $c_i$  contains  $N_L$  diameters. In the present work, candidates are represented as a string of natural numbers, which is used over a binary representation to improve the ease of encoding (Vairavamoorthy and Ali, 2000). The mutation operator replaces pipe diameters with a diameter from a uniform random distribution, where each link diameter has a probability of  $p_{mut}$  of mutating on each generation. The crossover operator randomly pairs individuals in the population with probability  $p_{xover}$  and performs a single-point crossover of the two individuals, where the point of crossover is randomly chosen. The fitness,  $f_i$ , of each candidate is assessed with penalties associated with the solution’s pipe cost,  $C_{pipe,i}$ , and hydraulic cost,  $C_{hyd,i}$ , which is assigned when violations of the pressure head requirements occur:

$$f_i = \frac{1}{C_{pipe,i} + C_{hyd,i}}. \quad (29)$$

The hydraulic cost is found for each individual by identifying nodes in which the pressure head is less than  $h_{min}$  and multiplying the total amount of head violation by a hydraulic penalty coefficient,  $a_{hyd}$ :

$$C_{hyd,iC} = a_{hyd} \sum_1^{N_L} (h_{min} - h_{iN}) |h_{iN} < h_{min}. \quad (30)$$

To allow for a hydraulic penalty coefficient to produce similar results in both small-scale (inexpensive) network and a large-scale (more expensive) cases, the hydraulic penalty coefficient is made directly proportional to the average solution cost. With each generation,  $a_{\text{hyd}}$  is updated by multiplying the normalized penalty coefficient,  $a_{\text{hyd, norm}}$ , by the average pipe cost of the population,

$$a_{\text{hyd}} = a_{\text{hyd, norm}} \frac{\sum_1^{N_c} C_{\text{pipe}, iC}}{N_c}. \quad (31)$$

The algorithm then selects candidates to be carried into the next generation with a tournament selection method, where  $N_c$  groups of  $s$  individuals are randomly assigned and the fittest candidate among each group is selected, thus replacing the previous population with an equally sized population of  $N_c$  individuals.

In this study, the genetic algorithm parameters used were  $N_c = 200$ ,  $p_{\text{mut}} = 0.05$ ,  $p_{\text{xover}} = 1$ ,  $N_{\text{gen}} = 500$ ,  $a_{\text{hyd, norm}} = 0.05$ , and  $s = 10$ . These parameters were chosen by systematically varying parameter values until the optimum cost of a network, case 2, could no longer be significantly improved. The first four of these values are in line with typically used values from Simpson et al. (1994) of  $N_c$  (30–200),  $p_{\text{mut}}$  (0.01–0.05),  $p_{\text{xover}}$  (0.7–1.0), and  $N_{\text{gen}}$  (100–1000).

## 5 Cases studied

Six cases were studied based on actual GDWN in Panama, Nicaragua, and the Philippines. Global characteristics of each network are presented in Table 1 and the details of each network are presented in Table 4a–f. Each network is a branching type without loops. The total lengths of the networks range from less than 1 to over 15 km. Two serial networks are tested to demonstrate the effect of a local high point on the algorithm solutions. Elevation plots for each case are shown in Fig. 5.

The choice of  $h_{\text{min}}$  is not standardized, and should appropriately balance the risk of negative pressure in pipes and the increase in network cost due to the requirement of using larger diameters. The choice of  $h_{\text{min}}$  in GDWN design is typically in the range of 5–20 m (Arnalich, 2010; Bouman, 2014; Swamee and Sharma, 2008). In the present study  $h_{\text{min}} = 7$  m, although this requirement was reduced at selected nodes at the beginning of a network where changes in elevation are still small (case 2, where the pressure head at node 2 is relaxed to 2 m). At the source node, the pressure head is fixed at atmospheric pressure. All cases assumed minor-lossless flow, although all algorithms (e.g., Eq. 18 for CB-Theor) are capable of handling minor loss coefficients through the equivalent length method as described above. All algorithms were run in a late-version of MATLAB (or Mathcad for CB) on an Intel i5 processor at 2.50 GHz.

## 6 Mapping the theoretical $D$ to discrete pipe sizes

The mapping between continuous diameters and the discrete nominal pipe sizes was accomplished in our solution in one of the following ways:

1. For small and moderate size networks, the designer may manually adjust the pipe sizes (downward, normally one pipe size) starting from the first link downstream from the source and continuing along the rest of the distribution main to the end in a step-by-step manner. A nearby plot of the pressure heads compared with the theoretical  $D_{ij}$  from the CB approach (e.g., on the same Mathcad page for our solution) will highlight the acceptability or unacceptability of any change. This exercise also gives the designer valuable understanding of the sensitivity of the design to small changes in pipe sizes.
2. Based on the theoretical  $D_{ij}$  from the CB approach, a split-pipe can be created for each link. That is, the lengths for the two discrete pipes sizes that bound the theoretical  $D_{ij}$  from above and below are calculated such that the pressure drop between two consecutive nodes in the distribution main matches between the composite pipeline and the CB approach. This also provides discrete pipe sizes that nearly match the CB solution in terms of cost.

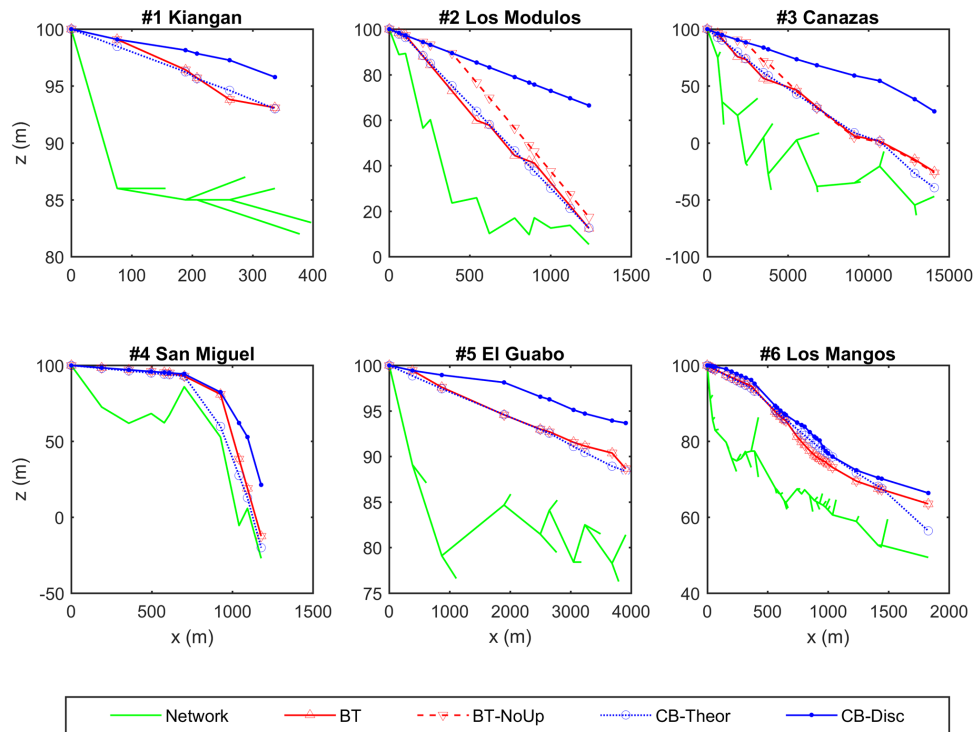
## 7 Results

The current study evaluated three types of algorithms that optimize the design of gravity-driven water networks (GDWN). The algorithms include the calculus-based (CB) algorithm (Eq. 18), a backtracking algorithm (BT) and its modified version (BT-NoUp), and a genetic algorithm (GA). The algorithms were applied to six test cases that are based on real GDWNs. Our results show that the CB, GA, and BT-NoUp algorithms could find solutions to the GDWNs within 25 % of the BT global optimum. All cases assume minor-lossless flow and a two-part pipe-cost model. Solution costs from each algorithm are shown in Table 2 and runtime statistics are shown in Table 3. BT could run to completion in < 1 min in all but the largest case (case 6 with 59 links), which did not complete after 7 days. As such, cost comparisons to BT are not made for case 6.

The CB algorithm based on Eq. (18), unlike the other algorithms in this work, finds a solution with theoretical diameters that are drawn from a continuous domain (CB-Theor). For all test cases, the costs of the CB-Theor solutions was less when compared with the BT discrete-diameter global optimum (5.5 to 2.6 % lower cost than BT). In fact, because of the discrete pipe sizes needed for an actual network, the continuous model will always produce the smallest theoretical network cost. The CB algorithm then maps this solution to a commercially-available discrete set (CB-Disc). The map-

**Table 1.** Characteristics of test cases.

Test case	Type	Number of diameter choices	Number of links	$Q_{tot}$ ( $L s^{-1}$ )	$L_{tot}$ (km)
(1) Kiangang, Philippines	Branching	8	9	4.37	0.82
(2) Los Modulos, Nicaragua	Serial	4	13	0.39	1.24
(3) Cañazas, Panama	Branching	10	23	6.29	15.2
(4) San Miguel, Nicaragua	Serial	9	10	0.40	1.18
(5) El Guabo, Nicaragua	Branching	12	17	17.7	4.71
(6) Los Mangos, Nicaragua	Branching	7	59	1.92	2.64

**Figure 5.** Network elevation ( $z$ ) and hydraulic grade lines (HGLs) of algorithm solution for main distribution links.

ping process used in this study simply mapped each theoretical diameter to the nearest available diameter of a larger size, thus producing a solution which still satisfies static head requirements but with a higher associated material cost. This tended to oversize the diameters, although the CB-Disc solutions were always within two diameters of the known global optimum solutions, as shown in Fig. 6. From all the combined test cases with known global optima, all but one (71 out of 72) of the diameter selections were within one diameter of the global optimum. More sophisticated mapping schemes, like independently adjusting  $D$  for each link in the distribution main in a step-by-step manner starting with the source while ensuring all pressure head constraints are satisfied, would be more likely to produce results identical to the global optimum (see Sect. 6). This was performed in the current study but the results are not presented because of space

constraints. The CB-Disc solution costs were, in all cases, larger than the global optimum, with costs ranging from 3.9 to 22.6 % above the global optimum. Thus, for all cases, the calculus-based algorithm bounded the cost of the global optima with a lower-cost CB-Theor solution and a higher-cost CB-Disc solution. This trend is a result of the additional constraints imposed by the finite set of diameter choices. If the algorithm is allowed a greater number of discrete diameter choices, i.e., through adding a less-common nominal diameter size to the available set, the cost of the CB-Disc solution would approach the CB-Theor solution. For all but case 6, the CB algorithm converged on a solution in 3 min or less.

BT-NoUp, a modified version of BT which does not consider smaller diameters upstream of large diameters, completed itself within 4 s for all cases, and found solutions which matched or came very close to the BT global opti-

**Table 2.** Solution costs for each algorithm.

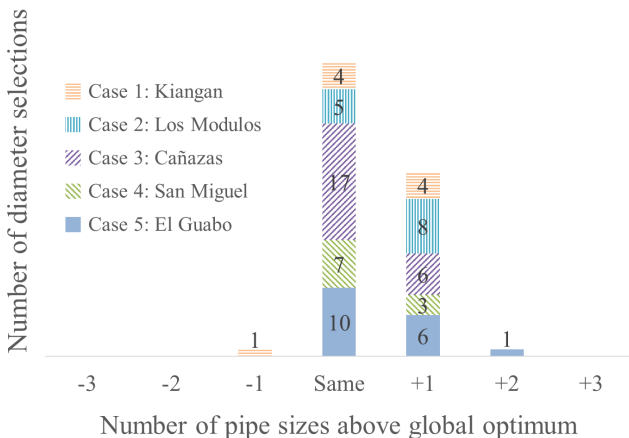
Case	Solution cost (USD)					Percentage cost increase over BT (global optimum)				Percentage cost increase over CB-Theor			
	BT	BT-NoUp	CB-Theor	CB-Disc	GA	BT-NoUp	CB-Theor	CB-Disc	GA	BT	BT-NoUp	CB-Disc	GA
(1) Kiangán, Philippines	2331	2331	2257	2594	2337	0	-3.2	11.3	0.3	3.3	3.3	14.9	3.5
(2) Los Modulos, Nicaragua	1441	1472	1404	1767	1445	2.1	-2.6	22.6	0.3	2.7	4.8	25.9	2.9
(3) Cañazas, Panama	72 190	72 443	68 245	84 441	73 964	0.4	-5.5	17.0	2.5	5.8	6.2	23.7	8.4
(4) San Miguel, Nicaragua	5418	5418	5172	5627	5422	0	-4.5	3.9	0.1	4.8	4.8	8.8	4.8
(5) El Guabo, Nicaragua	61 445	61 445	59 506	73 886	63 113	0	-3.2	20.2	2.7	3.3	3.3	24.2	6.1
(6) Los Mangos, Nicaragua	*	4082	3670	4405	4339	*	*	*	*	*	11.2	20.0	18.2

\* Note: BT did not complete after 7 days of runtime.

**Table 3.** Runtime and size of solution space for each algorithm.

Case	Runtime				Number of links	Number of diameter choices	Possible candidate solutions	Partial candidates considered	
	BT <sup>a</sup>	BT-NoUp <sup>a</sup>	CB	GA <sup>a</sup>				BT	BT-NoUp
(1) Kiangán, Philippines	0.2 s	0.05 s	<3 min	1 s	9	8	$1.3 \times 10^8$	269	126
(2) Los Modulos, Nicaragua	7 s	0.04 s	<3 min	2 s	13	4	$6.7 \times 10^7$	48 886	210
(3) Cañazas, Panama	40 s	0.1 s	<3 min	2 s	23	10	$1.0 \times 10^{23}$	433 210	2367
(4) San Miguel, Nicaragua	0.5 s	0.04 s	<3 min	2 s	10	9	$3.5 \times 10^9$	3671	244
(5) El Guabo, Nicaragua	0.5 s	0.05 s	<3 min	2 s	17	12	$2.2 \times 10^{18}$	3810	423
(6) Los Mangos, Nicaragua	>7 day <sup>b</sup>	2 s	94 min	5 s	59	7	$7.3 \times 10^{49}$	<sup>b</sup>	44 374

<sup>a</sup> BT, BT-NoUp, and GA algorithm run times do not include approximately 2 s of pre-processing time. <sup>b</sup> BT did not complete case 6 after 7 days of runtime.



**Figure 6.** Diameter sizes from calculus-based (CB-Disc) solutions compared with global optimum solutions (from backtracking, BT). A global optimum for case 6, Los Mangos, is not included since BT did not complete after 7 days of runtime.

mum. BT-NoUp missed the global optimum in cases 2 and 3, although by a small percentage increase in cost (2.1 and 0.4 % respectively). BT-NoUp, however, finished its search in a shorter amount of time in comparison to BT, a benefit that becomes relevant on problems with larger solution spaces, such as cases 3 ( $1.0 \times 10^{23}$  candidate solutions) and case 6 ( $7.3 \times 10^{49}$  candidate solutions).

GA was run on each case a total of 100 times, each run itself evolved 200 candidates for 500 generations. The lowest-cost candidate amongst the final population that did not violate the pressure head condition was chosen as the GA solution. Because GA is a stochastic search algorithm producing different results from run-to-run, the costs of the optima from all 100 runs were averaged, with this averaged value presented in Table 2. Overall, GA costs came close to the global optima (within 3 %) for cases 1–5 where the global optimum was known from BT. GA solution costs increased with larger network sizes, with its solution cost 18 % higher than CB-Theor for case 6, the largest case run. Each GA run finished consistently within 1–5 s, not including about 2 seconds of pre-processor time. We note that variations of GAs have been reported in the literature and several of these may improve upon the GA results obtained in this study. Potential improvements to the GA a self-adapting penalty function (Wu and Walski, 2005), the use of elitism to preserve the best solutions (Kadu et al., 2008), and a reduction in the search space (Kadu et al., 2008). One reported improvement, the scaling of the fitness function to magnify the rewards towards slightly fitter candidates at later generations (Dandy et al., 1996), was attempted for case 2 but did not result in a noticeable effect on performance.

To visually compare the algorithm solutions, the hydraulic grade lines from BT, BT-NoUp, CB-Theor, and CB-Disc are presented in Fig. 5 along with the network elevation for each test case. For clarity, the hydraulic grade lines of branch links are omitted from the figure. In addition, the GA solutions

**Table 4.** Case network properties, diameter (D) results (inch nominal sizes, with CB-Theor in inches), and nodal *h* (in meters) for (a) Case 1, Kiangon, (b) Case 2, Los Modulos, (c) Case 3, Cañazas, (d) Case 4, San Miguel, (e) Case 5, El Guabo, and (f) Los Mangos.

		(a)											
		Network	Link	1-2	2-3	3-4	4-5	5-6	2-7	3-8	4-9	5-10	
	Network	Length (m)	76	113	19	54	75	80	99	170	135		
		$Q$ (L s <sup>-1</sup> )	4.37	3.68	2.94	1.46	0.69	0.69	0.74	1.48	0.77		
		$\Delta z$ (m)	14.0	1.0	0.0	0.0	-1.0	0.0	-2.0	3.0	2.0		
		<i>D</i> solutions	BT	3	2-1/2	2	1-1/2	1-1/2	1	1-1/4	1-1/2	1-1/4	
	<i>D</i> solutions	BT-NoUp	3	2-1/2	2	1-1/2	1-1/2	1	1-1/4	1-1/2	1-1/4		
		CB-Theor	2.751	2.562	2.141	1.830	1.356	1.062	1.376	1.584	1.128		
		CB-Disc	3	3	2-1/2	2	1-1/4	1-1/4	1-1/4	1-1/2	1-1/4		
		<i>h</i> (m)	Node	1	2	3	4	5	6	7	8	9	10
	<i>h</i> (m)	BT	0	13.09	11.43	10.72	8.81	7.10	7.27	7.21	7.57	7.58	
		BT-NoUp	0	13.09	11.43	10.72	8.81	7.10	7.27	7.21	7.57	7.58	
		CB-Theor	0	12.48	11.24	10.65	9.61	7.00	6.99	7.00	7.00	3.19	
		CB-Disc	0	13.09	13.15	12.85	12.27	9.78	11.51	8.94	9.70	11.04	

		(b)														
		Network	Link	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10	10-11	11-12	12-13	13-14
	Network	Length (m)	60	41	108	46	134	153	79	157	90	32	102	120	117	
		$Q$ (L s <sup>-1</sup> )	0.39	0.39	0.39	0.39	0.39	0.39	0.39	0.39	0.39	0.39	0.39	0.39	0.39	
		$\Delta z$ (m)	11.2	-0.5	32.8	-3.7	36.6	-2.3	15.7	-6.8	7.3	-7.4	4.5	-1.2	8.4	
		<i>D</i> solutions	BT	1	1	3/4	3/4	3/4	3/4	1	3/4	1	1	3/4	3/4	3/4
	<i>D</i> solutions	BT-NoUp	1	1	1	1	1	3/4	3/4	3/4	3/4	3/4	3/4	3/4	3/4	
		CB-Theor	0.987	0.984	0.849	0.849	0.849	0.849	0.849	0.849	0.849	0.849	0.849	0.849	0.849	
		CB-Disc	1	1	1	1	1	1	1	1	1	1	1	1	1	
		<i>h</i> (m)	Node	1	2	3	4	5	6	7	8	9	10	11	12	13
	<i>h</i> (m)	BT	0	9.55	7.94	31.59	24.00	49.25	33.99	47.55	27.45	32.32	24.05	19.91	8.55	7.04
		BT-NoUp	0	9.55	7.94	37.82	32.88	65.85	50.59	59.60	39.50	39.18	29.07	24.93	13.56	12.05
		CB-Theor	0	9.00	7.00	31.86	24.78	51.53	37.98	47.87	29.53	30.21	20.46	17.46	7.44	7.23
		CB-Disc	0	9.55	7.94	37.82	32.88	65.85	59.41	72.98	61.93	66.80	58.53	60.27	55.83	61.06

		(c)																								
		Network	Link	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10	10-11	11-12	12-13	2-14	3-15	4-16	5-17	6-18	7-19	8-20	9-21	10-22	11-23	12-24
	Network	Length (m)	646	275	957	509	1102	291	1764	1256	2320	1580	2170	1217	160	100	1250	110	570	180	1400	50	400	260	100	
		$Q$ (L s <sup>-1</sup> )	6.29	5.49	5.39	5.34	5.14	2.84	2.74	2.49	2.39	0.69	0.39	0.20	0.80	0.10	0.05	0.20	2.30	0.10	0.25	0.10	1.70	0.30	0.19	
		$\Delta z$ (m)	25.0	38.9	11.9	42.1	-22.9	32.3	-29.9	40.8	-3.0	-14.7	34.1	-7.6	-5.0	20.0	-15.0	2.0	-12.0	14.0	-6.0	5.0	-1.0	-13.0	9.0	
		<i>D</i> solutions	BT	4	3	3	4	3	3	3	2-1/2	2-1/2	2	1-1/4	1	1-1/4	1/2	1/2	1/2	2	1/2	1	1/2	1-1/2	1-1/4	1/2
	<i>D</i> solutions	BT-NoUp	4	4	4	4	3	3	2-1/2	2-1/2	2-1/2	2	1-1/4	1	1-1/4	1/2	1/2	1/2	1-1/2	1/2	1	1/2	1-1/2	1-1/2	1/2	
		CB-Theor	3.530	3.531	3.333	3.307	3.270	2.727	2.698	2.579	2.548	1.862	1.227	1.011	1.283	0.325	0.508	0.404	1.678	0.343	0.963	0.281	1.405	1.401	0.488	
		CB-Disc	4	4	4	4	4	3	3	3	3	2	1-1/4	1	1-1/4	1/2	1/2	1/2	2	1/2	1	1/2	1-1/2	1-1/2	1/2	
		<i>h</i> (m)	Node	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
	<i>h</i> (m)	BT	0	21.1	55.4	51.5	91.3	51.7	82.5	43.9	69.8	41.4	22.1	40.0	21.7	12.1	72.2	24.3	82.2	26.1	90.8	20.1	73.3	21.9	7.81	39.7
		BT-NoUp	0	21.1	58.8	66.4	106	66.6	97.4	42.8	68.8	40.4	21.0	39.0	20.7	12.1	75.6	39.2	97.1	9.5	106	19.0	72.2	20.9	7.43	38.7
		CB-Theor	0	17.8	54.3	55.6	91.9	56.7	86.3	40.3	69.0	44.1	21.9	27.9	7.64	6.99	8.02	7.70	7.98	7.72	7.79	7.70	8.18	7.68	7.69	7.71
		CB-Disc	0	21.1	58.8	66.4	106	78.8	110	70.9	106	94.4	75.1	93.0	74.7	12.1	75.6	39.2	97.1	53.1	118	47.1	110	74.9	61.4	92.7

Table 4. Continued.

		(d)										
		Link	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10	10-11
Network	Length (m)	189	168	139	81	32	92	225	115	52.3	85	
	$Q$ (L s <sup>-1</sup> )	3.60	3.60	3.60	3.60	3.60	3.60	3.60	3.60	3.60	3.60	
	$\Delta z$ (m)	27.4	10.7	-6.4	6.1	-5.2	-18.6	33.2	58.2	-11.3	32.9	
D solutions	BT	3	3	3	3	3	2-1/2	2	1-1/4	1-1/4	1-1/4	
	BT-NoUp	3	3	3	3	3	2-1/2	2	1-1/4	1-1/4	1-1/4	
	CB-Theor	2.939	2.929	2.929	2.929	2.929	2.929	1.671	1.462	1.462	1.368	
	CB-Disc	3	3	3	3	3	3	2	1-1/2	1-1/2	1-1/4	
h (m)	Node	1	2	3	4	5	6	7	8	9	10	11
	BT	0	25.88	35.20	27.68	33.13	27.70	7.02	28.27	43.86	13.19	14.60
	BT-NoUp	0	25.88	35.20	27.68	33.13	27.70	7.02	28.27	43.86	13.19	14.60
	CB-Theor	0	25.53	34.51	26.72	32.01	26.51	7.00	6.99	32.93	6.96	7.02
CB-Disc	0	25.88	35.20	27.68	33.13	27.70	8.37	29.62	67.54	47.02	48.43	

		(e)																	
		Link	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10	2-11	3-12	4-13	5-14	6-15	7-16	8-17	9-18
Network	Length (m)	383	486	1030	600	150	400	187	450	227	230	240	110	270	130	130	260	110	
	$Q$ (L s <sup>-1</sup> )	17.72	14.68	12.76	11.96	10.04	7.72	6.60	3.12	1.20	3.04	1.92	0.80	1.92	2.32	1.12	3.48	1.92	
	$\Delta z$ (m)	10.9	10.0	-5.6	3.2	-2.6	5.7	-4.1	4.2	-3.1	2.0	2.5	-1.2	2.0	-1.1	0.0	1.0	2.0	
D solutions	BT	8	6	6	6	6	5	5	4	2	2-1/2	1-1/2	1-1/2	2	3	1-1/4	3	1-1/2	
	BT-NoUp	8	6	6	6	6	5	5	4	2	2-1/2	1-1/2	1-1/2	2	3	1-1/4	3	1-1/2	
	CB-Theor	6.875	6.408	6.144	6.008	5.691	4.800	4.576	3.494	2.649	2.364	1.608	1.529	1.932	3.250	1.395	3.076	1.647	
	CB-Disc	8	8	8	6	6	5	5	4	3	2-1/2	1-1/2	1-1/2	2	4	1-1/2	4	2	
h (m)	Node	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	BT	0	10.34	18.50	9.91	11.53	8.61	13.16	8.65	12.11	7.25	8.48	7.23	7.35	8.84	7.03	7.16	7.68	7.80
	BT-NoUp	0	10.34	18.50	9.91	11.53	8.61	13.16	8.65	12.11	7.25	8.48	7.23	7.35	8.84	7.03	7.16	7.68	7.80
	CB-Theor	0	9.76	18.35	9.93	11.49	8.46	12.70	7.94	10.67	7.00	7.00	7.00	7.00	7.01	7.00	7.00	7.00	7.01
CB-Disc	0	10.34	19.85	13.47	15.09	12.17	16.72	12.21	15.67	12.27	8.48	8.58	10.91	12.40	10.94	13.84	12.67	15.76	

Table 4. Continued.

Link Length (m) $Q$ (L s <sup>-1</sup> ) $\Delta z$ (m)	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10	10-11	11-12	12-13	13-14	14-15	15-16	16-17	17-18	18-19	19-20	20-21	21-22	22-23	23-24	24-25	25-26	26-27	27-28	28-29	29-30	30-31		
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	24.3	20.2	20.3	91.0	45.3	42.0	37.0	38.4	45.0	27.5	17.0	16.1	27.4	31.4	14.1	92.6	37.3	25.0	42.4	38.8	16.2	16.2	27.0	28.4	20.8	19.2	38.4	197	180	30.0	386	
BT BT-NoUp CB-Theor CB-Disc	1.92	1.85	1.80	1.78	1.75	1.72	1.69	1.67	1.62	1.47	1.43	1.37	1.30	1.22	1.17	1.10	0.99	0.96	0.94	0.89	0.79	0.74	0.63	0.55	0.41	0.36	0.31	0.23	0.19	0.05		
	8.8	5.5	2.7	3.2	4.2	0.7	-0.7	-1.6	-0.4	0.0	9.7	0.6	1.2	1.6	0.5	-3.6	1.3	-1.1	1.5	1.4	0.3	0.5	0.5	0.5	0.7	1.3	1.7	6.3	0.0	3.2		
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Node BT BT-NoUp CB-Theor CB-Disc	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30		
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	
	0	8.38	13.58	15.97	17.76	21.28	21.41	20.24	18.17	17.20	16.22	20.09	20.15	20.55	21.38	21.54	13.80	13.73	11.76	11.80	11.95	11.83	11.69	11.71	11.97	12.08	12.51	10.62	14.93	14.69		
Link Length (m) $Q$ (L s <sup>-1</sup> ) $\Delta z$ (m)	2-32	3-33	4-34	5-35	6-36	7-37	8-38	9-39	10-40	11-41	12-42	13-43	14-44	15-45	16-46	17-47	18-48	19-49	20-50	21-51	22-52	23-53	24-54	25-55	26-56	27-57	28-58	29-59	30-60			
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A		
	20.0	17.0	29.0	22.0	40.5	35.0	38.0	45.0	59.1	31.2	12.0	15.0	20.0	22.0	18.3	26.6	25.0	27.0	25.0	27.0	26.7	23.0	25.0	33.5	17.5	29.0	20.0	31.0	48.8			
BT BT-NoUp CB-Theor CB-Disc	0.07	0.05	0.02	0.03	0.03	0.03	0.02	0.05	0.15	0.04	0.06	0.07	0.08	0.05	0.07	0.11	0.03	0.02	0.05	0.10	0.05	0.11	0.08	0.14	0.05	0.05	0.08	0.04	0.14			
	-1.1	-0.7	-0.3	1.7	3.5	-1.8	-2.0	3.9	-8.9	6.8	1.5	1.0	-0.5	2.3	1.4	0.0	-1.1	0.9	4.3	-0.9	2.1	-2.6	-1.1	-2.3	-1.0	-4.0	0.5	-6.9				
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	
Node BT BT-NoUp CB-Theor CB-Disc	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60		
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	
	14.12	7.23	12.67	15.67	19.42	24.69	19.48	18.18	21.64	7.31	22.91	21.38	20.92	19.60	23.45	22.65	12.81	12.57	12.65	15.83	10.15	13.65	8.25	10.07	7.85	10.87	8.23	8.79	15.32	7.09		
Link Length (m) $Q$ (L s <sup>-1</sup> ) $\Delta z$ (m)	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00		
	16.97	7.62	13.25	16.43	20.97	26.24	21.03	19.73	23.19	6.17	24.57	22.93	22.47	21.15	25.00	24.20	16.52	16.99	17.53	20.71	15.03	18.53	13.13	13.67	10.70	13.72	11.09	11.64	18.18	9.94		
	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	

N/A – not available.

**Table 5.** Optimization results from Bhave (1978) algorithm. LHS sum and RHS sum are the left and right sides of his Eq. (19), which should be equal.

Node	Main $b$	Main $C$ (USD)	Main head loss (m)	Branch $b$	Branch $C$ (USD)	Branch head loss (m)	LHS sum (USD m <sup>-1</sup> )	RHS sum (USD m <sup>-1</sup> )
2	1.19	77.73	0.428	1.16	231.86	11.20	215.43	241.13
3	1.16	62.87	0.335	1.16	15.53	0.36	217.19	265.77
4	1.16	62.53	0.335	1.16	5.51	5.87	216.11	216.33
5	1.16	278.99	1.499	1.16	6.69	8.73	215.25	215.66
6	1.16	138.01	0.743	1.16	5.13	12.37	214.77	214.60
7	1.16	127.12	0.687	1.16	10.07	17.62	214.12	213.93
8	1.16	111.23	0.603	1.16	9.17	12.31	213.27	213.21
9	1.16	114.87	0.626	1.16	8.87	10.90	212.35	212.13
10	1.16	133.07	0.729	1.16	14.81	14.62	211.19	98.10
11	1.16	67.55	0.806	1.16	69.63	0.70	96.93	212.18
12	1.16	414.08	4.973	1.16	8.42	15.63	96.31	96.70
13	1.16	38.53	0.464	1.16	3.05	14.95	96.07	95.96
14	1.16	64.36	0.778	1.16	4.32	14.63	95.73	95.50
15	1.16	72.08	0.876	1.16	6.67	13.50	95.15	95.33
16	1.16	31.88	0.389	1.16	5.84	17.06	94.76	94.77
17	1.16	204.81	2.510	1.16	5.39	16.31	94.38	93.18
18	1.16	79.31	0.989	1.16	12.07	8.77	92.79	93.44
19	1.16	52.49	0.661	1.16	6.71	8.00	91.85	91.98
20	1.16	88.23	1.121	1.16	6.20	8.28	91.01	91.16
21	1.16	79.12	1.014	1.16	7.47	11.98	90.30	89.01
22	1.16	31.58	0.414	1.16	12.47	7.09	88.29	89.36
23	1.16	51.36	0.680	1.16	8.48	9.99	87.33	85.74
24	1.16	50.84	0.694	1.16	11.52	5.10	84.76	85.51
25	1.16	35.40	0.494	1.16	10.59	6.41	82.90	80.51
26	1.16	29.28	0.431	1.16	20.15	5.27	78.60	82.14
27	1.16	55.92	0.832	1.16	5.52	6.74	77.72	75.66
28	1.16	270.13	4.182	1.16	11.61	4.29	74.71	75.75
29	1.16	222.58	3.545	1.16	8.79	4.42	72.62	73.87
30	1.16	34.71	0.561	1.16	9.56	9.10	71.57	49.43
31	–	–	–	1.16	47.15	1.13	–	–

are omitted since 100 solutions were obtained for each test case. Collectively, the hydraulic grade lines reveal a close alignment of the BT solution (the global optimum) with the CB-Theor solution which utilizes a continuous diameter set. Furthermore, the mapping scheme used to generate a CB-Disc solution is shown to increase pipe sizes in some cases far beyond the limit imposed by  $h_{\min}$ , which was set to 7 m in the present work.

We compared the CB results for the Los Mangos network with those from the Bhave (1978) optimization algorithm (see Table 5). Like Eq. (15) in the present work, Bhave's optimality equation (his Eq. 19) equates the sum of a weighted term for all links entering and leaving each internal node in the distribution main. In the present work the term is proportional to the hydraulic gradient and the weighting factor is proportional to flow rate. In Bhave's case the term is the ratio of pipe cost to head loss, where the weighting factor is pipe-cost exponent  $b$ . There are 60 nodes for this network, including 30 nodes in the distribution main. The rest are delivery

nodes (note there are 2 branches from node 30 of the distribution main). The terms required for the calculations include  $b$ , pipe cost, and head loss in the main and branches. The designation LHS refers to nodes in the distribution main entering, and RHS to those leaving, the node at the far-left side of Table 5. The exponent  $b$  comes from curve fitting pipe-cost data to the two-part pipe-cost model. Linear interpolation was used between diameters  $D_{Co}$  and  $D_{Co+1}$  to obtain  $b$  in this range. Except for a few nodes, agreement between the two CB algorithms is very good. Although Bhave's Eq. (19) and Eq. (18) in the present work, appear quite different due to the different ways each was developed, both produce optimality for the networks considered in this paper. The key distinction between the two developments is the assumption of constancy in terms that comprise the coefficient  $A$  in Bhave (his Eq. 13), mainly the exponent  $b$  ( $m$  in his paper). In general,  $b$  depends on pipe diameter, thus making  $b = b(D)$  for multi-part cost models. When taking derivatives to obtain the final algorithms in both works, this dependence must be in-



cluded, which produces additional terms in the optimization equation (see our Eq. 18 above). However, if the system of equations is solved by an iterative method, as Bhave proposed, the dependency may be neglected (though issues with convergence of the numerical solution may arise because of this). It is very important to note that if a non-iterative method is used to solve the system of equations as done in the present work (using a commercial program like Mathcad), all terms in the governing equations must be treated as continuous, not discrete, and the  $b = b(D)$  dependence must be explicitly included. It should also be noted that the optimization algorithm of Eq. (18) in this paper includes minor loss, which is not included in the Bhave (1978) work.

## 8 Conclusions

Algorithms to optimize the cost of branching gravity-driven water networks are evaluated on six test cases from real networks in the Philippines, Nicaragua, and Panama. A calculus-based algorithm produced a solution composed of theoretical diameters from a continuous set (CB-Theor), which are then mapped onto discrete commercially available diameters (CB-Disc). Backtracking (BT), a recursive algorithm, systematically searches discrete candidate solutions and, when run to completion, is guaranteed to find the global optimum by following rules that prune only higher-cost or hydraulically infeasible candidates. The BT algorithm was modified (BT-NoUp) to improve computational speed by rejecting all candidates that included a small diameter directly upstream of a larger diameter but allowed for the possibility of missing the global optimum. The third type of algorithm evaluated was a genetic algorithm (GA) that used single-point crossover and tournament selection.

BT could find the global optimum in most test cases with relatively little computational effort, although its poor scaling to larger networks is evidenced by its inability to find a solution to case 6, a network with 60 nodes and 59 links. The BT-NoUp completed its search in less time than BT and could find a solution to case 6. Based on case 1–5 results, the extra pruning condition adopted in BT-NoUp sacrificed only small cost increases. Both BT and BT-NoUp, however, could become prohibitively time-consuming when dealing with networks with significantly more links, diameter choices, or an unfavorable layout. While the test cases represent the range of GDWN sizes encountered in the authors' experience, future work would be needed to verify the suitability of the BT and BT-NoUp algorithms on larger GDWNs. The calculus-based algorithm produced consistently good results for the networks tested, although a more robust mapping scheme from theoretical diameters to discrete diameters would further improve on these results as discussed above. In potential future work, the CB-Theor solutions could be used to prune the BT search space, like Kadu et al. (2008), by only including the two diameters above and below the CB-Theor

diameters, producing four diameter choices per link. The calculus-based methodology provides an additional benefit to the designer by explicitly revealing the sensitivities to cost for a design. The calculus-based algorithm requires greater computational effort than backtracking for smaller networks, however, this effort scales more linearly with the number of network links, while backtracking scales exponentially. Furthermore, backtracking's computational time is sensitive to the number of available diameters. Still, when applied to the present study's GDWN test cases with a modest number of links (23), backtracking quickly found a global optimum. In addition, because it is guaranteed to find the global optimum, it can be useful for benchmarking the performance of other algorithms which scale better with more network links. While the genetic algorithm produced solutions with good proximity to the global optimum, its solution costs tended to be further from the global optimum in cases with more links.

For all test cases, the calculus-based algorithm's theoretical diameter solutions (CB-Theor) produced a lower cost than the discrete-domain global optimum. This result is made possible because it is not constrained to a discrete set of diameters. As such, the CB-Theor results represent a lower-bound on the optimum solution within the problem formulation, which could be approached with a finer selection of pipe diameters. We also demonstrated good agreement between the CB-based optimization algorithm developed here and that of Bhave (1978). Though Bhave's algorithm and Eq. (18) in the present work appear quite different due to the different ways each was developed, both produce optimality for the networks considered in this paper. The key distinction between the two developments is that Bhave assumed exponent  $b$  constant in the pipe-cost model, which was justified based on his iterative method of solution. In the present work, which uses a commercial program to solve the non-linear governing equations for  $D$  and  $h$ ,  $b(D)$  dependence is explicitly included for multi-part cost models. Contrasted with Bhave, minor losses are included in the CB optimization algorithm in the present work.

**Data availability.** All survey data from the network cases tested are available in Table 4.

**Competing interests.** The authors declare that they have no conflict of interest.

**Acknowledgements.** This work was partially supported by the Villanova Undergraduate Research Fellowship Program and the Goldwater Foundation.

Edited by: Luuk Rietveld

Reviewed by: three anonymous referees

## References

- Alperovits, E., and Shamir, U.: Design of optimal water distribution systems, *Water Resour. Res.*, 13, 885–900, <https://doi.org/10.1029/WR013i006p00885>, 1977.
- Arnalich, S.: How to design a Gravity Flow Water System, Arnalich – Water and Habitat, 2010.
- Bhave, P. R.: Optimization of Gravity-Fed Water Distribution Networks: Theory, *J. Environ. Eng.-ASCE*, 109, 189–205, 1983.
- Bhave, P. R.: Non-computer optimization of single source networks, *J. Environ. Eng.-ASCE*, 104, 799–814, 1978.
- Bhave, P. R.: Optimal design of water distribution networks, Alpha Science International Ltd., Pangbourne, UK, 2003.
- Bouman, D.: Hydraulic design for gravity based water schemes, Aqua for All, Den Haag, the Netherlands, 2014.
- Churchill, S. W.: Friction factor equation spans all regimes, *Chem. Eng. J.*, 84, 91–92, 1977.
- Colebrook, C. F. and White, C. M.: Experiments with fluid friction in roughened pipes, *P. R. Soc. London*, 161, 367–381, <https://doi.org/10.1098/rspa.1937.0150>, 1937.
- Dandy, G. C., Simpson, A. R., and Murphy, L. J.: An improved genetic algorithm for pipe network optimization, *Water Resour. Res.*, 32, 449–458, <https://doi.org/10.1029/95WR02917>, 1996.
- Dongre, S. R. and Gupta, R.: Discussion of ‘Recursive Design of Pressurized Branched Irrigation Networks’ by César González-Cebollada, Bibiana Macarulla, and David Sallán, *J. Irrig. Drain Eng.*, 138, 697–697, [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0000441](https://doi.org/10.1061/(ASCE)IR.1943-4774.0000441), 2011.
- Gessler, J.: Pipe Network Optimization by Enumeration, Proceedings of the Specialty Conference on Computer Applications in Water Resources, American Society of Civil Engineers, New York, USA, 572–851, 1985.
- González-Cebollada, C., Macarulla, B., and Sallán, D.: Recursive Design of Pressurized Branched Irrigation Networks, *J. Irrig. Drain Eng.*, 137, 375–382, [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0000308](https://doi.org/10.1061/(ASCE)IR.1943-4774.0000308), 2011.
- Jones, G. F.: Gravity-driven Water Flow in Networks: Theory and Design, Wiley, Hoboken, NJ, USA, 2011.
- Kadu, M. S., Gupta, R., and Bhave, P. R.: Optimal design of water networks using a modified genetic algorithm with reduction in search space, *J. Water Res. Plan. Man.*, 134, 147–160, [https://doi.org/10.1061/\(ASCE\)0733-9496\(2008\)134:2\(147\)](https://doi.org/10.1061/(ASCE)0733-9496(2008)134:2(147)), 2008.
- Kansal, A., Gupta, R., and Bhave, P. R.: Optimization algorithms for design of branching water distribution networks, *J. Indian Water Works Association*, 28, 135–140, 1996.
- Keedwell, E. and Khu, S.: Novel Cellular Automata Approach to Optimal Water Distribution Network Design, *J. Comput. Civil. Eng.*, 20, 49–56, [10.1061/\(ASCE\)0887-3801\(2006\)20:1\(49\)](https://doi.org/10.1061/(ASCE)0887-3801(2006)20:1(49)), 2006.
- Kessler, A. and Shamir, U.: Analysis of the Linear Programming Gradient Method for Optimal Design of Water Supply Networks, *Water Resour. Res.*, 25, 1469–1480, <https://doi.org/10.1029/WR025i007p01469>, 1989.
- Krapivka, A. and Ostfeld, A.: Coupled genetic algorithm – Linear programming scheme for least cost design of water distribution systems, *J. Water Res. Plan. Man.*, 135, 298–302, [https://doi.org/10.1061/\(ASCE\)0733-9496\(2009\)135:4\(298\)](https://doi.org/10.1061/(ASCE)0733-9496(2009)135:4(298)), 2009.
- Maier, H. R., Simpson, A. R., Zecchin, A. C., Foong, W. K., Phang, K. Y., Seah, H. Y., and Tan, C. L.: Ant colony optimization for design of water distribution systems, *J. Water Res. Plan. Man.*, 129, 200–209, [https://doi.org/10.1061/\(ASCE\)0733-9496\(2003\)129:3\(200\)](https://doi.org/10.1061/(ASCE)0733-9496(2003)129:3(200)), 2003.
- Martin, Q. W.: Optimal design of water conveyance systems, *J. Hydr. Eng. Div.-ASCE*, 106, 1415–1433, 1980.
- Mohan, S. and Jinesh Babu, K. S.: Water distribution network design using heuristics-based algorithm, *J. Comput. Civil. Eng.*, 23, 249–257, [https://doi.org/10.1061/\(ASCE\)0887-3801\(2009\)23:5\(249\)](https://doi.org/10.1061/(ASCE)0887-3801(2009)23:5(249)), 2009.
- Monbaliu, J., Jo, J., Fraise, C. W., and Vadas, R. G.: Computer aided design of pipe networks, Proc. Int. Symp. On Water Resource Systems Application, Friesen Printers, Winnipeg, Canada, 1990.
- Nicklow, J., Reed, P., Savic, D., Dessalegne, T., Harrell, L., Chan-Hilton, A., Karamouz, M., Minsker, B., Ostfeld, A., Singh, A., Zechman, E., and ASCE Task Committee on Evolutionary Computation in Environmental and Water Resources Engineering: State of the Art for Genetic Algorithms and Beyond in Water Resources Planning and Management, *J. Water Res. Plan. Man.*, 136, 412–432, [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000053](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000053), 2010.
- Prasad, T. D. and Park, N. S.: Multiobjective genetic algorithms for design of water distribution networks, *J. Water Res. Plan. Man.*, 130, 73–82, [https://doi.org/10.1061/\(ASCE\)0733-9496\(2004\)130:1\(73\)](https://doi.org/10.1061/(ASCE)0733-9496(2004)130:1(73)), 2004.
- Raad, D. N.: Multi-objective optimisation of water distribution systems design using metaheuristics, PhD thesis, University of Stellenbosch, South Africa, 2011.
- Saldarriaga, J., Páez, D., Cuero, P., and León, N.: Optimal Design of Water Distribution Networks Using Mock Open Tree Topology, World Environmental and Water Resources Congress, 19–23 May 2013, Cincinnati, Ohio, USA, 869–880, 2013.
- Samani, H. M. V. and Mottaghi, A.: Optimization of water distribution networks using integer linear programming, *J. Hydraul. Eng.*, 132, 501–509, [https://doi.org/10.1061/\(ASCE\)0733-9429\(2006\)132:5\(501\)](https://doi.org/10.1061/(ASCE)0733-9429(2006)132:5(501)), 2006.
- Simpson, A. R., Dandy, G. C., and Murphy, L. J.: Genetic Algorithms Compared to Other Techniques for Pipe Optimization, *J. Water Res. Plan. Man.*, 120, 423–443, [https://doi.org/10.1061/\(ASCE\)0733-9496\(1994\)120:4\(423\)](https://doi.org/10.1061/(ASCE)0733-9496(1994)120:4(423)), 1994.
- Streeter, V. L., Wylie, E. B., and Bedford, K. W.: Fluid Mechanics, McGraw-Hill, New York, NY, USA, 1998.
- Suribabu, C. R.: Heuristic-based pipe dimensioning model for water distribution networks, *J. Pipeline Syst. Eng.*, 3, 115–124, [https://doi.org/10.1061/\(ASCE\)PS.1949-1204.0000104](https://doi.org/10.1061/(ASCE)PS.1949-1204.0000104), 2012.
- Swamee, P. K. and Jain, A. K.: Explicit equations for pipe flow problems, *J. Hydr. Eng. Div.-ASCE*, 102, 657–664, 1976.
- Swamee, P. K. and Sharma, A. K.: Gravity low water distribution network design, *J. Water Supply Res. T.*, 49, 169–179, 2000.
- Swamee, P. K. and Sharma, A. K.: Design of Water Supply Pipe Networks, Wiley, Hoboken, NJ, USA, 2008.
- Taorimina, R. and Chau, K. W.: Data-driven input variable selection for rainfall–runoff modeling using binary-coded particle swarm optimization and Extreme Learning Machines, *J. Hydrol.*, 529, 1617–1632, 2015.

- Tospornsampan, J., Kita, I., Ishii, M., and Kitamura, Y.: Split-pipe design of water distribution network using simulated annealing, *International Journal of Computer, Information, and Systems Science, and Engineering*, 1.3, 153–163, 2007.
- Vairavamorthy, K. and Ali, M.: Optimal design of water distribution systems using genetic algorithms, *Comput. Aided Civ. Infrastruct. Eng.*, 15, 374–382, <https://doi.org/10.1111/0885-9507.00201>, 2000.
- Vasan, A. and Simonovic, S. P.: Optimization of water distribution network design using differential evolution, *J. Water Res. Plan. Man.*, 136, 279–287, [https://doi.org/10.1061/\(ASCE\)0733-9496\(2010\)136:2\(279\)](https://doi.org/10.1061/(ASCE)0733-9496(2010)136:2(279)), 2010.
- Wu, Z. Y. and Walski, T.: Self-adaptive penalty approach compared with other constraint-handling techniques for pipeline optimization, *J. Water Res. Plan. Man.*, 131, 181–192, [https://doi.org/10.1061/\(ASCE\)0733-9496\(2005\)131:3\(181\)](https://doi.org/10.1061/(ASCE)0733-9496(2005)131:3(181)), 2005.
- Yang, K. P., Liang, T., and Wu, I. P.: Design of conduit system with diverging branches. *J. Hydr. Eng. Div.-ASCE*, 101, 167–188, 1975.
- Zhao, W., Beach, T., and Rezgui, Y.: Optimization of Potable Water Distribution and Wastewater Collection Networks: A Systematic Review and Future Research Directions, *IEEE T. Syst. Man Cyb.*, 46, 659–681, <https://doi.org/10.1109/TSMC.2015.2461188>, 2016.
- Zheng, F., Simpson, A. R., and Zecchin, A. C.: A decomposition and multistage optimization approach applied to the optimization of water distribution systems with multiple supply sources, *Water Resour. Res.*, 49, 380–399, <https://doi.org/10.1029/2012WR013160>, 2013.