This article was downloaded by: [130.70.8.131] On: 04 April 2019, At: 12:35
Publisher: Institute for Operations Research and the Management Sciences (INFORMS)
INFORMS is located in Maryland, USA

INFORMS Journal on Applied Analytics

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

NeatWork: A Tool for the Design of Gravity-Driven Water Distribution Systems for Poor Rural Communities

Frédéric Babonneau, Gilles Corcos, Laurent Drouet, Jean-Philippe Vial

To cite this article:
Frédéric Babonneau, Gilles Corcos, Laurent Drouet, Jean-Philippe Vial (2019) NeatWork: A Tool for the Design of Gravity-Driven Water Distribution Systems for Poor Rural Communities. INFORMS Journal on Applied Analytics

Published online in Articles in Advance 03 Apr 2019

. https://doi.org/10.1287/inte.2018.0983

Full terms and conditions of use: https://pubsonline.informs.org/page/terms-and-conditions

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2019, INFORMS

Please scroll down for article—it is on subsequent pages

INFORMS is the largest professional society in the world for professionals in the fields of operations research, management science, and analytics.
For more information on INFORMS, its publications, membership, or meetings visit http://www.informs.org

informs
http://pubsonline.informs.org/journal/inte/

INFORMS JOURNAL ON APPLIED ANALYTICS
Articles in Advance, pp. 1–8
ISSN 0092-2102 (print), ISSN 1526-551X (online)

NeatWork: A Tool for the Design of Gravity-Driven Water Distribution Systems for Poor Rural Communities

Frédéric Babonneau,[a,b] Gilles Corcos,[c,d] Laurent Drouet,[e] Jean-Philippe Vial[a,f]

[a] Ordecsys, 1224 Chêne-Bougeries, Switzerland; [b] Business School, Adolfo Ibañez University, 7910000 Santiago, Chile; [c] University of California, Berkeley, Berkeley, California 94720; [d] Agua Para La Vida, Berkeley, California 94705; [e] Resources for the Future–Centro Euro-Mediterraneo sui Cambiamenti Climatici, European Institute in Economics and the Environment, 20144 Milan, Italy; [f] University of Geneva, 1211 Geneva, Switzerland

Contact: frederic.babonneau@uai.cl, https://orcid.org/0000-0002-3430-0314 (FB); gilcorc@gmail.com (GC); laurent.drouet@cmcc.it, https://orcid.org/0000-0002-4087-7662 (LD); jphvial@gmail.com (J-PV)

Received: October 13, 2017
Revised: November 22, 2017; June 22, 2018; October 8, 2018
Accepted: October 15, 2018
Published Online in Articles in Advance: April 3, 2019

https://doi.org/10.1287/inte.2018.0983

Copyright: © 2019 INFORMS

Abstract. NeatWork is an optimization and simulation tool for the design of purely gravity-driven water-distribution systems with the objective of delivering clean water to poor rural communities. NeatWork proposes a heuristic approach for the design of least-cost systems under stochastic intermittent water demands, in which devices, such as pumps and pressure regulators, which are operated by humans and commonly used to control pressures and flows, are excluded. The resulting designs are thus as simple and as robust as possible to operate, and the operating, maintenance, and investment costs are kept as low as possible, an important requirement in poor rural communities. We illustrate the application of NeatWork on a typical project implemented by Agua Para La Vida, the nongovernmental organization currently using the tool for its activities in Nicaragua.

History: This paper was refereed.

Keywords: gravity-driven distribution systems • design • simulation • optimization

Populations to whom access to potable water is almost completely denied are often inhabitants of small and destitute rural villages. It is unrealistic to expect people in poor countries to depend on their governments or for those in rural areas to depend on private water-distribution companies to remedy this situation. They cannot depend on government because their tax revenues are too meager, and local competence is often lacking; they cannot depend on private companies because these companies cannot justify the investments required based on potential revenues they might derive from exceedingly poor populations. Nongovernmental organizations (NGOs) such as Agua Para La Vida (APLV) and local initiatives can provide substitutes by building simple, dependable, and moderately inexpensive water-distribution systems. The choice of a particular method to bring water (e.g., wells, rain harvesting, desalination, and/or convection from distant sources) will depend substantially first on the geomorphology of the land surrounding the village, then on the local rainfall distribution, and, finally, although still importantly, on the relative initial and subsequent maintenance costs and on the simplicity of the system. If we admit that providing a separate and individual supply to each home is rarely a rational choice, no matter how water is conveyed to the site of the village, distribution will have to originate at an accumulation device such as a tank and end at the sites of consumption. Our paper addresses this aspect of the distribution problem.

We assume that (1) water must be distributed in the simplest and most economical way possible to a rural population that needs to make only the most elementary uses of that water (e.g., to fill containers or use the water at points of delivery for bathing or washing clothes); (2) maintenance should be performed by the beneficiaries and so must be as simple and as inexpensive as possible; and (3) any solution has to be dependable year-round and should deliver flow rates within prescribed limits to individual faucets under variable conditions of use. Therefore, we are looking for the most satisfactory means of engineering for that type of distribution.

For small NGOs with limited available funds (e.g., APLV), cost considerations are particularly important. Cost components consist of investment and operations. Investment costs are essentially material costs (primarily pipes), especially for the highly dispersed villages that are typical in Nicaragua: A village of 150 inhabitants may require more than 10 km of piping for its distribution system. The cost of manpower, which consists of the unpaid beneficiaries, is negligible. The recurrent cost of operations, however, should be minimized because fund providers to which NGOs such as APLV have access are reluctant to fund these costs. We thus conclude that a purely gravity-driven distribution system, excluding the use of devices operated by humans, such as pumps and pressure regulators, best meets the needs of these populations by eliminating most maintenance and operational costs.

We organized the remainder of the paper as follows. We first discuss the challenges of designing reliable self-regulated systems and the limits of existing approaches. We then present the heuristic approach we followed in NeatWork to address these challenges. Finally, we describe how NeatWork is used in the design and test phases of a typical project implemented by APLV. We note here that NeatWork's executable code can be downloaded freely from http://neatwork.ordecsys.com/, and its source code can be downloaded from http://github.com/lolow/neatwork.

Challenges of Self-Regulated Systems

Selecting a self-regulated system entails addressing the following critical issues: (1) the design of a self-regulated distribution system in which users randomly open or close their faucets, and (2) the sizes of the pipe diameters that are necessary to handle the stochastic demands. Conceivably, a poorly designed system may fail to deliver adequate pressures and flows at some faucets when certain open or closed faucet configurations prevail. A simple way to avoid this situation is to select very large inner pipe diameters to make friction headlosses negligible, but compensate for gravitational energy through orifices. Unfortunately, this result is achieved only at very high material costs. Balancing costs with service quality is thus mandatory to achieve affordable systems. We propose a heuristic approach to do so.

The optimal design of a self-regulated water-distribution system subject to randomly intermittent demands is not addressed in the literature. As we mention above, the material costs for the end-user section of the networks serving individual dwellings has not been deemed critical in practice. Oversized solutions are adopted and are often coupled with devices, such as pumps, to increase and control the pressure. Most studies deal with higher-level networks, including loops for safety reasons, and with delivery points that serve relatively stable aggregate demands. The models for water distribution reported in the literature rely on a mathematical formulation based on the continuity equations (i.e., conservation of masses) and on the continuity of the energy line to represent the steady-state flow in a network. The energy line follows the experimental Colebrook law of energy loss of a steady-state flow in a pipe. As early as 1936, Cross (1936) proposed a method to compute the steady-state flows in a general network (i.e., a network including loops) as the solution to a system of linear and nonlinear equations. Alternative and more efficient methods have since been proposed. For example, the freely distributed software EPANET (Rossman 1994, 2000) is based on the gradient method of Todini and Pilati (1987); this software analyzes and simulates the hydraulic behavior of large networks, including loops, but does not deal with the optimal design problem per se. Papers dealing with the optimal design problem consider networks in which the outlets serve deterministic aggregate demands. Pipe material costs are still the greatest, if not only, cost component in the objective function. Models and methods have been proposed to deal with the interaction between cost minimization and the mathematical representation of steady-state flows. Alperovits and Shamir (1977) formulate the joint problem of minimizing cost and serving a prescribed demand as a nonlinear programming problem. The variables they use are flows and pipe diameters, and the constraints are mass-conservation and energy-line equations. The latter are nonlinear: They involve the absolute value of flows and, as such, are nondifferentiable. This nonconvex, nondifferentiable problem is hard to solve, and the solution methods generate local, possibly suboptimal, optima; to this end, Alperovits and Shamir (1977) proposed the so-called linear programming gradient, and Eiger et al. (1994) developed an efficient branch-and-bound scheme to compute a global solution.

In a paper devoted to the related problem of gas transmission, Maugis (1977) noticed that integrating the Colebrook law with respect to the flow yields an expression whose interpretation is the total energy dissipation in a pipe as a function of the flow. This opens the possibility of computing steady-state flows in an existing network as the solution to a strongly convex programming problem whose objective is the steady-state total energy per unit of time and where the only constraint is mass conservation. This total-energy approach has been embedded into an optimal design formulation in Babonneau et al. (2012); however, whether it can be used in the context of stochastic open or closed faucet configurations is not clear. In this paper, we use the total-energy approach only to compute flows in a simulation module.

A Heuristic Approach

The water-distribution network under consideration is a tree represented by an oriented graph from a tank to faucets. Water flows along links through pipes whose inner diameters directly influence the amount of friction and thus contribute to the determination of the steady-state flows. Another influencing factor is pipe roughness, which depends on the pipe material (e.g., polyvinyl chloride [PVC] versus iron). APLV uses only PVC; however, NeatWork can accommodate other materials. Terminal nodes are equipped with faucets; all are of the same type and have two positions—open and closed. There are also orifices, devices consisting of a disk with a small hole at the center, which may be set close to and upstream of the faucet; they are used to enforce a pressure drop and regulate flow. The design problem consists then of choosing pipes, as characterized by cost and inner diameter, and orifice diameters. The choice of node location is dictated by local topography and is not part of the problem we are solving.

The main feature of the system at work is the random behavior of the users in opening or closing their faucets. The resulting intermittent demands must be anticipated in the design phase. Although the problem could be formulated as a two-stage stochastic programming problem, the numerical effort required to solve it would be formidable, if not prohibitive, and not justified in the present context. We propose instead a heuristic approach for the design problem; in our solution, we replace the unknown random flows by deterministic approximating values. The goal of this heuristic is to produce a design that will be satisfactory for a sufficiently large proportion of intermittent-demand configurations. The key feature in determining the approximating flows is a simple notion of quality of service, ensuring that, with a sufficiently large probability, the approximating flow is larger than the faucet target flow times the (random) number of open downstream faucets. In a second step, a module simulates flows based on the design of a sample of configurations. If the outcome of the simulation reveals troublesome statistics at some faucets, the initial design is modified. The suggested scheme to modify the initial design exploits a constraint that is implemented in the design module. That is, at each intermediate node, the total headloss from tank to that node times a safety factor, which must be greater than one, must be smaller than local gravity drop. A high safety factor enforces a smaller upstream headloss, which contributes to reducing faucet flow variability. We explain the scheme and the mathematical formulation of the heuristic in the appendix. The case study we discuss below illustrates this process.

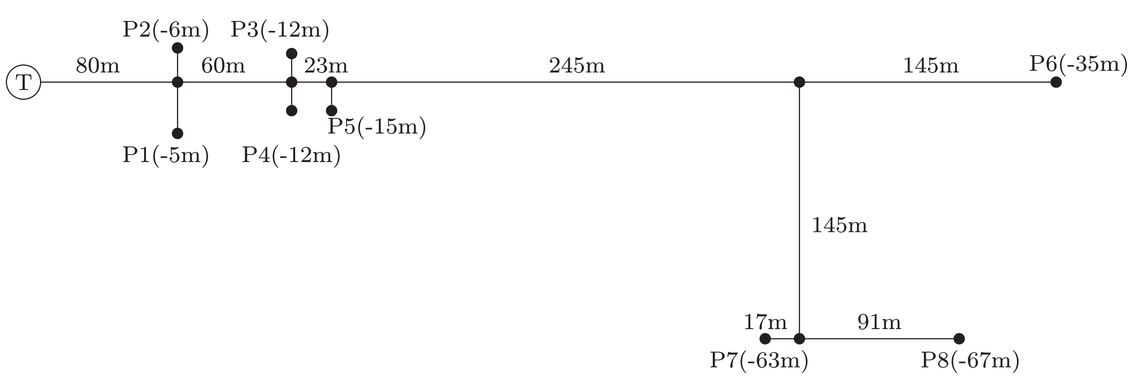
Case Study

We present the design phase of a typical project implemented by APLV. The project, La Luna, involves a network of 129 faucets used by over 700 hundred inhabitants. Faucet elevations relative to the tank range from −128.7 to −1.8 meters. There are 117 branching nodes and as many intermediary branches. The total network length is 4,837 meters. In view of the number of faucets, the network length, and the range of elevations, finding a satisfactory design was a challenge. These characteristics make it difficult to show a graphical representation of this network. Instead, to construe a network configuration, we use Figure 1, which represents a much smaller, but typical, network in Nicaragua.

The design for La Luna was produced by a local technician who was taught to use NeatWork. In a first run, the technician specified the three main design parameters: the target flow $\bar{\psi}$, the quality of service (which determines load factor λ via the design heuristic), and the proportion of open faucets at peak hours, according to the norms in use in the organization. In this first run, the performance of the design proposed by NeatWork is simulated. If the performance is not deemed to be satisfactory, another design is generated, mainly by modifying the safety factor in the headloss constraints, which we show in Equation (A.5c), at the branching nodes. (We do not change the headloss constraints at the terminal or faucet nodes in this process.)

The established norm at APLV is a target flow of 0.12 l/s, where l/s represents liters per second (i.e., filling a 1-l bottle in 8.3 s). The recommended fraction of open faucets is 0.28 at peak hours, and the quality of service is usually 65% (see the appendix for more details). Simulations produce an empirical distribution of the outflow for each faucet when that faucet is open. This distribution is summarized by

Figure 1. The Network of the El Anzuelo Community in Nicaragua



Notes. The elevations of the eight faucets (P1–P8) are relative to the tank T level. The distances between faucets and tank and the important differences between node elevations make the design of a low-cost and self-regulated system a nontrivial challenge, even for such a simple network.

statistics (e.g., minimum, mean, maximum, and various other fractiles). The most important statistic is the proportion of times the flow at a faucet is below a threshold value. The acceptability criterion at APLV is that the flow must be above 0.06 l/s at least 98% of the time. To capture information on the tail of the distributions, the sample size must be sufficiently large. In this study, we simulated 10,000 configuration scenarios.

We evaluated the cost in cordobas, the local Nicaraguan currency (100 cordobas, denoted c$100, is U.S. dollars [USD] $3.3). Using the default values for the parameters, NeatWork produced a first design with a total cost of c$71,553. The overall average flow in the first simulation run was 0.123, an adequate value. Unfortunately, the second goal (i.e., flows rarely below the threshold value of 0.06) was not met for 10 of the 129 faucets. For these critical faucets, flow fell below the threshold value more than 3% of the time. We show this in Table 1, which also includes the result for one near-critical faucet. A 1.2 safety factor applied to branching node constraints yielded a new design. A quick simulation of moderate sample size showed that, although the problems at the faulty faucets were eliminated, costs increased. Therefore, we tried lower values for the safety factor. Eventually, we deemed a design with a 1.05 safety factor to be satisfactory. A large sample-size simulation confirmed the observation. The cost of the final design was c$72,784, an increase of less than 2%.

Table 1 displays information on the faulty faucets in the initial design and the improvements that the final design showed. The table does not report on the other 118 faucet nodes, because none flowed below 0.06 in the final design.

The values support the argument that the safety factor modifies and can decrease faucet flow variability. For critical faucets, variability is, on average, more than halved. To further illustrate the benefit of our approach, we ran the same case with a safety factor of 10; that is, we divided the (negative) elevations of the intermediate nodes by 10. NeatWork proposed a design with almost perfect stability of faucet flows. Variability was at most 1%, and otherwise negligible. The total cost jumped to c$134,319, nearly twice the cost of the final design. The ability to make trade-offs between cost and variability, as shown in Table 1, illustrates NeatWork's benefits.

It is legitimate to wonder whether the implemented system would behave as predicted by NeatWork. Discrepancies between theory and practice may occur because of errors in data gathering or incorrect implementation. In principle, local technicians would be responsible for some basic in situ tests in which each faucet would be opened one at a time and flow would be compared with NeatWork's predictions. This is useful in detecting errors in the topographical data or in the implementation (e.g., incorrect placement of pipes or unauthorized local changes). In general, the gap is around 2%–3%; larger deviations suggest that investigations of the data and/or the implementation should be done; however, remedying such a situation is rarely possible, because testing the system under daily random usage is beyond the competence of local workers. The satisfaction criterion, which is retained instead, is the user feedback collected by a local water committee that APLV established. This ongoing evaluation during the project life was positive relative to NeatWork: The local water committee never mentioned encountering any problems with La Luna or with other

Table 1. The Results of Simulations with 10,000 Instances for the Initial and Final Designs

	Initial design		Final design	
Average flow	0.123 l/s		0.122 l/s	
Design cost	c$71,553		c$72,784	
	Percent in the simulation with flow < 0.06	Flow variability, %	Percent in the simulation with flow < 0.06	Flow variability, %
Faucet				
P89	**20.47**	59.46	0.52	23.55
P127	**12.96**	49.86	0.39	22.01
P121	**12.25**	44.81	0.04	19.12
P62	**10.91**	47.80	0	18.64
P109	**9.45**	41.31	2.39	24.60
P120	**8.64**	41.52	0.03	15.33
P83	**6.90**	38.29	0	16.84
P84	**4.17**	38.48	0	18.06
P125	**4.12**	36.72	0.17	19.66
P105	**3.14**	30.61	0	13.55
P88	2.66	36.44	0	19.24

Notes. The table focuses on the 11 faucets that we found to be critical or subcritical in the initial design. It also shows the performance improvement in the final design. A faucet was deemed critical if the observed flow fell below 0.06 l/s more than 3% of the time. Columns 1 and 3 report the percentage of time that faucet flow was below the threshold. Columns 2 and 4 report the coefficient of variability of the faucet flows, measured as the ratio between the empirical standard deviation and the empirical mean.

projects designed using NeatWork. The only real issue that might occur would be a significant and unforeseen increase in the number of inhabitants. This would lead to different operating conditions, typically because of a much larger fraction of open faucets at peak hours. Addressing this would require an enhancement of the present design, possibly with additional faucets and pipe modifications of some authorized links. NeatWork permits such an enhancement.

Finally, we note that APLV technicians have successfully designed significantly larger projects than La Luna. For example, a project for the Wany community involved 269 faucets and a total length above 25 km. The design process and the resulting performance results are similar to those for the La Luna project.

Conclusion

NeatWork is a computer tool for the design of self-regulated, gravity-driven water-distribution systems. The tool is designed to meet the very specific and restrictive requirements of water systems in hilly regions, which are populated by dispersed and poor inhabitants and lack local administrative support for adequate maintenance and where the total investment cost is the driving force. The design problem is far from trivial, because low-cost networks are more sensitive to open or closed faucet configurations. Moreover, the work must be performed by local technicians with limited training. The experience that APLV has gained in successfully implementing more than 70 projects over the past 17 years has proved the benefits of NeatWork. The systems installed are all still in use and serve clean, drinkable water to over 40,000 people in over 90 communities. In 2015, APLV was selected as the winner of the seventh edition of the Water and Sanitation Prize, which the Inter-American Development Bank and the FEMSA Foundation sponsored, for the successful use of the NeatWork software.

NeatWork may also be useful in other contexts. In less hilly and/or drier regions, gravity-actuated distribution systems between an elevated tank and a lower set of dwellings are common. Moreover, the ease with which a user can generate a design with NeatWork enables that user to make trade-offs between the elevation of the reservoir and the cost of pipe material—the higher the reservoir, the higher the pressure and less need for pipes with large diameters.

Acknowledgments

The authors thank Erling Andersen, who allowed the Mosek tool to be used in the free, nonprofit distribution of NeatWork; and the University of Geneva for its support in the early development phase of NeatWork.

Appendix

A Heuristic Approach

In this section, we describe the heuristic approach implemented in NeatWork, which comprises a design module, in which the unknown random flows are replaced by deterministic approximating values, and a simulation module, which rigorously and nonheuristically computes flows. We start with a description of this module.

The water-distribution network under consideration is a tree represented by the graph $G = (\mathcal{N}, \mathcal{L})$, where $\mathcal{N} = \{1, 2, \ldots, J\}$ denotes the node set and $\mathcal{L} = \{1, 2, \ldots, l\}$ is the set of links. A link $k \in \mathcal{L}$ joins the origin $i_k \in \mathcal{N}$ and the destination $j_k \in \mathcal{N}$. The node set is partitioned into $\mathcal{N}_o$ (tank), $\mathcal{N}_b$ (the branching nodes with no outflow), and $\mathcal{N}_f$ (terminal faucet nodes). The graph is oriented from tank to faucets. Graph attributes are node elevations h_i and length L_k of the links. Water flows along links (i, j) through pipes with inner diameters D_{ij}.

Flow Simulation Model. The objective of the simulation tool in NeatWork is to test the behavior of a system produced by the design model. It deals with stabilized steady-state flows. Transitory regimes are not considered in the analysis.

According to Maugis (1977) and Babonneau et al. (2012), steady-state flows satisfy the principles of minimum energy and mass conservation. They are solutions to the following energy minimization problem:

$$\min_{\phi,\psi} \sum_{(ij)\in\mathcal{L}} \gamma \frac{L_{ij}}{D_{ij}^q} \frac{\phi_{ij}^{p+1}}{p+1} + \sum_{t\in\mathcal{N}_f} \left(\frac{1}{\alpha} + \left[\frac{0.62}{d_t}\right]^4\right) \frac{\psi_t^3}{3} - \sum_{j\in\mathcal{N}_f} (h_0 - h_j)\psi_j, \tag{A.1a}$$

$$\text{s.t.} \sum_{(j,k)\in\mathcal{L}} \phi_{jk} - \sum_{(k,j)\in\mathcal{L}} \phi_{kj} = 0, \forall j \in \mathcal{N} \backslash (\mathcal{N}_0 \cup \mathcal{N}_f), \tag{A.1b}$$

$$-\phi_{it} + \psi_t = 0, \text{ for } (i, t) \in \mathcal{L} \text{ and } \forall t \in \mathcal{N}_f \tag{A.1c}$$

$$\psi_t \geq 0, \forall i \in \mathcal{N}_t. \tag{A.1d}$$

In this formulation, variable ψ_j is outflow at faucet $j \in \mathcal{N}_f$ and variable ϕ_{ij} is flow on the link $(i, j) \in \mathcal{L}$. We propose an interpretation of these three components as energy per unit of time. The first component of the objective function (A.1a) represents energy dissipated by friction from flow ϕ_{ij} in link (i, j), where coefficient γ depends on pipe characteristics (e.g., degree of roughness), and exponents p and q take values 1.781 and 4.781, respectively. The second component deals with pressure losses in orifices and in faucets as a function of outflow ψ, where parameter α is relative to the faucet (the default value is $\alpha = 1.83\,10^{-8}$), d_t is the inner diameter of the orifice, and constant 0.62 (relative to the orifice) is an experimental value. The last component of objective function (A.1a) pertains to the gravitational energy between the tank at elevation h_0 and faucets $j \in \mathcal{N}_f$ at elevations h_j. Because $h_0 > h_j$, gravitational energy contributes negatively to the total, inducing positive flows in the network. The objective function (A.1a) does not include the contribution of kinetic energy, which is negligible for the systems under consideration. For an extensive discussion of the components in objective function (A.1a) and the associated parameter values, we refer the reader to the *NeatWork User Guide* (http://neatwork.ordecsys.com/dl/neatworkuserguide.pdf). Equations (A.1b) and (A.1c) simply ensure flow conservation at branching nodes and faucets, respectively. Because the network is a tree and the outflows ψ_t are nonnegative, all flows are nonnegative. The minimization programming problem is thus strongly convex. Note that the optimal solution does not exclude zero flow at some faucets and in some links. The embedded commercial code MOSEK ApS (2015) solves it routinely.

In sketching the dual of Model (A.1), the dual constraint associated with a strictly positive flow $\phi_{ij} > 0$ on link (i, j) is expressed in term of dual variables π_i and π_j, and the derivative with respect to ϕ_{ij} in objective function (A.1a):

$$\gamma L_{ij} \phi_{ij}^{p} D_{ij}^{-q} = \pi_j - \pi_i. \tag{A.2}$$

For our purpose, Equation (A.2) is a satisfactory approximation of the Colebrook formula. The left side, the headloss term, reflects friction loss of energy per unit of time and per unit of flow, and π_i and π_j are the pressures at nodes i and j. Regarding the dual constraint associated with the outflow $\psi_t \geq 0$, we have

$$\left(\left[\frac{0.62}{d_t} \right]^4 + \frac{1}{\alpha} \right) \psi_t^2 + \pi_t \leq h_o - h_j, \tag{A.3}$$

with equality if $\psi_t > 0$. Given the strong convexity of the functions in Model (A.1), the necessary optimality conditions (A.2) and (A.3) are also sufficient. They univocally characterize steady-state flows. Moreover, we know that in network flow problems, the dual variables π are defined up to a constant (i.e., π_o can be set to zero). We can eliminate the dual variables altogether by summing Equations (A.2) and (A.3) along each path from the tank to a faucet. Denoting P_t the set of nodes that are on the path from the tank to faucet t, this gives:

$$\sum_{k \in P_t} \gamma \frac{L_k}{D_k} \phi_k^p + \sum_{t \in \mathcal{N}_f} \left(\frac{1}{\alpha} + \left[\frac{0.62}{d_t} \right]^4 \right) \psi_t^2 \leq h_o - h_t, \quad \forall t \in \mathcal{N}_f, \tag{A.4}$$

with equality if $\psi_t > 0$. This alternative formulation fully determines the steady-state flows. It is commonly used in the literature. Although it is not as efficient for numerical solutions as the minimization approach, it is helpful in the design problem.

Optimal Design Model. The design model aims to satisfy two goals: to minimize material cost and to ensure a certain quality of service. That is, at each faucet, the flow must be as close as possible to a given target outflow $\bar{\psi}$. In view of the uncertainty in the actual use of the network, the second goal is difficult to attain—faucets can be randomly opened or closed by users.

The formulation that comes to mind is that of two-stage stochastic programming with recourse: First-stage variables would be pipe diameters, and second-stage variables, or recourses, would be flows adjusted to the 0–1 (off–on) status of the faucets. This approach would rely on Inequality (A.4) to characterize the steady-state flows and would thus be a constraint for the cost-minimization approach. That inequality relates flows to pipe diameters through an explicit formula. To match the random configuration of open-closed faucets, recourse variables—outflows ψ_t and flows ϕ_{ij} in the links—would have to be indexed by configurations of faucet status. Other components of such a formulation would be an objective and some constraints to determine whether the probability distributions of the faucet flows are admissible. Conceivably, this approach could be implemented. However, computing the solution would be a disproportionate numerical challenge with regard to our problem of concern. Rather, we propose a heuristic approach in which recourse variables are approximated by a deterministic value, permitting the design problem to be formulated through the first-stage variables only (i.e., pipe diameters). The designer may want to test the quality of the heuristic by submitting the constructed design to simulation on a sample of open or closed faucet configurations.

Defining Approximated Flows. The proposed deterministic approximation of the random flow over a given link $(i, j) \in \mathcal{L}$ is a multiple $\lambda_{ij}\bar{\psi}$ of the target flow. The multiplicative factor λ_{ij} depends on the number of faucets downstream from the link. The goal is that the flow at an open faucet should be equal to a target outflow $\bar{\psi}$. Suppose this goal is achieved: In a tree, knowledge of flows at faucets connected to a segment determines the flow in that segment. Because of the random status of each faucet, the cumulated flow is random. To avoid an explicit treatment of randomness, the heuristic bases the design not on true stochastic flows in the segments, but on deterministic approximations of their random values: the more numerous the faucets downstream of a segment, the smaller the flow variability in that segment. More precisely, flow in a given segment is chosen to be large enough to ensure that target flows are achieved at all dependent faucets within an acceptable proportion of time.

In our probability model, individual faucets are open or closed according to an independent identically distributed binary process. If n is the number of faucets downstream of a given link, the heuristic assigns a flow $\lambda\bar{\psi}$ instead of $n\bar{\psi}$ with $\lambda < n$. We name λ the load factor. The higher λ, the greater the probability that the flow $\lambda\bar{\psi}$ is sufficient for delivering $\bar{\psi}$ at each faucet that is open downstream. The probability that the number of open faucets is less than λ is interpreted as quality of service. The main feature of the heuristic consists of assigning appropriate load factors to each link to achieve the same quality of service everywhere. We describe the formula for computing load factors in the last section of the appendix. NeatWork allows the user to set the desired quality of service, but recommends a value of 0.65.

A Linear Optimization Model. Pipe cost and headloss are nonlinear functions of the pipe diameters and linear functions of the pipe length. Because pipes must be chosen from among a finite set of available commercial pipes $\nu = 1, \ldots, n$ with known diameters D_ν and known cost-per-unit length C_ν, the nonlinear components are computed offline for each commercial item, and we can reduce the problem to finding the appropriate linear combination of pipe length on each segment.

By defining variable $\ell_{k\nu}$ as the length of commercial pipe ν to be installed on link $k \in \mathcal{L}$, the cost objective becomes linear in the $\ell_{k\nu}$ variables. The heuristic is based on the concept that when a faucet is open, Equation (A.4) prevails, when flows inside the equation are implied by the load factor $\lambda\bar{\psi}$. This constraint is again in the ℓ_{kv} variables alone. The optimal design associated with load factors $\lambda\bar{\psi}$ is thus the solution of the simple linear cost-minimization problem:

$$\min_{\ell} \sum_{k\in\mathcal{L}} \sum_{v=1,\ldots,n} \ell_{kv} C_v \tag{A.5a}$$

$$\sum_{k\in P_t}\sum_{v=1}^{n} \ell_{kv}\gamma \frac{(\lambda_k\bar{\psi})^p}{D_v^q} + \frac{1}{\alpha}\bar{\psi}^2 + \tau_t = h_o - h_t, \forall t\in\mathcal{N}_f \tag{A.5b}$$

$$\sum_{k\in P_i}\sum_{v=1}^{n} \ell_{kv}\gamma \frac{(\lambda_k\bar{\psi})^p}{D_v^q} \le \frac{h_o - h_i}{\sigma_i}, \forall i\in\mathcal{N}\backslash\mathcal{N}_f \tag{A.5c}$$

$$\sum_{v=1,\ldots,n} \ell_{kv} = L_k,\ \forall k\in\mathcal{L}, \tag{A.5d}$$

$$\ell_{kv} \ge 0, \forall k, v \text{ and } \tau_t \ge 0,\ \forall t\in\mathcal{N}_f. \tag{A.5e}$$

In this formulation, Equation (A.5b) constrain the total headloss to match gravity in Equation (A.4), whereas Equation (A.5c) aim to protect the network against damage caused by a leakage in some pipe. The parameter $\sigma_i \ge 1$ is a safety factor. In practice, it takes a value greater than 1 only at nodes that are direct predecessors of a faucet. The slack variables τ_t determine which orifices to use. If $\tau_t > 0$, the inner diameter of the orifice is $d_t = 0.62\tau_t^{-0.25}\bar{\psi}^{0.5}$.

Adjusting the Design. Simulation model (A.1) is used to test the performance of the design generated by the heuristic over a sufficiently large sample of configurations. If the resulting design is unsatisfactory at some faucets, typically because of too high of an occurrence of flows below a tolerance level, the user can rerun the design module with alternative parameters.

The procedure of choice is to make headloss constraints at intermediary nodes more stringent by stating that headloss should be less than a given fraction of the gravity pressure. In the design model, this consists of dividing the right side of the constraint in (A.5c) by the safety factor. The inverse of this fraction plays the role of a typical engineering safety factor. There are several ways to implement the safety factor. The most straightforward one is to impose the same safety factor at all intermediary nodes. As we explain in the Case Study section, a satisfactory factor is found after very few trial-and-error attempts. A slightly less simple way is to apply different safety factors at intermediary nodes to account for topographical and network characteristics. Unfortunately, after years of experience, no clear rules have emerged to anticipate which faucets might be critical. The user is also free to adjust other design parameters, such as quality of service, but any impact from doing so proved to be far less conclusive.

Probability Model and Load Factor. User water consumption is a certain stochastic process. Faucets are opened, then closed, a specific number of times during the peak hour. The opening and closing times are random according to independent, identical distributions, which are the same for all users. During the peak period, users open their faucets during some fraction r of the peak-period duration. The goal is that, during consumption time, flow will match the target flow $\bar{\psi}$. Under this scheme, we can say at any time during the peak period, there is a probability r that a given faucet will be open, independently of all other faucets.

Consider now a branching node with n downstream faucets. We are interested in the number N of faucets open downstream at some point of time. The random variable N is binomial with parameters (n, r). It can be used to describe the desired flow $N\bar{\psi}$ in the branch adjacent to the node upstream. Note that when a faucet is closed, the characteristic of the network is of no concern for this node. Therefore, the case $N = 0$, which occurs with probability $(1 - r^*)^n$, is not relevant in the computation of the load factor. The relevant quantity is the conditional random variable (i.e., the number of faucets open downstream under the condition that at least one downstream faucet is open). Using a conditional expectation argument, it is easy to show that the variable of interest is $\tilde{N} = 1 + X_{n-1,r}$, where $X_{n-1,r}$ is the binomial distribution and 1 stands for the contribution of an open faucet.

Let F be the cumulated distribution of $\tilde{N}$. Suppose we assign to the pipe the flow $\lambda\bar{\psi}$, with $\lambda < n$. The resulting quality of service is $F(\lambda) = \text{Prob}\{\tilde{N}\bar{\psi} \le \lambda\bar{\psi}\}$. Now, if the desired quality of service has a given value σ, the load factor is $F^{-1}(\sigma)$. Because $\tilde{N}$ takes integer values only, we typically have

$$F(F^{-1}(\sigma) - 1) < \sigma \le F(F^{-1}(\sigma)).$$

Instead of taking $\lambda = F^{-1}(\sigma)$, we interpolate between to the two values

$$\lambda = F^{-1}(\sigma) - \frac{F(F^{-1}(\sigma)) - \sigma}{F(F^{-1}(\sigma)) - F(F^{-1}(\sigma) - 1)}. \tag{6}$$

References

Alperovits E, Shamir U (1977) Design of optimal water distribution systems. *Water Resources Res.* 13(6):885–900.

Babonneau F, Nesterov Y, Vial JP (2012) Design and operations of gas transmission networks. *Oper. Res.* 60(1):34–47.

Cross H (1936) Analysis of flow in networks of conduits or conductors. Report, University of Illinois at Urbana–Champaign, Urbana.

Eiger G, Shamir U, Ben-Tal A (1994) Optimal design of water distribution systems with networks. *Water Resources Res.* 36(9): 2637–2646.

Maugis JJ (1977) Etude de réseaux de transport et de distribution de fluide. *Oper. Res.* 11(2):243–248.

MOSEK ApS (2015) The MOSEK optimization toolbox for MATLAB manual. Version 7.1 (Revision 28). Accessed October 15, 2018, https://docs.mosek.com/7.1/toolbox/index.html.

Rossman LA (1994) The EPANET water quality model. Coulbeck B, ed. *Integrated Computer Applications in Water Supply, Volume 2: Applications and Implementations for Systems Operations and Management* (John Wiley & Sons, Chichester, UK), 79–93.

Rossman LA (2000) *EPANET 2, User's Manual* (U.S. Environmental Protection Agency, Cincinnati).

Todini E, Pilati S (1987) A gradient algorithm for the analysis of pipe networks. Coulbeck B, Orr C-H, eds. *Computer Applications in Water Supply, Volume 1: Systems Analysis and Simulation* (Research Studies Press Ltd., Letchworth, UK), 1–20.

Frédéric Babonneau is assistant professor in operations at the Business School at Adolfo Ibáñez University in Chile, CEO-Founder of OR Decision in Chile and VP-Partner at ORDECSYS, an R&D consultancy based in Geneva. He holds a PhD in economic and social sciences from the University of Geneva in 2006. He is specialized in energy modeling, risk management, and nondifferentiable and large-scale optimization.

Gilles M. Corcos holds a PhD in aeronautical engineering from the University of Michigan and a Doctorat d'Etat in physics from the University of Grenoble, France. He is the cofounder of Dicon, a California fiberoptics components designing and manufacturing enterprise and of Agua Para La Vida, a California NGO helping poor villages of Nicaragua secure potable drinking water. He is a fellow of the American Physical Society known for his research in fluid mechanics.

Laurent Drouet is a senior researcher in integrated assessment modelling at the RFF-CMCC European Institute on Economics and the Environ in Italy. He holds a PhD in economics and social sciences from the University of Geneva. He currently coleads the development of the WITCH model. He was involved in the implementation of many model components: energy system, economy, land-use, climate, air pollution, mitigation options, and climate change impacts.

Jean-Philippe Vial holds PhDs in OR from the University of Louvain and a Doctorat d'Etat in mathematics from the University Paris-Dauphine. He authored one book and over 100 articles in operations research and mathematics, which appeared in major journals. He chaired the Mathematical Programming Society in 1998. He is the cofounder of ORDECSYS, a consulting company in economic and environmental impacts. He is on the board of APLV, the ONG for which he developed the software NeatWork.