Stochastic Design of Water Distribution Systems with Expected Annual Damages

Y. R. Filion[1]; B. J. Adams, M.ASCE[2]; and B. W. Karney, M.ASCE[3]

Abstract: This paper presents a stochastic design approach that quantifies the expected annual damages sustained by residential, commercial, and industrial users during low- and high-pressure hydraulic failures in a water network. The approach, which couples stochastic models of water demand, fire flow, and pipe breaks with Monte Carlo simulation, was used to solve part of the Anytown design problem. Results indicated that a significant proportion of low-pressure failures occurred during low-demand months in the last 10 years of the planning period. The timing and spatial distribution of demands observed during failure differed significantly from the demands assumed in conventional design (maximum hour and maximum day demand+fire). The results also indicated that including damages in design makes it possible to produce cost effective systems that yield a low level of expected annual damages. This gives force to framing the network design problem as a stochastic, multiobjective one to balance cost efficiency with system capacity/redundancy and provide a hedge against hard-to-anticipate temporal and spatial patterns of demand in networks.

DOI: 10.1061/(ASCE)0733-9496(2007)133:3(244)

CE Database subject headings: Water distribution systems; Damage; Stochastic processes; Design.

Introduction

The hydraulic design of water distribution systems is typically governed by the requirement to provide pressure above a minimum standard when the network is experiencing a worst-case loading—usually the greater of maximum hour peak demands, or maximum day demands and a fire at a critical node. The implicit objective of this design standard is to provide enough pressure and flow to control whatever fire that might erupt in the network, or to deal with an unforeseen public emergency during peak conditions, in order to minimize the probability of losing human life and property. However, because the design problem is posed as a deterministic one (the design loads are not treated as random variables), the designer has no means of knowing how frequently design loads recur, and more importantly, the frequency and damage of hydraulic failures that result if the minimum pressure requirement is not met under peak demands. Often, the supposition is that if certain design practices are adopted such as closing loops, including redundant pumping capacity and tank storage, then the network will provide pressures at or above the required minimum during peak demands and the probability of hydraulic failure will be small—albeit still unspecified—as will the damages in the system.

[1]Assistant Professor, Dept. of Civil Engineering, Queen's Univ., Kingston, Canada K7L 3N6. E-mail: yves.filion@civil.queensu.ca

[2]Professor, Dept. of Civil Engineering, Univ. of Toronto, Toronto ON, Canada M5S 1A4. E-mail: adams@ecf.utoronto.ca

[3]Professor, Dept. of Civil Engineering, Univ. of Toronto, Toronto ON, Canada M5S 1A4. E-mail: karney@ecf.utoronto.ca

Note. Discussion open until October 1, 2007. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on December 23, 2003; approved on April 19, 2006. This paper is part of the ***Journal of Water Resources Planning and Management***, Vol. 133, No. 3, May 1, 2007. ©ASCE, ISSN 0733-9496/2007/3-244–252/$25.00.

Although adopting design practices such as closing pipe loops increases the hydraulic capacity and ability of a system to provide drinking water at or above a minimum pressure under peak demands, they do not eliminate the possibility of hydraulic failure causing catastrophic damages (e.g., loss of life and property). In large systems with numerous pipes and pumps, there exists many combinations of peak demands, pipe breaks, and low tank levels that can cause low pressures at times when high pressures are needed to deal with emergencies. Despite this, constraints on time and resources allow a designer to consider only a very small subset of loads—usually the greater of maximum day+fire or maximum hour combined with a small number of contingencies (e.g., trunk main break, major pump outage). This conventional design regime makes it difficult for the design engineer to assess the potential for damages in a system once it has been built and put into operation.

To begin to address this shortcoming, researchers have developed chance-constrained optimization schemes to design water distribution systems at minimum cost under a wide range of demands and other loads. In these studies, demands and the hydraulic conductivity of pipes at the end of the design life of a system are treated as uncertain quantities, and modeled with probability density functions (PDF). The studies by Lansey et al. (1989), Xu and Goulter (1999), Kapelan et al. (2004), Tolson et al. (2004), and Babayan et al. (2005) represent a worthwhile research effort to conduct least-cost design of networks under random loading.

These previous studies have focused strictly on the hydraulic performance in systems by computing the proportion of time pressure or flow falls below some minimum level at a node. Here, residential, commercial, and industrial users are only included in the analysis insofar as they create a demand for water in the system. No consideration is given to how users are affected, and what damages they might suffer, during shortfalls in pressure or flow. The present work seeks to broaden the boundaries of the analysis to account for the economic damages sustained by residential, commercial, and industrial users during low-pressure and high-pressure hydraulic failures. Incorporating economic dam-

J. Water Resour. Plann. Manage. 2007.133:244-252.

ages in design has the potential of redefining what is optimal, and at a more practical level, of introducing some flexibility through the careful balancing of system cost with damages.

This paper also presents a new stochastic design approach that couples generating models of water demand, fire flows, and pipe breaks with a Monte Carlo simulation (MCS) algorithm. The MCS algorithm is used to generate stochastic demand and pipe break loads to compute expected annual damages over an extended planning period. In this paper, the stochastic design approach is applied to the Anytown design problem to begin to answer four important questions in relation to the long-term hydraulic behavior and design of systems: Do low pressures mostly occur during high-demand months as is assumed in conventional design? What loading factors most often cause low pressures and economic damages? Is the stochastic design approach better suited than the conventional approach to design systems under variable loads? What are the implications of considering economic damages in system design optimization and decision making?

This paper is organized in two parts. In the first part, a method for computing expected annual damages as well as a new stochastic design approach are presented. In the second part, the stochastic design approach is applied to the Anytown design problem to address the four research questions posed previously.

Failure Consequences and Damage Costs

Generally, three different types of failures are possible in a water distribution system—namely, structural, hydraulic, and water quality failures. For each failure type, there is a response variable(s) with minimum and maximum limits that delineates failure from nonfailure. For example, a hydraulic failure occurs when the response variable of pressure falls below a minimum level—usually for fire protection—or exceeds a maximum level (these limits are often fuzzy). Once failure has occurred, one or more consequences can materialize. For example, a low-pressure failure can inconvenience residential and commercial users, interrupt industrial production, cause pipes to collapse, and, in the event of a fire, cause the loss of property. A high-pressure failure can cause pipes to burst with flooding damage to surrounding property. Note that these consequences are not mutually exclusive and that for each consequence, market-based indicators are used to estimate their damages. For example, the loss of property to fire can be reduced to economic terms by means of the insured value of the buildings consumed by fire.

The use of market-based estimators is justified for damages to the system and damages to property that are directly associated with commodities (e.g., materials, land, and labor) for which market prices reflect their social value. Damages arising from death and human injury are not traded in the market and thus have no direct economic value. Nonmarket valuation techniques such as willingness-to-pay and contingent valuation developed in the field of economics may prove useful to estimate the value of human life and the costs of human injury. Despite this, one should never lose sight of the fact that these techniques will never fully account for the profound impact system failure can sometimes have on a human life.

Expected Annual Damages

A damage function that "maps" a response variable (pressure) level at a system node to a unique average level of damages [Fig. 1(b)] forms the basis for calculating expected annual damages. For clarity of exposition, the response variable of pressure is used in the discussion that follows. The function takes a "bathtub" shape as indicated in Fig. 1(b). The reasoning for this is as follows: At very low pressures (e.g., pressure below s_{low}) the possible consequences include inconvenience to residential and commercial users, interruption of industrial production, pipe collapse, and loss of life and property to uncontrolled fires. However, one can reasonably assume that damages are dominated by the loss of life and property resulting from an inability to suppress fire at a system node. At very high pressures (e.g., pressures above s_{high}), damages are dominated by pipe breaks and impairment to property. The middle pressure range between s_{low} and s_{high} yields negligible damages.

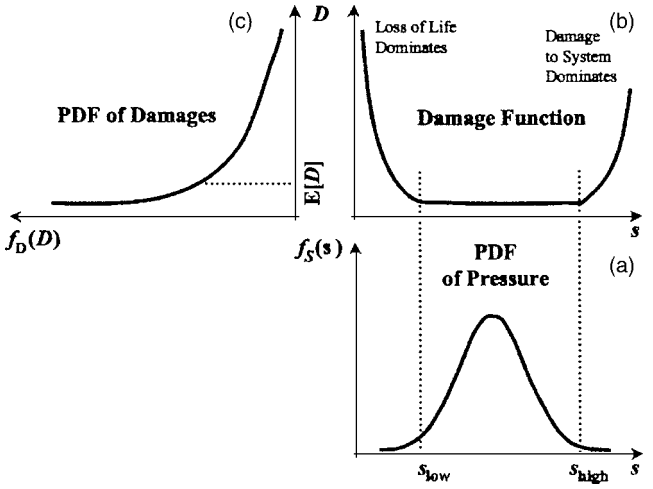

Fig. 1. PDF of pressure, continuous damage function, and derived PDF of damages

Two simplifications are necessary to develop an operational damage function. First, no attempt is made to associate a response variable level (pressure) to the precise level of damages observed in the field. In reality, if a particular response level is encountered frequently (e.g., pressures below s_{low}), damages will vary in each instance depending on the social and economic circumstances that exist at the time the system node experiences low pressures. Instead, the damage function only associates an average level of damages to each response variable level. And second, different types of damages must be scaled to a common unit (monetary value) if a damage function is to be amenable to numerical treatment. A pipe burst is amenable to numerical scaling because its economic or utility value can be defined, whereas the loss of a human life is not easily scalable.

Once a damage function is defined, an estimate of expected annual damages is calculated by integrating the product of a continuous damage function with an empirical PDF of pressure over the feasible range of pressures (Fig. 1) such as

$$E[D] = \int_{s_0}^{s_1} D(s) f_S(s) ds \tag{1}$$

where $E[D]$=expected annual damages (\$/year); s_0, s_1=lower and upper limits of response variable (pressure) observed (m); $D(s)$=damage function that associates response variable s to an average level of damages (\$); and $f_S(s)$=probability density function of response variable s.

J. Water Resour. Plann. Manage. 2007.133:244-252.

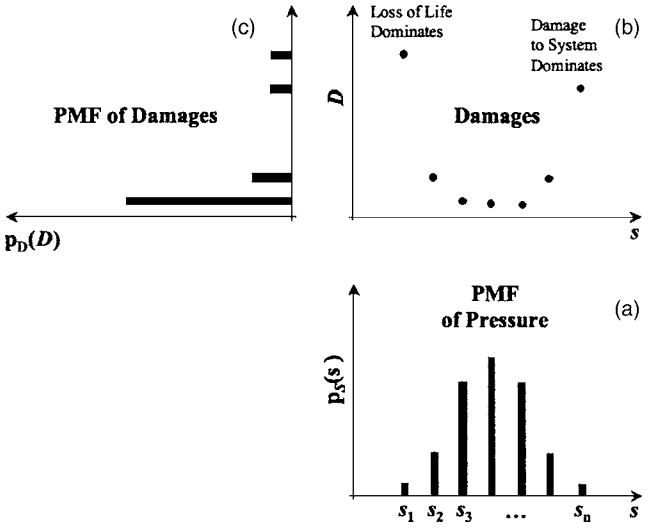

Fig. 2. Probability mass function (PMF) of pressure, discrete damages, and derived PMF of damages

In practice, the scarcity of field data with which to tabulate damages may constrain the engineer to estimate only a few points on a damage function. To accommodate this case, expected annual damages in Eq. (1) is recast in a discrete form. The probability density function of the response variable (pressure) is transformed into a probability mass function [Fig. 2(a)] by integrating probability densities over $s_1, s_2, s_3, \ldots, s_n$ ranges of the response variable. Each range of the response variable can be thought of as a variable "state" with a finite probability mass $p_S(s_i)$. Expected annual damages are estimated by summating damages over all variable states

$$E[D] = \sum_{i=1}^{S} D_i \, p_S(s_i) \tag{2}$$

where D_i=damages incurred when pressure state s_i encountered (\$); $p_S(s_i)$=probability mass function of variable state s; and S=number of pressure states.

The implicit assumption in the development of Eq. (2) is that damages D_i are incurred every time variable state s_i is encountered in a system. This is not necessarily true in real systems. When a system finds itself in pressure state s_i, damages may or may not arise. For example, fires do not erupt and cause property damage at a node every time pressure falls below a minimum level for fire protection. By the same token, pipes do not burst every time pressure rises above a maximum level. Expected annual damages in Eq. (2) is expanded to include a conditional probability mass function that accounts for the fact that damages are incurred only a fraction of the time is in a particular variable state s_i

$$E[D] = \sum_{i=1}^{S} \sum_{j=1}^{KK} D_{ij} \, p_{K_{ij}|s_i}(K_{ij}|s_i) \, p_S(s_i) \tag{3}$$

where $p_{K_{ij}|s_i}(K_{ij}|s_i)$=conditional probability that variable state s_i results in consequence j, K_{ij}, with damage level D_{ij} and KK=number of possible consequences when system is in variable state s_i. Expected annual damages as computed with Eq. (3) implies that each variable state s_i can produce $j=1,2,3,\ldots,KK$ different consequences. For example, a pressure head below 14.0 m H_2O at a node can result in the loss of property if there is a fire while the node is experiencing a low pressure. At the same time, a pressure head below 14.0 m H_2O can produce delays in industrial production resulting in economic losses. In this paper, it is assumed that each variable state, s_i, results in only one consequence K_i whose damages D_i dominate all other damages. This important assumption simplifies Eq. (3) to yield

$$E[D] = \sum_{i=1}^{S} D_i \, p_{K_i|s_i}(K_i|s_i) \, p_S(s_i) \tag{4}$$

where $p_{K_i|s_i}(K_i|s_i)$=conditional probability that variable state s_i results in consequence K_i with damage level D_i. Note that as the full PDF of damages is indicated in Figs. 1 and 2, expected annual damages in Eq. (4) only compute the mean of damages.

The damage functions indicated in Figs. 1 and 2 assume logical shapes but remain theoretical. It is beyond the scope of this paper to determine the specific shape of these damage functions. The shape of a damage function will vary from one system to the next, and can be ascertained through field data. However, so long as a damage function follows a logical shape, the stochastic design approach will yield a reasonable picture of damages.

Monte Carlo Simulation Algorithm

Expected annual damages as defined previously are computed with a MCS algorithm whose overall structure is outlined in the following. First a candidate design (e.g., pipe sizes) is chosen and simulated under a number of future loading scenarios. As indicated in Table 1, a simulation is organized in $l=1,2,3,\ldots,N$ independent runs (each initiated with a random seed) that are divided into $m=1,2,3,\ldots,T$ time steps lasting 1 day. At each time step in a run, water demands, fire flows, and pipe breaks are generated and "loaded" into the EPANET2 network solver (Rossman 2000) to compute pressure head at a node(s) of interest. For

Table 1. Structure of Monte Carlo Simulation Algorithm

	Time (day)							
Run	$m=1$	2	3	⋯	T	Annual cost (dollars/year)	Annual damages (dollars/year)	Annual damages squared (dollars/year)^2
$l=1$	$h_{1,1}$	$h_{1,2}$	$h_{1,3}$	⋯	$h_{1,T}$	C_1	D_1	D_1^2
$l=2$	$h_{2,1}$	$h_{2,2}$	$h_{2,3}$	⋯	$h_{2,T}$	C_2	D_2	D_2^2
$l=3$	$h_{3,1}$	$h_{3,2}$	$h_{3,3}$	⋯	$h_{3,T}$	C_3	D_3	D_3^2
⋮	⋮				⋮	⋮	⋮	⋮
$l=N$	$h_{N,1}$	$h_{N,2}$	$h_{N,3}$	⋯	$h_{N,T}$	C_N	D_N	D_N^2
						$E[C]$	$E[D]$	$E[D^2]$

J. Water Resour. Plann. Manage. 2007.133:244-252.

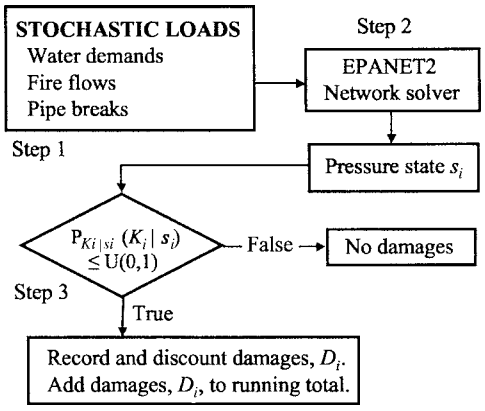

Fig. 3. Steps to compute damages in a MCS run

example, in Table 1, the pressure computed in the second time step of the first run is denoted by $h_{1,2}$. If the pressure computed exceeds some minimum or maximum level (see Figs. 1 and 2), then economic damages (if any) are recorded at that time step. This procedure is repeated over $m=1,2,3,\ldots,T$ time steps in a run. At the end of run l, annual damages, D_l, and annual damages squared, D_l^2, are computed by tallying damages recorded over the $m=1,2,3,\ldots,T$ time steps in that run. After reaching the end of the last run, expected annual damages $E[D]$ in Eq. (4) are estimated by taking the arithmetic mean of annual damages, D_l, across $l=1,2,3,\ldots,N$ runs. Similarly, the expectation of annual damages squared, $E[D^2]$, is estimated by taking the arithmetic mean of annual damages squared, D_l^2, across $l=1,2,3,\ldots,N$ runs in Table 1. Note that annual damages, D_l, are calculated over $m=1,2,3,\ldots,T$ time steps in a single run, whereas expected annual damages, $E[D]$, is computed as the arithmetic mean of annual damages, D_l, over $l=1,2,3,\ldots,N$ runs. The variance of annual damages, D_l, is calculated with the following equation through the well-known property of the variance:

$$V[D] = E[D^2] - E[D]^2 \tag{5}$$

The variance of annual damages in Eq. (5) is then used to compute the variance of expected annual damages, $E[D]$, in

$$V[E[D]] = V[D]/N \tag{6}$$

Also, at the end of run l, the capital and pumping costs associated with the candidate design are discounted and annualized over the length T of the run to compute annual cost, C_l. Annual cost is a random variable because pumping costs are based on stochastic demands and pressures. Expected annual cost $E[C]$ is computed by taking the arithmetic mean of annual cost, C_l, over $l=1,2,3,\ldots,N$ runs (Table 1). Once expected annual cost, $E[C]$, expected annual damages, $E[D]$, and the variance of expected annual damages, $V[E[D]]$, have been computed, the next candidate design is generated and the previous steps are repeated.

It is also worth describing the activities that unfold in each time step of a run. These activities are divided into three parts which are outlined in detail in Fig. 3. First, water demands, fire flows, and pipe breaks are generated with the stochastic models described in the next section. These loads are entered into the EPANET2 network solver (Rossman 2000). Second, the solver is run to compute nodal pressure head(s). The nodal pressure head(s) is associated with its appropriate pressure range, or pressure state s_i. In this paper, three pressure states are considered (Table 2). For example, pressure state s_2 in Table 2 corresponds to pressures that range between 14.0 and 26.0 m H_2O. In the third part, the MCS verifies if event K_i with damages D_i occurs (or does not occur) when the system is in pressure state s_i. This verification is performed by generating a uniform random variate $U(0,1)$ and comparing it to the conditional probability $p_{K_i|s_i}(K_i|s_i)$ in Table 2. If the uniform random variate $U(0,1)$ is greater than $p_{K_i|s_i}(K_i|s_i)$, then event K_i does not occur and no damages are incurred. Conversely, if the uniform random variate $U(0,1)$ is smaller than or equal to $p_{K_i|s_i}(K_i|s_i)$, then event K_i occurs and damages D_i are incurred. For example, if the system is in pressure state s_2 (low pressure between 14.0 and 26.0 m H_2O) in Table 2 and if $U(0,1)$ is smaller than or equal to the conditional probability value 1/10 probability/day, then event K_2 occurs (backup pumps on industrial properties fail when system is experiencing low pressures) and $20,000 in damages (interruption of industrial production) are sustained during that day of service. If damages are recorded in the third part, they are discounted to the beginning of the planning period, annualized, and added to the running total of damages. The three parts above are repeated at the next time step until the simulation reaches the end of the run.

Stochastic Water Demand

In this paper, municipal demand $q_k(t)$ is comprised of base, seasonal, and stochastic components. The base and seasonal components are denoted by $f_k(t)$, whereas the stochastic component is denoted by $x_k(t)$ in

$$q_k(t) = f_k(t) + x_k(t) \tag{7}$$

The base demand is the volume of water needed to satisfy minimal residential, commercial, and industrial needs in a city. It correlates strongly with indoor water use for these user classifications and it is calculated as the average water use during the winter months. Base demand tends to increase (decrease) over a period of years owing to structural changes like an increase in household income and new development. The seasonal component coincides with outdoor, residential water use and is driven by air temperature. It exhibits a sinusoidal pattern with high demands during the summer growing season when residential customers water their lawn and garden, and low demands during the winter season when there is little or no outdoor water use. Base and

Table 2. Damage Function for Anytown Example

Pressure range (m)	Pressure state (s_i)	Conditional probability $p_{K_i|s_i}(K_i|s_i)$ (prob)	Average damages (D_i)	Damage type
$s<14.0$	s_1	Pr(fire) ≈ 1/3,650	$4,000,000	Type 1: loss of life and property
$14.0\leq s<26.0$	s_2	1/10	$20,000	Type 2: interruption of industrial production
$s\geq 88.0$	s_3	1/25	$100,000	Type 3: damage to system pipes

J. Water Resour. Plann. Manage. 2007.133:244-252.

seasonal demands are modeled as time-varying, deterministic variables with a regression line and cosine function defined over $k=1,2,3,\ldots,n$ network nodes in

$$f_k(t) = \mu_k^{(0)} + \left(\frac{\mu_k^{(1)} - \mu_k^{(0)}}{T}\right)t + \alpha_k \cos\left(\frac{2\pi}{365}t + \tau_k\right) \quad (8)$$

where $\mu_k^{(0)}$, $\mu_k^{(1)}$=mean municipal demand at node k at start and end of run (lps); and α_k=demand amplitude at node k (lps). Amplitude accounts for the seasonal variation of demand above and below the mean level defined by $\mu_k^{(0)}$ and $\mu_k^{(1)}$ in Eq. (8); τ_k=phase angle at node k (rad); T=number of time steps in run (day); and t=time (day). Note that a single cosine function with a fundamental period of 365 days is used to model the seasonal component. Generally, periodic regression is used to simulate the seasonality of demand over a number of significant harmonics.

The municipal demand also exhibits a strong stochastic character owing to short-term fluctuations in rainfall and daily use. This stochastic component is denoted by $x_k(t)$ in Eq. (7). A lag-1 Markov model is used to simulate the stochastic component of demand and to account for lag-1 serial correlation between observations of demand. The lag-1 Markov model in the following equation has an underlying normal distribution (the superscript y denotes parameters associated with a normal distribution):

$$x_k^y(t) = \rho_k^{(1,y)} x_k^y(t-1) + \varepsilon_k(t)\sigma_k^{(y)}[1 - (\rho_k^{(1,y)})^2]^{0.5} \quad (9)$$

where $x_k^y(t)$, $x_k^y(t-1)$=normal demand deviate in the current and previous days at node k (lps). A normal demand deviate is the difference between normal demand, $q_k^y(t)$, and the base and seasonal normal demand, $f_k^y(t)$, or $q_k^y(t)-f_k^y(t)$; $\sigma_k^{(y)}$=standard deviation of normal demand at node k (lps); $\rho_k^{(1,y)}$=lag-1 serial correlation coefficient of normal demand at node k; $\varepsilon_k(t)$=standard normal variate.

In this paper, the lag-1 Markov model in Eq. (9) is coupled with a three-parameter lognormal distribution with a lower bound of 0 and an unbounded maximum to generate lognormal demands. Lognormal demands are generated with Eqs. (7)–(9) in a number of steps. First, the lognormal parameters $\mu_k^{(0)}$, $\mu_k^{(1)}$, α_k, τ_k, σ_k^2, and $\rho_k^{(1)}$ are computed from field data. The lognormal variance σ_k^2 and lag-1 serial correlation coefficient $\rho_k^{(1)}$ are transformed into their normal equivalent $\sigma_k^{2(y)}$ and $\rho_k^{(1,y)}$ via

$$\sigma_k^2 = \left[\frac{\mu_k^{(0)} + \mu_k^{(1)}}{2}\right]^2 [\exp[\sigma_k^{2(y)}] - 1] \quad (10)$$

$$\rho_k^{(1)} = \frac{\exp[\sigma_k^{2(y)}\rho_k^{(1,y)}] - 1}{\exp[\sigma_k^{2(y)}] - 1} \quad (11)$$

The lognormal parameters $\mu_k^{(0)}$, $\mu_k^{(1)}$, α_k, and τ_k are substituted into Eq. (8) to compute $f_k(t)$. The deterministic component of demand, $f_k(t)$, is then transferred to normal space $f_k^y(t)$ via

$$\sigma_k^2 = f_k(t)^2[\exp[\sigma_k^{2(y)}] - 1] \quad (12)$$

$$f_k(t) = \exp\left[f_k^y(t) + \tfrac{1}{2}\sigma_k^{2(y)}\right] \quad (13)$$

Note that the normal variance $\sigma_k^{2(y)}$ in Eq. (12) is calculated strictly as an intermediate value to update the normal parameter $f_k^y(t)$ in Eq. (13). The normal variance $\sigma_k^{2(y)}$ in Eq. (12) is different from that calculated with Eq. (10).

With the normal parameters $\sigma_k^{2(y)}$ and $\rho_k^{(1,y)}$, the lag-1 Markov model in Eq. (9) is used to generate a normal demand deviate $x_k^y(t)=q_k^y(t)-f_k^y(t)$. The normal, deterministic demand $f_k^y(t)$ calculated in Eq. (13) is then added to the normal demand deviate $q_k^y(t)-f_k^y(t)$ generated with the lag-1 Markov model to yield a normal demand $q_k^y(t)$. Once a normally distributed demand has been generated, its antilog is taken to yield a lognormal demand $q_k(t)$. The previous steps are repeated in each time step of a simulation.

Stochastic Fire Flow

A stationary Poisson process is used to simulate the eruption of fires at each node. Although the statistical behavior of fires in urban areas has not been widely studied, it is reasonable to assume that fires occur independently both in time and in space. The Poisson process, with its assumption of time independence is thus well suited to model the occurrence of fires in a distribution network. The discrete Poisson distribution is written as

$$p_X(x) = \beta^x \exp(-\beta)/x! \quad (14)$$

where x=number of individual fires in a time step; $p_X(x)$=probability that x fires erupt at a node during a single time step (probability/step); β=fire rate at a node (fires/year). The fire rate is calculated by dividing the number of fires at a node by the time period over which these fires were observed.

In water distribution network modeling, it is common to assign users of different classifications (e.g., residential, commercial, and industrial) to a single network node. The aggregation of users at a single network node creates the possibility (however unlikely) that one or more fires can erupt at a node during a single time step. In this chapter, a fire event is defined as the eruption of one or more fires at a network node within a time step. The probability of a fire event, or $\Pr[x \geq 1]$, is calculated by summating the discrete Poisson distribution in Eq. (14) from 1 to infinity to give $\Pr[x \geq 1] = 1 - \Pr[x=0] = 1 - \exp(-\beta)$. The MCS assumes that fire events are suppressed and dealt with in a single time step.

When a fire event does occur at a node, the fire flow required to suppress the x fires that constitute the fire event may vary according to some probability distribution. Here, fire flow is provisionally assumed to follow a normal distribution with mean $\mu_{k,f}^{(y)}$ and standard deviation $\sigma_{k,f}^{(y)}$

$$F_k^y(t) = \mu_{k,f}^{(y)} + \sigma_{k,f}^{(y)}\varepsilon_k(t) \quad (15)$$

where $F_k^y(t)$=normal fire flow required to suppress x fires at a node in a time step (lps); and $\varepsilon_k(t)$=standard normal variate.

In this work, a three-parameter lognormal distribution with a lower bound of 0 and an unbounded maximum is used in conjunction with Eq. (15) to generate lognormal fire flows in a number of steps. First, the mean $\mu_{k,f}$ and standard deviation $\sigma_{k,f}$ of lognormal fire flows are determined with field data. The lognormal parameters are then transformed into their normal equivalent $\mu_{k,f}^{(y)}$ and $\sigma_{k,f}^{(y)}$ with

$$\sigma_{k,f}^2 = \mu_{k,f}^2[\exp[\sigma_{k,f}^{2(y)}] - 1] \quad (16)$$

$$\mu_{k,f} = \exp\left[\mu_{k,f}^{(y)} + \tfrac{1}{2}\sigma_{k,f}^{2(y)}\right] \quad (17)$$

Once the normal distribution parameters $\mu_{k,f}^{(y)}$ and $\sigma_{k,f}^{(y)}$ are found, they are substituted into Eq. (15) to generate a normally distributed fire flow $F_k^y(t)$. The antilog of this normally distributed fire flow is taken to yield a lognormal fire flow $F_k(t)$.

J. Water Resour. Plann. Manage. 2007.133:244-252.

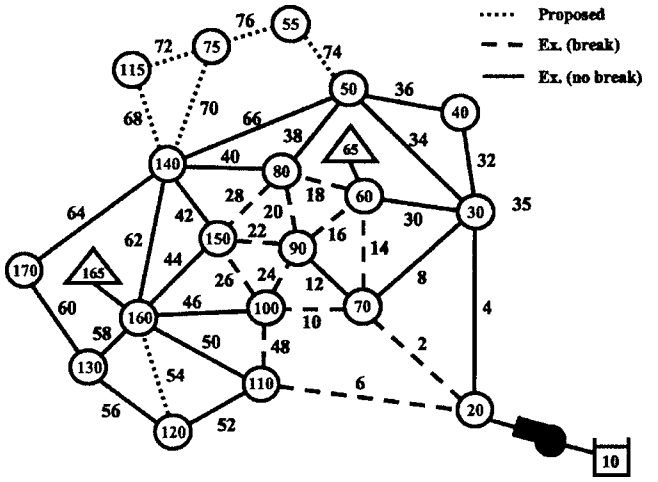

Fig. 4. Anytown network of Walski et al. (1987)

Stochastic Pipe Breaks

Pipe breaks are also modeled as a stationary Poisson process for the sake of simplicity. Although there is no consensus on which stochastic process best represents pipe breaks, a number of researchers have in the past used the Poisson process to model the occurrence of breaks in individual pipes (Guercio and Xu 1997; Shinstine et al. 2002). Given the great flexibility of MCS, it is eminently possible to substitute a more complex pipe-break model if appropriate. The discrete Poisson distribution for pipe breaks is

$$p_Y(y) = \xi^y \exp(-\xi)/y! \tag{18}$$

where y=number of individual pipe breaks in a time step; $p_Y(y)$=probability that y pipe breaks occur in a time step (probability/step); and ν=break rate in a pipe (breaks/km/year).

The notion of a pipe break *event* is useful when simulating breaks in a MCS. A pipe break event is defined as the occurrence of one or more breaks in a pipe, in a time step. The probability that a pipe break event occurs in a time step is calculated by summating the discrete Poisson distribution from 1 to infinity or $\Pr[y \geq 1] = 1 - \Pr[y = 0] = 1 - \exp(-\nu)$. The MCS assumes that one or more pipe breaks (pipe break event) are repaired within a time step.

Anytown Example

The Anytown system is chosen because it is well documented in the literature, and because it has a topological complexity typical of many real-world systems. In the example that follows, the stochastic design approach is applied to the Anytown problem to compute economic damages. The results are compared to those reported in Walski et al. (1987).

MCS Parameters

A set of 300 candidate designs are generated with a random number generator. In each design, Pipes 54, 68, 70, 72, 74, and 76 in Fig. 4 are each randomly assigned a commercially available diameter. Each design is subjected to a MCS comprised of 1,000 independent runs divided into 7,300 1 day time steps (1985–2005). As the time step length in the MCS is set to 1 day, demand and pressure are averaged over a single day (diurnal fluctuations not considered).

Node and Demand Parameters

The node and demand data for the Anytown system is taken from Walski et al. (1987). The beginning of the simulation, $t=0$, coincides with January 1, 1985. The mean of lognormal demand at the start and end of the simulation, $\mu_k^{(0)}$ and $\mu_k^{(1)}$, are set to the average day demand for the years 1985 and 2005. The phase angle τ_k is set to -3.14 rad to ensure that the maximum day demand is encountered on the 182nd day of each calendar year (July 1st). To produce a maximum day peaking factor of 1.3, as assumed in Walski et al. (1987), on July 1st of each calendar year, the demand amplitude is set to $\alpha_k = 0.3\ \mu_k^{(1)}$ in Eq. (8). The standard deviation of lognormal demand is set to 18% of the average day demand (year 2005), or $\sigma_k = 0.18\ \mu_k^{(1)}$. The lag-1 serial correlation coefficient of lognormal demand, $\rho_k^{(1)}$, is set to 0.1. The average fire rate β in Eq. (14) is set to 0.10 fires/year (or 2.74×10^{-4} fires/day), which means that a fire event occurs at a node once every 10 years on average.

Fire Flow Parameters

Mean lognormal fire flows $\mu_{k,f}$ are set to the deterministic values in Walski et al. (1987) (standard deviation of lognormal fire flow $\sigma_{k,f}$ is set to 0). In particular, Node 90 has a fire flow of 158 L/s, Nodes 75, 115, and 55 have a fire flow of 94.6 L/s, Nodes 120 and 160 have a fire flow of 63.1 L/s, and all other nodes have a fire flow of 31.6 L/s.

Pipe and Break Parameters

The pipe data (including commercially available diameter and unit cost) for the Anytown system is taken from Walski et al. (1987). Pipes 2, 6, 10, 14, 16, 18, 20, 22, 24, 26, 28, and 48 indicated with a dashed line in Fig. 4 can experience random breaks. These pipes have a finite break rate of $\xi = 0.3$ breaks/km/year (8.22×10^{-4} breaks/km/day) in Eq. (18).

Pump and Tank Parameters

There are three pumps in a parallel arrangement between Nodes 10 and 20. All three pumps have the same pump curve reported in Table 2 of Walski et al. (1987). As no on–off pump controls are specified, the pumps adjust to changing demand conditions by moving along their pump curve. The pumps are assumed to have an average pump efficiency of 55%. The price of electricity is set to \$0.12/kW h and the interest rate to 12%. The existing Tanks 65 and 165 in Fig. 4 have water levels that can vary between 65.6 and 77.7 m in Walski et al. (1987). In this paper, tank levels are fixed at a mean level of 71.7 m.

Damage Parameters

Damages at Nodes 55 and 170 are calculated with Eq. (4) and the data in Table 2. Three types of damages are considered in this paper. Type-1 damages occur when pressure head falls below 14.0 m at a node and a fire erupts at the time the node is experiencing this low pressure. Type-2 damages occur when pressure head is between 14.0 and 26.0 m at a node and backup pumps on industrial properties surrounding the node fail at the time the node experiences this low pressure. Type-3 damages occur when pressure head at a node rises above the structural capacity of pipes connected to the node and causes them to fail. Pipe breaks leading to Type-3 damages are assumed to have no effect on the hydraulic

J. Water Resour. Plann. Manage. 2007.133:244-252.

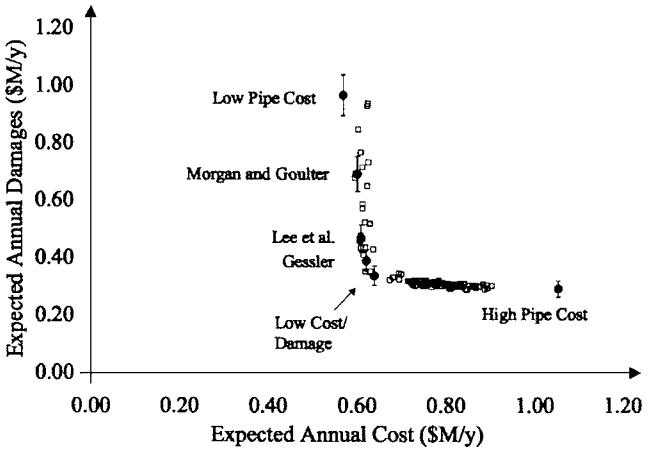

Fig. 5. Expected annual damages versus expected annual cost in Anytown network

conditions in the distribution system once they occur. This is a simplification since a pipe break inevitably leads to water loss and reduced pressures. Damages of all types are discounted at the rate of 12%. In real-world applications, it may be preferable to avoid discounting Type-1 damages sustained by humans (e.g., loss of life) as it implies that a human life is worth more now than at some later time.

It is also assumed that in Type-1 damages, fires erupt independently of pressure conditions at Nodes 55 and 170. In other words, the pressure head at a node at a particular time does not affect the probability that a fire will erupt at that time. The conditional probability $p_{K_1|s_1}(K_1|s_1)$ is set to the unconditional probability of a fire event Pr(fire)—calculated by summating Eq. (14) from 1 to infinity—in Table 2 to reflect this assumption. By contrast, Type-2 and Type-3 damages are contingent on the pressure conditions at a node. In the case of Type-3 damages, a pipe is at risk of breaking only once the pressure head at a node exceeds 88 m H_2O (Table 2).

Results of Analysis

Of the 300 designs run in the MCS, a total of 121 low-cost solutions are selected. The low-cost solutions exhibit a low expected annual cost for a given level of expected annual damages in Fig. 5. The plot in Fig. 5 shows clearly that choosing larger pipe sizes, and thus increasing the cost of a design, lowers expected annual damages. However, the cost–damages relationship is nonlinear and once expected annual cost rises above $0.63 million/year, expected annual damages are only marginally reduced. The design implication is that a significant reduction in expected annual damages is achievable with a modest capital investment. However, there is a point beyond which any additional capital investment will achieve only a marginal decrease in damages, as indicated by the flat segment of the cost–damages relationship in Fig. 5.

Confidence limits (95% level) are placed on expected annual damages in Fig. 5 to indicate their variability. (Note that limits are placed only on six representative solutions for clarity of exposition.) These limits are calculated with the variance of expected annual damages in Eq. (6). As expected annual damages, $E[D]$, is calculated as the arithmetic mean of annual damages across 1,000 runs, it is assumed that $E[D]$ follows a normal distribution $N(E[D], V[D]/N)$ by virtue of the Central Limit Theorem. The plot in Fig. 5 indicates that the statistical error of $E[D]$ increases (95% confidence limits are farther apart) as expected annual damages increase. This is owing to the fact that high-damage solutions experience a larger number of Type-1 failures which, by virtue of their high damage cost of $4 million, introduce more fluctuations in $E[D]$. Confidence limits are not placed on expected annual cost since pipe cost is deterministic and pumping cost exhibits little variability.

The solutions reported by Gessler, Lee et al., and Morgan and Goulter in Walski et al. (1987) are also run in the MCS and their cost and damages are indicated in Fig. 5. In the original Anytown problem of Walski et al. (1987), the solutions by Gessler, Lee et al., and Morgan and Goulter include new tanks, new source and booster pumps, new pipes, and the cleaning and lining of existing pipes. In this paper, these solutions are comprised of the new Pipes 54, 68, 70, 72, 74, and 76 only. The low cost/damage solution in Fig. 5 strikes a balance between expected annual cost and expected annual damages. The low pipe cost and high pipe cost solutions assign 152 and 762 mm diameters to the six new pipes and thus comprise the lower and upper bounds on cost in Fig. 5. The pipe sizes, expected annual cost, expected annual damages, and coefficient of variation of expected annual damages (calculated as $\Theta = V[E[D]]^{0.5}/E[D]$) of these six solutions are reported in Table 3.

The results in Table 3 indicate that Pipe 54 (near Node 170) in the low cost/damage solution is assigned a diameter of 305 mm which is 2–3 sizes larger than that found in the solutions of Gessler, Lee et al., and Morgan and Goulter. Equally significant, Pipe 74 that leads to Node 55 is assigned a diameter of 508 mm in the low cost/damage solution which is 3–6 sizes larger than that found in the three other solutions in Table 3. (Note that the large difference in sizes between Pipes 74 and 76 connected to Node 55 does not necessarily correspond to good design practice.) Pipes 68, 70, 72, and 76 in the low cost/damage solution are within 0–3 sizes of those found in the Gessler, Lee et al., and

Table 3. Pipe Sizes, Expected Annual Cost, Expected Annual Damages, and Coefficient of Variation of Expected Annual Damages for Anytown Design Solutions

Solution	P54 (mm)	P68 (mm)	P70 (mm)	P72 (mm)	P74 (mm)	P76 (mm)	$E[C]$ ($M/year)	$E[D]$ ($M/year)	Θ (%)
Gessler	203	305	305	152	356	152	0.62	0.39	4.9
Lee et al.	152	305	305	152	305	152	0.61	0.46	5.0
Morgan and Goulter	152	254	254	203	203	254	0.60	0.69	4.5
Low cost/damage	305	203	152	203	508	203	0.64	0.33	5.2
Low pipe cost	152	152	152	152	152	152	0.57	0.96	3.8
High pipe cost	762	762	762	762	762	762	1.05	0.29	4.9

J. Water Resour. Plann. Manage. 2007.133:244-252.

Table 4. Type-1 and Type-2 Failure Conditions Observed in Gessler, Lee et al., Morgan and Goulter, and Low Cost/Damage Solutions in Anytown Network

Row number	Failure details	Gessler (%)	Lee et al. (%)	Morgan and Goulter (%)	Low cost/damage (%)
1	Node 55 (Types 1 and 2)	6.3	9.9	26.9	7.5
2	Node 170 (Types 1 and 2)	93.7	90.1	73.1	92.5
3	Type 1 (N55 and N170)	0.4	0.4	0.5	0.3
4	Type 2 (N55 and N170)	99.6	99.6	99.5	99.7
5	Type 1 in high-demand months (% of 3)	56.6	63.3	41.8	61.2
6	Type 1 in low-demand months (% of 3)	43.4	36.7	58.2	38.8
7	Type 2 in high-demand months (% of 4)	51.7	51.6	50.7	56.8
8	Type 2 in low-demand months (% of 4)	48.3	48.4	49.3	43.2
Failure conditions during low-demand months					
9	Type 1: 1 fire, no breaks (% of 6)	92.5	98.0	100.0	97.0
10	Type 1: 1 fire, ≥1 break (% of 6)	7.5	2.0	0.0	3.0
11	Type 1: ≥2 fires (% of 6)	0.0	0.0	0.0	0.0
12	Type 2: no fires, no breaks (% of 8)	96.8	96.7	96.9	97.1
13	Type 2: 1 fire or 1 break (% of 8)	3.2	3.2	3.0	2.9
14	Type 2: 1 fire, 1 break (% of 8)	0.0	0.0	0.0	0.0
15	Type 2: ≥2 fires (% of 8)	0.0	0.0	0.0	0.0
16	Type 2: ≥2 breaks (% of 8)	0.1	0.1	0.1	0.0

Morgan and Goulter solutions. The difference in pipe sizes between the four solutions produces a marginal difference in expected annual cost (pipe cost). However, the larger sizes assigned to Pipes 54 and 74 in the low cost/damage solution "builds in" a large hydraulic capacity owing to the nonlinear relationship between pipe diameter and hydraulic resistance. A large hydraulic capacity in Pipes 54 and 74 produces surplus pressures at Nodes 55 and 170 and greatly lowers expected annual damages at these nodes. The results in Table 3 also suggest that the low cost/damage solution allocates most of its pipe cost to sizing Pipes 54 and 74 to yield low damages at Nodes 55 and 170 where damages are computed. This underscores the fact that which nodes are chosen to compute damages has a large influence on the sizing of pipes in a design. It also points to the importance of computing damages at all system nodes to remove this bias in future work.

The Gessler, Lee et al., Morgan and Goulter, and low cost/damage solutions are analyzed to uncover patterns of hydraulic failure and damages. Table 4 indicates that the majority of Type-1 and Type-2 failures occur at Node 170 (73.1–93.7%) which is located in an area with a high ground elevation (36.6 m) and which frequently experiences low pressures. Of the Type-1 and Type-2 failures that do arise at Nodes 55 and 170, only a small fraction of them are Type-1 (0.3–0.5%), whereas the rest are Type-2 (99.5–99.7%). Surprisingly, a large fraction of Type-1 (36.7–58.2%) and Type-2 failures (43.2–49.3%) occur on low-demand months (months other than June, July, and August when demand is at its highest level). Even more surprising is that the majority of Type-2 failures that arise in low-demand months do so in the absence of fires or pipe breaks in the system (96.7–97.1%). Further analysis showed that most of the low-demand, Type-2 failures occur during the spring and fall months of April, May, September, and October in the second half of the 20 year period when base demand is at its highest level. A random search of these failures revealed that many of them are caused by large demand fluctuations—demand is +0.5 to +3.5 standard deviations above the base and seasonal level $f_k(t)$—at high-use nodes (90 and 160) and at nodes located a distance from the pumping station and existing Tanks 65 and 165 (Nodes 40, 55, 75, 115, and 170). As defined earlier, Type-1 failures at Nodes 55 or 170 can only occur when there is a fire at the node in question. Therefore, all Type-1 failures that coincide with low-demand months also occur when there is at least one fire event either at Node 55 or 170 in Table 4. Type-3 failures are not recorded in any of the four solutions mentioned previously.

The water level in existing Tanks 65 and 165 is assumed fixed at 71.7 m. Fixed water levels likely yield more optimistic estimates of expected annual damages. This is because fixing tank levels at a mean level of 71.7 m yields higher pressures (and lower expected annual damages) at Nodes 55 and 170 during high-demand periods than would be observed if tank levels were allowed to fall below the mean level.

The averaging of demand and pressure over a 1 day time step "smoothes" instantaneous peak pressures. This "smoothing" of pressures further underestimates damages in the Anytown network during peak demand periods. Ideally, simulations should be performed with a 1-h time step to capture the diurnal pattern of demands and pressures. In practice, the selection of time step length is likely to be guided by constraints on computer resources, the time discretization of demand data from supervisory, control, and data acquisition (SCADA) systems, and whether steady-state or transient conditions are modeled.

Discussion of Results

Given the above results, the question remains: Is the stochastic approach better suited than the conventional approach to design systems under variable loads? While it is too early to make conclusive generalizations on this issue, the Anytown example has shown that many failures (particularly Type-2 failures) occur in low-demand months and often involve a peculiar spatial arrangement of nodal demands that cause low pressures at Nodes 55 and 170. What is significant here is that the timing and spatial configuration of the failure loads differ from those assumed in conventional design (greater of maximum hour and maximum day

J. Water Resour. Plann. Manage. 2007.133:244-252.

+fire in summer months). These results lend force to the argument that demands should be simulated as stochastic variables over an extended period to capture the sometimes unusual temporal and spatial patterns that can cause hydraulic failures.

The inclusion of expected annual damages in the Anytown design problem also raises another practical question: what are the implications of considering economic damages in system design optimization and decision making? Conventional design has focused on minimizing capital and operating cost while providing pressures above some nominal standard under specialized design loads. The inclusion of damages in the analysis implies a balancing of system cost with damages in design. The design problem then shifts from minimizing system cost alone to weighing system cost against additional capacity/redundancy to hedge against unexpected loads and reduce damages. Reframing the design problem in multiobjective terms affords a utility more flexibility in choosing a network design that best suits its capital and operational needs and that accords more closely with its tolerance for risk.

Shortcomings and Future Improvements

At present, the feasibility of the stochastic design approach is hamstrung by four factors: (1) the lack of data to calibrate the statistical parameters; (2) the complexity of the analysis required to compute these statistical parameters; (3) the high computational demands to perform a simulation; and (4) the aggregation of damages that prevents a designer from ascertaining the frequency of low-probability, high-damage failures. While small municipalities lack the necessary data needed to apply the stochastic approach, many large municipalities are now collecting demand and pipe break data with SCADA systems. Even more heartening is the fact that municipalities are deploying geographic information systems (GIS) to integrate user data (e.g., lot size, occupancy) with network models to compute demand at system nodes. GIS capabilities could be extended to connect property-value data with a network model to compute economic damages at network nodes. To be sure, the statistical analysis required to compute the moments of demand and pipe break rates in the stochastic approach discourages its widespread use in practice. However, as many engineers are entering the work force with postgraduate training and a wide exposure to statistical analysis, the writers are cautiously optimistic that in the near future, many practicing engineers will be both willing and able to apply the stochastic approach to real-world networks. Although the stochastic approach comes at a heavy computational price, the writers are also optimistic that future improvements in microcomputing and evolutionary algorithms will make the approach more feasible. And last, because expected annual damages are computed as a weighted average in Eq. (4), it is possible to record low damages and yet observe a significant number of low-probability, high-damage Type-1 failures. Thus in future work, the MCS code should be altered to count the number of Type-1 failures and weigh this last measure against expected annual damages.

Summary and Conclusions

Variations in demand, the occurrence of pipe breaks, and the eruption of fires can, on occasion, cause low-pressure failures in water networks with serious consequences and damages to the network users. The damages that are expected to be sustained by users in times of hydraulic failure should therefore be considered in design. This paper has presented a stochastic design approach to quantify the expected annual damages sustained by residential, commercial, and industrial users during low- and high-pressure hydraulic failures. The stochastic design approach coupled stochastic models of water demand, fire flow, and pipe breaks with MCS to compute expected annual damages in the well-known Anytown network. Results indicated that a significant proportion of low-pressure failures occurred during low-demand months in the second half of the 20 year planning period. Most low-pressure failures during low-demand months were caused by large fluctuations in demand at high-use nodes and at nodes located a distance from the Anytown pumping station and tanks. The timing and spatial distribution of demands was significantly different than those assumed in conventional design (maximum hour and maximum day demand+fire during summer months). The results also indicated that accounting for expected annual damages in design can help a utility strike a balance between cost and damages. In light of these findings, it was suggested that the design problem be recast as a multiobjective one that weighs cost efficiency against additional system capacity/redundancy to hedge against the surprisingly large number of temporal and spatial demand patterns that induce hydraulic failure in water networks.

Acknowledgments

The writers wish to thank the Natural Science and Engineering Research Council, Public Safety and Emergency Preparedness Canada, and the Canadian Water Network for their financial support of this research.

References

Babayan, A. V., Kapelan, Z., Savic, D. A., and Walters, G. A. (2005). "Least-cost design of water distribution networks under demand uncertainty." *J. Water Resour. Plann. Manage.*, 131(5), 375–382.

Guercio, R., and Xu, Z. (1997). "Linearized optimization model for reliability-based design of water systems." *J. Hydraul. Eng.*, 123(11), 1020–1026.

Kapelan, Z., Babayan, A. V., Walters, G. A., and Khu, S. T. (2004). "Two new approaches for the stochastic least cost design of water distribution systems." *Water Sci. Technol.*, 4(5–6), 355–363.

Lansey, K. E., Duan, N., Mays, L. W., and Tung, Y.-K. (1989). "Water distribution system design under uncertainties." *J. Water Resour. Plann. Manage.*, 115(5), 630–645.

Rossman, L. A., (2000). *EPANET users' manual*, U.S. Environmental Protection Agency, Cincinnati.

Shinstine, D. S., Ahmed, I., and Lansey, K. E., (2002). "Reliability/availability analysis of municipal water distribution networks: Case studies." *J. Water Resour. Plann. Manage.*, 128(2), 140–151.

Tolson, B. A., Maier, H. R., Simpson, A. R., and Lence, B. J., (2004). "Genetic algorithms for reliability-based optimization of water distribution systems." *J. Water Resour. Plann. Manage.*, 130(1), 63–72.

Walski, T. M., et al. (1987). "Battle of the network models: Epilogue." *J. Water Resour. Plann. Manage.*, 113(2), 191–203.

Xu, C., and Goulter, I., (1999). "Reliability-based optimal design of water distribution networks." *J. Water Resour. Plann. Manage.*, 125(6), 352–362.