Cross Correlation of Demands in Water Distribution Network Design

Yves Filion[1]; Barry Adams, M.ASCE[2]; and Bryan Karney, M.ASCE[3]

Abstract: The aim of water distribution design is to size and configure a system so that it meets existing and future demands while providing pressures above a minimum level for service and fire protection. Extended period simulation (EPS) is used in design to determine network pressures under varied diurnal demand patterns. In EPS, diurnal demands are almost invariably assumed to change in unison, or in statistical terms, to be strongly correlated in space. This paper first tests this common assumption by investigating the extent to which cross correlation in demand affects the mean and standard deviation of pressure heads in water networks, and then investigates how cross correlated demands can influence capital costs in network design. Preliminary findings from two examples indicate that the standard deviation of pressure head and capital costs can be sensitive to the level of cross correlation between nodal demands. Thus a realistic assessment of cross correlation in demand can lead to a more economical design.

DOI: 10.1061/(ASCE)0733-9496(2007)133:2(137)

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

Introduction

Modeling current and future water demand remains the most challenging activity in water distribution system design. To facilitate the process somewhat, geographic information systems (GIS) are increasingly being co-opted to assign water demand to network nodes based on user classifications (e.g., residential, commercial, industrial, institutional, etc.). When data on individual users is scarce, it is common to assign identical diurnal demand curves to all users of the same type. For example, all residential water users might be assigned a diurnal curve characterized by large peaks at the start and end of the work day, with slack periods during the early morning and mid-afternoon (Fig. 1). Similarly, all commercial users are assigned a diurnal curve that rises and falls sharply, and that is flat during the workday period (Fig. 1). In reality, each house, business, manufacturing plant, etc. has a unique diurnal curve (Boulos et al. 2004) that reflects the specific water needs and preferences of that user.

Recent studies have shown that residential water demand makes up the majority of water use in an urban water distribution network. Billings and Jones (1996) found that at the national level, residential demand accounts for 50–60% of total municipal water use in the United States. Studies by Flack (1982) and Raftelis Financial Consultants et al. (2000) found that residential water use can be as high as 75% in some systems. These statistics imply that the residential user type, and its diurnal pattern, often dominates the overall diurnal demand variation in many network models. This regime creates a situation in which the majority of nodal demands in a network model are synchronous. In statistical terms, nodal demands in a network model are often perfectly correlated in space.

Applying a single diurnal pattern to residential users in a network model implies that residential users—and thus the majority of users—react simultaneously and in exactly the same way to normal and peak demand conditions. In real systems, users respond to normal and peak conditions (e.g., hot summer day) based on their particular preferences, social habits, financial constraints, etc. which are partially independent of the preferences of others. Most of the time, users have little information about what is happening elsewhere or about the water-consumption patterns of others. Together, all these considerations imply that demands at individual consumption points in real water distribution networks are imperfectly correlated under normal and peak conditions. The term "imperfectly correlated" implies that demand pairs are not exactly synchronous.

The aim of this paper is to explore (and partly challenge) the common design assumption that most users draw water from a network in accordance with a single diurnal pattern, or in statistical terms, that demands are perfectly correlated in space during normal and peak conditions. Specifically, the paper explores to what extent cross correlation between demands influences the hydraulic performance of a system, namely in the mean and standard deviation of pressure head. This analysis is directed towards large systems (i.e., transmission and network trunk mains) where municipal demands dominate the design, and where fire flow considerations play only a minor role (fire flows are thus not considered in this paper). The second objective is to investigate to what extent cross correlated demands influence the pipe cost of a network to achieve a desired level of hydraulic reliability. Indeed, a clearer picture of the spatial correlation between pairs of demands at the design stage can lower network costs. It thus follows that

[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, Canada M5S 1A4. E-mail: adams@ecf.utoronto.ca

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

Note. Discussion open until August 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 May 7, 2004; approved on November 14, 2005. This paper is part of the ***Journal of Water Resources Planning and Management***, Vol. 133, No. 2, March 1, 2007. ©ASCE, ISSN 0733-9496/2007/2-137–144/$25.00.

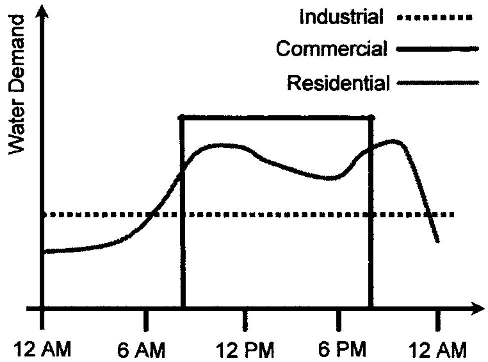

Fig. 1. Diurnal curves for different user classes

the reduction in design/rehabilitation costs can justify the time, cost, and resources needed to collect demand data in the field. The stated objectives are explored via two examples. In the first example, the relationship between cross correlated demands and hydraulic performance is investigated in a transmission main with no storage. In the second example, Monte Carlo simulation (MCS) is applied to investigate the relationship between cross correlation in demand and hydraulic performance in the realistic network of Alperovitz and Shamir (1977) that includes pumps and tank storage. The network of Alperovitz and Shamir (1977) is also used to explore how correlated demands can influence pipe sizes and network costs in design.

Definition of Cross Correlation

Since it is a major focus of this paper, the idea of cross correlation is described briefly. Cross correlation relates to the interdependence between two spatial variables X and Y, at a particular point in time. In the context of water distribution systems, X might refer to the value of demand at a particular node, and Y to the value of demand at an adjacent or distant node. The degree of interdependence, or correlation, between X and Y is measured by the covariance, or the average product of pairs of deviations of X and Y from their respective mean values μ_X and μ_Y, such that

$$\mathrm{Cov}(X,Y) = \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu_X)\cdot(y_i - \mu_Y) \quad (1)$$

As indicated in (1), the covariance of X and Y is calculated with m pairs of x_i, y_i values. The demands X and Y are said to be positively correlated when their values are large (or small) at the same time [Fig. 2(a)], or when the product of their respective deviations in (1) is positive. Similarly, demands X and Y are said to be negatively correlated [Fig. 2(b)] when large X values occur with small Y values (or vice versa). This occurs when the product

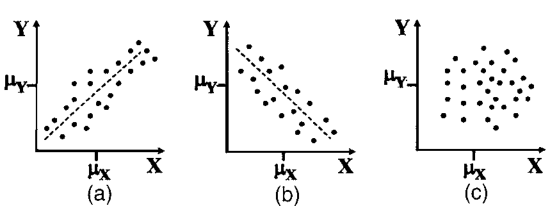

Fig. 2. Cross correlation between demands X and Y: (a) positive correlation; (b) negative correlation; and (c) independence

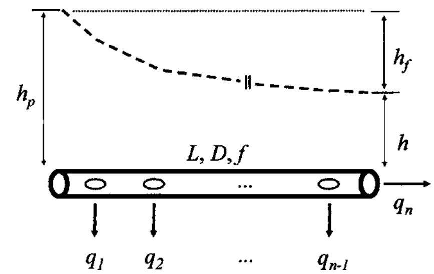

Fig. 3. Transmission main with n demand locations

of their respective deviations in (1) is negative. When X and Y are independent, no high–low relationship exists [Fig. 2(c)] and their covariance is zero. A zero covariance does not necessarily imply independence.

Cross Correlated Demands and Hydraulic Performance

Analytical Model for Serial Systems

In this first case, an analytical model is presented to elucidate the influence of cross correlation in demands on the mean and variance of downstream pressure head (measure of hydraulic performance) in a series pipe system. The prototypical system considered is illustrated in Fig. 3 comprises an upstream pumping station with fixed head, a single pipe with $n-1$ intermediate points of demand, and a terminal demand where pressure head is monitored. This downstream pressure head h at the pipeline terminus in Fig. 3 is calculated by subtracting the headloss Δh_i across each pipe segment from the pumping station head h_p as in (2). The headloss across pipe segment i is functionally related to its friction factor f_i, length L_i, diameter D_i, and the demand $\theta(i)$ downstream of the pipe as follows:

$$h = h_p - \sum_{i=1}^{n}\Delta h_i, \quad h = h_p - \sum_{i=1}^{n} k_i \cdot \theta(i)^{\alpha} \quad (2)$$

$$k_i = \frac{8}{\pi^2}\cdot\frac{f_i L_i}{g D_i^5} \quad (3)$$

$$\theta(i) = \sum_{j=i}^{n} q_j \quad q_j \geq 0 \quad (4)$$

where h_p=fixed pumping (or supply) head (m); n=total number of serial demands; k_i=pipe resistance ($\mathrm{s}^2/\mathrm{m}^5$); $\theta(i)$=operator that tallies downstream demand conveyed by pipe segment i (m^3/s); α=exponent in Darcy-Weisbach friction loss model (equal to 2); f_i=Darcy-Weisbach friction factor; L_i=length of pipe segment i (m); D_i=inner diameter of pipe segment i (mm); and q_i=individual demand at location i (m^3/s). Note that only steady state pressures are considered in this paper. Dynamic pressures caused by hydraulic transient or waterhammer conditions are not considered.

First Moment of Hydraulic Performance

The expressions in (2)–(4) are written in more general form as $h=G(q_1, q_2, \ldots, q_n)$ where $G(\cdot)$=mathematical function that "maps" an array of serial demands q_i to a single scalar value of downstream pressure head h. With this analytical convenience,

J. Water Resour. Plann. Manage. 2007.133:137-144.

one can use the mean-centered, first-order method (MCFOM) to estimate the first and second moments (mean and variance) of a downstream pressure head. In an engineering analysis, the MCFOM is employed to generate moments of a dependent variable in terms of a straightforward function of the first two moments of the independent variable, since often only the first and second moments of the independent variables are known (Benjamin and Cornell 1970). Therefore, the method is a practical alternative to the derived probability distribution theory that transforms probability density functions (PDFs) of independent variables to a complete PDF of the dependent variable. To determine the first moment or mean of the pressure head, the function $G(\cdot)$ is expanded as a multivariate Taylor series about a mean demand state, or the point at which all demands assume their mean value such that $q_i=\mu_i$ and the expected value of the expanded Taylor series is determined (Hahn and Shapiro 1967). The expression that results is

$$\mathrm{E}(h,n) = G(\mu_1,\mu_2,\ldots,\mu_n) + \frac{1}{2}\sum_{k=1}^{n}\left(\frac{\partial^2 h}{\partial q_k^2}\right)_{\mu_k}\sigma_k^2 + \sum_{k=1}^{n-1}\sum_{l=k+1}^{n}\left(\frac{\partial^2 h}{\partial q_k \partial q_l}\right)_{\mu_k,\mu_l}\sigma_{k,l}^2 \quad (5)$$

where $\mathrm{E}(h,n)$=mean of the downstream pressure head in a serial system with n imperfectly cross correlated demands (m); $G(\mu_1,\mu_2,\ldots,\mu_n)$=value of the downstream pressure head when demands assume their mean value (m); $\partial^2 h/\partial q_k^2$=second partial derivative of $G(\cdot)$ evaluated at $q_k=\mu_k$; $\partial^2 h/\partial q_k \partial q_l$=mixed second partial derivative of $G(\cdot)$ evaluated at $q_k=\mu_k$ and $q_l=\mu_l$; σ_k^2 =variance of the kth individual demand ($\mathrm{m}^6/\mathrm{s}^2$); and $\sigma_{k,l}^2$=cross covariance between pair of demands q_k and q_l ($\mathrm{m}^6/\mathrm{s}^2$). The analytical expression in (5) relates the mean of pressure head $\mathrm{E}(h,n)$ directly to the variance σ_k^2 of individual demands and the covariance $\sigma_{k,l}^2$ of pairs of demands. The variance of demand at node k, or σ_k^2, is calculated as the average squared deviation of demand from its mean, or $\Sigma(q_k-\mu)^2$ from a time series of demand values taken from a supervisory, control, and data acquisition (SCADA) system. Similarly, the cross covariance between demands q_k and q_l is calculated as the average product of the deviations from their respective means with (1) and with a time series of demand pairs q_k and q_l collected from a SCADA system.

Second Moment of Hydraulic Performance

The variance of the downstream pressure head is determined by expanding the "mapping" function $G(\cdot)$ about the mean demand state with a multivariate Taylor series approximation and then determining the variance of this expanded series. The expression that results is written as

$$\mathrm{V}(h,n) \simeq \sum_{k=1}^{n}\left(\frac{\partial h}{\partial q_k}\right)_{\mu_k}^2\sigma_k^2 + 2\sum_{k=1}^{n-1}\sum_{l=k+1}^{n}\left(\frac{\partial h}{\partial q_k}\right)_{\mu_k}\left(\frac{\partial h}{\partial q_l}\right)_{\mu_l}\sigma_{k,l}^2 \quad (6)$$

where $\mathrm{V}(h,n)$=variance of the downstream pressure head in a serial system with n imperfectly correlated demands (m^2); and $\partial h/\partial q_k$=first derivative of $G(\cdot)$ evaluated at $q_k=\mu_k$. Similar to the case of the mean, the analytical expression in (6) relates the variance $\mathrm{V}(h,n)$ of pressure head directly to the variance σ_k^2 of individual demands and the covariance $\sigma_{k,l}^2$ of pairs of demands. Note that the Taylor series estimate of the mean in (5) is truncated after the second-order term and the Taylor series estimate of the variance in (6) is truncated after the first-order term. Since the pressure-flow relationship $G(\cdot)$ in (2) is second-order accurate, the calculated mean in (5) is exact whereas the calculated variance in (6) is an approximation of the "true" variance of downstream pressure head.

Example 1: Transmission Main. The first example considers a transmission main (Fig. 3) with an upstream pumping station (acts as a fixed-head reservoir) and without tank storage. While it does not reflect the complexity of real systems, the transmission system is chosen expressly for its simplicity and to easily understand the link between cross correlated demands and hydraulic performance. The MCFOM is applied to the transmission main to calculate the first and second moments of downstream pressure head under different cross correlation conditions and different demand configurations (n=2, 4, 8, 16, and 32 individual demands). For each demand configuration, the cross correlation coefficient between pairs of demands is varied from 0.0 (uncorrelated) to 1.0 (strongly correlated). In the interest of simplicity, all pairs of demands are assumed to be identically cross correlated and are thus assigned identical cross correlation coefficients.

Users along the transmission main draw water from the upstream pumping station with a fixed head of 45.0 m to satisfy demands $q_1,q_2,\ldots,q_n$ at intermediate locations and the system terminus indicated in Fig. 3. The main has length L=1,300 m, diameter D=711 mm, and Darcy–Weisbach friction coefficient f=0.01. The entire length of the main is at a fixed datum of 0.0.

Since the mean and variance expressions in (5) and (6) do not explicitly account for the temporal variation of demand, any time step length can be selected for the transmission main example. The temporal variation of demand is implicitly defined in the coefficient of variation and the cross correlation coefficient defined below.

The statistical characteristics of each individual demand in Fig. 3 deserve further mention here. To provide a realistic comparison between demand configurations, the mean of the total demand μ (sum of individual demands $q_1,q_2,\ldots,q_n$) is held fixed at 0.75 m^3/s (750 lps). [Note that when demands are set to their mean level, the pipe velocities fall below the maximum design value of 1.9 m/s (Walski et al. 2003).] This implies that increasing the number of demand locations n produces a proportional decrease in the mean μ_i of each demand such that

$$\mu_i = \frac{\mu}{n} \quad (7)$$

where μ=mean of total demand in system (m^3/s); and μ_i=mean of individual demand q_i (m^3/s).

The variance of each individual demand q_i is reduced when the number of demands n is increased. To simplify the analysis, the n individual demands in the transmission main are assumed to vary in accordance with a single coefficient of variation set to 0.40. The variance of each individual demand is thus calculated with

$$\sigma_i^2 = \mathrm{Cv}^2 \cdot \mu_i^2 = \mathrm{Cv}^2 \cdot \frac{\mu^2}{n^2} \quad (8)$$

where σ_i^2=variance of individual demand ($\mathrm{m}^6/\mathrm{s}^2$); and Cv=coefficient of variation. It is clear in (8) that the variance σ_i^2 is inversely proportional to the square of n demands. Also, as n is increased, the pipe is divided into L/n segments each with hydraulic resistance calculated in (3).

Since the mean and variance of individual demands depends on the number of demands n in the system as in (7) and (8), it is clear that one cannot compare directly the hydraulic performance

J. Water Resour. Plann. Manage. 2007.133:137-144.

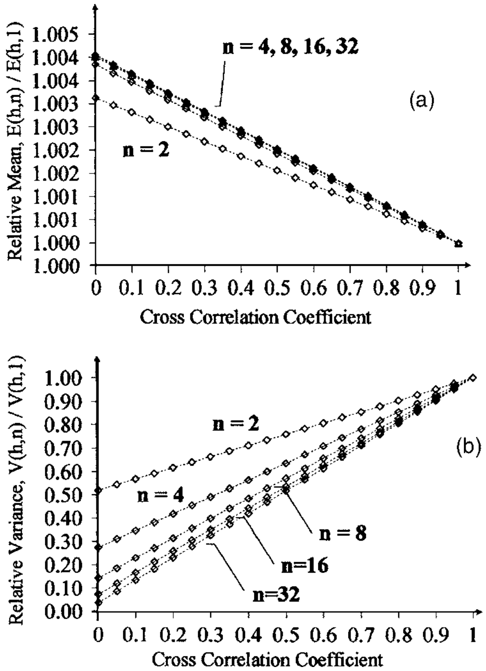

Fig. 4. Relative mean and variance of downstream pressure head versus cross correlation coefficient in transmission main with 2, 4, 8, 16, and 32 demands

of systems with different demand configurations. For example, a system with only two demand locations, equally spaced along the main, each with a mean of 0.375 m^3/s (375 lps), will generate headlosses that exceed those in a main with, say, four equally spaced demands each with a mean of 0.1875 m^3/s (187.5 lps). The reason for this is simple: Increasing the number of demand locations distributes the total demand into smaller parts, and each pipe segment carries a smaller flow and dissipates less energy. It is more convenient to compare two or more system configurations with a normalized parameter that accounts for only relative differences in hydraulic performance.

In this paper, the *relative* mean and variance of downstream pressure head are used to compare systems with different demand configurations. The relative mean is determined in a series of steps: (1) the mean downstream pressure head in the transmission main with n imperfectly cross correlated demands is calculated with the MCFOM in (5) and denoted as $E(h,n)$; (2) the mean downstream pressure head with n perfectly cross correlated demands is recalculated with the univariate version of the MCFOM in (5) and denoted as $E(h,1)$. Here, the n individual demands change in unison with one another according to a univariate probability distribution (hence the 1 inside parentheses); and (3) the *relative* mean is calculated by dividing the mean downstream pressure head with imperfectly cross correlated demands $E(h,n)$ with its counterpart $E(h,1)$. The resulting ratio of means (and variances) is dimensionless and can thus be used to assess fairly the hydraulic performance of systems across different configurations

$$g_1 = E(h,n)/E(h,1) \tag{9}$$

$$g_2 = V(h,n)/V(h,1) \tag{10}$$

The relationship between the cross correlation of demands and the mean of the downstream pressure head is illustrated in Fig. 4(a). In the transmission main, the relative mean downstream pressure head is insensitive to the cross correlation coefficient. In fact, varying the cross correlation coefficient from 0.0 to 1.0 leads to a decrease of only 0.3% (from 1.003 to 1.000) in the relative mean downstream pressure head when $n=2$. This insensitivity stems from the second partial derivative terms in (5) which have a small magnitude for all values of n. A sensitivity analysis confirmed that the relative mean downstream pressure head is insensitive to cross correlation in demand over a wide range of lengths and Darcy–Weisbach friction factors. The downstream pressure head becomes sensitive to the level of cross correlation in demand only for small pipe diameters.

The relative variance of downstream pressure head is sensitive to the level of cross correlation in demand. The plots in Fig. 4(b) clearly show that when the cross correlation coefficient is increased from 0.0 to 1.0, the relative variance of downstream pressure head $V(h,n)/V(h,1)$ increases by approximately 92% (from 0.52 to 1.0) when $n=2$ and by just over 2,600% (from 0.04 to 1.0) when $n=32$. The pronounced relationship between the relative variance of the downstream pressure head and the cross correlation coefficient is understood via the second derivative term in (6) which describes a first-order—and thus a significant—sensitivity to cross correlation between demands. The sensitive relationship between the relative variance of pressure and cross correlation in demand also has a physical basis. A cross correlation coefficient near zero means that demands are uncorrelated and vary independently from one another. This means that in many instances, high demands in one part of the system are balanced by low demands in other parts. This "demand-balancing" introduces a modest level of scatter in the pressure signal and yields a low variance of pressure. The opposite is true when demands are strongly correlated in space (cross correlation coefficient approaches 1.0) and demands are synchronous. Under these conditions, all demands are simultaneously high or low at any moment in time. Consequently, this high level of correlation between demand yields wide fluctuations in pressure and a correspondingly large variance in pressure.

The plot in Fig. 4(b) also indicates that increasing the number of demands from $n=2-32$ will lower the relative variance of the downstream pressure head at a fixed level of cross correlation. This is explained by two factors. First, increasing the number of demand locations leads to a decrease in their variance according to the square of the mean μ_i in (8). According to (6), decreasing the variance of demands will also decrease the variance of the downstream pressure head. Second, in a system with only two demands, both demands will be high on many occasions, and both demands will be low on a great many other occasions. This regime of demand variation will produce a large variance in pressure. By contrast, in a system with 32 demands, it is reasonable to expect that on only a small number of occasions will all 32 demands have a high or low value. Most of the time, there will be a number of demands that are low and a number of others that are high. This balancing of demands, or "staggering" effect, will tend to produce only small fluctuations in pressure and, consequently, a smaller variance in pressure than that observed in the system with two demands. A sensitivity analysis showed that the relative variance of pressure is highly sensitive to the level of cross correlation in demand and that this sensitivity remains high over a broad range of pipe diameters, friction factors, and pipe lengths.

J. Water Resour. Plann. Manage. 2007.133:137-144.

Simulation Model for Networks

Monte Carlo Simulation Algorithm

Unlike transmission mains and other serial systems, the flows and pressures in looped distribution networks can only be computed by solving a system of equations with a numerical model. Here, a popular network model is coupled with MCS to compute statistics on the hydraulic performance of a looped network. The MCS algorithm is divided into four steps: (1) at the start of each time step (time step length is 1 d) a stochastic model generates n cross correlated network demands; (2) correlated demands are entered into the network solver EPANET2 (Rossman 2000) to compute pressure heads; (3) nodal pressure heads are used to update estimators of the mean and standard deviation in (11) and (12); and (4) the first three steps are repeated over $1,2,\ldots,n_r$ time steps until n_r pressures have been generated

$$\bar{h} = \frac{1}{n_r}\sum_{i=1}^{n_r} h_i \tag{11}$$

$$s_h = \left\{ \frac{1}{n_r - 1}\sum_{i=1}^{n_r} (h_i - \bar{h})^2 \right\}^{0.5} \tag{12}$$

where h_i=pressure head at a node calculated in run i (m); $\bar{h}$=estimate of mean pressure head at a node (m); and s_h=estimate of standard deviation of pressure head at a node (m).

Synthetic Demand Generation

The multivariate streamflow model proposed by Fiering (1964) is used to generate demands. This model was originally formulated to account for the lag-zero and lag-one cross correlation between streamflows at gauged sites. In the absence of any lag-one cross correlations, the general multivariate model then reduces to the simplified form

$$\mathbf{X}_t = \mathbf{B} \times \epsilon_t \tag{13}$$

where $\mathbf{X}_t=\{x_1(t),x_2(t),\ldots,x_n(t)\}=n\times 1$ vector whose kth element $x_k(t)=q_k(t)-\mu_k$=demand deviate (relative to mean) observed at node k; $\mathbf{B}$=coefficient matrix that produces a set of synthetic demands at node k that is statistically similar to the historical set of demands at node k in terms of the first three moments as well as the cross correlation structure between demands observed at pairs of nodes k and l; $\epsilon_t=\{\epsilon_1(t),\epsilon_2(t),\ldots,\epsilon_n(t)\}=n\times 1$ vector of identical, independently distributed (IID) standard normal random variates with mean zero and unit variance (without loss of generality, variates with other distributions may be similarly accommodated). In this model, individual demands $q_k(t)$ in the pipe network are unbounded (again without loss of generality since bounds may be readily applied).

The coefficient matrix **B** is found through statistical analysis of demand data gathered in the field (e.g., from a SCADA system). First, an $n\times n$ cross covariance matrix is defined to account for the cross correlation between pairs of demands in a water distribution system with n nodes. The general composition of this matrix is

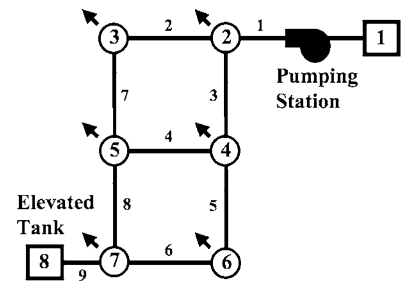

Fig. 5. Schematic of test network with six demands

$$\mathbf{M}_0 = \begin{bmatrix} \sigma_{1,1}^2 & \sigma_{1,2}^2 & \cdots & \sigma_{1,n}^2 \\ \sigma_{2,1}^2 & \sigma_{2,2}^2 & \cdots & \sigma_{2,n}^2 \\ \vdots & & & \\ \sigma_{n,1}^2 & \cdots & \cdots & \sigma_{n,n}^2 \end{bmatrix} \tag{14}$$

The elements of the cross covariance matrix in (14) are covariance terms that measure the degree of cross correlation between pairs of demands. For example, the term $\sigma_{1,2}^2$ denotes the covariance between demands at Nodes 1 and 2. Each term in the cross covariance matrix is easily calculated with (1) and a time series of demand pairs from a data acquisition system. Since the cross covariance matrix is symmetrical, terms with indices in reverse order are equal (i.e., $\sigma_{1,2}^2=\sigma_{2,1}^2$). Diagonal terms such as $\sigma_{1,1}^2$ denote the variance of demand at a particular node (in this case Node 1).

Fiering (1964) showed that the lag-zero cross covariance matrix $\mathbf{M}_0$ is related to the coefficient matrix **B** through the following relationship:

$$\mathbf{B} \times \mathbf{B}^T = \mathbf{M}_0 \tag{15}$$

Once the covariance terms in $\mathbf{M}_0$ have been calculated, the **B** matrix is solved recursively with a solution proposed by Young (1968) in which **B** is assumed to be lower triangular.

Example 2: Network of Alperovitz and Shamir (1977). The simple network of Alperovitz and Shamir (1977) in Fig. 5 comprises nine pipes. Pipes 2–8 have a diameter of 457 mm, a length of 1,000 m, and a Hazen–Williams "C" factor of 130. Pipes 1 and 9 have a diameter of 711 mm and Hazen–Williams "C" factor of 130. Pipe 1 has a length of 1,000 m while Pipe 9 has a length of 100 m. When mean demand conditions prevail in the network, the stipulated pipe sizes produce fluid velocities that range between 0.01 and 0.98 m/s—well below the maximum design velocity of 1.9 m/s (Walski et al. 2003).

The node parameters in Table 7a of Alperovitz and Shamir (1977) are used in this paper. The mean of each nodal demand is set to its deterministic value in Table 7a of Alperovitz and Shamir (1977). All demands are assumed to follow a normal distribution with a coefficient of variation (Cv) of 0.15 and to be cross correlated according to (14). In practice, most municipalities have only a few meters that record flow to large service areas or pressure zones. The skeletonized network of Alperovitz and Shamir (1977) represents such a situation, whereby each of the six nodes supplies a metered demand to a large service area. Node 1 is a clearwell with bottom elevation at 150.0 m and fixed water depth of 3.05 m (10 ft). Two pumps draw water from Node 1, each with a pump curve defined by the single point [42.0 m, 0.182 m^3/s].

A series of numerical experiments are run in which the cross correlation coefficient of all pairs of demands is systematically

J. Water Resour. Plann. Manage. 2007.133:137-144.

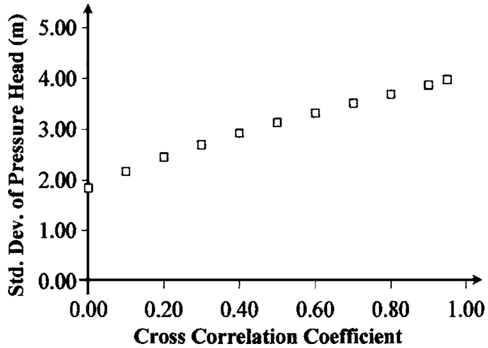

Fig. 6. Standard deviation of the pressure head at Node 6 versus cross correlation coefficient in network with 90% confidence bounds

increased from 0.0 to 0.95. (Negative cross correlation between demands is not investigated in this work.) Each experiment is comprised of 10,000 time steps, each lasting 1 d, in which the pressure at Node 6 (critical node at high elevation and furthest from pumping station) is calculated. The mean and standard deviation of pressure head at Node 6 are based on the 10,000 estimates of pressure calculated in each experiment. For this first set of experiments, a valve in Pipe 9 is closed and Tank 8 is hydraulically disconnected from the network.

The simulations revealed a number of interesting points. First, the mean of pressure head at Node 6 is insensitive to the level of cross correlation between demands (not shown). However, a sensitivity analysis suggests that the mean of pressure head is sensitive to cross correlation levels only for small pipe diameters. The sensitivity of the mean pressure head to cross correlation is low over a broad range of pipe lengths and values of Hazen–Williams "C" factor.

Second, Fig. 6 indicates that the standard deviation of pressure head at Node 6 is highly sensitive to cross correlation between demands in the network when no storage is considered. As the cross correlation between demands is increased from 0.00 to 0.95, the standard deviation of pressure head at Node 6 increases by just under 115% (from 1.85 to 3.96 m). The physical reasoning for this sensitivity is similar to that of the transmission main example. It is acknowledged that most real systems include tank storage to equalize pressures during normal and peak times. Nevertheless, the results in Fig. 6 suggest that the standard deviation of pressure might still be sensitive to demand correlation at nodes located far away from tanks and pumping stations (and thus not under their hydraulic influence). Note that since a small coefficient of variation is used for all demands in this example (Cv=0.15), the absolute value of standard deviation of pressure a Node 6 is correspondingly small. Kolmogorov–Smirnov testing indicated that the pressure head h_i at Node 6 can reasonably be expected to follow a normal distribution. The standard deviation s_h is thus assumed to follow a chi-squared distribution and 90% confidence limits are thus placed on it in Fig. 6. The statistical errors associated with the data points in Fig. 6 are so small that the 90% confidence limits are not visible in this figure. A sensitivity analysis shows that the standard deviation of pressure remains high over a broad range of pipe diameters, Hazen–Williams "C" coefficients, and pipe lengths.

A second set of numerical experiments is performed—this time assuming that the valve in Pipe 9 is open and Tank 8 is hydraulically connected to the system. The top water level in Tank 8 assumes a range of values during normal and peak conditions. To reflect this, the water level in Tank 8 thus follows a normal probability density function with mean of 195.5 m and standard deviation of 0.26 m.

Connecting Tank 8 to the network affects the hydraulic network performance in sundry ways. First, it is the fluctuating water level in Tank 8—and not the fluctuations in correlated nodal demands—that governs the variation in pressure at Node 6. This is not surprising given the proximity of Tank 8 to Node 6. Second, the small tank level fluctuations cause only small variations in pressure at Node 6. This is evident in the small standard deviation of pressure of 0.27 m (not shown). More important, the sensitive relationship between the standard deviation of pressure head and cross correlation in demand disappears when Tank 8 is connected to the network. Here, the standard deviation of pressure head at Node 6 increases by only 8% (from 0.067 to 0.073 m) when cross correlation between nodal demands is increased from 0.0 to 0.95 (not shown).

Cross-Correlated Demands and Network Cost

Recent research has investigated how uncertainty in demands can affect the hydraulic reliability and capital cost of distribution networks. Lansey et al. (1989), Xu and Goulter (1999), and Kapelan et al. (2004) found that a large scatter (or standard deviation) in nodal demands tends to lower network reliability and increase the size and cost of pipes, pumps, and tanks needed to achieve an acceptable level of reliability. More recently, Tolson et al. (2004), Babayan et al. (2005), and Kapelan et al. (2005) found that increasing the cross correlation level between nodal demands leads to higher capital costs for a given level of reliability. However, these researchers do not discuss how costs are linked to cross correlation in demand. From this starting point, a second set of numerical experiments is performed to investigate to what extent cross correlation in demand can affect network costs, at different reliability levels.

In these experiments, the diameter of Pipes 2–8 are uniformly increased from 305 to 711 mm. The cost of each network design is calculated with the pipe cost in Table 6c of Alperovitz and Shamir (1977). Pipes with a diameter ranging from 559 to 711 mm have a unit pipe cost that is linearly extrapolated from the Alperovitz and Shamir data. Each network design is simulated over 10,000 time steps, each lasting 1 d. Reliability is measured as the proportion of times the pressure head at Node 6 is above the minimum allowable limit of 30.0 m reported in Alperovitz and Shamir (1977).

The plot in Fig. 7 graphs pipe costs (y-axis) against reliability (x-axis) for when demands are uncorrelated (cross-correlation coefficient $\rho_{ij}=0.0$), moderately correlated ($\rho_{ij}=0.5$); and strongly correlated ($\rho_{ij}=0.95$) in space. Note that the valve in Pipe 9 is closed and Tank 8 is disconnected from the network. The plot shows that, at high levels of reliability, an increasing strength of cross correlation between demands leads to a significant increase in pipe cost. The reason being that a high level of coordination (or correlation) between demands leads to larger fluctuations in pressure, more frequent violations of the 30.0 m minimum, and thus larger pipe sizes to maintain a certain level of reliability.

Interestingly, the opposite is true when reliability is at a low level (below 0.5) in Fig. 7. Here, pipe costs are higher when demands are weakly correlated than when they are strongly correlated in space. Put differently, for the same pipe cost, reliability is higher in a network with strongly correlated demands than in a network with weakly correlated demands. The reason for this

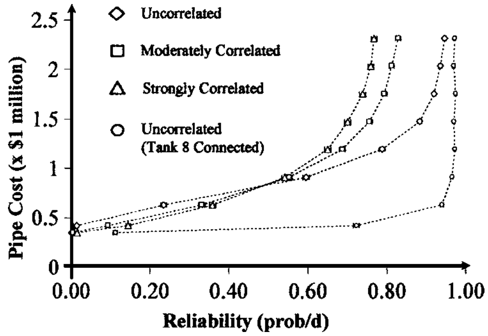

Fig. 7. Pipe cost versus hydraulic reliability for: (a) uncorrelated demands (Tank 8 disabled); (b) moderately correlated demands (Tank 8 disabled); (c) strongly cross correlated demands (Tank 8 disabled); and (d) uncorrelated demands (Tank 8 connected)

counterintuitive result is traced to the mean and standard deviation of pressure head at Node 6. The network with 406 mm diameter pipes is used as a test example here. In Fig. 7, this network has a pipe cost of \$0.63 million, a reliability of 0.23 when demands are weakly correlated, and a reliability of 0.36 when demands are strongly correlated. Testing showed that when demands are weakly correlated ($\rho_{ij}=0.0$) in this network, the mean pressure is 28.4 m and the standard deviation is 2.14 m. This is indicated in Fig. 8(a) which plots pressure head at Node 6 for the first 50 time steps of the simulation. When demands are strongly correlated ($\rho_{ij}=0.95$), the mean pressure sits slightly lower at 28.2 m but the standard deviation is larger at 4.45 m [Fig. 8(b)]. Since in both cases the mean of pressure is below the minimum of 30.0 m, strongly correlated demands [Fig. 8(b)] will cause larger fluctuations in pressure (higher standard deviation) and increase

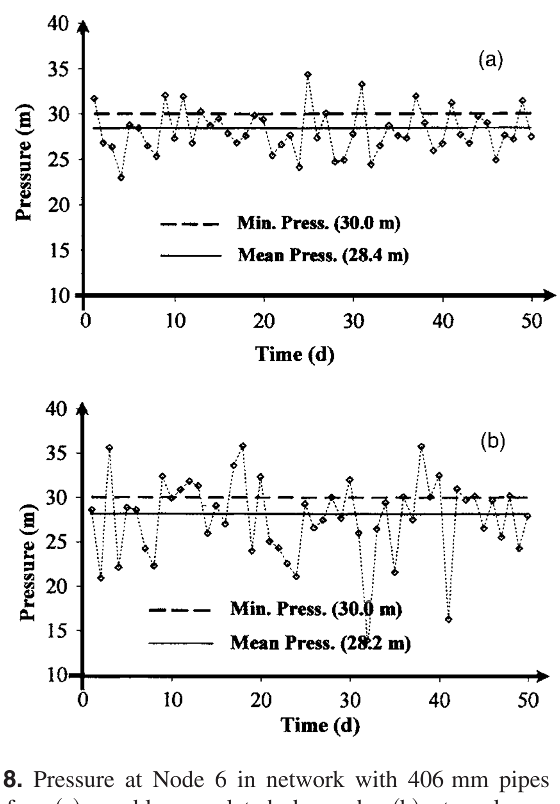

Fig. 8. Pressure at Node 6 in network with 406 mm pipes versus time for: (a) weakly correlated demands; (b) strongly correlated demands

the proportion of times the pressure at Node 6 rises above the 30.0 m minimum. This means that for a fixed level of cost, increasing the coordination, or cross correlation, between demands will actually improve the network reliability. Tolson et al. (2004) and Kapelan et al. (2005) have encountered this phenomenon in their work but have made no attempts to describe it.

An additional numerical experiment is run to explore how tank storage can lower pipe cost in the network. In this experiment, the valve in Pipe 9 is opened and Tank 8 is connected to the network. The diameter of Pipes 2–8 are uniformly increased from 305 to 711 mm and the pipe cost and reliability of each design is indicated on the curve labeled "Uncorrelated (Tank 8 Connected)" in Fig. 7. (In this experiment, demands are assumed to be uncorrelated in space ($\rho_{ij}=0.0$), and the tank level fluctuations follow a normal distribution with a mean of 195.5 m and standard deviation of 0.26 m.) Not surprising, Fig. 7 suggests that connecting Tank 8 to the network dramatically lowers pipe cost at all reliability levels at Node 6. The reason being that Tank 8 shortens the source-to-consumption path to Node 6 and eliminates the need for large pipes (and large pipe costs) to minimize losses between the pumping station and Node 6. The cost of Tank 8 is not included in Fig. 7.

A final experiment is run to see whether the level of cross correlation in demand changes pipe cost for a fixed level of reliability when Tank 8 is connected to the network. The results (not shown) indicate only marginal changes in pipe cost when demands are uncorrelated ($\rho_{ij}=0.0$), moderately correlated ($\rho_{ij}=0.5$); and strongly correlated ($\rho_{ij}=0.95$) in space. The reason being that it is the water level variation in Tank 8—and not the cross correlation between demands—that determines pressure at Node 6 and the reliability of the system.

The estimate of reliability $\hat{R}$ in Fig. 7 is calculated as $\hat{R}=\Sigma Z_i/n_r$. The term Z_i is a Bernoulli random variable that is equal to 1 if pressure at Node 6 is at or above the minimum level of 30.0 m, and equal to 0 if the pressure is below the 30.0 m minimum. The term n_r denotes the number of 1 d time steps in the simulation. By the central limit theorem, the estimate of reliability $\hat{R}$ is normally distributed with mean R (Milton and Arnold 1995). The 90% confidence limits on $\hat{R}$ indicated in Fig. 7 are thus

$$\hat{R} \pm z_{0.05}\sqrt{\hat{R}(1-\hat{R})/n_r} \tag{16}$$

The results above raise a number of important points about assumptions made in water distribution design. In extended period simulation (EPS) it is common to assume that diurnal demands—whether they are taken across different user classes or within a single user class—are very strongly correlated in space when they may be weakly or moderately correlated. Consequently, the assumption of strong correlation may yield wide fluctuations in pressure (especially in networks without storage), and create a need for larger pipe sizes—and higher costs—to ensure that a minimum pressure constraint is not violated. If a chance-constrained design optimization model is used, then enforcing a strong cross correlation between demands may lead to the oversizing of pipes—with higher costs—to achieve a desired reliability, especially in networks without storage. In networks with existing storage, the assumed level of cross correlation between demands may have little impact on pressure variation, except at distant nodes that are only slightly influenced by the hydraulic behavior of pumps and tanks.

The two examples of this paper suggest that collecting data on

J. Water Resour. Plann. Manage. 2007.133:137-144.

demand might help a utility reduce the cost of rehabilitating its network. Currently, many utilities collect demand flows at key locations in a network (e.g., pumping stations, boundary of pressure zones) with SCADA systems. Time series data at different network locations can be used to determine the spatial relationship between pairs of demands so that more realistic demand patterns are used to design/rehabilitate networks. The savings in network costs might more than justify the time, cost, and resources needed to collect data on demand. In light of these considerations, utilities may wish to reassess how the data collected with SCADA systems already in place is used in the future.

Summary and Conclusions

This paper preliminarily investigates how cross correlated demands can affect the hydraulic performance of a network, particularly in the mean and standard deviation of steady-state pressure heads. A second objective is to explore how cross correlated demands influence the capital cost of networks to achieve a desired level of hydraulic reliability. These questions are explored via a transmission main (without tank storage) and a skeletonized network with pumping capacity and tank storage. For the systems investigated, a number of conclusions are drawn:

1. The mean of pressure is insensitive to the strength of cross correlation between demands, whether or not Tank 8 is connected to the network.
2. The standard deviation of pressure is highly sensitive to the strength of cross correlation between demands in the transmission main and in the network when no tank storage is considered. When Tank 8 is connected to the test network, the strength of cross correlation between demands has little effect on the standard deviation of pressure.
3. When Tank 8 is disconnected from the network, increasing the strength of cross correlation between demands dramatically increases the pipe size—and thus cost—needed to achieve a desired hydraulic reliability. When Tank 8 is connected to the network, increasing the strength of cross correlation has little effect on pipe cost.
4. The experiments confirm the well-established conclusion that including tank storage decreases pipe sizes (and costs) required to attain a desired level of hydraulic reliability.

Taken in a broader context, the preliminary results point to the importance of gaining a clear picture of diurnal demand across different user classes in order to design more cost-efficient networks. This gives force to the argument for collecting time series data at different network locations to better determine the spatial relationship between diurnal demands in extended period simulation (EPS). Indeed, the savings in network costs might more than justify the time, cost, and resources needed to collect data on demand.

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

Alperovitz, E., and Shamir, U. (1977). "Design of optimal water distribution systems." *Water Resour. Res.*, 13(6), 885–900.

Babayan, A. V., Savic, D. A., and Walters, G. A. (2005). "Multiobjective optimization for the least-cost design of water distribution system under correlated uncertain parameters." *Proc., EWRI 2005: Impacts of Global Climate Change*, ASCE, Reston, Va.

Benjamin, J. R., and Cornell, C. A. (1970). *Probability, statistics, and decision for civil engineers*, McGraw-Hill, New York.

Billings, R. B., and Jones, C. V. (1996). "Forecasting urban water demand." AWWA, Denver.

Boulos, P. F., Lansey, K. E., and Karney, B. W. (2004). *Comprehensive water distribution systems analysis handbook for engineers and planners*, MWH Soft, Pasadena, Calif.

Fiering, M. (1964). "Multivariate technique for synthetic hydrology." *J. Hydr. Div.*, 90(5), 43–60.

Flack, J. E. (1982). *Urban water conservation: Increasing efficiency-in-use residential water demand*, ASCE, New York.

Hahn, G. J., and Shapiro, S. S. (1967). *Statistical models in engineering*, Wiley, New York.

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.

Kapelan, Z., Savic, D. A., Savic, A., and Walters, G. A. (2005). "An efficient sampling-based approach for the robust rehabilitation of water distribution systems under correlated nodal demands." *Proc., EWRI 2005: Impacts of Global Climate Change*, ASCE, Reston, Va.

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

Milton, J. S., and Arnold, J. C. (1995). *Introduction to probability and statistics*, McGraw-Hill, New York.

Raftelis Financial Consultants, Maddaus Water Management, Weber Group. (2000). *Town of Cary water conservation and peak demand management plan*, Charlotte, N.C.

Rossman, L. A. (2000). *EPANET 2: Users' manual*, U.S. EPA National Risk Management Research Laboratory, Cincinnati.

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., Chase, D. V., Savic, D. A., Grayman, W. M., Beckwith, S., and Koelle, E. (2003). *Advanced water distribution modeling and management*, Haestad Methods, Waterbury, Conn.

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

Young, G. K. (1968). "Discussion of 'Mathematical assessment of synthetic hydrology'." *Water Resour. Res.*, 4(3), 681–683.

J. Water Resour. Plann. Manage. 2007.133:137-144.