

Probabilistic Estimation of Water Conservation Effectiveness

David E. Rosenberg¹

Abstract – An analytical method is derived to describe the distribution of water quantity saved among customers within a water-use sector who adopt a water conservation action. Analytical results tend towards lognormal distributions with long tails, quantifying a smaller subset of customers that show potential to achieve large savings. Example effectiveness distributions are shown for seven long-term conservation actions potentially implemented by urban, domestic water users in Amman, Jordan. Monte-Carlo simulations verify the analytical derivations. The probabilistic outputs contrast with common methods that estimate conservation action effectiveness as a product of typical (average) characteristics for disaggregated customer groups. Implications to size water conservation programs to meet conservation objectives and target customers to adopt conservation actions are discussed.

CE Database Terms: Municipal water; conservation; probability distribution; Jordan

INTRODUCTION

Water consumption and the effective quantity of water conserved by implementing conservation actions vary significantly among customers with important effects related to various geographic, demographic, technological, behavioral, and temporal factors (Mayer

¹ Doctoral Candidate, Department of Civil and Environmental Engineering, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA. E-mail: deroenberg@ucdavis.edu

et al. 1999; Optiz and Dziegielewski 1998; Vickers 2001; Walski et al. 1985).

Conventional approaches to estimate water conservation action effectiveness commonly disaggregate water use by sectors and estimate effectiveness within a sector as a simple product of single parameter values representing average customer characteristics (Optiz and Dziegielewski 1998; Vickers 2001; Walski et al. 1985). For example, Vickers (2001, p. 25) presents typical values of 15 liters (4 gallons) per flush for residential toilets manufactured before 1994, 6 liters (1.6 gallons) per flush for low-volume toilets manufactured after 1997, 5.1 flushes per person per day, 2.64 persons per residence, and 365 days per year to show that a U.S. residential customer installing a low-flow toilet should typically conserve $(15 - 6)(5.1)(2.64)(365)/(1000) = 44 \text{ m}^3 \text{ year}^{-1}$. The number of customers needed to meet a conservation objective is then found by dividing the water conservation objective by the typical savings per customer. Sector-wide effectiveness is also estimated by multiplying parameters for total unrestricted water use, fractional water use reduction, coverage (fraction of customers adopting the action), and interaction with other conservation actions (Optiz and Dziegielewski 1998; Walski et al. 1985).

Conventional estimation approaches work well for homogenous customer populations where customers within each water-use sector have nearly identical unrestricted water uses, similar reduction potentials, and both factors can be quantified as singular values. In such cases multiplying typical customer effectiveness by the number of customers in the water-use sector likely to adopt the action readily yields the sector-wide effectiveness. However, when a customer population is heterogeneous, shows multiple water use behaviors and reduction fractions, or the likely coverage is uncertain, effectiveness calculated solely from typical values can prove problematic for several reasons. First, parameter values are uncertain and differ for different customers. The uncertainties propagate and also make the resultant effectiveness uncertain. Second, customers facing (extreme) situations represented by one or more parameters taking values at the lower

end of their feasible ranges should have little or no water savings. These customers may have insufficient financial or other incentives to adopt a conservation action. Third, data gathering, computing, and analysis efforts increase multiplicatively as the analyst further disaggregates the customer population to form homogenous sub-sectors (Walski et al. 1985). The analyst also must set separation points by trial and error. And fourth, effectiveness parameters are multiplied together so the uncertainties interact rather than cancel. Effectiveness will not necessarily be normally (i.e., evenly) distributed above and below the simple product of average parameter values. Thus, a single effectiveness value does not show how water savings may be distributed among the customers under study.

This paper presents an alternative, probabilistic approach to describe the likely distribution of effectiveness among a sector of customers considering adopting a water conservation action. First, probabilistic information is developed to describe the range and likelihood of values possible for each parameter influencing effectiveness (Jaynes 2003; Tribus 1969). Second, the uncertainties are propagated analytically—and verified numerically with Monte Carlo simulations—to develop the distribution of effectiveness. Because parameters are multiplied together, effectiveness tends to a lognormal distribution (Aitchison and Brown 1957). And third, the continuous effectiveness distribution is used to select and size water conservation programs to meet conservation objectives. The approach is demonstrated for seven long-term conservation actions that are potentially implemented by urban, residential water users in Amman, Jordan. Probabilistic treatment achieves a continuous disaggregating of a customer population and suggests the minimum number of participants needed to meet a specific water conservation objective. The approach is useful to planners who understand the ranges of potential values for customer demographic, behavioral, and technological factors influencing effectiveness but who cannot measure effectiveness directly.

PROBABILISTIC METHOD

The probabilistic method to describe the likely distribution of effectiveness among customers considering adopting a water conservation action is summarized as follows:

1. Define how effectiveness is calculated from its component parameters,
2. Estimate a probability distribution (pdf) for each uncertain parameter,
3. Propagate uncertainties to calculate a composite probability distribution for conservation effectiveness,
4. Note statistics for the composite distribution, and
5. Use distribution properties to size conservation programs or estimate aggregate water savings.

These steps are further described as follows.

Functional form of conservation action effectiveness

Engineering estimates of the expected quantity of water conserved in a particular place over a specific period of time by implementing a conservation action are often calculated as a simple product of single parameter values (Optiz and Dziegielewski 1998; Vickers 2001; Walski et al. 1985). Although the effectiveness function is specific to each conservation action, the general form is

$$W = f_{conv} \cdot \prod_{j=1}^m (X_j)^{r_j} \prod_{k=m+1}^n (Z_k - Y_k)^{r_k} . \quad (1)$$

Here, W is the uncertain water conservation effectiveness or volume conserved per customer per unit time when a customer implements the conservation action; X_j , Z_k , and Y_k are uncertain parameters in units specific to the conservation action; r_j or r_k are fixed powers to which those parameters may be raised; f_{conv} is a unit conversion factor; m is the number of individual-termed uncertain parameters; and $n - m$ is the number of paired terms. (The capital letters X , Y , Z , and W reflect notation common to the probability

literature where a capital letter, i.e. X , means the parameter is uncertain. The lower case counterpart, i.e., x , refers to a particular value that the uncertain parameter may take.)

The paired parameters Z_k , and Y_k have the same units and occur together as a difference term when effectiveness is a function of change in state. For example, the effectiveness of installing a water-conserving fixture depends in part on the difference between the flow rate of the existing fixture (i.e., Z [l min^{-1}]) and flow rate of the water-conserving fixture (i.e., Y [l min^{-1}]). Both flow rates are often uncertain; therefore, their difference is also uncertain and must be considered explicitly. (Dividing the difference between the average existing flow rate and average conserving flow rate by the average existing flow rate gives the sector-wide reduction parameter used by Walski et al. (1985)).

Estimate probability distributions for parameters

The second step is to estimate probability distributions for each uncertain parameter. Distributions will depend on the prior information known about the parameter. They can be specified from detailed, statistically sampled, empirical information concerning customer demographics, water appliances, water-related behaviors and consumption [for example, see Mayer (1999)]. Distributions can also be fit to empirical data. Or, absent detailed information, distributions may be estimated using the method of maximum entropy. This method minimizes information content (maximizes entropy) to suggest the most simple distribution shape that completely encapsulates the limited prior information known for the parameter (such as upper bound, lower bound, and/or average value) [see Jaynes (2003) or Tribus (1969, pp. 128-130) for details]. Rows 1 and 2 of Table 1 summarize the likely distribution forms or pdfs for different cases of prior known information. Cases are discussed further in the sections below. Methods to estimate distribution forms for difference terms common to water conservation actions are

summarized in rows 1 and 2 of Table 2. These resultant distributions depend on the distributions of the component parameters and are also discussed below.

Known lower and upper bounds

When only the lower and upper bounds for a parameter are known, the principle of maximum entropy suggests that parameter values should be uniformly (rectangular) distributed. The parameter should have an equally likely (or constant) probability to take any value in the feasible range.

Known lower bound and mean

When only the lower bound and mean for a parameter are known, the principle of maximum entropy suggests that the lower bound value is most likely to occur. However, the occurrence probability should decay exponentially as the potential value the parameter may take increases. The initial value (λ_0) and rate of decay (λ_1) are calculated analytically from the prior known lower bound and mean.

Known frequencies for discrete ranges of parameter values

Results from empirical surveys are often summarized as frequencies for discrete ranges of parameter values (histograms). Frequencies can be used as-is, or fitted with a continuous functional form. In addition, any analytical probability density function may be approximated as a set of frequencies for discrete ranges of parameter values when the ranges chosen are sufficiently small.

Difference of two parameters

The difference of two uncertain parameters is also uncertain, and will be distributed according to the *convolution* of the uncertain parameters (Jaynes 2003, p. 677). For example, the uncertain difference $U = Z - Y$ has the probability distribution,

$$h(u) \equiv \int_{-\infty}^{\infty} pdf_z(x) \cdot pdf_y(x-u) \cdot dx. \quad (2)$$

Here, pdf_z and pdf_y are, respectively, the probability density functions of the component uncertain parameters Z and Y . For example, when Z is the uncertain flow rate of the existing fixture [l min^{-1}] and Y is the uncertain flow rate of the water conserving fixture [l min^{-1}], $h(u)$ will represent the distribution of reduced flow (hereafter, the convoluted distribution). The convoluted distribution may exist for some or all of the negative range ($u < 0$) depending on the lower and upper bounds (if any) of Z and Y . The convolution distribution will depend on the distribution forms of the component parameters (see results in rows 1 and 2 of Table 2 for example distributions and differences common to water conservation actions). Convolution allows us to transform a term with two uncertain parameters into a term with one uncertain parameter and further generalize the functional form of conservation action effectiveness to

$$W = f_{conv} \cdot \prod_{j=1}^n (X_j)^{r_j}. \quad (3)$$

Propagate uncertainties

With distributions specified or derived for each component parameter, the next step is to propagate uncertainties to determine the composite probability distribution of effectiveness among customers in the water use sector. Uncertainty can be propagated analytically or by Monte Carlo simulation.

Analytical propagation

The logarithm of the generalized effectiveness equation (3) gives

$$\log(W) = \log(f_{conv}) + \log(X_1) + \log(X_2) + \dots + \log(X_n). \quad (4)$$

Sampling from the right hand side of equation (4) and applying the Central Limit Theorem yields a sum that will be normally distributed about a composite mean value, $\mu_{(n)}$. This observation applies irrespective of the distributions of the log-adjusted component parameters. Thus, the logarithm of the composite conservation effectiveness W is normally distributed, meaning that W is lognormal distributed with a probability density function of

$$pdf(w) = \partial \Lambda(\mu, \sigma^2) = \begin{cases} 0, & w \leq 0 \\ \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-0.5\left(\frac{\log w - \mu}{\sigma}\right)^2\right] dw, & w > 0. \end{cases} \quad (5)$$

Equivalently, we may write W is distributed as $\Lambda(\mu, \sigma^2)$. Here μ and σ^2 are, respectively, the mean and variance of the normal distribution describing the log-transformation of W (and are different than the mean and variance of W) (Aitchison and Brown 1957). To determine the composite mean and variance indicators, Aitchison and Brown (1957, p.

14), find that the product $\prod_{j=1}^N X_j$ is asymptotically distributed as $\Lambda(\mu_{(n)}, \sigma_{(n)}^2)$ when:

- Each $\{X_j\}$ is an independent, positive variate, (6a)

- $\mu_{(n)} = \sum_{j=1}^N \mu_j$ and $\sigma_{(n)}^2 = \sum_{j=1}^N \sigma_j^2$, and (6b)

- $\mu_j = E\{\log X_j\}$ and $\sigma_j^2 = D^2\{\log X_j\}$. (6c)

Here, $E\{\}$ and $D^2\{\}$ denote, respectively, the expectation and variance operators.

For the more general function $f_{conv} \cdot \prod_{j=1}^N (X_j)^{r_j}$ that describes water conservation

effectiveness, the multiplicative and additive properties of the natural logarithm can be used to recast (6b) as

$$\mu_{(n)} = \sum_{j=1}^N r_j \mu_j + \log f_{conv} \quad \text{and} \quad \sigma_{(n)}^2 = \sum_{j=1}^N r_j^2 \sigma_j^2 \quad . \quad (6d)$$

The log-weighted first and second moments of parameter X_j are calculated as

$$\begin{aligned} \mu_j &= E\{\log X_j\} = \int_a^\infty \log x \cdot pdf_j(x) \cdot dx \\ \sigma_j^2 &= D^2\{\log X_j\} = \int_a^\infty (\log x)^2 \cdot pdf_j(x) \cdot dx - \mu_j^2 \end{aligned} \quad (7a)$$

and can be evaluated analytically or numerically depending on the distribution form of parameter X_j (rows 4 through 7 of Table 1). For these cases, the lower limit of integration, a , is the lower bound of the parameter distribution.

The method also applies to convolution distributions (rows 4 and 5 of Table 2) with two modifications. These modifications avoid integrating over negative ranges for which the convolution distribution may exist but for which the logarithm operation is not defined,

$$\begin{aligned} \mu_j &= \int_c^\infty \ln x \cdot \frac{pdf(x)}{1 - p_c} dx \\ \sigma_j^2 &= \int_c^\infty (\ln x)^2 \cdot \frac{pdf(x)}{1 - p_c} dx - \mu_j^2 \end{aligned} \quad (7b)$$

First, the analyst must specify the cutoff value c – the lower limit of integration – as greater than zero ($c > 0$). This cutoff value represents the analyst's best estimate of the value below which customers will not implement the conservation action because the reduced flow (or consumption) will be either negligible or negative (i.e., increased flow or consumption). Second, the analyst must re-weight the convolution pdf by a divisor $1 -$

p_c so that the cumulative proportion of customers above the cutoff value who participate in the conservation action sum to unity

$$1 = \int_c^{\infty} \frac{pdf(x)}{1 - p_c} \cdot dx. \quad (8a)$$

Rearranging (8a) and switching the integration limits show that p_c is just the proportion of customers below the cutoff value who do not implement the conservation action

$$p_c = cdf(c) = \int_{-\infty}^c pdf(x) \cdot dx. \quad (8b)$$

This fraction is also the *cumulative density function* (cdf) evaluated at c .

Because selecting a cutoff value amounts to censoring the portion of customers that do not implement the conservation action, the distribution of conservation action effectiveness must likewise reflect censoring [$\Lambda(\mu, \sigma^2)$ is insufficient]. A censored lognormal distribution, $\Lambda(\delta, \mu, \sigma^2)$ can be defined (Aitchison and Brown 1957, p. 95) as:

$$pdf(z) = \partial \Lambda(\delta, \mu, \sigma^2) = \begin{cases} 0, & z < 0 \\ \delta, & z = 0 \\ \delta + (1 - \delta) \Lambda(z | \mu, \sigma^2) dz, & z > 0 \end{cases} \quad (9)$$

Where δ is the fraction of the population that tends towards zero (or negative) values. In specifying the censored pdf for a conservation action, substitute p_c from equation (8b) for δ . When $\delta = 0$, equation (9) simplifies to (5).

In summary, when all uncertain parameters are independent, have values greater than zero, and are multiplied together to determine conservation action effectiveness, equations (5), (6d), and (7a) together define the analytical probability density function, mean, and variance for the lognormal-distributed conservation action effectiveness.

When one of the parameters can have negative values, the analyst must specify a cutoff

value, and equations (9), (6d), (7b), and (8b) define the analytical lognormal distribution of effectiveness for customers implementing the action.

Monte Carlo propagation

Uncertainties also can be propagated with Monte Carlo simulation. The general method is: a) generate random variates from the distributions of the component parameters [see (Law and Ketton 1991) for details]. b) Combine instantiations of the random variates according to the effectiveness function. c) Repeat steps (a) and (b) for a large number of samples. And (d) Sort effectiveness samples from smallest to largest and report the fraction (frequency) of samples falling within discrete ranges of water conservation action effectiveness. Together, the frequencies will form a histogram. Divide each frequency by the width of the range from which values were aggregated to obtain the Monte Carlo simulated pdf of water conservation action effectiveness.

Statistics of the composite distribution

When the composite distribution is lognormal distributed, the mean and quantiles are:

$$\begin{aligned} \text{Mean} = \bar{w} &= (1 - \delta) e^{\mu_{(n)} + 0.5 \cdot \sigma_{(n)}^2} \\ \text{Quantile}_q = w_q &= \begin{cases} 0 & q \leq \delta \\ e^{\mu_{(n)} + z_{q'} \cdot \sigma_{(n)}^2} & q > \delta \end{cases} \end{aligned} \quad (10)$$

where $z_{q'}$ is the z-value associated with the normal distribution N(0,1) for the quantile $q' = (q - \delta)/(1 - \delta)$ (Aitchison and Brown 1957, pp. 95-6). With no censoring ($\delta = 0$), the mean, median, and mode are simply $e^{\mu_{(n)} + 0.5 \cdot \sigma_{(n)}^2}$, $e^{\mu_{(n)}}$, and $e^{\mu_{(n)} - \sigma_{(n)}^2}$, and are successively decreasing indicating significant positive skew.

For an effectiveness distribution generated by Monte Carlo simulation, the mean is best estimated by the average of the entire sample of effectiveness calculations. The quantile q

can be approximated by the value of the $(k \cdot q)^{\text{th}}$ sample in the list of simulated effectiveness sample results sorted from lowest to highest (k = number of simulations). The mode will correspond to the effectiveness range with the largest frequency.

Size conservation programs

The final step is to use the derived effectiveness distribution and its common properties to size a conservation program to meet an overall water conservation objective. Program sizing can be done by several methods. The first method, blanket application, as used by typical engineering approaches requires just

$$s_{\text{blanket}} = \frac{t}{\hat{w}}. \quad (11)$$

where s_{blanket} is the estimated number of customers required to implement the conservation action, t is the program-wide conservation objective [m^3 per year], and \hat{w} is the average savings per customer [m^3 per customer per year] generally calculated as a point estimate using (typical) average parameter values. The conservation objective t represents the desired annual water savings and can correspond to the projected shortfall between future water supplies and future water demand or some portion of that shortfall that the utility wants to meet by encouraging customers to adopt conservation actions. Blanket application assumes customers adopt with uniform effectiveness.

The second sizing method focuses on market segmentation and targeting customers that show potential to achieve large water savings. A targeted approach makes use of the probabilistic distribution of effectiveness.

The targeted customers should have large values for effectiveness w . The sizing task is to determine the threshold effectiveness level, w_t , so that water saved by the customer with the largest effectiveness plus the water saved by the customer with the next largest

effectiveness, and so on down to the water saved by the customer with effectiveness at the threshold level sum to meet the conservation goal. This sum is the integral of the first moment distribution of W (i.e., the customer effectiveness level w weighted by its probability of occurrence) evaluated from the threshold w_t through infinity, or

$$t = s_{sect} \int_{w=w_t}^{w=\infty} w \cdot \partial\Delta(w | \mu_{(n)}, \sigma_{(n)}^2, p_c). \quad (12)$$

Here, s_{sect} is the sector size (number of customers) potentially available to adopt the conservation measure and is required to scale customer effectiveness, w [m^3 per customer per year], to the absolute conservation objective, t [m^3 per year]. Equation (12) is solved for w_t using two identities. First, the integral of the first moment of $\Delta(w | \mu, \sigma^2, \delta)$ over the entire feasible range of W is, by definition, the mean effectiveness,

$$\bar{w} = \int_{w=0}^{w=w_t} w \cdot \partial\Delta(w | \mu_{(n)}, \sigma_{(n)}^2, p_c) + \int_{w=w_t}^{w=\infty} w \cdot \partial\Delta(w | \mu_{(n)}, \sigma_{(n)}^2, p_c). \quad (13a)$$

Second, the first moment of $\Delta(w | \mu, \sigma^2)$ is lognormal distributed as $\Delta(w | \mu + \sigma^2, \sigma^2)$ (Aitchison and Brown 1957, p. 12). This identity also applies to the censored distribution $\Delta(w | \mu, \sigma^2, \delta)$, so

$$\frac{1}{\bar{w}} \int_{w=0}^{w=w_t} w \cdot \partial\Delta(w | \mu_{(n)}, \sigma_{(n)}^2, p_c) = \int_{w=0}^{w=w_t} \partial\Delta(w | \mu_{(n)} + \sigma_{(n)}^2, \sigma_{(n)}^2, p_c). \quad (13b)$$

Rearranging and then substituting (13a) and (13b) into (12) gives

$$1 - \frac{t}{s_{sect} \cdot \bar{w}} = \int_0^{w_t} \partial\Delta(w | \mu_{(n)} + \sigma_{(n)}^2, \sigma_{(n)}^2, p_c) = CDF_{\Delta(\mu_{(n)} + \sigma_{(n)}^2, \sigma_{(n)}^2, p_c)}(w_t). \quad (14)$$

Here, $CDF_{\Delta(\mu + \sigma^2, \sigma^2, p_c)}$ is the cumulative density function of $\Delta(\mu + \sigma^2, \sigma^2, p_c)$. The left hand side of (14) is a fraction between 0 and 1 ($0 \leq t \leq s_{sect} \cdot \bar{w}$). Since $CDFs$ monotonically increase, they are invertible. Thus,

$$w_t = CDF_{\Delta(\mu_{(n)} + \sigma_{(n)}^2, \sigma_{(n)}^2, p_c)}^{-1} \left(1 - \frac{t}{s_{sect} \bar{w}} \right), \quad 0 \leq t \leq s_{sect} \bar{w}. \quad (15)$$

Finally, the targeted conservation program size is determined by multiplying the sector size by the fraction of the sector having effectiveness above the threshold w_t

$$s_{targeting} = \left(1 - CDF_{\Delta(\mu_{(n)}, \sigma_{(n)}^2, p_c)}(w_t)\right) \cdot s_{sect} . \quad (16)$$

Equation (14) may also be rearranged to express the fraction t / s_{sect} as a function of w_t ,

$$\frac{t}{s_{sect}} = \bar{w} \cdot \left(1 - CDF_{\Delta(\mu_{(n)} + \sigma_{(n)}^2, \sigma_{(n)}^2, p_c)}(w_t)\right). \quad (17)$$

Varying the threshold w_t (or the fraction of the community represented by w_t) will identify the average conservation expected per customer. This formula determines the sizing curve for the conservation action and is demonstrated below.

EXAMPLE APPLICATION

We now develop distributions of water savings for seven conservation actions available to urban, residential water users in Amman, Jordan. The actions include rainwater harvesting from roofs, installing spray nozzles on garden hoses (rather than using open hoses), installing carpets on floors (to replace floor washing with water), and retrofitting showerheads, bathroom faucets, kitchen faucets, or toilets with water saving devices. These actions represent some of the many long- and short-term water supply enhancement and demand management actions that can help residential, urban customers cope with water shortages. Probabilistic analysis is readily applied to each action; here, we demonstrate the approach for seven long-term water conservation actions.

The Amman water utility serves about 1,940,000 residents through 306,000 residential connections and reported 52.4 million cubic meters (Mm³) of residential billed water use in 2004. Customers face severe water shortages: water is typically available through the

distribution network for only 12 – 60 hours per week. Jordan is starting to implement water demand management programs but there is scarce empirical documentation showing the effectiveness of water conservation actions. Thus, probabilistic statements describing potential effectiveness can help guide conservation program planning.

The seven functions for conservation effectiveness are:

$$W_{Rainfall Catch} = \frac{1}{1000}(A)(B)^{-1}(C), \quad (18a)$$

$$W_{Showerhead Retro} = \frac{52}{1000}(G - T)(D)(L)(M), \quad (18b)$$

$$W_{Bath Faucet Retro} = \frac{365}{1000}(H - U)(O)(D), \quad (18c)$$

$$W_{Kitchen Faucet Retro} = \frac{365}{1000}(H - U)(P), \quad (18d)$$

$$W_{Toilet Retro} = \frac{365}{1000}(I - V)(N)(D), \quad (18e)$$

$$W_{Spray Nozzle} = 0.429(J)^2(F)^{0.5}(X)(Q)(R)(E), \text{ and} \quad (18f)$$

$$W_{Carpet Install} = \frac{52 \cdot 3.785}{1000}(K)(S)(Y). \quad (18g)$$

The letters A through V, X, and Y represent the uncertain parameters influencing effectiveness and are further described in Table 3 (DOS 1999; DOS 2004; IdRC 2004; JMD 2000; Snobar 2003; WEPIA 2000).

The following details are also important. The 78-year record of rainfall at the Amman Airport (JMD 2000) was fitted with a Gamma distribution by estimating the shape and scale parameters from the mean and variance of the observed annual rainfalls. The water conserved by installing a spray nozzle on a garden hose was estimated by the reduction of flow through an open-ended hose. This flow is related to the square of the hose diameter,

square root of the customer water pressure, and time for which the nozzle restricts wastage flow. In the example, the distribution of water pressure was assumed to correlate directly to the distribution of households sharing a building. In Amman, rooftop tanks are the primary regulator of residential water pressure; thus, pressure depends on head differential between roof and point of use. This difference is also the number of floors (or apartments, i.e. households) in the building. This conservation action is only available to the approximately 15.4% of households that garden outdoors (DOS 1999).

Limited information is available concerning several of the parameters, and in some cases, distributions were derived from engineering estimates of maximum upper and lower limits. These estimates rely on the author's experiences living and working in Jordan and were verified by others with significant experience in the Jordan residential water sector (Tawarneh, pers. comm., 2004; Abdul Al-Khalaq, pers. comm., 2004).

Some parameters may co-vary. For example, more single-family residences may be located in West Amman where elevation differences result in higher rainfall. With better data, we could segment Amman households into classes and subclasses (such as by geographic location and building type within a location) to eliminate covariance. Then, calculate effectiveness distributions for each subclass using parameter distributions specific to the subclass. While further disaggregating the population requires increased data gathering, computation, and analysis effort, the probabilistic approach can achieve continuous disaggregating (within the sub-classes) which is not possible with point estimate approaches. Based on the data readily available and for demonstration purposes, the population of Amman residential customers was not disaggregated.

Parameter uncertainties were propagated both analytically and with Monte Carlo simulation (10,000 simulations for each conservation action). In analytical derivations,

numerical integrations of the log-weighted exponential decay functions were made with central differences and approximately 10,000 steps over the feasible parameter range. Figures 1 and 2 compare the analytically derived distribution of effectiveness to the Monte Carlo simulation results for the first two conservation actions.

Both actions show a preponderance of the population with effectiveness close to the lower limit, but also a large tail stretching towards a small proportion of customers who show potential to realize large water savings by adopting the conservation actions. Both distributions have positive skew with $\text{mean} > \text{median} > \text{mode}$. This behavior is also seen in the effectiveness distributions derived for the other conservation actions (Figure 3).

A chart for sizing targeted conservation programs (Figure 4) was calculated using equation (17). The chart shows water conservation level as a function of the coverage or fraction of total customers who adopt the action. This fraction is explicitly ordered from left to right by customers with potentials to conserve the largest down to the smallest volumes of water. The sizing curves are fastest rising for small program sizes as customers with the most effectiveness adopt first. As coverage reaches 100%, the curves become flat and approach the mean value of the effectiveness distribution. This value defines an upper bound for the savings when all customers adopt.

The chart is used as follows: First, set the overall water conservation objective (in volume per year) and community size or number of customers that can potentially adopt the conservation action. Second, divide the conservation objective by the community size to figure the average water volume conserved per customer. Third, find this volume on the vertical axis. Fourth, use the sizing curve to find the corresponding targeted coverage. Finally, multiply the coverage by the community size to determine the number of

customers required to meet the conservation objective (when customers with the largest potential to conserve are targeted to participate in the program).

The sizing chart can also help identify water efficient conservation actions. Actions with faster rising curves require smaller number of customers to meet a specified conservation objective. Thus, retrofitting showerheads or kitchen faucets are more effective than installing carpets or spray nozzles on garden hoses. For example, to meet a water conservation objective of 6.5 Mm^3 per year (12% of Amman's billed residential water use), the Amman water utility need only target 8% of its 306,000 residential customers to retrofit kitchen faucets (should the utility identify its customers with the potential to conserve 124.8 m^3 per year or more; 60% of customers are needed with a blanket approach). Alternatively, the utility need only target the largest 10%, 26%, and 49% of customers that show potential to conserve more than 106.8, 38.8, or 20.6 m^3 per year by, respectively, retrofitting showerheads, toilets, or collecting rainwater (Table 4). The utility will likely not meet the conservation object even if all customers (100%) retrofit bath faucets, install carpets or spray nozzles on outdoor hoses. The sizing chart also shows these infeasibilities: these actions never reach an average water conservation level of 21.2 m^3 per customer per year (6.5 Mm^3 per year / 306,000 customers).

Including average retrofit costs for each conservation action identifies the cost-effective actions (Table 4). Here, costs reflect estimates for customers to purchase water saving devices (author's estimates; IrDC, 2004) and exclude utility costs to implement a program. However, utility costs would likely be similar for each conservation action. In the Amman, Jordan example, retrofitting kitchen faucets appears as the most cost effective conservation action to meet the annual conservation objective.

DISCUSSION

Although Table 4 shows that average conservation action effectiveness values calculated with typical point estimates and the proposed probabilistic approach are often similar, the implications for sizing conservation programs differ substantially. In the Amman, Jordan example to achieve annual water savings of 6.5 Mm^3 per year, targeted conservation programs to retrofit kitchen faucets, showerheads, and toilets sized using the probabilistic approach can be much smaller than blanket application programs sized using point estimates of average effectiveness. These targeted conservation programs can reduce implementation costs by factors of 2.5 to 8 over typical blanket application approaches. These differences are most pronounced when the annual water conservation objective is small compared to the maximum savings achievable when the entire community adopts the conservation action. Differences are less pronounced as the conservation objective approaches or exceeds the maximum savings.

An outstanding issue concerns how to expeditiously identify and target the customers with the most potential to conserve (where they are located and what characteristics distinguish them from low-effective customers). Three customer identification methods—use of surrogate indicators, customer surveys, and water audits—are introduced below and their relative advantages and disadvantages are discussed. These methods represent public awareness, education, and targeted marketing approaches typical for water conservation programs (Baumann et al. 1998; Vickers 2001). The single difference is using the probabilistic-determined threshold effectiveness level to determine which customers to contact and suggest to adopt the conservation action. Discussion also emphasizes that no one method to identify customers can efficiently and precisely demarcate all customers with high effectiveness from customers with low effectiveness. Rather, a combination of approaches is likely needed.

Surrogate indicators of effectiveness

Geographic information systems and databases offer the water utility or conservation program coordinator a wealth of customer-specific information related to conservation action effectiveness. Example data include water-billing records (indicating customer water consumption), land assessments (indicating building size and age, i.e., a further surrogate indicator of water appliance age and flow rates), satellite or digital orthographical photos (showing landscaped areas), or census records (indicating household size), among others. In fact, the coordinator may have used such data to estimate distributions for some component parameters. Linking and joining multiple data sources provides a powerful tool to identify the subset of customers with co-occurrence of multiple factors that suggest high conservation action effectiveness. If data sources are not linked, low indicator values can still flag customers with low effectiveness. This analysis can beneficially shrink the customer pool on which to focus more costly or labor-intensive identification approaches.

Customer surveys

A utility can also telephone or distribute written questionnaires to each customer to learn more about the customer's demographic makeup, water use behaviors, and other factors that influence water conservation action effectiveness. The utility can use responses to project the customer's likely effectiveness if they adopt and then follow up with customers that show effectiveness larger than the threshold effectiveness level. And while telephone surveys and written questionnaires are quick and relatively inexpensive to implement, customer response rates may be low. However, positive customer response can also indicate strong willingness to adopt the conservation action.

Water Audits

A utility can also dispatch staff to visit each customer, solicit the information that bears on the customer's water conservation effectiveness, and then instantaneously estimate the effectiveness. If the estimated effectiveness exceeds the threshold effectiveness level, staff can then immediately recommend or proceed with retrofits. Although water audits are costly in terms of time, staff, and materials, they still serve as beneficial screening tools. Identifying customers for which no follow-up action is taken can save the utility resources required to implement the conservation action and time required for follow-up visits to verify continued implementation and actual water savings.

Together, surrogate indicators, customer surveys, and water audits can help identify customers with potential to achieve large water savings. After adoption, these methods can also help verify that estimated effectiveness translates into actual effectiveness.

CONCLUSIONS

Water conservation program planners can probabilistically describe water conservation effectiveness by understanding the ranges of values for customer demographic, behavioral, and technological parameters influencing water savings. Probabilistic treatment achieves a continuous disaggregating of a customer population but avoids the time and costs of additional data gathering, computation, and analysis associated with common point estimates and blanket application that further disaggregate the population into smaller, homogenous sub-sectors. Because effectiveness is a product of uncertain parameter values, it tends towards a lognormal distribution with significant positive skew towards a small population of customers that show potential to achieve large savings by implementing a conservation action.

Effectiveness distributions are readily used to suggest cost efficient conservation actions, the minimum number of customers needed to meet specific water conservation objectives, or the threshold effectiveness levels on which to target customer adoption. Seven example distributions for urban, residential water users in Jordan show that a small subset of customers can achieve significant annual water savings by retrofitting showerheads or kitchen faucets. Also, that targeting consumers with the largest potential to conserve can significantly reduce the size and cost of programs to meet water conservation objectives compared to blanket application approaches. To realize these size and cost savings, planners must develop targeted marketing, public awareness, and education campaigns to first identify the customers with high conservation effectiveness and then persuade or encourage them to adopt. Follow-up work is also needed to verify that estimated effectiveness translates to actual effectiveness.

ACKNOWLEDGEMENTS

The author is supported by a National Science Foundation graduate research fellowship. The author thanks Jay Lund and three anonymous reviewers for their comments.

REFERENCES

- Aitchison, J., and Brown, J. A. C. (1957). *The Lognormal Distribution: with special reference to its uses in economics*, Cambridge University Press, London, U.K.
- Baumann, D. D., Boland, J. J., and Hanemann, W. M. (1998). *Urban Water Demand Management and Planning*, McGraw-Hill, Inc., New York.
- Department of Statistics (DOS). (1999). "Urban Agriculture Survey." Amman, Jordan. http://www.dos.gov.jo/sdb/env/env_e/home.htm [Accessed on: June 26, 2004].

- Department of Statistics (DOS). (2004). "The Preliminary Results of The Population And Housing Census." Amman, Jordan. <http://www.dos.gov.jo> [Accessed on: January 2, 2006].
- Interdisciplinary Research Consultants (IdRC). (2004). "A Survey of Home Appliances and Toilets in the Jordanian Markets." Academy for Educational Development: Water Efficiency for Public Information and Action (WEPIA) Program, Amman, Jordan, 50 pp.
- Jaynes, E. T. (2003). *Probability Theory: the logic of science*, Cambridge University Press, Cambridge, United Kingdom.
- Jordan Meteorological Department (JMD). (2000). "Rainfall." <http://www.jmd.gov.jo/climate-rain.html> [Accessed on: January 23, 2005].
- Law, A. M., and Ketton, W. D. (1991). "Chaper 8. Generating Random Variates." *Simulation Modeling and Analysis*, McGraw-Hill, New York.
- Mayer, P. W., DeOreo, W. B., Opitz, E. M., Kiefer, J. C., Davis, W. Y., Dziegielewski, B., and Nelson, J. O. (1999). "Residential End Uses of Water," American Water Works Association, Denver, Colorado. <http://www.aquacraft.com/index.htm>.
- Optiz, E. M., and Dziegielewski, B. (1998). "Demand Management Planning Methods." *Urban Water Demand Management and Planning*, W. M. Hanemann, ed., McGraw-Hill, Inc., New York, 237-281.
- Snobar, A. H. (2003). "Feasibility of Domestic Graywater Recycling," Senior Thesis, Jordan University of Science and Technology, Irbid.
- Tribus, M. (1969). *Rational Descriptions, Decisions and Designs*, Pergamon Press, New York.
- Vickers, A. L. (2001). *The handbook of water conservation*, WaterPlow Press, Amherst, Massachusetts. www.waterplowpress.com.

Walski, T. M., Richards, W. G., McCall, D. J., Deb, A. K., and Morgan, J. M. (1985).

"Calculating Effectiveness of Water Conservation Measures." *Journal of Environmental Engineering-Asce*, 111(2), 209-221.

Water Efficiency and Public Information for Action (WEPIA). (2000). "Capacity building project in Fuhies." Amman, Jordan, 14 pp.

LIST OF TABLES

- Table 1. Parameter Distributions and Methods to Calculate Log-Weighted Moments for Different Types of Known Information.
- Table 2. Probability Distributions and Methods to Calculate Log-Weighted Parameter Moments for Different Types of Parameter Convolutions.
- Table 3. Description of parameters influencing water conservation effectiveness.
- Table 4. Indicators of Effectiveness for Conservation Actions

LIST OF FIGURES

- Figure 1. Distribution of rainfall catchment among households in Amman, Jordan.
- Figure 2. Distribution among households of water conserved by retrofitting showerheads.
- Figure 3. Analytically derived distributions of conservation action effectiveness
- Figure 4. Chart for sizing targeted water conservation programs

Table 1. Parameter Distributions and Methods to Calculate Log-Weighted Moments for Different Types of Known Information

Parameter information known	Lower bound a Upper bound b	Lower bound a Mean m	Frequencies, f_i , for discrete ranges of parameter values
PDF est. method PDF formula	Maximum entropy (rectangular distribution) $pdf(x) = \begin{cases} 0, & x \leq a \\ \frac{1}{b-a}, & x \leq b \\ 0, & x \geq b \end{cases}$	Maximum entropy (exponential decay distribution) $pdf(x) = \begin{cases} 0, & x \leq a \\ e^{\lambda_0 + \lambda_1 \cdot x}, & x \geq a \end{cases}$ $\lambda_1 = \frac{1}{a-m}, \lambda_0 = \ln(-\lambda_1) - \lambda_1 \cdot a$	Empirical results (histogram) $P(x_i \leq x \leq x_{i+1}) = f_i$
Log-weighted parameter moments	$\mu_j = \int_{x=a}^b \ln x \cdot \frac{1}{b-a} dx$ $\sigma_j^2 = \int_{x=a}^b (\ln x)^2 \frac{1}{b-a} dx - \mu_j^2$	$\mu_j = \int_{x=a}^{\infty} \ln x \cdot e^{\lambda_0 + \lambda_1 \cdot x} dx$ $\sigma_j^2 = \int_{x=a}^{\infty} (\ln x)^2 \cdot e^{\lambda_0 + \lambda_1 \cdot x} dx - \mu_j^2$	$\mu_j = \sum_i \int_{x=x_i}^{x_{i+1}} \ln x \cdot f_i \cdot dx$ $\sigma_j^2 = \sum_i \int_{x=x_i}^{x_{i+1}} (\ln x)^2 \cdot f_i \cdot dx - \mu_j^2$
Integration method Formula for 1st moment Formula for 2nd moment	Analytic integration by parts $\mu_j = \left[x(\ln x - 1) \right]_{x=a}^{x=b} \cdot \frac{1}{b-a}$ $\sigma_j^2 = \left[x((\ln x)^2 - 2 \ln x + 2) \right]_{x=a}^{x=b} \cdot \frac{1}{b-a} - \mu_j^2$	Numerical integration $\mu_j \approx \sum_{i=0}^N \left\{ \ln x_i \cdot e^{\lambda_0 + \lambda_1 x_i} \cdot (x_{i+1} - x_i) \right\}$ $\sigma_j^2 \approx \sum_{i=0}^N \left\{ (\ln x_i)^2 \cdot e^{\lambda_0 + \lambda_1 x_i} \cdot (x_{i+1} - x_i) \right\} - \mu_j^2$	Analytic integration by segments $\mu_j = \sum_i \left\{ \left[x(\ln x - 1) \right]_{x=x_i}^{x=x_{i+1}} f_i \cdot (x_{i+1} - x_i) \right\}$ $\sigma_j^2 = \sum_i \left\{ \left[x((\ln x)^2 - 2 \ln x + 2) \right]_{x=x_i}^{x=x_{i+1}} \cdot f_i \cdot (x_{i+1} - x_i) \right\} - \mu_j^2$

Table 2. Probability Distributions and Methods to Calculate Log-Weighted Parameter Moments for Different Types of Parameter Convolutions

Parameter information known	Uniform dist. with range $[a_2, b_2]$ subtracted from uniform distribution with range $[a_1, b_1]$.	Uniform dist. with range $[a_2, b_2]$ subtracted from exponential decay distribution with lower bound a_1 and mean m_1 .	Exponential decay distribution with lower bound a_2 , mean m_2 subtracted from exp. decay dist. with lower bound a_1 , mean m_1 .
PDF est. method	Convolution to give trapezoid dist. with fixed lower bound d_1 and upper bound d_2 .	Convolution to give quasi-triangle dist. with fixed lower bound and exponential decaying upper bound	Convolution to give exponential rising limb and exponential falling limb centered at $a_1 - a_2$.
PDF formula	$pdf(x) = \begin{cases} 0, & x < d_1, \\ m \cdot (x - d_1), & x \leq d_2 \\ 1/d_3, & x \leq d_3 \\ -m \cdot (x - d_4), & x \leq d_4 \\ 0, & x > d_4 \end{cases}$ $d_1 = a_1 - b_2, d_4 = d_2 + d_3,$ $d_2 = d_1 + \min(b_1 - a_1, b_2 - a_2),$ $d_3 = d_1 + \max(b_1 - a_1, b_2 - a_2),$ $\text{and } m = 1/(d_2 d_3)$	$pdf(x) = a \cdot \begin{cases} 0, & x \leq a_1 - b_2 \\ e^{\lambda_1(x+b_2)} - e^{\lambda_1 a_1}, & x \leq a_1 - a_2 \\ e^{\lambda_1 x} (e^{\lambda_1 b_2} - e^{\lambda_1 a_2}), & x \geq a_1 - a_2 \end{cases}$ <p>Where $a = \frac{e^{\lambda_0}}{(b_2 - a_2) \cdot \lambda_1}$, $\lambda_1 = \frac{1}{a_1 - m_1}$, and $\lambda_0 = \ln(-\lambda_1) - \lambda_1 \cdot a_1$</p>	$pdf(x) = \frac{-1}{\lambda_{1,1} + \lambda_{1,2}} e^{\lambda_0 + \lambda_1 x}$ <p>where for $x < a_1 - a_2$ $\lambda_1 = -\lambda_{1,2}$ and $\lambda_0 = \lambda_{0,1} + \lambda_{0,2} + a_1 \cdot (\lambda_{1,1} + \lambda_{1,2})$ and for $x > a_1 - a_2$ $\lambda_1 = \lambda_{1,1}$ and $\lambda_0 = \lambda_{0,1} + \lambda_{0,2} + a_2 \cdot (\lambda_{1,1} + \lambda_{1,2})$</p>
Log-weighted parameter moments	$\mu_j = \int_c^{d_4} \ln x \cdot \frac{pdf(x)}{1 - p_c} dx$ $\sigma_j^2 = \int_c^{d_4} (\ln x)^2 \cdot \frac{pdf(x)}{1 - p_c} dx - \mu_j^2$ <p>where $p_c = \int_{-\infty}^c pdf(x) \cdot dx$</p>	$\mu_j = \int_{x=c}^{\infty} \ln x \cdot \frac{pdf(x)}{1 - p_c} dx$ $\sigma_j^2 = \int_{x=c}^{\infty} (\ln x)^2 \cdot \frac{pdf(x)}{1 - p_c} dx - \mu_j^2$ <p>where $p_c = \int_{-\infty}^c pdf(x) \cdot dx$</p>	$\mu_j = \int_{x=c}^{\infty} \ln x \cdot \frac{pdf(x)}{1 - p_c} dx$ $\sigma_j^2 = \int_{x=c}^{\infty} (\ln x)^2 \cdot \frac{pdf(x)}{1 - p_c} dx - \mu_j^2$ <p>where $p_c = \int_{-\infty}^c pdf(x) \cdot dx$</p>
Integration method	Analytic integration by parts from cutoff value $c > 0$ to d_4	Numerical integration from cutoff value $c > 0$ to infinity.	Numerical integration from cutoff value $c > 0$ to infinity.

Table 3. Description of parameters influencing water conservation effectiveness

Effectiveness parameter	Units	Low value	High value	Average	St. Dev	Distribution ¹	Reference (sample size)
Geographic							
A. Annual rainfall	mm/yr	110	550	270	94	FG	JMD, 2000 (78 years)
Demographic							
B. Households sharing building	#/building	1	--	2.7	--	ED	DOS, 2004 (383,000 households)
C. Roof area of building	m2	100	--	206	--	ED	DOS, 1999 (1,800 households)
D. Household size	persons	3	--	5.1	--	ED	DOS, 2004 (383,000 households)
E. Households that garden outdoors	unitless, 1 = yes	0	1	0.15	--	BI	DOS, 1999 (1,800 households)
Technologic - existing infrastructure							
F. House water pressure	bar	0.29	--	0.80	--	ED	Linearly correlated to parameter (b)
G. Shower flow rate - current device	l/min	6	20	--	--	UN	Engineering estimate (10 devices)
H. Faucet flow rate - current device	l/min	5.5	20	--	--	UN	Engineering estimate (10 devices)
I. Toilet tank volume - current device	l/flush	5.5	15	--	--	UN	IrDC, 2004 (31 devices)
J. Hose diameter	cm	1.3	3.8	--	--	UN	Engineering estimate, 0.5 to 1.5 inches
K. Bucket size	l	11.4	26.5	--	--	UN	Engineering estimate, 3 to 7 gallons
Behavioral - existing water uses							
L. Length of shower - current	min	3	--	8.4	--	ED	Engineering estimate
M. Shower frequency	#/week	1	--	3.6	--	ED	Engineering estimate (28 persons)
N. Toilet flushes	#/person/day	1	--	4.03	--	ED	Snobar, 2003 (30 households)
O. Faucet use - bathroom	min/day/person	0.1	--	0.6	--	ED	Snobar, 2003 (30 households)
P. Faucet use - kitchen	min/day	1	--	14.36	--	ED	Snobar, 2004 (30 households)
Q. Irrigation frequency	#/week	0.22	--	1.45	--	ED	Engineering estimate (23 households)
R. Irrigation season	weeks/year	20	40	--	--	UN	Engineering estimate
S. Floor wash frequency	#/week	1	7	--	--	UN	Engineering estimate
Technologic - potentially adopted							
T. Shower flow rate - retrofit device	l/min	6	9	--	--	UN	Engineering estimate (10 devices)
U. Faucet flow rate - retrofit device	l/min	5.5	6.5	--	--	UN	Engineering estimate (10 devices)
V. Toilet flush rate - retrofit, full	l/flush	5.5	6.5	--	--	UN	IrDC, 2004 (16 devices)
Behavioral - potential modifications							
X. Reduced irrigation time - nozzle	minutes/use	0.5	--	3	--	ED	Engineering estimate
Y. Bucket application to floor	buckets/wash	1	--	5	--	ED	Engineering estimate

¹Distributions: BI=binary; FG=fitted gamma; ED=exponential decay; UN=uniform.

Table 4. Indicators of Effectiveness for Conservation Actions

Effectiveness Indicator	Collect Rainwater ^b	Retrofit Showerhead	Retrofit Bath Faucets ^b	Retrofit Kitchen Faucets	Retrofit Toilets	Spray Nozzle on Hoses ^b	Install Floor Carpets ^b
Average effectiveness			[m ³ per customer per year]				
Point estimate (typical)	20.4	43.8	8.0	35.4	31.7	7.7	19.7
Probabistic analytical estimate (proposed)	27.4	44.7	8.2	44.7	31.7	6.6	19.9
Required program size			[percentage of customers] ^a				
Point estimate (blanket application)	104 %	48 %	265 %	60 %	67 %	275 %	108 %
Probabilistic analytical estimate (targeted application)	49 %	10 %	NA	8 %	26 %	NA	NA
Threshold effectiveness level from probabistic est.			[m ³ per customer per year] ^a				
Target customers with effectiveness larger than	20.6	106.8	NA	124.8	38.8	NA	NA
Retrofit cost			[\$US per customer]				
Average customer expenditure	\$282	\$82	\$4	\$4	\$117	\$4	\$4,442
Required program expenditure			[\$US Million] ^a				
Blanket application (point estimate)	NA	\$12.1	NA	\$0.8	\$24.0	NA	NA
Targeted application (probabilistic approach)	\$42.7	\$2.5	NA	\$0.1	\$9.3	NA	NA

Notes: a. Percentage based on potential customer population of 306,000 households and conservation objective of 6.5 Mm³/year

b. Percentage greater than 100% or NA means not possible to achieve conservation objective

Figure 1. Distribution of Rainfall Catchment among Households

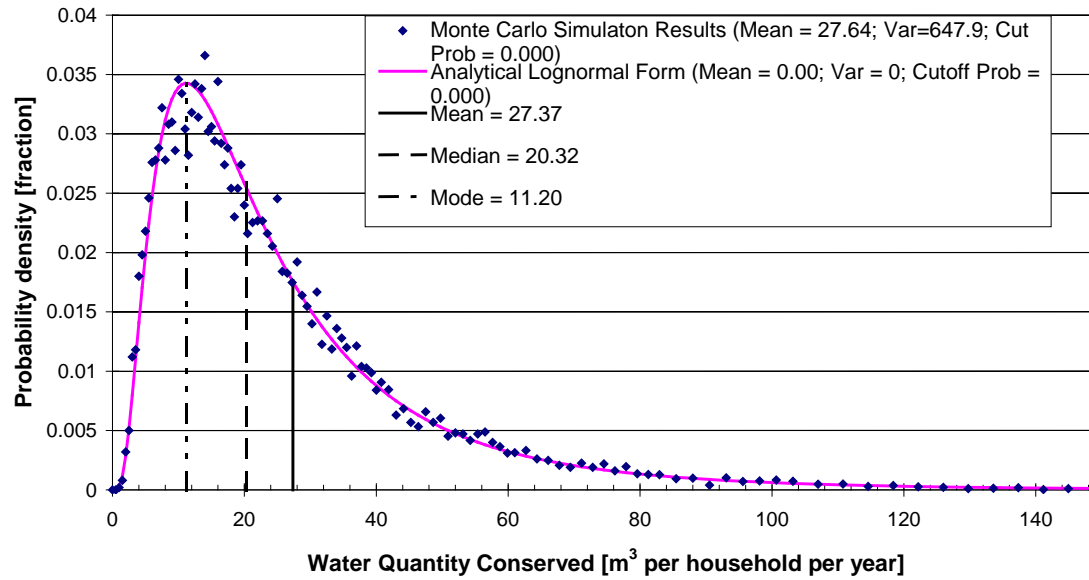


Figure 2. Distribution among households of water conserved by retrofitting showerheads

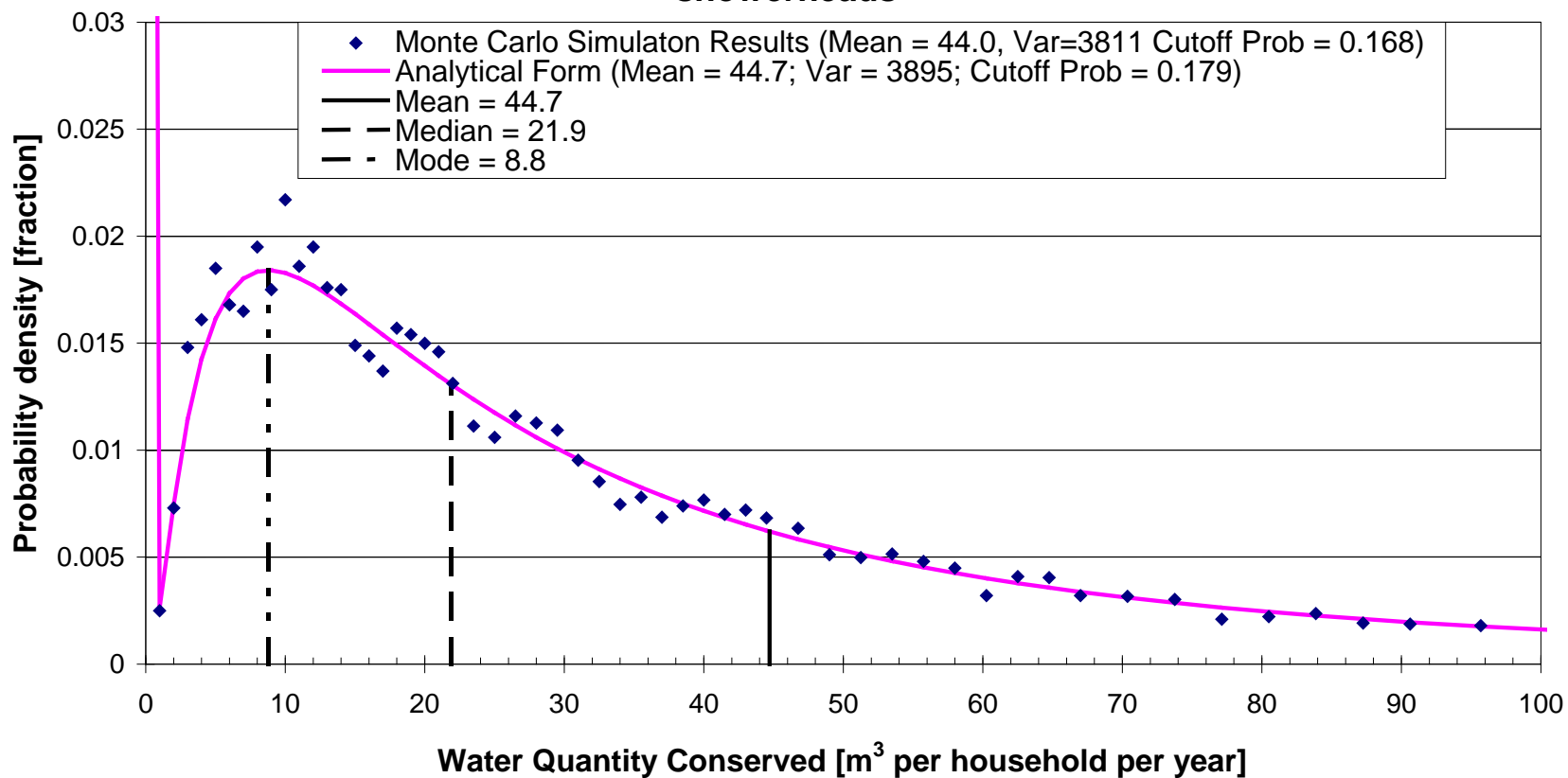


Figure 3. Analytically derived distributions of conservaton action effectiveness

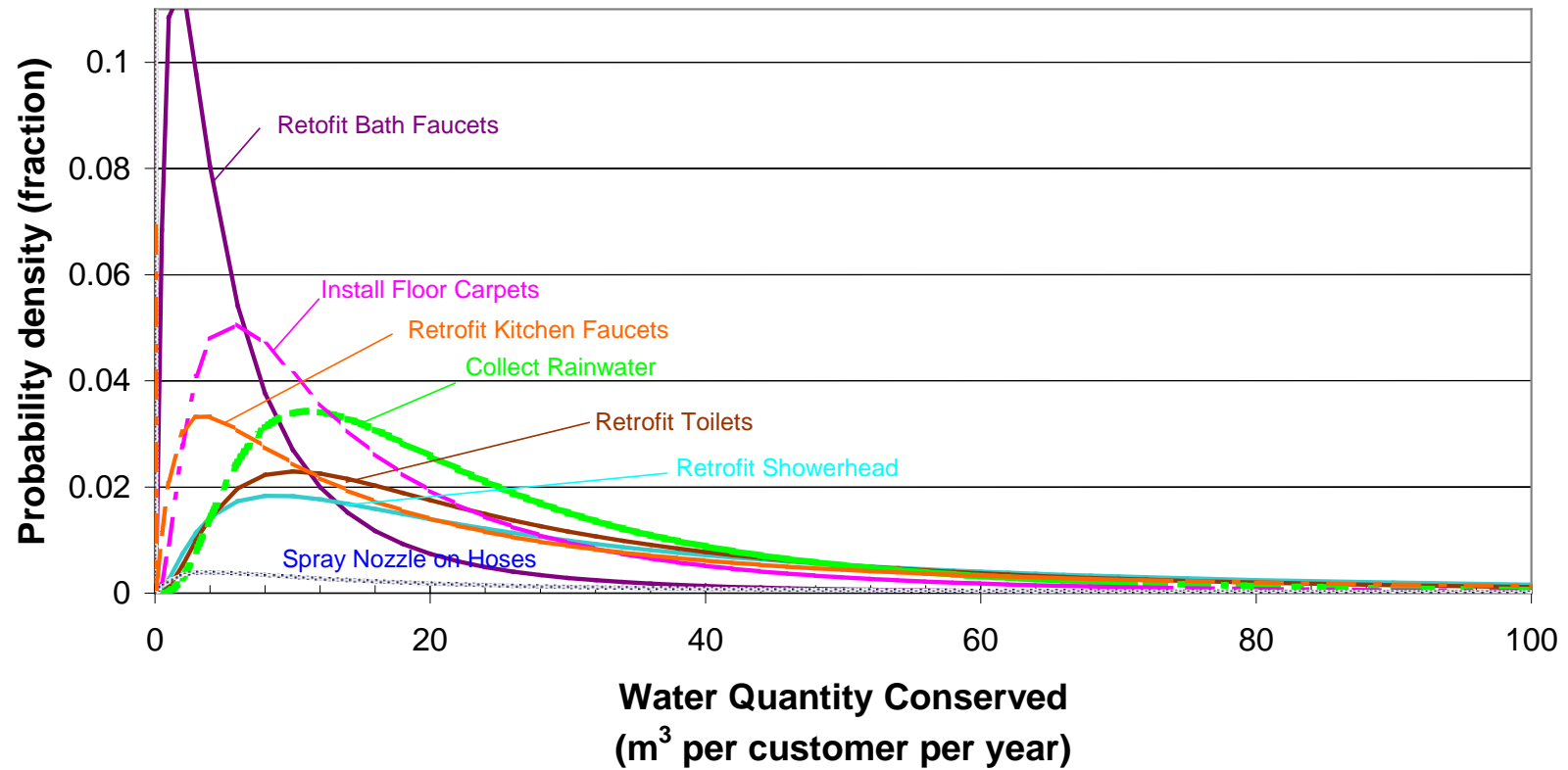


Figure 4. Chart for sizing targeted water conservation programs

