$\mathcal{I}^2\mathcal{SDS}$

The Institute for Integrating Statistics in Decision Sciences

Technical Report TR-2012-5

March 12, 2012

Urban water demand forecasting: A state-of-the-art review

Emmanuel A. Donkor
Department of Engineering Management and Systems Engineering
The George Washington University

Thomas A. Mazzuchi
Department of Engineering Management and Systems Engineering
The George Washington University

Refik Soyer
Department of Decision Sciences
The George Washington University

J. Alan Robertson
American Water Works Association
Washington DC

Urban Water Demand Forecasting: A Review of Methods and Models

Emmanuel A. Donkor, SM.ASCE[1]; Thomas A. Mazzuchi[2];

Refik Soyer[3]; and James Alan Roberson, P.E.[4]

ABSTRACT: We review the literature on urban water demand forecasting published from 2000 to 2010 in order to identify methods and models useful for specific water utility decision making problems. Results show that although a wide variety of methods and models have attracted attention, applications of these models differ, depending on the forecast variable, its periodicity and the forecast horizon. Whereas Artificial Neural Networks are more likely to be used for short-term forecasting, Econometric Models, coupled with simulation or scenario-based forecasting, tend to be used for long-term strategic decisions. If utilities are to make decisions that incorporate uncertainty in future demand forecasts, then much more attention needs to be given to probabilistic forecasting methods.

Key words: Urban Water Demand; Forecasting; Time Series; Uncertainty; Decision making

---

[1] Doctoral candidate, Engineering Management and Systems Engineering, School of Engineering and Applied Sciences, The George Washington University, 1776 G Street NW, Suite 101., Washington, D.C. 20052, United States(corresponding author). E-mail: eadonkor@gwu.edu.
[2] Professor, Engineering Management and Systems Engineering, School of Engineering and Applied Sciences, The George Washington University, 1776 G Street NW, Suite 101., Washington, D.C. 20052, United States.
[3] Professor, Decision Sciences, School of Business, The George Washington University, 2201 G Street NW, Funger Hall, Suite 415C., Washington, D.C. 20052. United States.
[4] Director of Security and Regulatory Affairs, American Water Works Association, 1300 Eye Street NW, #701W., Washington, D.C. 2005, United States.

Introduction

Accurate urban water demand forecasting provides the basis for making operational, tactical and strategic decisions for drinking water utilities (Billings & Jones, 2008; Gardiner & Herrington, 1990) and is critical for several reasons. For instance, utilities need to know what the water demand for today and tomorrow will be in order to operate their treatment plants and wells appropriately to meet these demands. Utilities also need to accurately predict the water demand 20-30 years in the future in order to develop new water sources and/or expand their treatment plants. Accounting for the uncertainty in these forecasts will help utilities optimize their operational and investment decisions.

Although forecasting is not a new discipline, its application in the water sector for demand estimation is fraught with many problems to the extent that it is known to be notoriously difficult due, probably, to the nature and quality of data available, the numerous variables that are hypothesized to affect water demand (Arbués, García-Valiñas, & Martínez-Espiñeira, 2003) and the multiplicity of forecast horizons and periodicities involved. These characteristics have engendered a plethora of studies in an attempt to improve forecast reliability. Despite this effort, water demand forecasting practice undertaken by utilities and their consultants currently differ extensively with respect to the models used, causing past research and practice to serve as prologues to our study.

As a contribution towards improving future forecasts, this paper surveys the recent literature on water demand forecasting. The intended purpose is to provide a guide to the literature for water utility managers looking for guidance on improving the practice of forecasting for effective decision making, and for water researchers seeking to extend the current knowledge in the field. In concert with this objective, we focus attention on four main areas. Firstly, we provide a framework for demand forecasting by characterizing the utility decision problems that engender the need for demand forecasting in the water sector and by identifying the decision variables and predominant determinants used. Secondly, we create an inventory of forecasting methods and models as used by drinking water utility managers and demand forecasters. Here, we differentiate between forecast models, methods and approaches for clarity. We will refer to forecast models when discussing specific mathematical formulations for predicting demand from time series data. Forecast *method* and *technique* will be used interchangeably and will refer to the class of qualitative and quantitative means by which forecast models are developed. We use the term *forecast approach* to refer to a collection of methods, tools, and processes for estimating future water demand. Thirdly, we synthesize the literature, classifying it by method and forecast periodicity, in order to identify what the main focus of research has been. It is important to note that *forecast horizon* refers to how far into the future demand is to be predicted, a terminology different from *forecast periodicity* which refers to the time span between consecutive forecasts, e.g hourly, monthly. Finally, we make proposals on how the practice of water demand forecasting can be improved.

Arbués et al., (2003) has reviewed the literature on residential water demand modeling, where the focus was on cross-sectional data for pricing purposes. While cross sectional models are important for identifying determinants of water demand for pricing and/or demand management purposes, they are inadequate for the kinds of planning decisions that utilities make when future demand is uncertain. We therefore differentiate our study by focusing on methods and models for time series data.

Framework for Water Demand Forecasting

Basis for Forecasting

As in other industries, planning for decision making forms the basis for forecasting in the water sector. A set of water demand forecasting literature differentiates forecast practice by the level of planning associated with the forecast (Gardiner & Herrington, 1990), or in accordance with the forecast horizon (Billings & Jones, 2008). In terms of planning level, all water demand forecasting exercises can be used either for strategic, tactical or operational decision making. These respectively concern decisions for capacity expansion, investment planning and system operation, management and optimization. In terms of forecast horizon, water demand forecasting can be categorized as either long-term, medium-term or short-term with these horizons being reflective of the general purpose of the forecast. Long-term, medium-term and short-term forecasts are prepared for strategic, tactical and operational decisions respectively (Alvisi, Franchini, & Marinelli, 2007; Ghiassi, Zimbra, & Saidane, 2008; Jain, Varshney, & Joshi, 2001).

No generally accepted time frame exists for these horizons. For instance, Billings and Jones (2008) contain different time horizons regarding what constitutes long-term, medium-term and short-term forecasts. One such definition classifies forecasts spanning more than 2 years as *long-term*, those from 3 months to less than 2 years as *medium-term*, and forecasts for 1 to 3 months as *short-term*. This contrasts with Gardner and Herrington (1990) who classify these categories as annual forecasts for 10 years or more, annual forecasts for 1 to less than 10 years, and hourly to monthly forecasts up to a year. In terms of application, Ghiassi et al. (2008) prepared monthly demand forecasts for 2 years, weekly demand forecasts for 6 months and daily demand forecasts for 2 weeks, and characterize these as long-term, medium-term and short-term respectively. We follow Gardiner and Herrington in categorizing urban water demand forecasts and present a summary of the relationship between the planning level, decision problem, forecast horizon and forecast periodicity in Table 1.

Forecast Variables and Determinants

Apart from the various planning levels and horizons that tend to complicate urban water demand forecasting, the forecast variable of interest and the determinants of water demand are two features that add to the complexity. In decreasing order of frequency, results of a survey through the American Water Works Association (AWWA) portray the kind of variables that are of interest to water utilities in urban water demand forecasting: *peak day* (73.9%), *daily total system demand* (65.9%), *monthly total system demand* (65.6%), *annual per capita demand* (65.4%), *annual demand by customer class* (58.0%) and *revenue* (57.9%) (Billings & Jones, 2008). These results clearly show that urban water demand forecasting can involve different variables measured at different periodicities. Appreciating this fact is important because the question regarding which method to use for urban water demand forecasting cannot be adequately answered without specifying the forecast variable and its periodicity. For example, interest in the annual variation of *per capita demand* makes the related variable a candidate for models amenable to medium to long-term forecasting. Forecasting this variable in the long-term may require completely different determinants as compared to its short-term equivalent. Similarly, because *total system demand* can be measured on hourly up to yearly basis, different forecasting models may be required to forecast it, depending on its periodicity.

In the drinking water community, many variables are considered influential in determining water demand. These range from socio-economic to various derivatives of weather-related variables. Examples of these weather-related variables and how they are used can be found in Coomes et al.,(2010) and Brekke, Larsen, Ausburn, & Takaichi (2002). Although "*A good understanding of the factors influencing demand and reliable estimates of the parameters describing demand behavior and consumption patterns are prerequisites* [to a good forecast]" (Burney, Mukhopadhyay, Al-Mussallam, Akber, & Al-Awadi, 2001), the enormity of these variables can create frustration for water utility mangers. As an illustration of the size and variability of the variables that can be considered, we refer to a report prepared for the Water Research Foundation by Coomes, Rockaway, Rivard, & Kornstein

(2010), in which the authors tested the effect of 26 variables on *average daily water use* for 293 residential customers of the Louisville Water Company. Thus, the use of regression models, as in the Coomes et al.(2010) paper, tend to exert the greatest demand on data collection and management, due to the many factors that are postulated to influence water demand.

The choice of independent variables can also influence the forecasting method used. For instance, whereas population projections and per capita demand are the drivers for unit rate models, these have no consideration when exponential smoothing methods or Box-Genkins models are used. Similarly, the mean and standard deviation of water demand time series data are essential for Geometric Brownian Motion (GBM) models but play no role in time series decomposition models.

Evaluating Forecast Accuracy

The accuracy of forecasts is evaluated by comparing them with observed demand. This evaluation provides insights in recommending changes to existing models in order to reduce deviations in future forecasts. Prior to observing future forecasts, however, such evaluations can form the basis for model selection. The general approach is to consider competing forecasting models in a sequence of steps: (i) divide the data set into an estimation period and a hold-out period; (ii) use the estimation period to model demand; (iii) evaluate the accuracy of the models by comparing the forecasts with the observed values for both the estimation period and the hold-out period and (iv) select the best model based on its performance, as measured by any of the loss functions specified in equations (1) to (4). For N time periods, Y_t and $\hat{Y}_t$ respectively represent the actual observation and the forecast value at time *t*.

$$\text{Mean Absolute Deviation (MAD)} = \frac{1}{N}\sum_{t=1}^{N}\left|Y_t - \hat{Y}_t\right| \qquad (1)$$

$$\text{Mean Absolute Percentage Error (MAPE)} = \frac{100}{N}\sum_{t=1}^{N}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right| \quad (2)$$

$$\text{Mean squared Error (MSE)} = \frac{1}{N}\sum_{t=1}^{N}\left(Y_t - \hat{Y}_t\right)^2 \quad (3)$$

$$\text{Root Mean Squared Error (RMSE)} = \sqrt{MSE} \quad (4)$$

In each of these functions, the forecast error is measured by the difference between the observed and forecast values, represented by $Y_t - \hat{Y}_t$. When comparing forecast performance for a given sample data, the model with the least value of a chosen loss function is deemed the most accurate. It is important, however, to note that these functions measure different characteristics of forecast error and hence sometimes model ranking could be different for each loss function. The MAD prefers models with the least deviations on average. The MAPE is similar but unitless. The MSE and its associated RMSE penalizes models that have large deviations and hence is used to select models that fit the data well.

Deterministic forecasts are not accurate most of the time and hence it is better to speak of their reliability. This calls for establishing threshold values on the loss functions in order to judge how reliable forecasting has been for a given utility. The non-existence of such threshold values makes it problematic for comparing the performance of models across utilities, since demand levels differ. Of the various metrics, MAPE might be the only one for which a threshold value can be used to compare forecasting performance among utilities, since it is independent of system capacity.

Forecasting Methods and Models

Several methods and models are available for urban water demand forecasting. These range from the simple to the complex and can be qualitative or quantitative in nature. A selection of the predominant quantitative forecasting methods and their associated mathematical models are presented here. We discuss each method, making reference to their appearance in the literature where possible.

Judgmental and Unit Rate Methods

Judgmental or qualitative methods include the use of heuristics or rule-based methods to forecast the value of a variable of interest. According to Gardiner & Herrington, "… [these] approaches rely upon the experience of an individual or, less likely, a group, and may be either entirely subjective in nature or a modification of more objective results derived from other approaches" (Gardiner & Herrington, 1990). In practice, the use of qualitative methods is necessitated by a desire for rudimentary forecasts for purposes of simplicity. Billings & Jones (2008) describe the method as applied by water utilities while Jentgen, Kidder, Hill, & Conrad (2007) reports of specific cases where utilities in the USA (Jacksonville Electric Authority, San Diego Water Department, Colorado Springs Utilities and Las Vegas Valley Water District) use heuristics, regression and neural networks, to prepare short-term forecasts for optimizing pumping schedules.

Unlike judgmental methods, demand forecasting based on what Brekke et al. (2002) terms "unit water demand analysis" employs the consumption per unit of a customer category, example *per capita water demand*, and the number of units of that category, example *population/size of domestic customers*, to forecast water demand. Here, the demand for a given future time period *t* ($Q_{i,t}$) is computed by taking the product of the unit consumption ($q_{i,t}$) and the number of units ($N_{i,t}$). Where a utility's customer mix includes other categories, the method requires disaggregating demand by *customer segment*, preparing forecasts for each, and then adding these forecasts to generate the total. The mathematical expression for this *Sectoral Forecasting* model is presented in equation (5), where for *C* customer categories indexed by *i*, the demand forecast at a future time *t* is given by:

$$\sum_{i=1}^{C} Q_{i,t} = \sum_{i=1}^{C} q_{i,t} * N_{i,t} \tag{5}$$

Although unit rate models do not feature extensively in the literature, Billings & Jones (2008) and Jentgen et al. (2007) have noted that in practice, it is the simplest model used by most utilities. All that is required is to estimate $q_{i,t}$ and $N_{i,t}$ in order to obtain a forecast. As an example, the Washington Metropolitan Area (WMA) water studies have consistently adopted a "unit use coefficient approach…[since] 1990…[because]…it is a transparent and easily understandable method…and was judged to provide the right balance between data needs and accuracy" (Hagen, Holmes, Kiang, & Steiner, 2005). It is therefore not surprising to observe that 65.4% of the utilities surveyed in Billings & Jones (2008) spend resources in forecasting per capita water demand. The reliability of these forecasts is questionable when simple rules of thumb or expect judgment, instead of empirical analysis, form the basis for estimating $q_{i,t}$ and $N_{i,t}$ for each customer category.

Forecasting by Time Series Analysis

Time series models, or what is technically known as extrapolation forecasts in the drinking water community (Billings & Jones, 2008; Gardiner & Herrington, 1990), forecast future water demand on the basis of past observations and associated error terms. They rely on the fundamental assumption that past trends will be repeated in the future. Their failure to take into consideration the effects of changes in demographic, economic and technological variables, as well as water demand management strategies (such as public awareness campaigns and/or price adjustments) in influencing future water demand turns out to be the main criticism. They may be more useful for short- to medium-term forecasts where variations in the determinants of the demand variable are expected to be negligible.

Moving Average and Exponential Smoothing Models

Moving averages and exponential smoothing models are simple deterministic time series models whose identification can be effected by understanding the eight generic time series profiles depicted in Fig. 1. Each of these profiles can have one or more of three main components: a level component (L_t), a trend component (T_t), and a seasonal component (S_t). The underlying mathematical model for each of these profiles depends on which components are present as well as the nature of the variation (constant or changing), observable in the series. These models are formulated as equations (6) to (13) in Table 2, where m = number of periods in the forecast lead-time, Y_t = observed value of demand at time t; $\hat{Y}_t$ = forecast of demand for period t; and $\hat{Y}_{t+m}$ = forecast of demand for m periods ahead from period t.

There are four profiles that exhibit no seasonal variation, as depicted in Figures 1a to 1d. Where the variation is constant and the series has no trend (Fig 1a), it is best modeled by a Single Moving Average as formulated in equation (6), where k = number of historical periods used in calculating the moving average. In the case where the variation increases or decreases with time as in Fig 1b, the m period ahead forecast is modeled with a Single Exponential Smoothing function as in equation (7), where α = smoothing parameter for the level of the series. In Fig 1c, the series has a trend and a constant but non-seasonal variation around the trend. This is modeled as a Double Moving Average for which the m period ahead forecast is given by equation (8), where M_t = 1^{st}/single moving average and D_t = 2^{nd}/double moving average. Finally, where the non-seasonal profile is characterized by a trend and a changing variation as depicted in Fig 1d, the forecast is obtained by using a Double Exponential Smoothing model through equation (9), where γ = smoothing parameter for the trend.

The remaining four profiles exhibit seasonal variation: a constant seasonal variation without a trend (Fig 1e); an increasing seasonal variation without a trend (Fig 1f); a constant seasonal variation with a trend (Fig 1g); and an increasing seasonal variation with a linear trend (Fig 1h). Respectively, the mathematical formulations of the underlying exponential smoothing models are termed Seasonal Additive [equation (10)], Seasonal Multiplicative [equation (11)], Holt-Winters Additive [equation (12)] and Holt-Winters Multiplicative [equation (13)], where δ = smoothing parameter for seasonal indices and p = number of periods in the seasonal cycle (p = 4 for quarterly data and 12 for monthly data). It is required that $p \geq m$.

In practice, time-dependent water demand data do not have, although may contain semblances of, such nice profiles. A typical example can be found in Brekke et al.(2002) and in Figure 3a. Under such circumstances, the time series can be decomposed into its components and then composite forecasting methods can be used. Figure 2 is an example of using additive seasonal decomposition to disaggregate monthly water production into its level (random), trend and seasonal components.

Discrete and Continuous –time Stochastic Process Models

Sometimes, time series data can exhibit more complex profiles for which the predominant exponential smoothing models cease to be adequate. Stochastic process models, which can be formulated in discrete- or continuous-time, are more advanced alternatives that can be used to model these complex structures. These models are mathematical formulations of processes that obey specific probabilistic and statistical laws and thus result in a series of outcomes for each period over a given time span.

Discrete-time models, typically known as Box-Genkins models, are categorized in equations (14) to (17), where ϕ = autoregressive or damping parameter; θ = moving average parameter; μ = mean value of the process; σ = standard deviation of the process; and ϵ_t = forecast error at time t. Because of their stochastic nature, these models can be simulated if their parameters are known. Here, the error term (ϵ_t) is assumed to follow a Normal (0,σ) distribution and the coefficients ϕ and θ take values between 0 and 1 for stationary time series. As presented in equations (14) to (16), the models are for processes that are assumed to respectively follow an autoregressive process of order p, symbolized by AR(p), a moving-average process of order q, symbolized by MA(q), and an autoregressive moving-average process of order (p,q), symbolized by ARMA(p,q). Where the time series has to be differenced by order *d* in order to make it stationary, a generalized formulation, using the stationary AR backshift operator $\phi_p(B) = (1- \phi_1 B - \ldots - \phi_p B^p)$ and the invertible MA backshift operator $\theta_q(B) = (1- \theta_1 B - \ldots - \theta_p B^q)$, can be used to represent an ARIMA (p,d,q) process as shown in equation (17).

$$Y_t = \mu + \sum_{k=1}^{p} \phi_k Y_{t-k} + \epsilon_t \tag{14}$$

$$Y_t = C + \epsilon_t + \sum_{k=1}^{q} \theta_k \epsilon_{t-k} \tag{15}$$

$$Y_t = \mu + \sum_{k=1}^{p} \phi_k Y_{t-k} + \epsilon_t + \sum_{k=1}^{q} \theta_k \epsilon_{t-k} \tag{16}$$

$$\phi_p(B)(1-B)^d Y_t = \theta_q(B)\epsilon_t \tag{17}$$

As an example, Alhumoud (2008) uses an MA (1) process to model freshwater demand in Kuwait. Model formulation requires the creation and understanding of what is known as the autocorrelation function (ACF) and partial autocorrelation function (PACF) in order to specify the number of historical/lagged components of the demand variable. Unfortunately, some regression models have incorporated lagged variables of demand without considering the structure of the underlying ACF and PACF, but rather on simple correlation between demand at time *t* and a couple of its previous values.

Continuous-time stochastic process models are typically formulated as Geometric Brownian Motion (GBM) models. Historical data is used, when it is available, to estimate parameters such as the mean (μ), volatility (σ) and jump size (θ) of the underlying process. Where historical data are not available, estimates can be made using comparable data or by making assumptions about the expected values of these parameters. These stochastic processes can be used to forecast univariate time series variables, example *population*, which in turn may serve as independent variables in an econometric model.

Several GBM models are available. For illustrative purposes, we present formulations for the most frequently encountered models in equations (18) to (20). Let S = the previous value of a demand variable, δS = the change in the variable's value from one step to the next, δt = change in time; ϵ = the white noise error term; μ = the periodic growth or drift term and σ = the periodic volatility. Then equation (18) models a random walk process. Similarly, equation (19) models a mean reversion process where $\bar{S}$ = the long-term value the process reverts to; and η = the rate of reversion to the mean. Finally, equation (20) is used to model a jump diffusion process, for which ω = the jump size of S; $F(\lambda)$ = cumulative distribution function of a Poisson process and λ = the jump rate of S. In both (19) and (20), e = the exponent term.

$$\frac{\delta S}{S} = \mu(\delta t) + \sigma\epsilon\sqrt{\delta t} \tag{18}$$

$$\frac{\delta S}{S} = \mu(\delta t) + \sigma\epsilon\sqrt{\delta t} + \eta\left(\bar{S}e^{\mu(\delta t)} - S\right)\delta t \tag{19}$$

$$\frac{\delta S}{S} = \mu(\delta t) + \sigma\epsilon\sqrt{\delta t} + \eta\left(\bar{S}e^{\mu(\delta t)} - S\right)\delta t + \omega F(\lambda)\delta t \tag{20}$$

In general, very little is known about the application of stochastic process models in the water demand forecasting literature, despite their potential value, which resides in the ability of these models to provide multiple paths of the future value of a variable. It is different from scenario-based approaches in that a large number of possible paths can result from the underlying process and thus allows for uncertainty analysis in forecasting. Billings and Jones (2008) refer to this method as *Risk Simulation* and presents an example on five possible paths of an AR(1) process for *population*, with a mean growth rate of 2.4% and a range of 0.5% - 5%. As an alternative, we present a continuous time analog in Fig. 3, which shows the profile for 11 years (2000-2010) of monthly water production data from a US utility, juxtaposed with five paths generated from equation (19). Thousands of these paths can be generated, allowing for a probability distribution of the demand variable to be assessed at each point in time.

Time-series Regression or Econometric Models

Regression models produce forecasts on the basis of the relationship between water demand and its determinants. These can be built using cross-sectional data, time series data or panel data. Cross-section models allows for the evaluation of the impact of policy variables that can be used to mange demand. Example, knowing the impact of price on water demand can and has informed pricing in order to conserve water (Arbués et al., 2003). Along this line of research, Arbués and Villanua (2006) used demand modeling to test the sensitivity of residential demand to price as a means of influencing water pricing for demand management in Zaragoza, Spain. A quick reference on the variables, model specification, data sets and the commonly encountered estimation problems in modeling water demand for demand management purposes can be found in Arbués(2003).

In time series regression, the determinants vary temporally instead of cross-sectional. Sometimes, *time* serves as the main determinant. Mathematical models are presented in Table 3, where $\hat{Y}_t$ is the forecast value of the demand variable at time *t*. An example of its application for water demand forecasting can be found in Polebitski & Palmer (2010). The generalized formulation in equation (21) represents a multiple linear regression and translates into equations (22) to (27), depending on the nature and number of determinant. Here, β_o is the regression intercept and *i* is an index of the *i*th independent variable, for a total of *n* such variables. Respectively, β_i and $x_{i,t}$ represent the coefficient and observed value of the *i*th independent variable. When $n = 1$ the generalized model becomes a simple linear regression model as in equation (22). Equation (23) has *time* and its polynomial derivatives as determinants, where *v* is the power of the polynomial function. Equation (24) uses dummies/indicator variables (S = 0,1) to model seasonal variation in the demand variable, where s = 3 or 11 for quarterly or monthly demand respectively. Equation (25) is an extension of (23) and (24), where *time* is used to model an identifiable trend. Non-linearities are accommodated in equations (26) and (27), where constant elasticities are assumed for the former and variable elasticities assumed for the latter. Sometimes, a mix of demand determinants, time and its polynomials, as well as indicator variables, are used in mixed time-series regression models as exemplified in Brekke et al. (2002).

Among other criteria, regression models require the residuals or error terms to be independent of each other. Time-series regression models are however known for the serial correlation of the error terms. If the residuals are correlated, the appropriate ARIMA model must be used to model them for subsequent integration with the parent model. For example, Burney et al. (2001) uses equation (26) and models the error term with an AR (1) process. However, serial correlation of the error terms can result in correlation structures more complicated than the usual 1st order autocorrelation assumed. The model in Burney et al is of special concern because the authors assumed a 1st order autocorrelation, unlike the expected correct approach of explicitly examining the ACF and PACF and modeling the residuals with the appropriate ARIMA (p,d,q) process.

Scenario-based Approaches and Decision Support Systems

Scenario-based approaches are basically regression models that determine the level of demand for long-term forecasts given specific scenarios. They are used when there is a need to account for uncertainty in demand forecasts brought on by a limited number of discrete combinations of the independent variables. The idea here is to determine the impact on water demand of various future scenarios of the determinants. Examples of this approach for urban water demand forecasting can be found in Burney et al. (2001), Polebitski, Palmer, & Waddell (2011), Wei et al. (2010) and Williamson, Mitchell, & McDonald (2002). Strzepek et al.(1999) presents a similar approach for linking climate change scenarios with water planning for agriculture.

Sometimes, the need for automating scenario-based approaches results in the development of custom-made decision support programs. The forecasts obtained from such programs are normally driven by system-defined models derived from different methods (see Feng, Li, Duan, & Zhang (2007), Froukh (2001) and Jain & Ormsbee (2001)), allowing the decision maker to select the most appropriate combination of methods that satisfies assumptions concerning future scenarios. Currently, IWR-MAIN, developed by the U.S. Army Corps of Engineers' Institute for Water Resources, and Demand-Side Management Least-Cost Planning Decision Support System (DSS), created by Maddaus Water Management, Alamo, California, are two well-known DSS packages for water demand forecasting and demand management. The use of these packages have been demonstrated by Mohamed & Al-Mualla (2010a, 2010b) and Levin, Maddaus, Sandkulla, & Pohl (2006) respectively.

Artificial Neural Networks

Artificial neural networks (ANN) and fuzzy logic techniques of forecasting water demand are advanced methods classified as *nonparametric* in Billings & Jones (2008). They can be used for both regression and time series models yet they do not require adherence to the assumptions that form the basis for these methods. However, identifying the optimal architecture first requires the determination of the structure of a univariate time series or regression model (Pulido-Calvo, Montesinos, Roldan, & Ruiz-Navarro, 2007). The structural components in the dataset are determined using a training data-set (estimation period) while forecasts are produced and compared with a hold-out data set (hold-out period).

The appearance of neural network models for water demand forecasting in the literature normally involves a comparative assessment of the performance between different neural network models and conventional regression models (Adamowski & Karapataki, 2010; Firat, Yurdusev, & Turan, 2009; Herrera, Torgo, Izquierdo, & Perez-Garcia, 2010; Jentgen, Kidder, Hill, & Conrad, 2007; Pulido-Calvo, Montesinos, Roldan, & Ruiz-Navarro, 2007) or with time series models (Ghiassi et al., 2008; Pulido-Calvo & Gutierrez-Estrada, 2009) or with both (Bougadis, Adamowski, & Diduch, 2005; Jain & Ormsbee, 2002; Jain et al., 2001; Pulido-Calvo, Roldán, López-Luque, & Gutiérrez-Estrada, 2003).

Hybrid Models or Composite Forecasts

Finally, an approach that has found relatively widespread application in water demand forecasting is classified as hybrid models (Jentgen et al., 2007; Pulido-Calvo & Gutierrez-Estrada, 2009). These models use more than one method and/or model to arrive at a composite forecast and usually involve some form of combination of forecasts from models via simple or weighted averages (Caiado, 2010; Wang, Sun, Song, & Mei, 2009) or by applying a mix of methods and models to forecast the decomposed components (see Fig. 3) of a time series (Alvisi et al., 2007; Aly & Wanakule, 2004; Gato, Jayasuriya, & Roberts, 2007a, 2007b; Zhou, McMahon, Walton, & Lewis, 2000). In the case of combining forecasts from different models to obtain a composite forecast, the following expression is used:

$$\hat{Y}_t = \beta_0 + \sum_{i=1}^{n} \beta_i \hat{Y}_{i,t} \tag{28}$$

In equation (28), $\hat{Y}_{i,t}$ is the predicted value of the time series at time *t* using the *i*th model. The β_i coefficients are determined by optimization or least squares regression to minimize the mean squared error (see equation (3)) between the composite forecast $\hat{Y}_t$ and the actual data.

The urban water demand forecasting literature tends to favor composite forecasts developed by decomposition. An example of this practice is found in Wu & Zhou (2010), where the authors used linear regression to model the deterministic component of demand and ANN to model the cyclical component. They subsequently compared their results with forecasts obtained from separately using conventional regression and ANN to model the deterministic and cyclic components respectively.

In the paper "*Combining Forecasts: A review and annotated bibliography*", Clemen provides the rationale for the use of these hybrid models: "… [the] idea of combining forecasts implicitly assumed that one could not identify the underlying process, but that different forecasting models were able to capture different aspects of the information available for prediction" Clemen (1989). In most instances, these composite forecasts are reported to have led to better forecasting performance for water demand (Wang et al., 2009).

Empirical Studies on Water Demand Forecasting: Current Emphasis

In this section we present and discuss results on what the emphasis of water demand forecasting research has been, in terms of the methods used and the periodicity of the demand variable. In Table 5, we classify the forecasting literature, as inventoried in previous sections, by *periodicity of demand variable* and by *method*.

Results show a considerable variation in the occurrence of methods in the literature. Note from Table 5 the coupling of neural network and conventional methods, and the paucity of stand-alone Box-Genkins models. Very little focus has been placed on the latter and much less considered are judgmental methods and the more advanced GBM models. However, two less known methods, micro-simulation as examined in Williamson et al.(2002) and space-time forecasting with Bayesian Maximum Entropy (MBE) as in Lee, Wentz, & Gober (2010), have been used for scenario-based approaches.

Results presented in Table 5 indicate that there is a shift from pure conventional methods to a focus on three approaches: (1) scenario-based and DSS models: approaches which accommodate some amount of uncertainty in demand forecasting; (2) comparative assessment of performance between neural nets and conventional methods; and (3) recognition of the need to improve forecast accuracy by using hybrid models.

Over the analysis period examined here, time-series regression models have been used extensively. In most instances they are not stand-alone but compared with neural networks and/or in combination with univariate time series models. The general conclusion from this line of work is that models developed from neural networks perform better than those developed by time-series regression or univariate time series models. This superiority in performance is attributed to the ability of neural networks to efficiently capture non-linearities that may exist in the structure of time-series regression and univariate time series models. However, the much touted better performance of ANN over these conventional methods refers mainly to short-term forecasts with very little research conducted on how they compare over medium-to-long-term forecasts.

Table 5 shows that in terms of periodicity of the demand variable, the literature is skewed towards daily and annual variation of water demand, reflecting an emphasis on satisfying the mandate of water utilities, which is to maintain a reliable supply of potable water to consumers and to ensure that this level of reliability is maintained in future years. Thus, forecasting for operations and strategic planning purposes seems to have been the emphasis in recent times.

A few characteristic features of Table 5 are worth mentioning here: (1) Stand-alone regression models have focused on monthly variation of demand probably to accommodate the seasonal variation of weather variables which are best modeled using econometrics (2) Annual variation in demand variables have attracted research employing scenario-based and DSS models, ostensibly to model uncertainties and (3) Long-term demand forecasting has not benefited much from the comparative performance assessment between neural networks and conventional methods. The table also seems to suggest that less emphasis has been placed on hourly variation of demand. Contrary to this perception, hourly variation of demand has been accommodated in the papers that examined multiple periodicities, except Zhou et al. (2002).

Overall, improving forecast accuracy, accounting for uncertainty in long-term forecasts and maintaining system reliability now and in the future seem to have provided the impetus for current research in urban water demand forecasting. As evident in Table 5, it is difficult to answer the question "*Which model is best for water demand forecasting?*" without specifying the periodicity of the demand variable.

The literature also reveals that neural networks and hybrid models are more appropriate for short-term forecasts but for extended ones, where incorporating future scenarios of a variable might be important, scenario-based and DSS models are more suitable. However, the use of regression in modeling monthly demand follows the generally held view that short-to-medium-term demand is typically influenced by weather variables while long-term forecasts are more determined by socio-economic factors. Ghiassi et al., emphasizes this view by noting that "…when analyzing or forecasting water demand over a longer time horizon such as decades, economic or demographic factors may be more effectively included in models" (Ghiassi et al., 2008). This view makes it imperative to use econometric models for long-term demand forecasting and apparently influenced the model developed by Burney et al. (2001). The economic and demographic factors, as well as the influences of demand management strategies, technological change and climate change on future demand, do not change quickly in the short-term and their long-term estimates can take one of several values. Thus, although weather variables are not predominantly included in long-term demand forecasting, the realities of uncertainty in climate change have resulted in papers that include various climate scenarios in long-term forecasts. The procedure may follow either Burney et al.(2001), Goodchild (2003), Polebitski et al (2011) or Wei et al.(2010) or may be incorporated in a

DSS model similar to Froukh (2001), Levin et al. (2006) and Mohamed & Al-Mualla (2010a, 2010b).

Improving the Practice of Water Demand Forecasting

The need to improve the practice of urban water demand forecasting calls for paying attention to some modeling issues observed in the literature. The first concerns the practicality of some models proposed in the literature. In all its forms, the search for models for practical application should lead to models whose input variables can be collected, monitored and used by the utility. Models that contain many variables, as is Coomes et al. (2010), and those that utilize derivatives such as “days since 2mm of rainfall” (Goodchild, 2003) pose the greatest challenge to practice in terms of collecting and keeping track of the data. Operationalizing such models will be practically difficult when regression is used for short-term forecasting. For determinants that are not under the control of water utilities, considerable difficulty in acquiring reliable forecasts of such variables will preclude the use of econometric models in favor of time series analysis models. Researchers should therefore take into consideration the ability of utilities to acquire and monitor predictors if the models they propose are to be used for forecasting. Models should be as parsimonious as possible without compromising on forecast quality.

Related to the problem of numerous determinants is what seems to be a naïve and baseless selection of autoregressive terms to model time series data as in Bougadis et al.(2005) and Jain et al.(2001), and the inclusion of lagged variables of both demand and weather derivatives in regression models (see for example Jain & Ormsbee (2002), Jentgen et al.(2007), Pulido-Calvo et al.(2007) and Wei et al.(2010)). The arbitrary use of various lags of the demand variable provides practicing water utilities with a less rational basis for their inclusion in forecasting models. Contrary to such practices, the use of autoregressive terms should be informed by the structure of the autocorrelation and partial autocorrelation functions to account for serial autocorrelation in the data, if any. This will require not a cursory exposure to, but a clear understanding of the Box-Genkins methodology. Where lagged values of other variables are included, the proper unit root tests must be conducted to justify their inclusion in order to avoid the problem of spurious correlations that can exist in time series regression models. In this regard, we make reference to Burney et al. (2001) and Martinez-Espineira (2007) who adequately handled the concept of co-integration and provided the right approach to including indicator variables in econometric forecasting of water demand.

An opportunity exists for conducting ex-post evaluation of the accuracy of implemented models. None of the papers reviewed for this study report the results of such research, although it is one of the key recommendations made in Billings and Jones (2008). From the literature, it is difficult to evaluate if past forecasts turned out to be accurate and reliable. It will be useful to have a retrospective evaluation of selected methods, by comparing how forecast values compared with actual demand. For future research, we propose post-evaluation of forecast methods for selected exiting utilities. This could be done along the lines proposed by Fischer, Herrnstadt, & Morgenstern (2009) and Shlyakhter, Kammen, Broido, & Wilson (1994), where the authors evaluated the errors in energy demand projections in the US. The value of this line of study is to contribute in improving the accuracy of future probabilistic forecasts as exemplified in the US energy sector. A good source of data for starting such a research in the water sector is the Interstate Commission on the Potomac River Basin's (ICPRB) initiative, where every five years, the institution prepares annual water demand forecasts for a 20-year horizon (Hagen et al., 2005).

To date, the quest for a generalized forecasting model that can be used by all utilities has eluded researchers. For each conventional method, there is currently no acceptable model for forecasting water demand, irrespective of the planning level involved. For instance, for univariate time series models, the probabilistic structure which generates *total monthly demand* is not known for certain. This contrasts the airline model described by Box et al. (1987) and models used for forecasting stock price returns. Such industry models are known to follow specific stochastic processes, making it possible for analysts to focus on determining the parameters of these models for a given data set. The existence of these stochastic industry models creates a need for similar models to be developed for the drinking water community. This line of research can concentrate on identifying the probabilistic structures that generate the series for short, medium and long-term water demand forecasting. The validation of such models will help utilities estimate and better manage the uncertainty in demand forecasts and will result in optimized operations and investment decisions.

Conclusion

This review has presented an overview of the water demand forecasting literature appearing from 2000 to 2010 in an attempt to identify which forecasting approaches, methods and models are appropriate for urban water utility management decision problems that are dependent on future levels of demand. We conclude that (1) the basis for urban water demand forecasting is enshrined in utility management decision problems that are dependent on uncertain /stochastic future levels of demand, and that the forecast horizon and periodicity are key drivers to method and model selection; (2) although in practice, unit rate models are predominantly preferred, these have not attracted research attention in recent times, rather (3) there has been a shift to scenario-based forecasting and approaches that use decision support systems, a comparison of the performance of Neural Networks against conventional methods, and the use of hybrid models, all in an attempt to account for uncertainty and to improve forecast accuracy; (4) if utilities are to account for uncertainty and make decisions that incorporate this uncertainty in their forecasts, then the neglected probabilistic forecasting methods will require greater attention than currently received, as an advanced step beyond scenario-based forecasting approaches.

Acknowledgement
This study was funded by the National Oceanic and Atmospheric Administration's (NOAA) Climate Program Office through the American Water Works Association, as part of the project "Decreasing climate induced risk for municipal water demand forecasting". We appreciate the feedback received from participants of two workshops in 2011 sponsored by AWWA designed to discuss outputs of this study.

References

Adamowski, J., & Karapataki, C. (2010). Comparison of Multivariate Regression and Artificial Neural Networks for Peak Urban Water-Demand Forecasting: Evaluation of Different ANN Learning Algorithms. *Journal of Hydrologic Engineering*, *15*(10), 729–743. doi:10.1061/(ASCE)HE.1943-5584.0000245

Alhumoud, J. (2008). Freshwater consumption in Kuwait: analysis and forecasting. *Journal of Water Supply Research and Technology-AQUA*, *57*(4), 279–288. doi:10.2166/aqua.2008.036

Alvisi, S., Franchini, M., & Marinelli, A. (2007). A short-term, pattern-based model for water-demand forecasting. *Journal of Hydroinformatics*, *9*(1), 39–50. doi:10.2166/hydro.2006.016

Aly, A., & Wanakule, N. (2004). Short-term forecasting for urban water consumption. *Journal of Water Resources Planning and Management-ASCE*, *130*(5), 405–410. doi:10.1061/(ASCE)0733-9496(2004)130:5(405)

Arbués, F., Garcıa-Valiñas, M. Á., & Martınez-Espiñeira, R. (2003). Estimation of residential water demand: a state-of-the-art review. *Journal of Socio-Economics*, *32*(1), 81. doi:10.1016/S1053-5357(03)00005-2

Arbues, F., & Villanua, I. (2006). Potential for pricing policies in water resource management: Estimation of urban residential water demand in Zaragoza, Spain. *Urban Studies*, *43*(13), 2421–2442. doi:10.1080/00420980601038255

Billings, B., & Jones, C. (2008). *Forecasting Urban Water Demand, 2e* (2nd ed.). American Waterworks Association.

Bougadis, J., Adamowski, K., & Diduch, R. (2005). Short-term municipal water demand forecasting. *Hydrological Processes*, *19*(1), 137–148. doi:10.1002/hyp.5763

Box, G. E. P., Pierce, D. A., & Newbold, P. (1987). Estimating Trend and Growth Rates in Seasonal Time Series. *Journal of the American Statistical Association*, *82*(397), 276.

Brekke, L., Larsen, M., Ausburn, M., & Takaichi, L. (2002). Suburban water demand modeling using stepwise regression. *Journal American Water Works Association*, *94*(10), 65–75.

Burney, N., Mukhopadhyay, A., Al-Mussallam, N., Akber, A., & Al-Awadi, E. (2001). Forecasting of freshwater demand in Kuwait. *Arabian Journal for Science and Engineering*, *26*(2B), 99–113.

Caiado, J. (2010). Performance of Combined Double Seasonal Univariate Time Series Models for Forecasting Water Demand. *Journal of Hydrologic Engineering*, *15*(3), 215–222. doi:10.1061/(ASCE)HE.1943-5584.0000182

Clemen, R. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, *5*, 559–583.

Coomes, P., Rockaway, T., Rivard, J., & Kornstein, B. (2010). *North America residential water usage trends since 1992*. Denver Colo.: Water Research Foundation.

Feng, S., Li, L., Duan, Z., & Zhang, J. (2007). Assessing the impacts of South-to-North Water Transfer Project with decision support systems. *Decision Support Systems*, *42*(4), 1989–2003. doi:10.1016/j.dss.2004.11.004

Firat, M., Yurdusev, M., & Turan, M. (2009). Evaluation of Artificial Neural Network Techniques for Municipal Water Consumption Modeling. *Water Resources Management*, *23*(4), 617–632. doi:10.1007/s11269-008-9291-3

Fischer, C., Herrnstadt, E., & Morgenstern, R. (2009). Understanding errors in EIA projections of energy demand. *Resource and Energy Economics*, *31*(3), 198–209. doi:16/j.reseneeco.2009.04.003

Froukh, M. (2001). Decision-support system for domestic water demand forecasting and management. *Water Resources Management*, *15*(6), 363–382.

Gardiner, V., & Herrington, P. (1990). *Water Demand Forecasting* (1st ed.). Spon Press.

Gato, S., Jayasuriya, N., & Roberts, P. (2007a). Temperature and rainfall thresholds for base use urban water demand modelling. *Journal of Hydrology*, *337*(3-4), 364–376. doi:10.1016/j.jhydrol.2007.02.014

Gato, S., Jayasuriya, N., & Roberts, P. (2007b). Forecasting residential water demand: Case study. *Journal of Water Resources Planning and Management-ASCE*, *133*(4), 309–319. doi:10.1061/(ASCE)0733-9496(2007)133:4(309)

Ghiassi, M., Zimbra, D., & Saidane, H. (2008). Urban water demand forecasting with a dynamic artificial neural network model. *Journal of Water Resources Planning and Management-ASCE*, *134*(2), 138–146. doi:10.1061/(ASCE)0733-9496(2008)134:2(138)

Goodchild, C. (2003). Modelling the impact of climate change on domestic water demand. *Journal of the Chartered Institution of Water and Environmental Management*, *17*(1), 8–12.

Hagen, E., Holmes, K., Kiang, J., & Steiner, R. (2005). Benefits of iterative water supply forecasting in the Washington, DC, metropolitan area. *Journal of the American Water Resources Association*, *41*(6), 1417–1430.

Herrera, M., Torgo, L., Izquierdo, J., & Perez-Garcia, R. (2010). Predictive models for forecasting hourly urban water demand. *Journal of Hydrology*, *387*(1-2), 141–150. doi:10.1016/j.jhydrol.2010.04.005

Jain, A., & Ormsbee, L. (2001). A decision support system for drought characterization and management. *Civil Engineering and Environmental Systems*, *18*(2), 105–140.

Jain, A., & Ormsbee, L. (2002). Short-term water demand forecast modeling techniques - Conventional methods versus AI. *Journal American Water Works Association*, *94*(7), 64–72.

Jain, A., Varshney, A., & Joshi, U. (2001). Short-term water demand forecast modelling at IIT Kanpur using artificial neural networks. *Water Resources Management*, *15*(5), 299–321.

Jentgen, L., Kidder, H., Hill, R., & Conrad, S. (2007). Energy management strategies use short-term water consumption forecasting to minimize cost of pumping operations. *Journal American Water Works Association*, *99*(6), 86–94.

Lee, S., Wentz, E., & Gober, P. (2010). Space-time forecasting using soft geostatistics: a case study in forecasting municipal water demand for Phoenix, Arizona. *Stochastic Environmental Research and Risk Assessment*, *24*(2), 283–295. doi:10.1007/s00477-009-0317-z

Levin, E., Maddaus, W., Sandkulla, N., & Pohl, H. (2006). Forecasting wholesale demand and conservation savings. *Journal American Water Works Association*, *98*(2), 102–111.

Martinez-Espineira, R. (2007). An estimation of residential water demand using co-integration and error correction techniques. *Journal of Applied Economics*, *10*(1), 161–184.

Mohamed, M., & Al-Mualla, A. (2010a). Water Demand Forecasting in Umm Al-Quwain (UAE) Using the IWR-MAIN Specify Forecasting Model. *Water Resources Management*, *24*(14), 4093–4120. doi:10.1007/s11269-010-9649-1

Mohamed, M., & Al-Mualla, A. (2010b). Water demand forecasting in Umm Al-Quwain using the constant rate model. *Desalination*, *259*(1-3), 161–168. doi:10.1016/j.desal.2010.04.014

Polebitski, A., & Palmer, R. (2010). Seasonal Residential Water Demand Forecasting for Census Tracts. *Journal of Water Resources Planning and Management-ASCE*, *136*(1), 27–36. doi:10.1061/(ASCE)WR.1943-5452.0000003

Polebitski, A., Palmer, R., & Waddell, P. (2011). Evaluating Water Demands under Climate Change and Transitions in the Urban Environment. *Journal of Water Resources Planning and Management-ASCE*, *137*(3), 249–257.

Pulido-Calvo, I., & Gutierrez-Estrada, J. C. (2009). Improved irrigation water demand forecasting using a soft-computing hybrid model. *Biosystems Engineering*, *102*(2), 202–218. doi:10.1016/j.biosystemseng.2008.09.032

Pulido-Calvo, I., Montesinos, P., Roldan, J., & Ruiz-Navarro, F. (2007). Linear regressions and neural approaches to water demand forecasting in irrigation districts with telemetry systems. *Biosystems Engineering*, *97*(2), 283–293. doi:10.1016/j.biosystemseng.2007.03.003

Pulido-Calvo, I., Roldán, J., López-Luque, R., & Gutiérrez-Estrada, J. C. (2003). Demand Forecasting for Irrigation Water Distribution Systems. *Journal of Irrigation and Drainage Engineering*, *129*(6), 422. doi:10.1061/(ASCE)0733-9437(2003)129:6(422)

Shlyakhter, A., Kammen, D., Broido, C., & Wilson, R. (1994). Quantifying the credibility of energy projections from trends in past data - The United-States energy sector. *Energy Policy*, *22*(2), 119–130.

Strzepek, K., Major, D., Rosenzweig, C., Iglesias, A., Yates, D., Holt, A., & Hillel, D. (1999). New methods of modeling water availability for agriculture under climate change: The US Cornbelt. *Journal of the American Water Resources Association*, *35*(6), 1639–1655.

Wang, X., Sun, Y., Song, L., & Mei, C. (2009). An eco-environmental water demand based model for optimising water resources using hybrid genetic simulated annealing algorithms. Part II. Model application and results. *Journal of Environmental Management*, *90*(8), 2612–2619. doi:10.1016/j.jenvman.2009.02.009

Wei, S., Lei, A., & Islam, S. (2010). Modeling and simulation of industrial water demand of Beijing municipality in China. *Frontiers of Environmental Science & Engineering in China*, *4*(1), 91–101. doi:10.1007/s11783-010-0007-6

Williamson, P., Mitchell, G., & McDonald, A. (2002). Domestic water demand forecasting: A static microsimulation approach. *Journal of the Chartered Institution of Water and Environmental Management*, *16*(4), 243–248.

Wu, L., & Zhou, H. (2010). Urban water demand forecasting based on HP filter and fuzzy neural network. *Journal of Hydroinformatics*, *12*(2), 172–184. doi:10.2166/hydro.2009.082

Zhou, S., McMahon, T., Walton, A., & Lewis, J. (2000). Forecasting daily urban water demand: a case study of Melbourne. *Journal of Hydrology*, *236*(3-4), 153–164.

Zhou, S., McMahon, T., Walton, A., & Lewis, J. (2002). Forecasting operational demand for an urban water supply zone. *Journal of Hydrology*, *259*(1-4), 189–202.

Table 1. Relationship among planning level, water utility decision problems and forecast attributes

Planning level	Decision problem	Forecast Horizon	Forecast Periodicity
Operational	System operation management and optimization	Short-term (Less than 1 year)	Hourly Daily Weekly Monthly
Tactical	Revenue forecast; Investment planning; Staging system improvement	Medium term (1 – 10 years)	Monthly Annual
Strategic	Capacity expansion	Long term: (More than 10 years)	Annual

Table 2. Moving average and exponential smoothing models

Model Description	Mathematical formulation	Equation
Single moving average	$\hat{Y}_t = (Y_{t-1} + Y_{t-2} + \ldots + Y_{t-k})/k$	(6)
Single exponential smoothing	$\hat{Y}_{t+m} = L_t$	(7)
	where $L_t = \alpha Y_t + (1-\alpha)L_{t-1}$	
Double moving average	$\hat{Y}_{t+m} = L_t + mT_t$	(8)
	where $L_t = 2M_t - D_t$	
	$T_t = 2(M_t - D_t)/(k-1)$	
	$M_t = (Y_t + Y_{t-1} + \ldots + Y_{t-k+1})/k$	
	$D_t = (M_t + M_{t-1} + \ldots + M_{t-k+1})/k$	
Double exponential smoothing	$\hat{Y}_{t+m} = L_t + mT_t$	(9)
	where $L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + T_{t-1})$	
	$T_t = \gamma (L_t - L_{t-1}) + (1 - \gamma)T_{t-1}$	
Seasonal additive	$\hat{Y}_{t+m} = L_t + S_{t+m-p}$	(10)
	where $L_t = \alpha(Y_t - S_{t-p}) + (1 - \alpha)L_{t-1}$	
	$S_t = \delta (Y_t - L_t) + (1 - \delta)S_{t-p}$	
Seasonal multiplicative:	$\hat{Y}_{t+m} = L_t \; S_{t+m-p}$	(11)
	where $L_t = \alpha(Y_t/S_{t-p}) + (1 - \alpha)L_{t-1}$	
	$S_t = \delta (Y_t/L_t) + (1 - \delta)S_{t-p}$	
Holt-Winters Additive	$\hat{Y}_{t+m} = L_t + mT_t + S_{t+m-p}$	(12)
	where $L_t = \alpha(Y_t - S_{t-p}) + (1 - \alpha)(L_{t-1} + T_{t-1})$	
	$T_t = \gamma (L_t - L_{t-1}) + (1 - \gamma) L_{t-1}$	
	$S_t = \delta(Y_t - L_t) + (1 - \delta)S_{t-p}$	
Holt-Winters Multiplicative	$\hat{Y}_{t+m} = (L_t + mL_t)S_{t+m-p}$	(13)
	where $L_t = \alpha(Y_t /S_{t-p}) + (1 - \alpha)(L_{t-1} + T_{t-1})$	
	$T_t = \gamma (L_t - L_{t-1}) + (1 - \gamma) T_{t-1}$	
	$S_t = \delta (Y_t/L_t) + (1 - \delta)S_{t-p}$	

Table 3. Time series regression models

Model Description	Model Formulation	Equation
Generalized regression model	$\hat{Y}_t = \beta_0 + \sum_{i=1}^{n} \beta_i x_{i,t} + \epsilon_t$	(21)
Simple linear	$\hat{Y}_t = \beta_0 + x_t + \epsilon_t$	(22)
Non-linear extrapolative	$\hat{Y}_t = \beta_0 + \sum_{j=1}^{k} \beta_i x_t^j + \epsilon_t$	(23)
Seasonal regression without trend	$\hat{Y}_t = \beta_0 + \sum_{i=1}^{s} \beta_{Si} S_i + \epsilon_t$	(24)
Seasonal regression with trend	$\hat{Y}_t = \beta_0 + \sum_{j=1}^{v} \beta_i x_t^j + \sum_{i=1}^{s} \beta_{Si} S_i + \epsilon_t$	(25)
Multiple regression with constant elasticities	$\ln(\hat{Y}_t) = \beta_0 + \sum_{i=1}^{n} \beta_i \ln(x_{i,t}) + \epsilon_t$	(26)
Multiple regression with variable elasticities	$\ln(\hat{Y}_t) = \beta_0 + \sum_{i=1}^{n} \beta_i x_{i,t} + \epsilon_t$	(27)

Table 4. Water demand forecasting literature by forecast periodicity and method

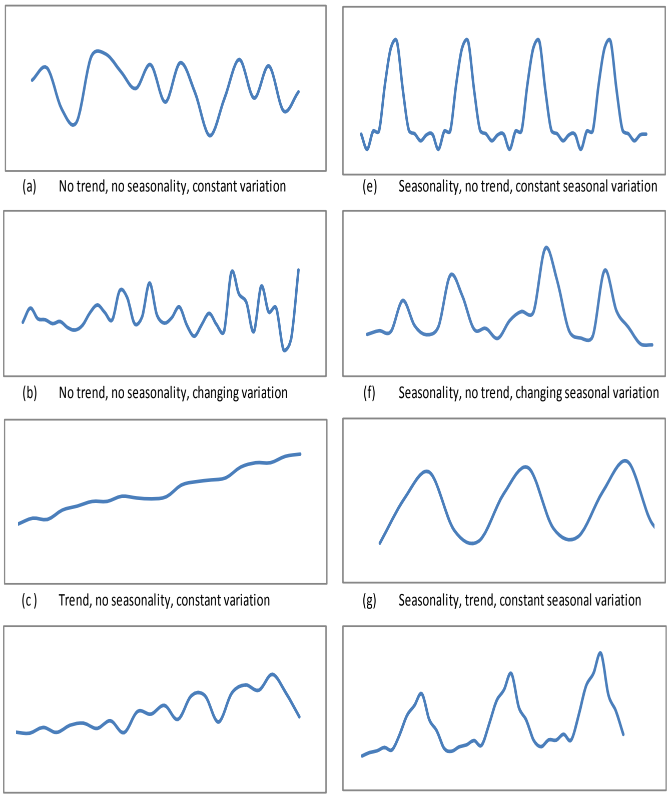
Periodicity of demand variable	**Water demand forecasting approach**						
	Box-Genkins	Regression	Scenario-based	DSS	Neural networks with conventional methods	Hybrid	**Total**
Hourly					Herrera et al. (2010)		**1**
Daily			Goodchild (2003)	Froukh (2001) Jain et al. (2001)	Jain et al. (2002) Pulido-Calvo et al. (2009) Pulido-Calvo et al. (2007) Pulido-Calvo et al. (2003)	Caiado (2010) Gato et al. (2007) Gato et al. (2007b) Zhou et al. (2000)	**11**
Weekly					Adamowski et al. (2010) Bougadis et al. (2005) Jain et al. (2001)		**3**
Monthly		Brekke et al.(2002) Martinez-Espineira (2007) Polebitski et al. (2010)			Firat et al. (2009)		**4**
Annual	Alhumoud (2008)		Burney et al. (2001) Lee et al. (2010) Wei et al. (2010) Williamson et al. (2002)	Feng et al. (2007) Levin et al. (2006) Mohamed et al. (2010b) Mohamed et al. (2010)		Wang et al. (2009) Wu et al. (2010)	**11**
Multiple periodicities					Ghiassi et al. (2008) Jentgen et al. (2007)	Alvisi et al. (2007) Aly et al. (2004) Zhou et al. (2002)	**5**
Total	**1**	**3**	**5**	**6**	**11**	**9**	**35**

Figure Captions

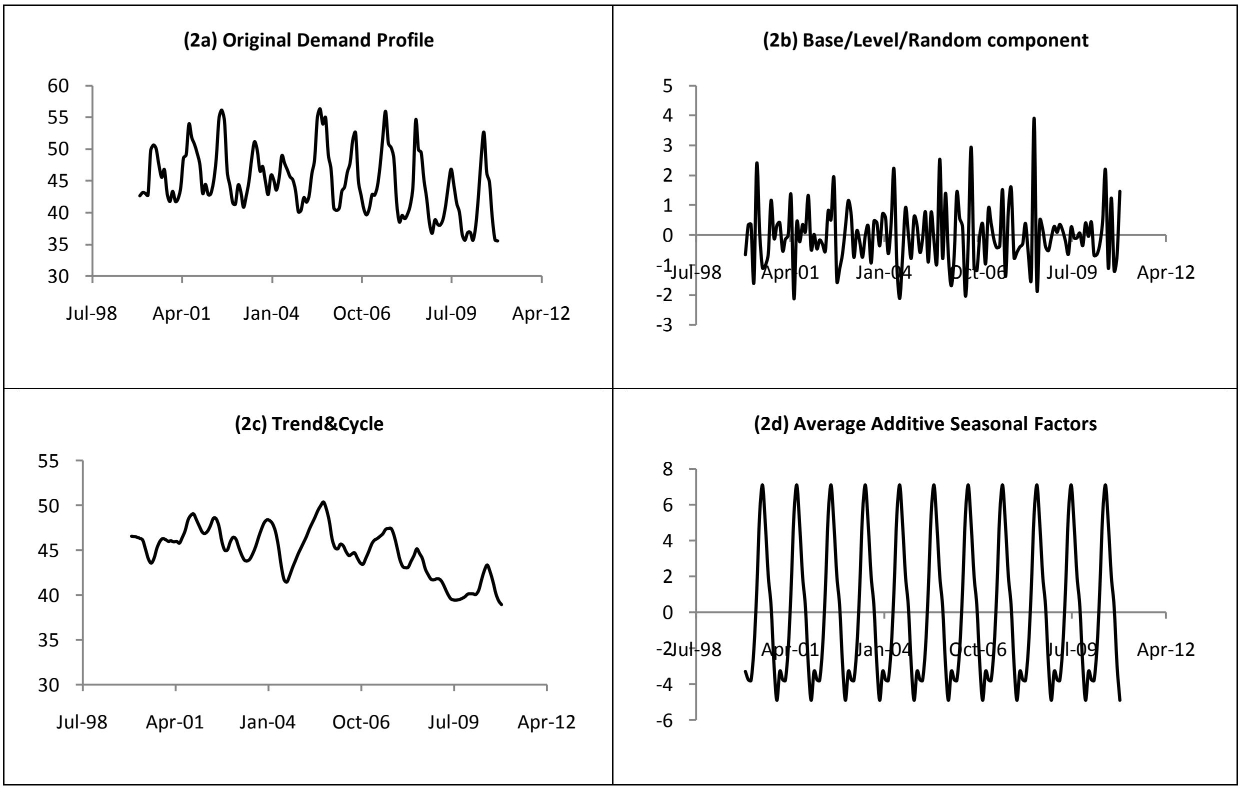
Figure 1. Generic time series profiles

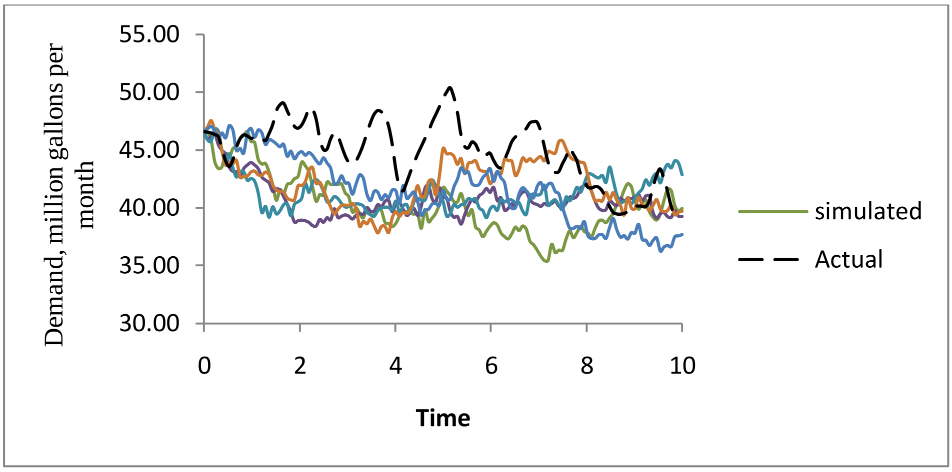
Figure 2. Time series decomposition

Figure 3. Actual series and five simulated paths of a mean reversion process of 10 years of monthly water demand data.



(a) No trend, no seasonality, constant variation
(b) No trend, no seasonality, changing variation
(c) Trend, no seasonality, constant variation
(d) Trend, no seasonality, changing variation
(e) Seasonality, no trend, constant seasonal variation
(f) Seasonality, no trend, changing seasonal variation
(g) Seasonality, trend, constant seasonal variation
(h) Seasonality, trend, changing seasonal variation


(2a) Original Demand Profile
60
55
50
45
40
35
30
Jul-98
Apr-01
Jan-04
Oct-06
Jul-09
Apr-12
(2b) Base/Level/Random component
5
4
3
2
1
0
-1
-2
-3
Jul-98
Apr-01
Jan-04
Oct-06
Jul-09
Apr-12
(2c) Trend&Cycle
55
50
45
40
35
30
Jul-98
Apr-01
Jan-04
Oct-06
Jul-09
Apr-12
(2d) Average Additive Seasonal Factors
8
6
4
2
0
-2
-4
-6
Jul-98
Apr-01
Jan-04
Oct-06
Jul-09
Apr-12



Demand, million gallons per month
55.00
50.00
45.00
40.00
35.00
30.00
0
2
4
6
8
10
Time
simulated
Actual

*Sizing worksheet (.xls)

Click here to download Sizing worksheet (.xls): MSizingWebExcel.xlsx

ASCE Worksheet for Sizing Technical Papers & Notes

*****Please complete and save this form then email it with each manuscript submission.*****

Note: The worksheet is designed to automatically calculate the total number of printed pages when published in ASCE t format.

Journal Name:	**f Water Resources Planning and Ma**	**Manuscript # (if known):**	
Author Full Name:	Emmanuel A. Donkor	**Author Email:**	eadonkor@gwu

The maximum length of a technical paper is 10,000 words and word-equivalents or 8 printed pages. A technical note should not exceed 3,50(word-equivalents in length or 4 printed pages. Approximate the length by using the form below to calculate the total number of words in the t adding it to the total number of word-equivalents of the figures and tables to obtain a grand total of words for the paper/note to fit ASCE form Overlength papers must be approved by the editor; however, valuable overlength contributions are not intended to be discouraged by this proc

1. Estimating Length of Text

A. Fill in the four numbers (highlighted in green) in the column to the right to obtain the total length of text.

NOTE: Equations take up a lot of space. Most computer programs don't count the amount of space around display equations. Plan on counting 3 lines of text for every simple equation (single line) and 5 lines for every complicated equation (numerator and denominator).

Estimating Length of Text

Count # of words in 3 lines of text:		40
Divided by 3		3
Average # of words per line		13
Count # of text lines per page		24
# of words per page		320.00
Count # of pages (don't add references & abstract)		24
Title & Abstract		500
Total # refs	47	1139
Length of Text is		9319
		664
		9983
		8

2. Estimating Length of Tables

A. First count the longest line in each column across adding two characters between each column and one character between each word to obtain total characters.

1-column table = up to 60 characters wide

2-column table = 61 to 120 characters wide

B. Then count the number of text lines (include footnote & titles)

1-column table = up to 60 characters wide by:
17 lines (or less) = 158 word equiv.
up to 34 lines = 315 word equiv.
up to 51 lines = 473 word equiv.
up to 68 text lines = 630 word equiv.

2-column table = 61 to 120 characters wide by:
17 lines (or less) = 315 word equiv.
up to 34 lines = 630 word equiv.
up to 51 lines = 945 word equiv.
up to 68 text lines = 1260 word equiv.

C. Total Characters wide by Total Text lines = word equiv. as shown in the table above. **Add word equivalents** for each table in the column labeled "**Word Equivalents.**"

3. Estimating Length of Figures

A. First reduce the figures to final size for publication.

Figure type size can't be smaller than 6 point (2mm).

B. Use ruler and measure figure to fit 1 or 2 column wide format.

1-column fig. = up to 3.5 in.(88.9mm)

2-col. fig. = 3.5 to 7 in.(88.9 to 177.8 mm) wide

C. Then use a ruler to check the height of each figure (including title & caption).

1-column fig. = up to 3.5 in.(88.9mm) wide by:
up to 2.5 in.(63.5mm) high = 158 word equiv.
up to 5 in.(127mm) high = 315 word equiv.
up to 7 in.(177.8mm) high = 473 word equiv.
up to 9 in.(228.6mm) high = 630 word equiv.

2-column fig. = 3.5 to 7 in.(88.9 to 177.8 mm) wide by:
up to 2.5 in.(63.5mm) high = 315 word equiv.
up to 5 in.(127mm) high = 630 word equiv.
up to 7 in.(177.8mm) high = 945 word equiv.
up to 9 in.(228.6mm) high = 1260 word equiv.

D. Total Characters wide by Total Text lines = word equiv. as shown in the table above. **Add word equivalents** for each table in the column labeled **"Word Equivalents."**

Estimating Length of Tables &

Tables	Word Equivalents	**Figures**
Table 1	158	Figure 1
2	315	2
3	158	3
4	630	4
5		5
6	0	6
7	0	7
8	0	8
9	0	9
10	0	10
11	0	11
12	0	12
13	0	13
14	0	14
15	0	15
Please double-up tables/figures if additional space is needed (ex. 20+21).		16
		17
		18
		19
		20 and 21

Total Tables/Figures:	2521	(word equivalents)
Total Words of Text:	9983	

Total words and word equivalents:	**12504**
printed pages:	**10**

updated 1/16/03

wo-column

.edu

) words and
ext and
at.
cedure.

subtototal
plus headings
TOTAL words
printed pages

Figures:
Word Equivalents
315
630
315
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0