WATER RESOURCES RESEARCH, VOL. 24, NO. 2, PAGES 201-206, FEBRUARY 1988

A Model of Daily Municipal Water Use for Short-Term Forecasting

JAMES A. SMITH

Interstate Commission on the Potomac River Basin, Rockville, Maryland

A time series model of daily municipal water use is developed. The model is termed a conditional autoregressive process and can be interpreted as an autoregressive process with randomly varying mean. The randomly varying mean accounts for changes in water use that result from the complex interaction over time of "structural features" of the water use system. These features may include the price of water, total service area connections, plumbing code provisions, and customer income, among many others. The modeling approach is semiparametric. The model can be split into a component that is treated in a nonparametric framework and a component that is treated parametrically. The random mean process, which represents long-term trend in water use, is treated in a nonparametric framework. Conditional on the random mean water use, the model reduces to a Gaussian autoregressive process with a modest number of parameters. The water use model is the core of a forecast system which is used to schedule releases from two water supply reservoirs which serve the Washington, D. C., Metropolitan Area. Model structure dictates that the key step in producing a water use forecast is an updating step in which a revised estimate of current mean water use is computed.

1. INTRODUCTION

In this paper a forecast system for daily municipal water use is developed. The forecast system is based on a time series model of water use and is used to schedule releases from two water supply reservoirs that serve the Washington, D. C., Metropolitan Area (WMA). The model is termed a conditional autoregressive process and can be viewed as an autoregressive process with randomly varying mean. The randomly varying mean accounts for changes in water use that result from the complex interaction over time of "structural features" of the water supply system. These features may include the price of water, total service area connections, plumbing code provisions, and customer income, among many others.

The modeling approach is semiparametric. The model can be split into a component that is treated in a nonparametric framework and a component that is treated parametrically. The random mean process, which represents long-term trend in mean water use, is treated in a nonparametric framework. Conditional on the random mean water use, the model reduces to a Gaussian autoregressive process with a modest number of parameters. Time series models with randomized parameters have been used in a variety of applications, including economic forecasting [e.g., *Nicholls and Quinn*, 1982; *Swamy*, 1982] and hydrologic modeling [*Klemes*, 1974; *Potter*, 1976; *Smith and Karr*, 1983].

In the WMA, trend in mean water use (see Figure 1) is tied to a number of factors beyond the control of water managers (such as sectors of growth in the local economy and the influence of interest rates on housing development). The determining factors of water use are not all beyond the control of water managers. Numerous publications have appeared assessing the possibility (and subsequent successes and failures) of manipulating the water use values appearing in Figure 1 through pricing and conservation measures (see, for example, *Howe and Linaweaver* [1967], *Davis and Hanke* [1973], and *Carver and Boland* [1980]). *Carver and Boland* [1980] report that seasonal price elasticities of water use for Washington, D. C., are "not significantly different from zero." Their results contrast with those of *Howe and Linaweaver* [1967] in which a seasonal price elasticity of -1.6 is reported. Carver and Boland note that "the elasticity of seasonal water use may have fallen in the WMA during the interval which separates the two studies (1963–1969). Present attitudes toward the environment and resource conservation differ considerably from attitudes of the early 1960s." The complex interaction of changing attitudes, prices, and interest rates all contribute to the random fluctuations over time of mean water use.

Parameter estimation and forecasting procedures developed for the random mean model are nonstandard (compare, for example, *Salas-LaCruz and Yevjevich* [1972] and *Maidment and Parzen* [1984]). A notable feature of the parameter estimation procedure is inclusion of a "state estimation" step (for a similar estimation procedure, see *Smith and Karr* [1985]). Model structure dictates that the key step in producing a water use forecast is an updating step in which a revised estimate of current mean water use is produced (updating algorithms for water use forecasting models are also considered by *Kher and Sorooshian* [1986]).

Contents of the sections are as follows. Model development is the topic of section 2. In section 3 we develop estimators for unknown parameters of the model. State estimation techniques necessary for implementing the forecast system are also developed in section 3. The forecast system is applied to WMA water use in section 4. A summary and conclusions are presented in section 5.

2. MODEL DEVELOPMENT

In this section we present a model for daily municipal water use. We denote daily water use on day t of year i by $X_i(t)$. A year consists of $T = 7J$ days (the number of weeks in the year is J). To facilitate modeling day-of-week features of water use, the first day of each year is taken to be a Sunday. In section 4, for example, a model for the period May–September is developed; the first day of the year is taken to be the first Sunday in May.

The random mean water use for day t of year i is assumed to be the product of two terms: Y_i, the random mean daily water use for year i and $m(t)$, the "unit demand function," which does not vary from year to year (we adopt the notation-

Paper number 7W0870.
0043-1397/88/007W-0870$05.00

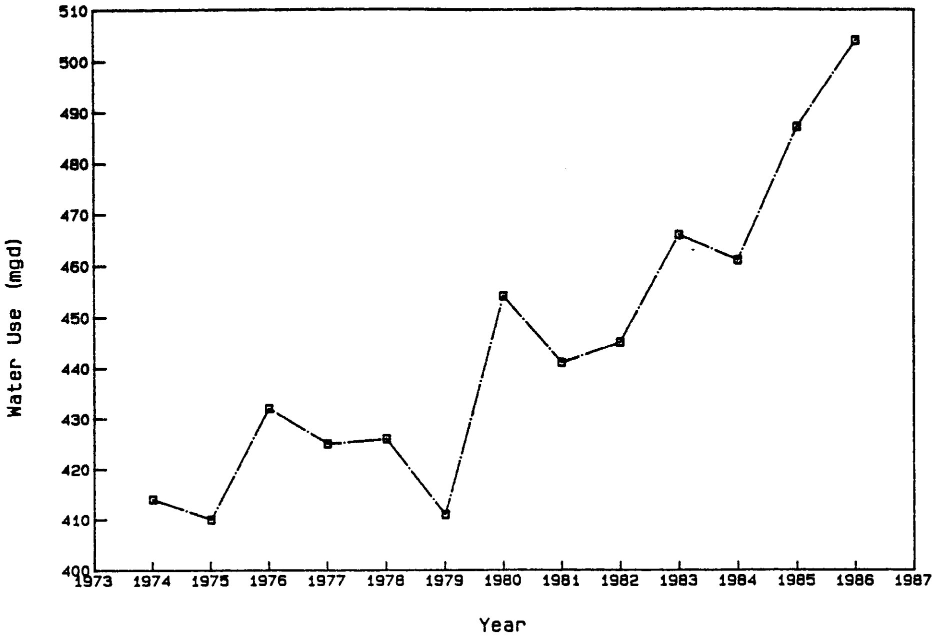

Fig. 1. Mean daily WMA summer water use, 1974–1986. The summer season extends from May through September.

al convention that uppercase symbols refer to random processes and random variables, e.g., Y_i, while lowercase symbols are used for deterministic functions and parameters, e.g., $m(t)$). The random variables $\{Y_i\}$ are assumed to represent structural attributes of the service area that vary slowly over time and interact in complex fashion. These attributes may include the price of water (and price history), total service area connections, plumbing code provisions, and customer income, among many others. The actual mean daily water use for year i will differ from Y_i due to the influence of random factors, such as climatological conditions, which operate on a rapidly fluctuating time scale relative to the factors which influence Y_i. The random variables $\{Y_i\}$ are not necessarily independent or indentically distributed. Indeed, trend and correlation are likely to be important features of the process. Distributional assumptions are not made on the random process $\{Y_i\}$.

The model is specified by

$$X_i(t) = m(t)Y_i + a[X_i(t-1) - m(t-1)Y_i] + Y_i^{1/2}A_i(t) \quad (1)$$

where

$$E[X_i(t)|Y_i] = m(t)Y_i \quad (2)$$

$$E[(X_i(t) - m(t)Y_i)^2|Y_i] = vY_i \quad (3)$$

a is a real-valued parameter, v is a nonnegative parameter, and

$$m(t) = [E[X_i(t)]]/T^{-1}\sum_{k=1}^{T} E[X_i(k)] \quad (4)$$

Equation (2) states that conditional on Y_i (that is, if we know Y_i), $m(t)Y_i$ is the mean of $X_i(t)$; equation (3) states that vY_i is the conditional variance of $X_i(t)$.

The error process $\{A_i(t)\}$ is assumed to be an independent and identically distributed sequence of Gaussian random variables, with mean zero and variance s^2. The error process is also assumed to be independent of $\{X_i(t)\}$ and $\{Y_i\}$.

It follows from (1) and (3) that the conditional variance parameter v must satisfy the equation

$$v = a^2v + s^2 \quad (5)$$

implying that for $|a| < 1$,

$$v = s^2/(1 - a^2) \quad (6)$$

Furthermore, it follows from (1) and the Gaussian assumption on $\{A_i(t)\}$ that conditional on Y_i, $X_i(t)$ has a Gaussian distribution with mean $m(t)Y_i$ and variance $(s^2/(1 - a^2))Y_i$; we will write

$$X_i(t) \overset{D}{\sim} N(m(t)Y_i, [s^2/(1 - a^2)]Y_i) \quad (7)$$

We include a lagged water use term in our model to reflect persistence in daily water use. Persistence may result from several causes, including limitations of the distribution system and meteorological conditions. To the extent that there is persistence in meteorological conditions that affect water use, this persistence is translated to water use. We do not directly incorporate meteorological variables in our model due to difficulty in forecasting these variables.

The unit demand function at time t, $m(t)$, is the ratio of mean water use on day t to mean daily water use over the course of the year. The unit demand function does not vary from year to year even if long-term trend in mean water use is present. This assumption implies that although mean water use may exhibit trends over time, seasonal and day-of-week structure of water use do not.

We denote the J weekly average values of the unit demand function by $q_1, \cdots, q_J$, that is,

$$q_j = 7^{-1}\sum_{t=1}^{7} m(7(j-1) + t) \qquad j = 1, \cdots, J \quad (8)$$

Structure of the unit demand function is determined by one final assumption. We assume that day-of-week coefficients $p_1, \cdots, p_7$ exist such that

$$m(t) = q_j p_k \qquad \text{for } 7(j-1) + k = t \quad (9)$$

Mean daily water use can vary by day of week; day-of-week effects cannot, however, vary seasonally or from year to year.

We denote the water use data set available on day t of year n by

$$H_n(t) = \{X_i(k);\ i = 1, \cdots, n-1;\ k = 1, \cdots, T;\ X_n(1), \cdots, X_n(t)\} \tag{10}$$

The data set $H_n(t)$ contains $n-1$ consecutive years of complete daily water use data and the first t days of data for year n. For a random variable X we use the notation $E[X|H_n(t)]$ as shorthand for the conditional expectation of X given $\{X_i(k);\ i = 1, \cdots, n-1, k = 1, \cdots, T;\ X_n(1), \cdots, X_n(t)\}$.

The model is described in the Introduction as a conditional autoregressive process (more specifically, an autoregressive process with random mean). It follows from (1) that

$$E[X_n(t+1)|H_n(t), Y_n] = E[X_n(t+1)|X_n(t), Y_n] = m(t+1)Y_n + a[X_n(t) - m(t)Y_n] \tag{11}$$

so that conditional on Y_n, the process $\{X_n\}$ has autoregressive structure. The model is not, however, an autoregressive process. It is not even a Markov Process; note that

$$E[X_n(t+1)|H_n(t)] \neq E[X_n(t+1)|X_n(t)] \tag{12}$$

Intuitively, the left side of (12) differs from the right because the data set $H_n(t)$ contains information about the random (and unknown) mean Y_n, as well as the correlation information contained in the previous day's observation $X_n(t)$. This point is further illustrated in deriving the forecast equations below.

The forecast that we will use on day $t+1$ for water use on day $t+k$ is the conditional expectation of water use on day $t+k$ given observations up to and including day t, $E[X_n(t+k)|H_n(t)]$. The conditional expectation is obviously a function of the data in $H_n(t)$. It is, perhaps most notably, that function of the data which minimizes the expected squared error with $X_n(t+k)$. The following result tells us how to construct our forecasts.

$$E[X_n(t+k)|H_n(t)] = a^k X_n(t) + [m(t+k) - a^k m(t)]E[Y_n|H_n(t)] \tag{13}$$

The result can be proven as follows. For $k = 1$,

$$\begin{aligned}
E[X_n(t+1)|H_n(t)] &= m(t+1)E[Y_n|H_n(t)] + a\{X_n(t) - m(t)E[Y_n|H_n(t)]\} \\
&\quad + E[Y_n^{1/2}A_n(t+1)|H_n(t)] \\
&= m(t+1)E[Y_n|H_n(t)] \\
&\quad + a\{X_n(t) - m(t)E[Y_n|H_n(t)]\} \\
&\quad + E[Y_n^{1/2}|H_n(t)]E[A_n(t+1)|H_n(t)] \\
&= m(t+1)E[Y_n|H_n(t)] + aX_n(t) - am(t)E[Y_n|H_n(t)] \\
&= aX_n(t) + [m(t+1) - am(t)]E[Y_n|H_n(t)]
\end{aligned} \tag{14}$$

Assume the result is true for $k-1$.

$$\begin{aligned}
E[X_n(t+k)|H_n(t)] &= m(t+k)E[Y_n|H_n(t)] \\
&\quad + a\{E[X_n(t+k-1)|H_n(t)] - m(t+k-1)E[Y_n|H_n(t)]\} \\
&\quad + E[Y_n^{1/2}A_n(t+k)|H_n(t)] \\
&= m(t+k)E[Y_n|H_n(t)] \\
&\quad + a\{a^{k-1}X_n(t) + m(t+k-1)E[Y_n|H_n(t)] \\
&\quad - a^{k-1}m(t)E[Y_n|H_n(t)] - m(t+k-1)E[Y_n|H_n(t)]\} \\
&= a^k X_n(t) + [m(t+k) - a^k m(t)]E[Y_n|H_n(t)]
\end{aligned} \tag{15}$$

The result follows by induction.

Note that

$$E[X_n(t+k)|H_n(t)] \approx m(t+k)E[Y_n|H_n(t)] \tag{16}$$

for "large" k. Equation (13), and especially (16), emphasize our interest in accurately modeling mean water use (see also section 4). It follows from (13) that for short forecast lead times (roughly, 1–3 days) the estimate of mean water use $(m(t+k)E[Y_n|H_n(t)])$ is an important component. It follows from (16) that for "long" lead times the forecast is virtually identical to the estimate of mean water use.

To conclude this section we note that in some situations it may be desirable to allow the parameters a, v, and s to depend on time (as the unit demand function $m(t)$ does). In this case, (6), which relates the parameters a, v, and s, is changed to the recursive equation

$$v(t) = a(t)^2 v(t-1) + s(t)^2 \tag{17}$$

The forecast equation (13) becomes

$$E[X_n(t+k)|H_n(t)] = \left[\prod_{j=1}^{k} a(t+j)\right] X_n(t) + \left[m(t+k) - \left(\prod_{j=1}^{k} a(t+j)\right) m(t)\right] E[Y_n|H_n(t)] \tag{18}$$

Extensions of the parameter and state estimation procedures developed in the following sections to the extended model are also straightforward but are not pursued.

It is also conceptually straightforward to extend the model to a "conditional autoregressive moving average" or form. The forecast equation (13), however, is not generalized in a straightforward fashion. Computational tractability is a major reason for restricting consideration to the conditional autoregressive model of (1).

3. PARAMETER AND STATE ESTIMATION

To implement a forecast system based on (13), we need to estimate the unknown parameters $m(t)$, a, and s and estimate the unknown random mean water use Y_n. The second problem is one of "state estimation," that is, the optimal prediction of an unobserved random variable. Nearly always the optimal predictor is (as in (13)) the conditional expectation of the unobservable random variable given the observations.

Because the random mean process is treated in a nonparametric framework, it is especially difficult to separate the problems of parameter and state estimation. Our approach to parameter estimation is to replace the random mean water use Y_i, for each year i, by the state estimator

$$\hat{Y}_i = T^{-1} \sum_{k=1}^{T} X_i(k) \tag{19}$$

and proceed as though no error were involved. In other words, to estimate the parameters $m(t)$, a, and s, we replace the random variables $Y_1, \cdots, Y_{n-1}$ by the sample means $\hat{Y}_1, \cdots, \hat{Y}_{n-1}$ and treat the random mean as known. We begin with the unit demand function $m(t)$.

From (9) it is clear that we need to estimate the weekly coefficients $q_1, \cdots, q_J$ and day-of-week coefficients $p_1, \cdots, p_7$.

Based on (4), (8), and (19), we choose our estimator of the weekly demand coefficients to be

$$\hat{q}_j = (n-1)^{-1} \sum_{i=1}^{n-1} \left[7^{-1} \sum_{k=1}^{7} X_i(7(j-1)+k)/\hat{Y}_i \right] \qquad j = 1, \cdots, J \tag{20}$$

The estimator $\hat{q}_j$ is the average value of "scaled" daily demand for week j, with daily values scaled by the average daily demand for the year.

Our estimators of the day-of-week coefficients are given by

$$\hat{p}_k = (n-1)^{-1} \sum_{i=1}^{n-1} \left[J^{-1} \sum_{j=1}^{J} X_i(7(j-1)+k)/(\hat{q}_j \hat{Y}_i) \right] \qquad k = 1, \cdots, 7 \tag{21}$$

The estimator of the Monday day-of-week coefficient, for example, is the average of all Monday values; each value must be scaled by the product of the estimated weekly coefficient $\hat{q}_k$ and yearly sample mean $\hat{Y}_i$. It is straightforward at this point to construct our estimator for the unit demand function. From (9) we have

$$\hat{m}(t) = \hat{q}_j \hat{p}_k \qquad \text{for } 7(j-1)+k = t \tag{22}$$

Centered water use values $\tilde{X}_i(t)$ are given by

$$\tilde{X}_i(t) = X_i(t) - \hat{m}(t)\hat{Y}_i \tag{23}$$

Our estimator for the autoregressive parameter a is given by

$$\hat{a} = \sum_{i=1}^{n-1} \sum_{t=2}^{T} \tilde{X}_i(t)\tilde{X}_i(t-1) \bigg/ \sum_{i=1}^{n-1} \sum_{t=2}^{T} \tilde{X}_i(t)^2 \tag{24}$$

The estimator $\hat{a}$ is a standard least squares estimator except that $\hat{m}(t)\hat{Y}_i$ is a state estimator for the unknown random mean.

The estimator for the standard error coefficient s is given by

$$\hat{s}^2 = (n-1)^{-1} \sum_{i=1}^{n-1} (T-1)^{-1} \sum_{t=2}^{T} [[\tilde{X}_i(t) - \hat{a}\tilde{X}_i(t-1)]/(\hat{Y}_i)]^2 \tag{25}$$

We now turn our attention to the state estimation problem of computing $E[Y_n|H_n(t)]$. Because $\{Y_i\}$ is treated in a nonparametric framework, we will not be able to explicitly compute the conditional expectation of Y_n given $H_n(t)$. As is often the case in state estimation problems (see, for example, *Karr* [1986]), we will retreat to estimators that are linear combinations of our forecast data (or nearly so). The estimator of $E[Y_n|H_n(t)]$ will be denoted by $M_n(t)$.

In constructing the estimator $M_n(t)$ we wish to exploit structure of the data set $H_n(t)$, which divides naturally into two components: data from previous years $H_{n-1}(T)$ and observations from the current year $X_n(1), \cdots, X_n(t)$. To forecast mean water use it is natural to condense information from previous years to the sample means $\hat{Y}_1, \cdots, \hat{Y}_{n-1}$. Similarly, observations for the current year are condensed to the "partial sample means"

$$\hat{Y}_n(t) = \sum_{k=1}^{t} X_n(k) \bigg/ \sum_{k=1}^{t} m(k) \tag{26}$$

Nonparametric trend techniques developed by *Hirsch et al.* [1982] are used to forecast Y_n from previous years' sample means, $\hat{Y}_1, \cdots, \hat{Y}_{n-1}$. The forecast is of the form $\hat{Y}_{n-1} + \hat{b}_n$, where $\hat{b}_n$ is the Kendall slope estimator obtained from $\hat{Y}_1, \cdots, \hat{Y}_{n-1}$. The Kendall slope estimator is the median value of the slope random variables $\{S_{ij}, i < j < n\}$, where

$$S_{ij} = (\hat{Y}_j - \hat{Y}_i)/(j-i) \qquad \text{for } i < j < n \tag{27}$$

Using (1) and (26), it is straightforward to show that

$$\hat{Y}_n(t) = Y_n + \tilde{a}(t)\left[\sum_{k=1}^{t} m(k)\right]^{-1} (X_n(0) - m(0)Y_n) + Y_n^{1/2}\left[\sum_{k=1}^{t} m(k)\right]^{-1} \sum_{k=1}^{t}\sum_{j=1}^{k} a^{k-j}A_n(j) \tag{28}$$

where

$$\tilde{a}(t) = \sum_{k=1}^{t} a^k = (1-a^{t+1})(1-a)^{-1} - 1 \tag{29}$$

It follows from (28) that conditional on Y_n,

$$\hat{Y}_n(t) \overset{D}{\sim} N(Y_n, c(t)Y_n) \tag{30}$$

where $c(t)$ is a function of the parameters a, $m(t)$, and s, which, most notably, is decreasing in t. It follows from (30) that for each t, $\hat{Y}_n(t)$ is an unbiased state estimator of Y_n and that for t greater than s, $\hat{Y}_n(t)$ is a better estimator than $\hat{Y}_n(s)$.

Based on the preceding discussion, we take our state estimator to be

$$M_n(t) = w(t)\hat{Y}_n(t) + (1-w(t))(\hat{Y}_{n-1} + \hat{b}_n) \tag{31}$$

where

$$w(t) = 1 - [(T+1-t)/T]^2 \tag{32}$$

The weight function is chosen to have the following properties

$$0 \le w(t) \le 1 \qquad t = 1, \cdots, T+1 \tag{33}$$

$$w(1) = 0 \tag{34}$$

$$w(T+1) = 1 \tag{35}$$

$$w(T/2+1) > 0.50 \tag{36}$$

On the first day of year n we have only observations from preceding years, so all of the weight must be on preceding years, that is, $w(1)$ must equal zero. At the other extreme we have all of the data from year n available. The assumption that $w(T+1)$ equals 1 implies that year-to-year dependence in the random mean process is weak. Specifically, the assumption implies that for estimating Y_n from $H_n(T)$, previous years' data provide no further information once $\hat{Y}_n$ is available. Equation (36) implies that much of the information about Y_n is available at the midpoint of the year. Relatively less "new" information about Y_n should be expected as the year progresses. The second half of the year, for example, contains less new information than the first (provided, of course, that the first half is observed).

4. Application of the Water Use Model

The water use forecast system is used to schedule releases from two water supply reservoirs located in the Potomac River basin upstream of Washington, D. C. (see *Palmer et al.* [1982] for a detailed discussion of water supply management for the WMA). A large reservoir, located far from the WMA, can provide water to the WMA with a travel time of approximately 5 days. Releases from a small local reservoir reach

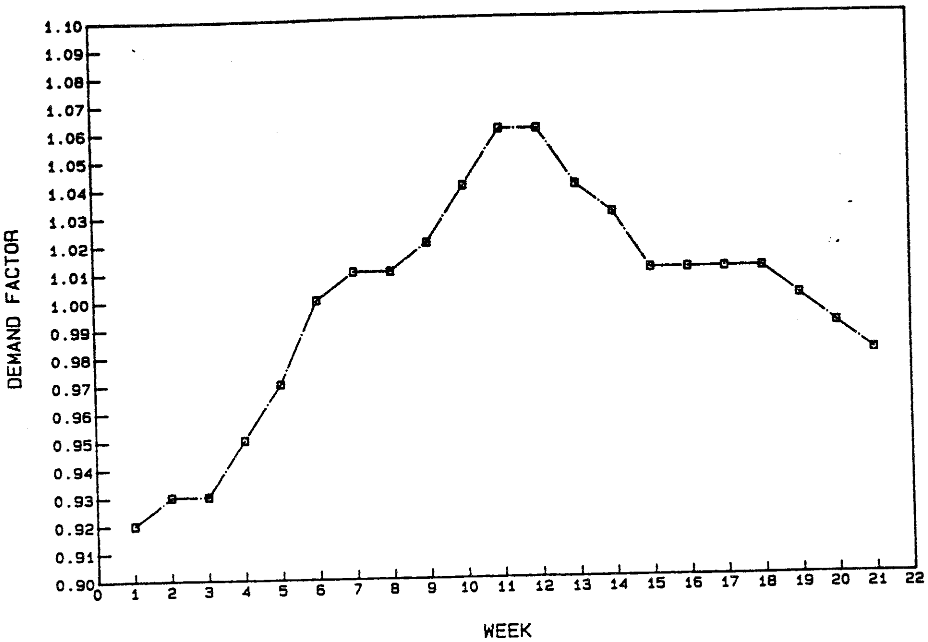

Fig. 2. Weekly demand factors for WMA water use for the 21 weeks beginning in May and ending in September.

water utility intakes within the day of release. The water use model is developed for the summer season (May–September), during which forecast information is needed for scheduling water supply releases.

For the WMA water supply system the large upstream reservoir is operated to meet "average" water demand; the local reservoir meets shortfalls arising from "extreme" demands. Average demand is clearly a moving target. The estimated Kendall slope estimator $\hat{b}_n$ obtained from 1974 to 1986 WMA water use data is 8 mgd/yr (mgd stands for million gallons per day; 1 mgd = 5680 m^3 per day). The estimated weekly coefficients (see Figure 2) obtained from (21) range from a minimum of 0.92 at the beginning of May to a maximum of 1.06 in mid-July, from which they decrease below 1 by the end of September. For an annual mean water use of 500 mgd (the 1986 value for the WMA) the seasonal variation in mean water use is 70 mgd (from 460 mgd in early May to 530 mgd in mid-July). The estimated day-of-week coefficients $\hat{p}_1, \cdots, \hat{p}_7$ (equation (22)) range from a maximum of 1.02 on Wednesday to a minimum of 0.97 on Sunday (the estimated day-of-week coefficients, beginning with Sunday, are 0.97, 1.00, 1.01, 1.02, 1.01, 1.01, and 0.99). Differences in day-of-week water use result in part from the fact that there is a larger population in the service area during the week than on weekends (a significant number of people work in the region served by WMA water utilities but live outside of the area served by WMA water utilities). For a weekly mean water use of 500 mgd the range in mean water use associated with day of week is 25 mgd.

For scheduling an upstream release on day t, average water use (for day $t + 5$) is estimated by $\hat{m}(t + 5)M_n(t)$. It follows from (31) that the forecast is of the form

$$\hat{m}(t + 5)M_n(t) = \hat{m}(t + 5)[w(t)\hat{Y}_n(t) + (1 - w(t))(\hat{Y}_{n-1} + \hat{b}_n)] \tag{37}$$

To operate the small local reservoir, forecasts of 1-day-ahead water use are required. To apply the forecast equation (13), we need only specify the autoregressive parameter a. The estimate obtained from (25) is 0.76. The forecast obtained from (13) for 1-day-ahead water use is

$$\begin{aligned}\hat{X}(t + 1) &= \hat{a}X_n(t) + [\hat{m}(t + 1) - \hat{a}\hat{m}(t)]M_n(t)\\ &= \hat{a}X_n(t) + [\hat{m}(t + 1) - \hat{a}\hat{m}(t)]\\ &\quad \cdot [w(t)\hat{Y}_n(t) + (1 - w(t))(\hat{Y}_{n-1} + \hat{b}_n)]\end{aligned} \tag{38}$$

Figure 3 shows errors of 1-day-ahead forecasts for WMA water use during the summer of 1986. Note that the errors are weakly correlated and that variability of the estimators decreases as the year proceeds. The percent bias for 1986 1-day-ahead forecasts is -0.1%; the standard error is 29 mgd.

5. SUMMARY AND CONCLUSIONS

A time series model of daily municipal water use is developed. Emphasis in model development is placed on long-term trend, seasonality, and day-of-week effects. The model, which is termed a conditional autoregressive process, can be interpreted as an autoregressive process with randomly varying mean. The random mean process, which represents long-term trend in mean water use, is treated in a nonparametric framework. Conditional on the random mean water use, the model reduces to a Gaussian autoregressive process with a modest number of parameters.

Seasonality and day-of-week effects are captured in the model through the unit demand function. The unit demand function at time t, $m(t)$, is the ratio of mean daily water use on day t of the year to mean daily water use over the course of the year. An important model assumption is that the unit demand function does not vary from year to year, even if long-term trend in mean water use is present.

An attractive feature of the water use model is computational tractability. The forecast equation derived in section 2 (equation (13)) can be easily implemented provided that two estimation problems are solved. To implement the forecast

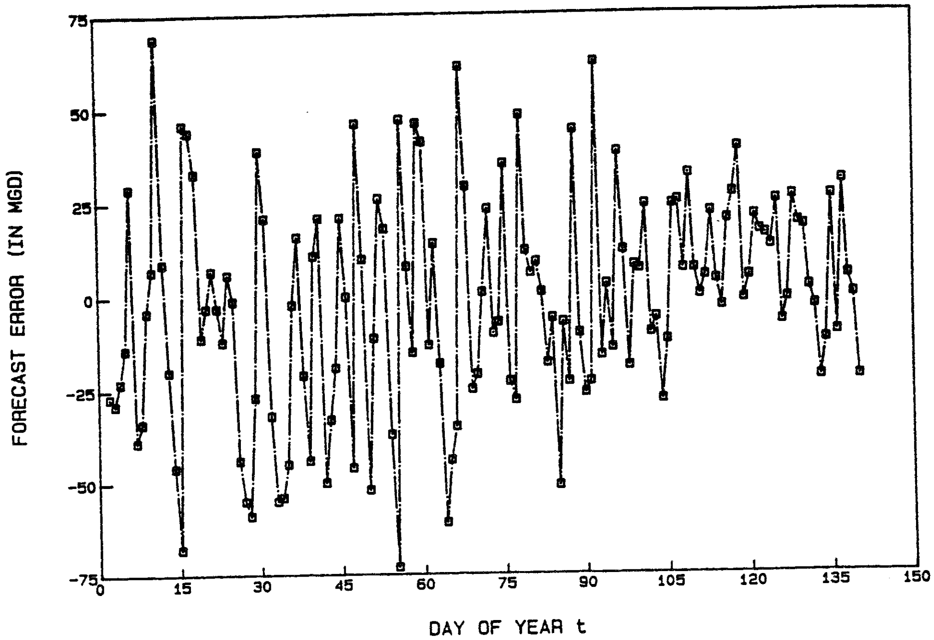

Fig. 3. One-day-ahead WMA water use forecast errors for the summer of 1986.

equation, parameter estimators for the model parameters $m(t)$ and a are needed. Also, state estimators for the random (and unknown) mean Y_n are required. A notable feature of the parameter estimation procedure developed in section 3 is inclusion of a state estimation step. Parameter estimates for $m(t)$ and a are obtained after first applying a "detrending" procedure which is based on simple state estimators for the random mean process. The state estimator that is developed for use in the forecast equation is nonparametric and exploits special structure of the water use data set.

Parameter and state estimation procedures are applied to summer season water use in the Washington, D. C., Metropolitan Area. The estimation results confirm that long-term trend, seasonality, and day-of-week effects are prominent features of WMA water use. The water use forecast system developed for the WMA is used to schedule releases from two water supply reservoirs. A large upstream reservoir is operated to meet "average" water demands. For operation of this reservoir the state estimator of current mean water use (equation (37)) provides the necessary forecast information. One-day-ahead forecasts, obtained from (13), are used to operate a small local reservoir, which covers shortfalls due to "extreme" water demands.

A potentially useful extension to the forecast system involves incorporation of precipitation data. This could be accomplished by expanding the model equation (1) to explicitly include precipitation variables. An alternative is to incorporate precipitation variables into the random mean process $\{Y_n\}$. With the second approach, precipitation forecasts are not required for implementing a forecast system. Instead, observed precipitation data are used, along with observed water use data, to update the estimate of current mean water use.

Acknowledgments. This research was supported in part by grant 14-08-0001-G-1145 of the Water Resources Research Program of the U.S. Geological Survey. The author gratefully acknowledges comments of J. Boland, R. Steiner, and J. Whitcomb.

REFERENCES

Carver, P., and J. Boland, Short and long-run effects of price on municipal water use, *Water Resour. Res., 16*(4), 609–616, 1980.

Davis, R., and S. Hanke, Conventional and unconventional alternatives for water supply management, *Water Resour. Res., 9*(4), 861–870, 1973.

Hirsch, R., J. Slack, and R. Smith, Techniques of trend analysis for monthly water quality data, *Water Resour. Res., 18*(1), 107–121, 1982.

Howe, C., and F. Linaweaver, The impact of price on residential water demand and its relation to system design and price structure, *Water Resour. Res., 3*(1), 13–22, 1967.

Karr, A., *Point Processes and their Statistical Inference*, Marcel Dekker, New York, 1986.

Kher, L., and S. Sorooshian, Identification of water demand models from noisy data, *Water Resour. Res., 22*(3), 322–330, 1986.

Klemes, V., The Hurst phenomenon: A puzzle?, *Water Resour. Res., 10*(4), 675–688, 1974.

Maidment, D., and E. Parzen, Cascade model of monthly municipal water use, *Water Resour. Res., 20*(1), 15–23, 1984.

Nicholls, D. F., and B. G. Quinn, *Random Coefficient Autoregressive Models: An Introduction, Lecture Notes in Statistics*, vol. 11, Springer-Verlag, New York, 1982.

Palmer, R., J. Smith, J. Cohon, and C. ReVelle, Reservoir management in the Potomac River basin, *ASCE J. Water Resour. Planning Manage. Div., WR1*, 47–66, 1982.

Potter, K. Evidence for nonstationarity as a physical explanation of the Hurst phenomenon, *Water Resour. Res., 12*(5), 1047–1052, 1976.

Salas-LaCruz, J., and V. Yevjevich, Stochastic structure of water use time series, *Hydrology Paper 52*, Colo. State Univ., Fort Collins, 1972.

Smith, J. A., and A. F. Karr, A point process model of summer season rainfall occurrences, *Water Resour. Res., 19*(1), 95–103, 1983.

Smith, J. A., and A. F. Karr, Statistical inference for point process models of rainfall, *Water Resour. Res., 21*(1), 73–79, 1985.

Swamy, P. A. V. B., *Statistical Inference in Random Coefficient Regression Models, Lecture Notes in Operations Research and Mathematical Systems*, vol. 55, Springer-Verlag, New York, 1982.

J. A. Smith, Interstate Commission on the Potomac River Basin, 6110 Executive Boulevard, Suite 300, Rockville, MD 20852.

(Received April 24, 1987;
revised October 27, 1987;
accepted November 4, 1987.)