Peak Daily Water Demand Forecast Modeling Using Artificial Neural Networks

Jan Franklin Adamowski[1]

Abstract: Peak daily water demand forecasts are required for the cost-effective and sustainable management and expansion of urban water supply infrastructure. This paper compares multiple linear regression, time series analysis, and artificial neural networks (ANNs) as techniques for peak daily summer water demand forecast modeling. Analysis was performed on 10 years of peak daily water demand data and meteorological variables (maximum daily temperature and daily rainfall) for the summer months of May to August of each year for an area of high outdoor water usage in the city of Ottawa, Canada. Thirty-nine multiple linear regression models, nine time series models, and 39 ANN models were developed and their relative performance was compared. The artificial neural network approach is shown to provide a better prediction of peak daily summer water demand than multiple linear regression and time series analysis. The best results were obtained when peak water demand from the previous day, maximum temperature from the current and previous day, and the occurrence/nonoccurrence of rainfall from five days before, were used as input data. It was also found that the peak daily summer water demand is better correlated with the rainfall occurrence rather than the amount of rainfall itself, and that assigning a weighting system to the antecedent days of no rainfall does not result in more accurate models.

DOI: 10.1061/(ASCE)0733-9496(2008)134:2(119)

CE Database subject headings: Artificial intelligence; Forecasting; Municipal water; Regression analysis; Time series analysis; Water demand; Water resources.

Introduction

Water supply systems around the world have become stressed in recent years due to rapid population growth and increased per capita water consumption. In Ottawa, Canada, it has been predicted that both the residential and employment population in the West Center region of the city will increase substantially in the next 25 years due to the rapid development of this suburban area. In addition, recent trends in the West Center region pressure zone indicate that both average and peak water demand have been increasing for the summer period between May and August. Studies have shown that a major fraction of the water consumption in the summer in the Ottawa West Center (OWC) pressure zone in Ottawa can be attributed to outdoor water use, which essentially consists of the watering of lawns and gardens. The water demand process in such situations is usually mainly driven by the maximum air temperature with the rainfall occurrences interrupting the process to cause transient drops in the water use. The water consumption can be expected to be high on consecutive dry days with high temperatures and low on rainy days. However, the water demand may not depend on the amount of rainfall since it may actually be a function of the occurrence of rainfall instead. This can be attributed to the fact that people may not want to water their lawns or gardens on a rainy day regardless of the amount of rainfall.

Fig. 1 shows the diurnal water demand pattern for the day of the highest peak demand (MXDY) for the entire OWC zone record period for the summers between 1993 and 2002. It can be seen that the demand rises significantly after 4 p.m., reaches a peak just before 9 p.m., and then gradually decreases. For comparative purposes, the diurnal pattern of a typical rainy summer day is also shown in Fig. 1. It can be seen that on such a day, when there is likely little or no outdoor water use, there is no significant rise in demand between the hours of 4 and 11 p.m. This illustrates the significance of outdoor water use on peak water demand in the summer. The high rates of outdoor water demand in the summer can be attributed to the fact that the OWC zone consists largely of neighborhoods with many homes with large landscaped areas that need to be irrigated.

In 2002, a typical winter day (September to April) water demand in the OWC zone was 21.3 ML/day, a typical low summer day (rainy day) water demand was 24.8 ML/day, and an average summer (May to August) day water demand was 32.4 ML/day (ROMC 2003). The maximum peak daily summer water demand has increased from 67.8 ML/day in 1993 to 109.3 ML/day in 2002. This illustrates the great variability in water demand in the OWC zone.

As a result of the projected population and employment growth in the OWC zone, coupled with increasingly high peak water use in the summer due to outdoor water demand, it has been determined that the existing water supply infrastructure will most likely not be able to meet future water demands (Bougadis et al.

[1]Postdoctoral Researcher, Cyprus Institute Program for Energy, Environment, and Water Resources, Massachusetts Institute of Technology, 77 Massachusetts Avenue, E40-469, Cambridge, MA 02139; formerly, Institute of Environmental Engineering Systems, Warsaw Univ. of Technology, ul. Nowowiejska 20, Warsaw, Poland 00-653. E-mail: jfadamowskica@yahoo.ca

Note. Discussion open until August 1, 2008. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on January 25, 2006; approved on May 21, 2007. This paper is part of the ***Journal of Water Resources Planning and Management***, Vol. 134, No. 2, March 1, 2008. ©ASCE, ISSN 0733-9496/2008/2-119–128/$25.00.

J. Water Resour. Plann. Manage. 2008.134:119-128.

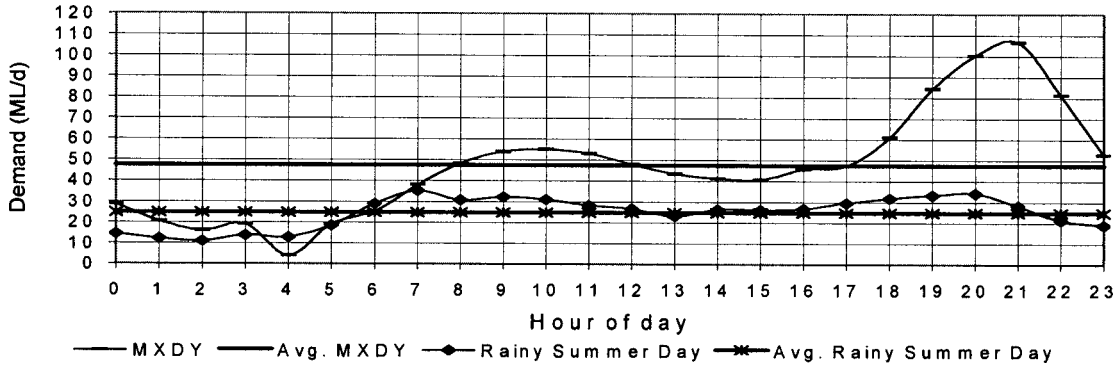

Fig. 1. Typical diurnal summer water usage in Ottawa West Center zone (MXDY=day of highest peak demand for entire 1994–2002 record; average MXDY=average peak demand for entire record)

2005). In order to address this issue, it will be necessary to pursue a course of action that might involve a combination of the following: optimizing the water supply system through real-time control by a hybrid expert system, imposing effective water use restrictions, and developing a sustainable and least-cost infrastructure expansion strategy. A critical aspect of each of these initiatives is the accurate forecasting of short-term water demands, and in particular, peak daily water demands.

Despite the relative importance of peak daily water demand, limited detailed research has been devoted to this topic, including factors driving peak daily water demand and forecasting methods (Day and Howe 2003). The motivation for this research was, therefore, to study three important issues that have not been explored in the short-term water demand literature: (1) to investigate the use of artificial neural networks for forecasting peak (as opposed to average or total) daily water demand in the summer months in an area of high outdoor water usage; (2) to determine whether rainfall occurrence or rainfall amount is a more significant variable in modeling peak summer water demand forecasts; and (3) to test the hypothesis of whether assigning a weight system to the antecedent days of no rainfall would result in more accurate models.

It has been shown that the peak summer water demand process is stochastic and nonlinear because outdoor water use, the major component of peak summer water demand, depends on the duration and intensity of rainfall and the characteristics of temperature (Gutzler and Nims 2005). As such, the forecasting of peak summer water demand is complex, and thus the use of artificial neural networks (ANNs), which are capable of modeling nonlinear systems, needs to be explored. The issue of whether rainfall occurrence or rainfall amount is a more significant variable in modeling short-term water demand has been investigated by Jain et al. (2001) and Bougadis et al. (2005) but they arrived at opposite conclusions. Therefore, this issue was further investigated in this study for peak daily water demand.

In this research, three methods for peak daily water demand were developed and compared. Multiple linear regression was used in this study because it is one of the most widely used techniques for water demand forecasting and as such is ideal for comparative purposes with the newer ANN technique. Multiple linear regression was also used because time series analysis, the other most widely used technique for water demand forecasting, involves only present and past water demand, whereas with multiple linear regression one can include additional parameters such as climatic ones. Time series analysis was also explored as a potential technique because regression models do not illuminate the inherent autocorrelation structure of a water use pattern over time. A time series analysis is capable of revealing the autocorrelation structure of a short-term water demand pattern over time. ANN analysis was explored as a potential technique because, as mentioned earlier, the use of ANNs to forecast peak daily water demand has not been explored in great detail in the literature and because of the high potential of the ANN approach due to its ability to handle nonlinear relationships.

Previous Research

A variety of techniques have been used in short-term water demand forecasting. Examples of short-term water demand forecast modeling using regression analysis include: Howe and Linaweaver (1967), Oh and Yamauchi (1974), Hughes (1980), Anderson et al. (1980), and Maidment and Parzen (1984). Maidment et al. (1985) used short-term time series models for daily municipal water use as a function of rainfall and air temperature. Maidment and Miaou (1986) applied this model to the water consumption from nine cities in the United States. Some other examples of short-term water demand forecast modeling using time series analysis include Smith (1988), Miaou (1990), and Zhou et al. (2000).

Artificial neural networks have recently begun to be used for short-term water demand forecasting. Jain et al. (2001) developed ANN models by using weekly maximum air temperature, weekly rainfall amount, weekly past water demand, and the occurrence or nonoccurrence of rainfall as parameters for their models and compared them to regression and time series models. It was found that the occurrence of rainfall was a more significant variable than the amount of rainfall itself in the modeling of short-term water demand, and that the ANN models outperformed both the regression and time series models. One of their main recommendations for future study was to conduct a similar study, but using a longer and more continuous set of data since their data was scattered with a lot of missing records. This study uses a long record of continuous data with no missing records.

Jain and Ormsbee (2002) examined regression, time series analysis, and ANN models for daily water demand forecasting. Only one simple ANN model was investigated in this study, which was a function of the daily water demand from the previous day and the daily maximum air temperature of the current day. It was found that this ANN model was slightly better than the time series and the regression–disaggregation models.

J. Water Resour. Plann. Manage. 2008.134:119-128.

Pulido-Calvo et al. (2003) examined regression, time series, and ANN models for total daily water demand for Fuente Palmera, Spain. It was determined that the best model (an ANN model) was a function of the water demand and maximum temperature from the two previous days.

Bougadis et al. (2005) explored regression, time series, and ANN models for weekly water demand. They determined that the ANN models consistently outperformed the regression and time series models, and that the best results were obtained when employing previous weekly demand along with the current week's rainfall and temperature. They found, in contrast to Jain et al. (2001), that the weekly water demand is better correlated with the amount of rainfall rather than the rainfall occurrence. As such, this issue is explored again in this study.

Ottawa West Center Zone Water Supply System

The City of Ottawa is responsible for providing potable water services to 750,000 customers. It operates two purification plants, 20 pumping stations, 14 reservoirs, and four communal well systems. There are seven water supply zones in the city. The OWC pressure zone investigated in this study is the third largest water supply zone in Ottawa and serves a residential population of approximately 63,000 and an employment population of approximately 17,500. The land use of the OWC zone is 53% residential, 5% commercial/institutional, 21% park/recreational, and 21% vacant (ROMC 2003). The OWC zone receives all of its water from the Ottawa River via a purification plant. No groundwater is used.

Data

Many variables influence water demand, most of which can be grouped into two classes: socioeconomic and climatic variables. Studies have demonstrated that socioeconomic variables are responsible for the long-term effects on water demand, while climatic variables are mainly responsible for short-term seasonal variations in water demand (Miaou 1990).

This study used climatic variables, past water demand and population. More specifically, the data used in this study consisted of daily total rainfall (mm), maximum daily temperature (°C), peak daily water demand (ML/day), and population. The peak daily water demand for a specific day was the peak hour water demand for that specific day.

The peak daily water demand data was obtained from the City of Ottawa for the OWC pressure zone in Ottawa, Canada. Daily maximum temperature (°C) and daily total rainfall (mm) were obtained from Environment Canada for the Ottawa region. The water demand series record was available from 1994 to 2002. Only the summer months (May to August) were used in the analysis since peak demand usage occurs then for a given year. There were no special events (such as a huge pipe break or tournaments) that could have invalidated the data.

All the data series were divided into a training/calibration set and a testing set (split sample study). The training set began in 1994 and ended in 2001. The performance of all statistical models was analyzed by comparing the known peak water demand values in 2002 with the forecasted peak water demand values for 2002 obtained from the different models.

Model Performance Tests

The performance of developed models can be evaluated using several statistical tests that describe the errors associated with the model. After each of the model structures is calibrated using the calibration/testing data set, the performance can then be evaluated in terms of these statistical measures of goodness of fit. In order to provide an indication of goodness of fit between the observed and forecasted values, the average absolute relative error (AARE), the maximum absolute relative error (Max ARE), and the coefficient of determination (R^2) can be used.

The AARE is a quantitative measure of the average error in one step ahead forecasts from a particular model and is defined by (Jain et al. 2001)

$$\mathrm{AARE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{O_i - D_i}{O_i}\right| \times 100\% \tag{1}$$

where O_i=observed peak water demand; and D_i=forecasted peak water demand found from regression, time series, and ANN models, respectively. The smaller the value of AARE, the better is the performance of the model.

The Max ARE is the maximum of the absolute relative error among all of the forecasted data points and is a measure of the robustness of the model. The smaller the value of the Max ARE, the better is the performance of the model. The coefficient of determination (R^2) is a measure of the strength of the model in developing a relationship among input and output variables. It measures the degree of correlation among the observed and forecasted values.

Model Development

Multiple Linear Regression Analysis

Thirty-nine multiple linear regression (MLR) models were developed for daily peak water demand forecasts and can be seen in Table 1. Models MLR-1–MLR-20 were designed to assess the significance of various input variables on peak daily water demand and to corroborate the cross correlation and simple linear regression analyses. Models MLR-20–MLR-39 were developed to test whether the occurrence or nonoccurrence of rainfall (using both binary and weighted systems) was a more significant variable than the amount of rainfall.

The first group of models was MLR-1–MLR-6, which was developed to explore the importance of temperature from the previous day (T_{t-1}) on peak daily water demand. An example of one of these models is MLR-3, which is a function of daily peak demand from the previous day, the maximum temperature of the current day, and the total rainfall of the current day, and is shown by

$$D_t = B_0 + B_1 D_{t-1} + B_3 T_t + B_5 R_t \tag{2}$$

The second group of models was MLR-7–MLR-10, which was developed to explore the importance of peak daily water demand from two days ago (D_{t-2}) on peak daily water demand. The third group of models was MLR-11–MLR-20, which was developed to test the effect of population on models MLR-1–MLR-10. The fourth group of models was MLR-21–MLR-26, which was developed to test the effect of rainfall amount from the current day to five days before.

J. Water Resour. Plann. Manage. 2008.134:119-128.

Table 1. Performance Statistics for All MLR Models

Model	Parameters	AARE	Max ARE	R^2 (training)	R^2 (testing)
MLR-1	D_{t-1}, T_t	16	81	0.56	0.51
MLR-2	D_{t-1}, T_t, T_{t-1}	16	80	0.58	0.54
MLR-3	D_{t-1}, T_t, R_t	16	76	0.58	0.53
MLR-4	D_{t-1}, T_t, T_{t-1}, R_t	15	81	0.59	0.55
MLR-5	D_{t-1}, T_t, CR_t	15	70	0.58	0.54
MLR-6	D_{t-1}, T_t, T_{t-1}, CR_t	14	64	0.59	0.57
MLR-7	D_{t-1}, D_{t-2}, T_t	15	80	0.57	0.52
MLR-8	D_{t-1}, D_{t-2}, T_t, T_{t-1}	14	70	0.59	0.55
MLR-9	D_{t-1}, D_{t-2}, T_t, T_{t-1}, R_t	15	67	0.60	0.56
MLR-10	D_{t-1}, D_{t-2}, T_t,T_{t-1}, CR_t	1	60	0.61	0.58
MLR-11	MLR-1+population	15	86	0.57	0.51
MLR-12	MLR-2+population	16	81	0.59	0.55
MLR-13	MLR-3+population	15	82	0.58	0.53
MLR-14	MLR-4+population	16	77	0.60	0.55
MLR-15	MLR-5+population	15	74	0.60	0.55
MLR-16	MLR-6+population	16	71	0.61	0.57
MLR-17	MLR-7+population	16	88	0.57	0.52
MLR-18	MLR-8+population	14	70	0.60	0.55
MLR-19	MLR-9+population	14	67	0.61	0.56
MLR-20	MLR-10+population	14	60	0.62	0.58
MLR-21	D_{t-1}, T_t, T_{t-1}, R_{t-1}	14	71	0.59	0.56
MLR-22	D_{t-1},T_t, T_{t-1}, R_{t-2}	15	72	0.58	0.54
MLR-23	D_{t-1}, T_t, T_{t-1}, R_{t-3}	15	72	0.58	0.54
MLR-24	D_{t-1}, T_t, T_{t-1}, R_{t-4}	15	72	0.58	0.54
MLR-25	D_{t-1}, T_t, T_{t-1}, R_{t-5}	15	72	0.58	0.54
MLR-26	MLR-21+population	15	79	0.59	0.56
MLR-27	D_{t-1}, T_t, T_{t-1},CR_{t-1}	15	62	0.61	0.59
MLR-28	D_{t-1}, T_t, T_{t-1}, CR_{t-2}	15	61	0.61	0.58
MLR-29	D_{t-1}, T_t, T_{t-1}, CR_{t-3}	14	59	0.61	0.59
MLR-30	D_{t-1}, T_t, T_{t-1}, CR_{t-4}	14	61	0.60	0.58
MLR-31	D_{t-1}, T_t, T_{t-1}, CR_{t-5}	**14**	**56**	0.61	**0.59**
MLR-32	MLR-31+population	14	61	0.62	0.59
MLR-33	D_{t-1},T_t, T_{t-1}, DR_t	14	60	**0.61**	0.59
MLR-34	D_{t-1}, T_t, T_{t-1}, DR_{t-1}	15	74	0.59	0.54
MLR-35	D_{t-1}, T_t, T_{t-1}, DR_{t-2}	15	73	0.58	0.54
MLR-36	D_{t-1}, T_t, T_{t-1}, DR_{t-3}	15	73	0.58	0.54
MLR-37	D_{t-1}, T_t, T_{t-1}, DR_{t-4}	15	73	0.58	0.55
MLR-38	D_{t-1}, T_t, T_{t-1}, DR_{t-5}	15	73	0.58	0.54
MLR-39	MLR-33+population	17	67	0.61	0.59
Average		15	71	0.59	0.56

Note: MLR=multiple linear regression; AARE=average absolute relative error; R^2=correlation coefficient; D_t=demand at time t; T_t=temperature at time t; R_t=rainfall at time t; CR_t=occurrence or nonoccurrence of rainfall; and DR_t=weighted occurrence or nonoccurrence of rainfall.

The fifth group of models was MLR-27–MLR-32, which was developed to test whether the occurrence of rainfall is a more significant variable than the rainfall amount itself in modeling the peak daily water demand forecasts. In these models, the actual rainfall amount was replaced by the occurrence or nonoccurrence of rainfall, which is denoted by the *CR* coefficient in the multiple linear regression equations. If the sum of rainfall is greater than 2.5 mm for the given time period, then the *CR* coefficient is equal to 1; if the sum of rainfall is less than 2.5 mm for the given time period, then the *CR* coefficient is equal to zero, respectively. An example of this type of model is MLR-27, which is a function of the peak daily demand from the previous day, the maximum temperature of the current day, the maximum temperature of the previous day, and the occurrence/nonoccurrence of rainfall value of the previous day. It is shown by

$$D_t = B_0 + B_1 D_{t-1} + B_3 T_t + B_4 T_{t-1} + B_{12} CR_{t-1} \tag{3}$$

The sixth group of models was MLR-33–MLR-39, which was developed in order to test whether assigning a weight system to the antecedent days of no rainfall (i.e., no rainfall above 2.5 mm) would result in better models. For example, if the sum of the rainfall in the preceding 10 days was less than 2.5 mm, then a value of 10 would be assigned to the *DR* coefficient shown in the

J. Water Resour. Plann. Manage. 2008.134:119-128.

Table 2. Performance Statistics for All Time Series Models

Model	AIC	AARE	Max ARE	R^2 (training)	R^2 (testing)
ARIMA (1,1,1)	8459	15	94	0.53	0.44
ARIMA (2,1,1)	8446	15	89	0.54	0.45
ARIMA (2,1,2,)	**8443**	15	94	**0.54**	0.45
ARIMA (1,1,0)	8567	15	106	0.53	0.45
ARIMA (2,1,0)	8537	**15**	**62**	0.53	**0.46**
ARIMA (3,1,0)	8515	15	119	0.53	0.45
ARIMA (0,1,1)	8542	19	127	0.46	0.36
ARIMA (0,1,2)	8501	20	129	0.47	0.36
ARIMA (0,1,3)	8483	19	123	0.47	0.37
Average		17	105	0.55	0.40

Note: ARIMA=autoregressive integrated moving average; AIC=Akaike criterion; AARE=average absolute relative error; and R^2=correlation coefficient.

Eq. (4). However, if there was a day with rainfall greater than 2.5 mm, then a value of 0 would be assigned to the *DR* coefficient and the process would restart. An example of this type of model is MLR-33, which is a function of the peak demand from the previous day, the temperature of the current day, the temperature of the previous day, and the weighted occurrence or nonoccurrence of rainfall of the current day. It is shown by

$$D_t = B_0 + B_1 D_{t-1} + B_3 T_t + B_4 T_{t-1} + B_{17} DR_t \tag{4}$$

All of the MLR models were first trained (to determine the regression coefficients) using the data in the training set (1994–2001) and then tested using the testing data set (2002), and compared using the three statistical measures of goodness of fit.

Time Series Analysis

The S-plus statistical software program was used for time series analysis in this study. The first stage in developing the time series models was to determine if the data are stationary. For this, the autocorrelation coefficient function (ACF) was used. It was found that there is a time dependence in the series for lags up to 15 at a 95% confidence level, which implies that the data series is not stationary. The data were, therefore, differenced to convert it into a stationary process. An ACF plot was then performed on the differenced series, and it was determined that the new ACF plot contained one significant lag while the remaining lags were not significant. Stationary conditions were, therefore, satisfied.

Since the data were transformed into a stationary model through differencing, autoregressive integrated moving average (ARIMA) models of order p, d, and q were used. Using the differenced data, the ACF and partial ACF (PACF) plots were then used to identify the order of the ARIMA models. In the models that were developed, the number of autoregressive parameters (p) varied from 0 to 3 and the number of moving average parameters varied from 0 to 3. One difference of the data set ($d=1$) was required to transform the series into a stationary process.

A total of nine ARIMA models were identified to fit the peak water demand series and the Akaike information criterion (AIC) was used to verify each of the models. The parameters of these nine models are shown in Table 2. All of the time series models were first trained using the data in the training set (1994–2001) and then tested using the testing data set (2002) and compared using the three statistical measures of goodness of fit.

ANN Analysis

Using the Tiberius software package, back-propagation ANNs with the "generalized delta rule" as the training algorithm were used to develop all the ANN models. The Tiberius 2.0.0 neural network modeling software package for Excel is a feed-forward multilayer perceptron trained with the back propagation algorithm and is completely written in visual basic for applications (VBA), the macrolanguage of Microsoft Excel. To develop an ANN model, the primary objective is to arrive at the optimum architecture of the ANN that captures the relationship between the input and output variables. The task of identifying the number of neurons in the input and output layers is normally simple as it is dictated by the input and output variables considered to model the physical process. The number of neurons in the hidden layer has to be optimized using the available data through the use of a trial and error procedure. In addition, optimal values for the learning coefficients and the momentum correction factor have to be determined.

In this study, ANN networks consisting of an input layer with 2–6 input nodes, one single hidden layer composed of between three and five nodes, and one output layer consisting of one node denoting the predicted peak water demand were developed. Each ANN model was tested on a trial and error basis for the optimum number of neurons in the hidden layer (three, four, and five neurons in the hidden layer were tested for each model), and for the optimum learning coefficient, which was assumed to lie within the range of 0–0.2 (Jain et al. 2001).

Thirty-nine separate ANN models were identified in this study. The neurons in the input layer of each of these 39 different ANN models represented different combinations of the various physical variables considered and can be seen in Table 3. Each of the 39 ANN models used the exact same input variables as each of the MLR models that were developed. As for the corresponding MLR models, Models ANN-1–ANN-20 were designed to assess the significance of the various input variables on peak daily water demand and to corroborate the graphical, cross-correlation, and simple linear regression analyses performed in this study. Models ANN-20–ANN-39 were developed to test whether the occurrence or nonoccurrence of rainfall (using both binary and weighted systems) was a more significant variable than the amount of rainfall.

All of the ANN models were first trained using the data in the training set (1994–2001) to obtain the optimized set of connection strengths and then tested using the testing data set (2002) and compared using the three statistical measures of goodness of fit.

Results

Cross-Correlation Analysis

To investigate the dependency between variables that influence water demand, cross-correlation coefficients between peak daily demand and each variable were calculated and are shown in Table 4. This information was used to aid in selecting input variables for multiple linear regression and ANN models. It can be seen from Table 4 that the peak daily water demand series at time t is strongly correlated with the peak demand from the previous day (with a correlation value of 0.73); the temperature of the current day (with a correlation value of 0.45); and the temperature of the previous day (with a correlation value of 0.34). It can also be observed that the daily peak water demand is more strongly correlated with the occurrence or nonoccurrence of rain-

J. Water Resour. Plann. Manage. 2008.134:119-128.

Table 3. Performance Statistics for All ANN Models

Model	Parameters	Learning Coefficient	AARE	Max ARE	R^2 (training)	R^2 (testing)
ANN-1	D_{t-1}, T_t	0.06	15	82	0.57	0.54
ANN-2	D_{t-1}, T_t, T_{t-1}	0.07	14	73	0.59	0.57
ANN-3	D_{t-1}, T_t, R_t	0.07	14	67	0.61	0.61
ANN-4	D_{t-1}, T_t, T_{t-1}, R_t	0.08	14	60	0.63	0.63
ANN-5	D_{t-1}, T_t, CR_t	0.08	14	58	0.63	0.62
ANN-6	D_{t-1}, T_t, T_{t-1}, CR_t	0.06	13	55	0.65	0.66
ANN-7	D_{t-1}, D_{t-2}, T_t	0.08	14	79	0.58	0.56
ANN-8	D_{t-1}, D_{t-2}, T_t, T_{t-1}	0.09	14	72	0.60	0.58
ANN-9	D_{t-1}, D_{t-2}, T_t, T_{t-1}, R_t	0.08	13	64	0.65	0.66
ANN-10	D_{t-1}, D_{t-2}, T_t,T_{t-1}, CR_t	0.08	14	63	0.66	0.63
ANN-11	MLR-1+population	0.08	15	88	0.59	0.54
ANN-12	MLR-2+population	0.08	15	79	0.60	0.57
ANN-13	MLR-3+population	0.07	15	83	0.62	0.58
ANN-14	MLR-4+population	0.08	14	72	0.63	0.62
ANN-15	MLR-5+population	0.08	13	64	0.64	0.63
ANN-16	MLR-6+population	0.08	14	61	0.65	0.64
ANN-17	MLR-7+population	0.08	15	87	0.59	0.54
ANN-18	MLR-8+population	0.07	14	78	0.61	0.57
ANN-19	MLR-9+population	0.06	14	70	0.64	0.61
ANN-20	MLR-10+population	0.08	13	72	0.65	0.63
ANN-21	D_{t-1},T_t,T_{t-1},R_{t-1}	0.08	14	71	0.60	0.60
ANN-22	D_{t-1}, T_t, T_{t-1}, R_{t-2}	0.08	14	73	0.59	0.57
ANN-23	D_{t-1}, T_t, T_{t-1}, R_{t-3}	0.08	14	71	0.60	0.58
ANN-24	D_{t-1}, T_t, T_{t-1}, R_{t-4}	0.07	14	73	0.59	0.58
ANN-25	D_{t-1}, T_t, T_{t-1}, R_{t-5}	0.08	14	75	0.60	0.58
ANN-26	MLR-21+population	0.08	14	81	0.62	0.59
ANN-27	D_{t-1}, T_t,T_{t-1}, CR_{t-1}	0.06	12	**40**	0.66	0.69
ANN-28	D_{t-1}, T_t, T_{t-1}, CR_{t-2}	0.06	12	41	0.65	0.67
ANN-29	D_{t-1}, T_t, T_{t-1}, CR_{t-3}	0.06	13	42	0.65	0.67
ANN-30	D_{t-1}, T_t, T_{t-1}, CR_{t-4}	0.08	12	43	0.63	0.67
ANN-31	D_{t-1}, T_t, T_{t-1}, CR_{t-5}	0.05	**12**	41	**0.66**	**0.69**
ANN-32	MLR-31+population	0.05	12	55	0.65	0.68
ANN-33	D_{t-1}, T_t, T_{t-1}, DR_t	0.07	13	49	0.63	0.66
ANN-34	D_{t-1}, T_t, T_{t-1}, DR_{t-1}	0.08	13	75	0.61	0.60
ANN-35	D_{t-1}, T_t, T_{t-1}, DR_{t-2}	0.08	14	74	0.60	0.58
ANN-36	D_{t-1}, T_t, T_{t-1}, DR_{t-3}	0.08	14	74	0.60	0.58
ANN-37	D_{t-1}, T_t, T_{t-1},DR_{t-4}	0.08	13	76	0.60	0.58
ANN-38	D_{t-1}, T_t, T_{t-1}, DR_{t-5}	0.08	14	73	0.60	0.56
ANN-39	MLR-33+population	0.1	14	64	0.66	0.65
Average			14	67	0.62	0.61

Note: ANN=artificial neural network; AARE=average absolute relative error; R^2=correlation coefficient; D_t=demand at time t; T_t=temperature at time t; R_t=rainfall at time t; CR_t=occurrence or nonoccurrence of rainfall; and DR_t=weighted occurrence or nonoccurrence of rainfall.

fall (up to and including five days before) in comparison with the actual rainfall amount. The cross-correlation coefficients between the peak demand and rainfall amount ranged from −0.12 to −0.23, whereas the cross-correlation coefficients between the peak demand and occurrence or nonoccurrence of rainfall ranged from −0.22 to −0.46. The direction of correlation is negative with rainfall, which means that a high occurrence or amount of rainfall is responsible for decreasing amounts of peak daily water demand.

Seasonal or Periodic Component Analysis

To determine whether any seasonal patterns exist in the water demand series, Fourier analysis was used. The results of the Fourier analysis indicate that there is no predominant single periodicity component in the peak daily water demand series. The strongest periodic components were recorded for the 55 and 123 day periods (which correspond to 2 and 4 month periods, respectively), although they account for only 9% of the total variance.

J. Water Resour. Plann. Manage. 2008.134:119-128.

Table 4. Cross-Correlation between Demand and Variables

Variable	Peak demand (t)
Temperature (t)	**0.45**
Temperature ($t-1$)	**0.34**
Peak demand (t)	1
Peak demand ($t-1$)	**0.73**
Rainfall (t)	−0.17
Rainfall ($t-1$)	−0.23
Rainfall ($t-2$)	−0.18
Rainfall ($t-3$)	−0.16
Rainfall ($t-4$)	−0.13
Rainfall ($t-5$)	−0.12
Occurrence of rain (t)	−0.22
Occurrence of rain ($t-1$)	−0.35
Occurrence of rain ($t-2$)	−0.42
Occurrence of rain ($t-3$)	**−0.46**
Occurrence of rain ($t-4$)	**−0.45**
Occurrence of rain ($t-5$)	**−0.46**

This indicates that, for predictive purposes, Fourier analysis cannot be used since it does not account for much of the variance. As a result, ARIMA models were developed.

Simple Linear Regression Analysis

In order to help select input variables for multiple linear regression and ANN models, simple linear regression was performed. The results from simple linear regression analysis are shown in Table 5. Strong dependence was found in peak demand (t) versus peak demand ($t-1$) with an R^2 value of 0.5268. Adequate dependence was found in peak demand (t) versus temperature (t) and ($t-1$), with R^2 values of 0.1960 and 0.1119, respectively. Weak dependence was found in peak demand (t) versus rainfall (t) and rainfall ($t-1$) with R^2 values of 0.0307 and 0.0507, respectively.

Multiple Linear Regression Analysis

Table 1 shows goodness of fit statistics for the testing of all MLR models. The first group of models, MLR-1–MLR-6, demonstrate that including T_{t-1} with T_t improved the coefficient of determination by between 3 and 5% and resulted in generally lower AARE and Max ARE values. This corroborates the cross-correlation analysis findings. As a result, it was decided to use T_{t-1} in addition to T_t in models MLR-21–MLR-39.

The second group of models, MLR-7–MLR-10, demonstrate that including D_{t-2} with D_{t-1} improved the coefficient of determination by only l.l–1.6%, but resulted in generally lower AARE and Max ARE values. However, it was decided to not include D_{t-2} in Models MLR-21–MLR-39 since the improvement was negligible.

The third group of models, MLR-11–MLR-20 and MLR-26, MLR-32, and MLR-39, demonstrate that including population as an input variable improved the coefficient of determination by only a negligible amount (less than 1% for each of the 13 models) and, in almost all cases, resulted in higher AARE and Max ARE values. As a result, models that included population as a variable were not considered in the determination of the best overall model. This corroborates the conclusions of several other studies that found that including socioeconomic variables such as population has a negligible impact on short-term water demand forecasting. This is understandable since it is highly unlikely that overall zone population would have an affect on daily fluctuations in water demand.

From the cross-correlation and simple linear regression analyses in addition to the results of Models MLR-l–MLR-20, it can be seen that D_{t-1}, T_t, and T_{t-1} are significant variables in terms of forecasting peak daily water demand in the OWC zone. It can also be seen that D_{t-2} and population are not significant. As such, it was decided to use D_{t-1}, T_t, and T_{t-1} as the basis for Models MLR-21–MLR-39.

Models MLR-21–MLR-39 show that the daily peak water demand series in the OWC zone is, in fact, better described with the use of occurrence or nonoccurrence of rainfall rather than actual rainfall amount. Models including the occurrence or nonoccurrence of rainfall (MLR-27–MLR-32) produced testing R^2 coefficients ranging from 0.58 to 0.60, AARE values ranging from 14 to 15, and Max ARE values ranging from 59 to 62. Models using the rainfall amount (MLR-21–MLR-26) produced R^2 coefficients ranging from 0.54 to 0.56, AARE values ranging from 14 to 15, and Max ARE values ranging from 71 to 79. As well, the best MLR model (discussed below) included the occurrence or nonoccurrence of rainfall as opposed to the amount of rainfall. All of this suggests that the water demand process in the OWC zone is better correlated with the occurrence of rainfall rather than the amount of rainfall.

Models using the weighted system of occurrence or nonoccur-

Table 5. Simple Linear Regression Results

Parameter	Versus parameter	Slope	R^2
Rainfall (t)	Time (t)	−0.0006	0.0007
Rainfall (t)	Rainfall ($t-1$)	0.0561	0.0031
Temperature (t)	Rainfall (t)	−0.0442	0.0034
Temperature (t)	Time (t)	0.0019	0.0143
Temperature (t)	Temperature ($t-1$)	0.7311	0.5403
Peak demand (t)	Time (t)	0.0161	0.1008
Peak demand (t)	Temperature (t)	1.4344	**0.196**
Peak demand (t)	Temperature ($t-1$)	1.0778	**0.1119**
Peak demand (t)	Peak demand ($t-1$)	0.7229	**0.5268**
Peak demand (t)	Rainfall (t)	−0.4316	0.0307
Peak demand (t)	Rainfall ($t-1$)	−0.5546	0.0507

Note: R^2=correlation coefficient.

J. Water Resour. Plann. Manage. 2008.134:119-128.

rence of rainfall did not perform as well as the models using the binary system of occurrence or nonoccurrence of rainfall. Models using the weighted occurrence or nonoccurrence of rainfall (MLR-33–MLR-39) produced R^2 coefficients ranging from 0.54 to 0.60, AARE values ranging from 14 to 17, and Max ARE values ranging from 60 to 74. Although these models did not perform as well as the models with a binary system of occurrence of rainfall, they nevertheless did perform better than the models using the rainfall amount.

In terms of the coefficient of determination, the MLR model that performed the best out of all models in training (R^2 of 0.61) was MLR-33. However, the model performing the best in the testing data set was MLR-31 (R^2 of 0.60), which was a function of the peak daily water demand of the previous day, the maximum temperature of the current and previous day, and the occurrence or nonoccurrence of the rainfall five days before. In terms of AARE and Max ARE, the MLR-31 model also performed the best among all the MLR models with values of 14 and 56, respectively. As such, overall, the MLR-31 model performed the best among all MLR models (excluding the models with population for reasons explained above). This model is given by

$$D_t = 11.2224 + 0.6132D_{t-1} + 1.035T_t - 0.6359T_{t-1} - 6.9573CR_{t-5} \tag{5}$$

Time Series Analysis

Table 2 shows the performance statistics for the testing of all time series models. A total of nine ARIMA (p, d, q) models were postulated to model daily water demand series. In terms of the AIC, the ARIMA (2,1,2) model had the best fit with the observed data as indicated by the lowest AIC value (8,443).

In terms of the coefficient of determination, all models had relatively low R^2 values in training. These values ranged from 0.46 to 0.54, with the ARIMA (2,1,2) model having the highest coefficient of determination value in the training data set. However, during the testing, it was found that the ARIMA (2,1,0) model performed the best in terms of the three statistical measures of goodness of fit. This model had two autoregressive components ($p=2$) and no moving average component ($q=0$). Like all the other ARIMA models, the data set used for this model was differenced only once ($d=0$) in order to transform the series into a stationary process.

The ARIMA (2,1,0) model had the highest R^2 value (0.46) for the testing data set. In terms of AARE and Max ARE, the ARIMA (2,1,0) model also performed the best among all ARIMA models with an AARE of 15 and a Max ARE of 62. This model is shown by

$$D_t = -0.2787(D_{t-1} - D_{t-2}) - 0.1498(D_{t-2} - D_{t-3}) \tag{6}$$

ANN Analysis

Table 3 shows the performance statistics for the 39 ANN models developed in this study and includes the learning coefficients for all the models that were found to produce the lowest root-mean-square error between the observed and forecasted water demand in the training session for each model. The optimized learning coefficients ranged from 0.05 to 0.10 for the 39 ANN models.

For each of the 39 models, three, four, and five neurons in the hidden layer were tested. It was determined that four neurons in the hidden layer produced the highest coefficients of determination. Each of the 39 models used four neurons in the hidden layer.

Models ANN-l–ANN-6 indicate that including T_{t-1} in addition to T_t improved the coefficient of determination by between 4 and 6% and resulted in lower AARE and Max ARE values. This confirms the significance of T_{t-1} an input variable. This also corroborates the graphical and cross-correlation analysis results in addition to the MLR results. As a result, it was decided to include both T_t and T_{t-1} in Models ANN-21–ANN-39.

Models ANN-7–ANN-10 indicate that including D_{t-2} with D_{t-1} can result in marginally better coefficients of determination (+3%), but it can also result in lower coefficients of determination (−5%). It was also found that including D_{t-2} with D_{t-1} generally resulted in higher AARE and Max ARE values. As such, it was decided to not include D_{t-2} in Models ANN-21–ANN-39.

Models ANN-11–ANN-20 indicate that in some cases, including population as an input variable can result in marginally better coefficients of determination. However, these models also indicate that including population as an input variable can also result in significantly lower coefficients of determination. Models ANN-26, ANN-32, and ANN-39 indicate that the addition of population as an input variable in the best ANN models actually results in lower coefficients of determination in all three cases. The lack of significance of population as an input variable corroborates the MLR analysis performed in this study.

It can be seen from Models MLR-21–MLR-39 that the peak daily water demand series in the OWC zone is better described with the use of the occurrence or nonoccurrence of rainfall rather than the actual rainfall amount. Models including the binary system of occurrence or nonoccurrence of rainfall (ANN-27–ANN-32) produced testing R^2 coefficients ranging from 0.67 to 0.69, AARE values ranging from 12 to 13, and Max ARE values ranging from 40 to 55. Models using the rainfall amount (ANN-21–ANN-26) produced R^2 coefficients ranging from 0.57 to 0.60, AARE values ranging from 13.5 to 14, and Max ARE values ranging from 71 to 81. As well, the best ANN model (discussed below) included the occurrence or nonoccurrence of rainfall as opposed to the amount of rainfall. Like the MLR results, this suggests that the water demand process in the OWC zone is better correlated with the occurrence of rainfall rather than the amount of rainfall.

Models using the weighted system of occurrence or nonoccurrence of rainfall did not perform as well as the models using the binary system of occurrence or nonoccurrence of rainfall. These models (ANN-33–ANN-39) produced R^2 coefficients ranging from 0.56 to 0.66, AARE values ranging from 13 to 14, and Max ARE values from 49 to 76.

In terms of the coefficient of determination, the best ANN model overall for both training and testing was ANN-31 with R^2 values of 0.66 and 0.69, respectively. ANN-31 is a function of the peak daily demand from the previous day, the temperature from the current day and the previous day, and the occurrence or nonoccurrence of rainfall from five days before (D_{t-1}, T_t, T_{t-1}, and CR_{t5}). In terms of AARE the best model was also ANN-31 with an AARE of 12. However, the model with the best Max ARE (with a value of 40) was the ANN-27 model. The ANN-31 model had the second best Max ARE (with a value of 41). It should be pointed out that some of the other ANN models (especially ANN-27) had performance statistics that were almost as good as the ANN-31 model.

J. Water Resour. Plann. Manage. 2008.134:119-128.

Table 6. Comparison Analysis of the Three Forecasting Techniques

Mode type	AARE	Max ARE	R^2 (training)	R^2 (testing)
(a) Average values for each forecasting technique				
MLR	15	71	0.59	0.56
ARIMA	16	105	0.55	0.39
ANN	**14**	**67**	**0.62**	**0.61**
(b) Best model for each type of technique				
MLR-31	14	56	0.61	0.60
ARIMA (2,1,0)	15	62	0.53	0.46
ANN-31	**12**	**41**	**0.66**	**0.69**

Note: MLR=multiple linear regression; ARIMA=autoregressive integrated moving average; ANN=artificial neural network; AARE=average absolute relative error; and R^2=correlation coefficient.

Comparative Analysis

A comparative analysis was performed to evaluate the relative performance of each modeling technique investigated in this study. Average values of the statistical measures of goodness of fit were calculated from all the models employing a particular technique and the results for all three techniques investigated in this study are presented in Table 6.

It is quite clear that based on the average statistical measures of goodness of fit, the ANN models outperformed the regression and time series models, recording the lowest AARE and Max ARE statistics, while recording the highest R^2 values in training and testing. It is also clear that the regression models outperformed the time series models.

The time series models had relatively high coefficients of determination during training, however, they were not able to perform well during testing, with very low testing R^2 values, high AARE values, and very high Max ARE values. A possible reason as to why the time series models did not perform well could be due to the fact that this technique did not consider climatic variables during the modeling process. In ANN and multiple linear regression, the relationship between present and past water demand data and climatic data was examined, while the time series models calculated the relationship between present and past water demand data only. This suggests the importance of climate on peak daily water demand and that the daily water demand process in the OWC zone is mainly governed by the maximum air temperature and interrupted by occurrences of rainfall.

Based on the statistical measures of goodness of fit calculated in this study, it can be observed that Model ANN-31 performed the best out of all models. Table 6 shows that it had the highest R^2 value in training (0.66), the highest R^2 value in testing (0.69), the lowest AARE (12), and the second lowest Max ARE (41) out of all MLR, ANN, and time series models. In terms of the value of R^2 in testing, the ANN-31 model was found to be 14% more accurate than the best MLR model (MLR-31), and 34% more accurate than the best time series model [ARIMA (2,l,0)].

Discussion

The best predictive variables for peak daily water demand found in this study were D_{t-1}, T_t, T_{t-1}, and CR_{t-5}. These variables were used for both the best MLR model and the best ANN model. For these variables, it was determined that there is a time dependency (autocorrelation) between D_t and D_{t-1}, to a lesser extent a time dependency between T_t and T_{t-1}, as well as a time dependency in the occurrence of rainfall between CR_t and CR_{t-5}. The reason that the occurrence or nonoccurrence of rainfall from five days before was significant can perhaps be explained in terms of a "lag of effect" phenomenon. It was also found that for both the MLR and ANN models, the peak daily water demand series in the OWC zone was better described with the use of the occurrence or nonoccurrence of rainfall rather than the actual rainfall amount. This supports the findings of Jain et al. (2001) but is opposite to the findings of Bougadis et al. (2005). Another finding was that models using the weighted system of occurrence or nonoccurrence of rainfall did not perform as well as the models using the binary system of occurrence or nonoccurrence of rainfall. As well, the lower R^2 and higher AARE values obtained in this study for peak water demand compared to some of the studies that investigated average or total daily and weekly water demand (Jain et al. 2001; Jain and Ormsbee 2002), indicate that the forecasting of peak summer water demand is less accurate perhaps due to the very high variations in peak demands found in areas of high outdoor water usage.

In this study it was shown that the ANN technique is marginally better than the multiple linear regression and time series analysis techniques in forecasting peak daily summer water demand in an area of high outdoor water demand. There are several reasons that might explain why the ANN technique was somewhat better capable of handling the data that were used in this study in order to forecast peak daily water demand. The time series models might not have performed well compared to the ANN (and multiple linear regression) models because this technique did not consider climatic variables during the modeling process. In ANN (and multiple linear regression) analysis, the relationship between present and past water demand data and climatic data is examined, while in the time series analysis performed in this study, only the relationship between present and past water demand data was calculated. However, the time series technique is, nevertheless, very useful for water supply systems where climatic data are not available.

The multiple linear regression models might have performed worse than the ANN models because MLR equations can only capture relationships of a prespecified functional form, and as such they may not always be sufficient to accurately predict the nonlinear nature of the variables involved. On the other hand, ANNs make no assumptions about the nature of the relationships between input and output variables.

An overall conclusion of this study is that none of the methods performed very well in predicting peak water demand. This seems to indicate that there were problems with the data used or the wrong driving variables were used. The present study focused on the modeling of peak water demand forecasts using climatic variables in addition to past water demands. The work could potentially be improved if other variables that affect water demand were to be examined. Examples of socioeconomic variables that could be investigated include housing characteristics (number of bathrooms, number of rooms, size of garden, household size, and the number of people in the house); property values; land use (residential, commercial or industrial); economic status (house income); day of the week (including weekday, weekend and holidays); and water price. Of these variables, it is most likely that the day of the week (weekday, weekend, or holiday) and water price could potentially improve the short-term water demand forecasts. Examples of climatic variables not used in this research that could be investigated in future studies include: evaporation; evapotranspiration; wind speed; relative humidity; cloud amount;

J. Water Resour. Plann. Manage. 2008.134:119-128.

and sunshine amount. Unfortunately, not all of the above data are readily available, and often do not exist at all. Nevertheless, if the above-mentioned socioeconomic and climatic variables are available, it is possible that different combinations of driving variables could potentially improve the forecasting ability of the various techniques explored in this study. Another possible improvement would be to explore the use of different ANN training algorithms such as radial basis functions, genetic algorithms, or self-organizing networks.

Conclusions

The motivation for this study was to investigate three important issues that have not been investigated in the literature concerning short-term peak water demand. From the results of this study, the following can be concluded: (1) the use of artificial neural networks for use in forecasting peak daily water demand in the summer months in an area of high outdoor water demand is marginally better than multiple linear regression and time series analysis; (2) peak daily water demand is better correlated with the rainfall occurrence rather than the rainfall amount itself; and (3) assigning a weighting system to the antecedent days of no rainfall does not result in more accurate models.

Acknowledgments

This study formed part of a thesis submitted for the degree of Master of Philosophy in Engineering for Sustainable Development at the Cambridge-MIT Institute at the University of Cambridge. Funding for part of these studies was provided by the Cambridge Commonwealth Trust of the University of Cambridge and this is gratefully acknowledged. The writers also wish to thank Dr. Richard Fenner of the Center for Sustainable Development at the University of Cambridge for his enthusiastic support and valuable advice throughout the course of this research. Dr. Fenner was also responsible for the idea of testing the usefulness of a weighted system of antecedent days of rainfall. Data were provided by Mr. John Bougadis of Delcan Corporation in Ottawa, Canada.

Notation

The following symbols are used in this paper:

B = regression coefficient;
CR = occurrence or nonoccurrence of rainfall;
D = demand;
DR = weighted occurrence or nonoccurrence of rainfall;
N = number of observations;
R = amount of rainfall;
T = temperature;
d_j = desired demand at node j;
w_{ij} = weight from hidden node i or from an input to node j at time t;
y_j = actual output at node j;
α = momentum correction factor;
δ = error term; and
η = learning coefficient.

References

Anderson, R., Miller, T., and Washburn, M. (1980). "Water savings from lawn watering restrictions during a drought year in Fort Collins, Colorado." *Water Resour. Bull.*, 16(4), 642–645.

Bougadis, J., Adamowski, K., and Diduch, R. (2005). "Short-term municipal water demand forecasting." *Hydrolog. Process.*, 19(1), 137–148.

Day, D., and Howe, C. (2003). "Forecasting peak demand—What do we need to know?" *Water Sci. Technol.: Water Supply*, 3(3), 177–184.

Gutzler, D. S., and Nims, J. S. (2005). "Interannual variability of water demand and summer climate in Albuquerque, New Mexico." *J. Appl. Meteorol.*, 44(12), 1777–1787.

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

Hughes, T. (1980). "Peak period design standards for small western U.S. water supply." *Water Resour. Bull.*, 16(4), 661–667.

Jain, A., and Ormsbee, L. (2002). "Short-term water demand forecast modeling techniques—Conventional methods versus Al." *J. Am. Water Works Assoc.*, 94(7), 64–72.

Jain, A., Varshney, A., and Joshi, U. (2001). "Short-term water demand forecast modeling at IIT Kanpar using artificial neural networks." *Water Resour. Manage.*, 15(5), 299–321.

Maidment, D., and Miaou, S. (1986). "Daily water use in nine cities." *Water Resour. Res.*, 22(6), 845–851.

Maidment, D., Miaou, S., and Crawford, M. (1985). "Transfer function models of daily urban water use." *Water Resour. Res.*, 21(4), 425–432.

Maidment, D., and Parzen, E. (1984). "Monthly water use and its relationship to climatic variables in Texas." *Water Resour. Bull.*, 19(8), 409–418.

Miaou, S. (1990). "A class of time series urban water demand models with non-linear climatic effects." *Water Resour. Res.*, 26(2), 169–l78.

Oh, H., and Yamauchi, H. (1974). "An economic analysis of the patterns and trends in water consumption within the service area of the Honolulu Board of Water Supply." *Rep. No. 84*, Univ. of Honolulu, Honolulu.

Pulido-Calvo, I., Roldan, J., Lopez-Luque, R., and Gutierrez-Estrada, J. (2003). "Demand forecasting for irrigation water distribution systems." *J. Irrig. Drain. Eng.*, 129(6), 422–431.

Regional Municipality of Ottawa–Carleton (ROMC). (2003). *Community profile, planning, and projects*, Ottawa, Ont., Canada.

Smith, J. (1988). "A model of daily municipal water use for short-term forecasting." *Water Resour. Res.*, 24(2), 201–206.

Zhou, S., McMahon, T., Walton, A., and Lewis, J. (2000). "Forecasting daily urban water demand: A case study of Melbourne." *J. Hydrol.*, 236(3), 153–164.

J. Water Resour. Plann. Manage. 2008.134:119-128.