WATER CONSUMPTION FORECASTING TO IMPROVE ENERGY EFFICIENCY OF PUMPING OPERATIONS

Harold Kidder, P.E., Lawrence Jentgen, P.E., Robert Hill, Ph.D., P.E, Steve Conrad
EMA, Inc.
8910 Bevlyn
Houston, Texas 77025

ABSTRACT

This paper summarizes the results of a joint AwwaRF and California Energy Commission Research Project, Water Consumption Forecasting to Improve Energy Efficiency of Pumping Operations.

Short Term Consumption Forecasting (STCF) is required for water utilities to proactively optimize pumping and treatment operations to minimize energy, water supply, and treatment costs while maintaining a reliable and high quality product for their customers. This paper provides information on various techniques, performance data, benchmarks, and selection criteria to assist utilities in evaluating and selecting the best option for consumption forecasting to support optimized operations with SCADA, advanced operational software, and new business practices.

The paper also defines operational results from water utilities currently forecasting consumption on a daily basis and prototype consumption forecasters that were tested at utilities not currently forecasting. The operational results of different forecasting methods are documented. Operational data were collected from nine utilities across North America (Colorado Springs Utilities, East Bay Municipal Utility District, Greater Vancouver Regional District, JEA, Las Vegas Valley Water District, San Diego Water Department, Seattle Public Utilities, Toronto Water, and Washington Suburban Sanitary Commission). These utilities represent a broad diversity in climatic conditions and customer types. STCF performance data were analyzed for all seasons of the year to provide peak, off-peak, and average day consumption data.

Results of the research included in this paper include a summary of STCF methodologies and tools, STCF benchmarks, operational experience from water utilities, case study results, and recommendations for system selection and implementation.

The value of accurate forecasts to reduce cost associated with energy, water supply, and treatment management in system operations is presented.

INTRODUCTION

Pumping operations at water utilities typically follow consumption. That is, wells and booster pumps are automatically controlled based on reservoir levels and distribution pressures. As consumption increases, levels and pressures fall and pumps are turned on. As consumption decreases, reservoir levels rise, distribution pressures increase, and wells and booster pumps turn off. The pumps follow consumption during the day to maintain reservoir levels and system pressures within normal operating ranges.

While this reactive mode of operation meets operating criteria from a reliability perspective, it does not leverage the opportunity to reduce operating costs associated with proactive system operations. To move to proactive system operations requires a means to accurately forecast consumption on a daily (Figure 1) and/or hourly (Figure 2) basis and then schedule water supplies, treatment, and pumping to minimize cost and maximize quality.

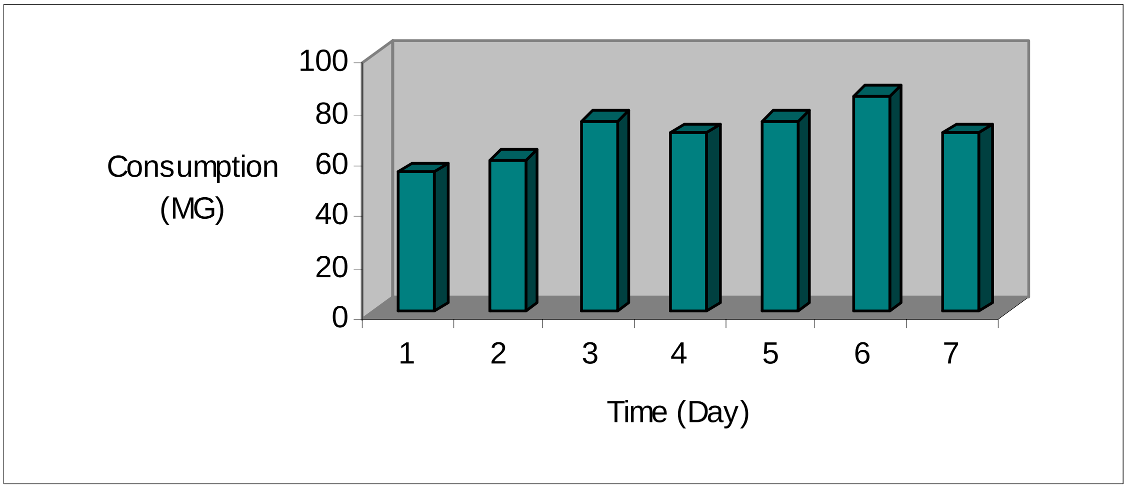

Figure 1: Daily Consumption Profile

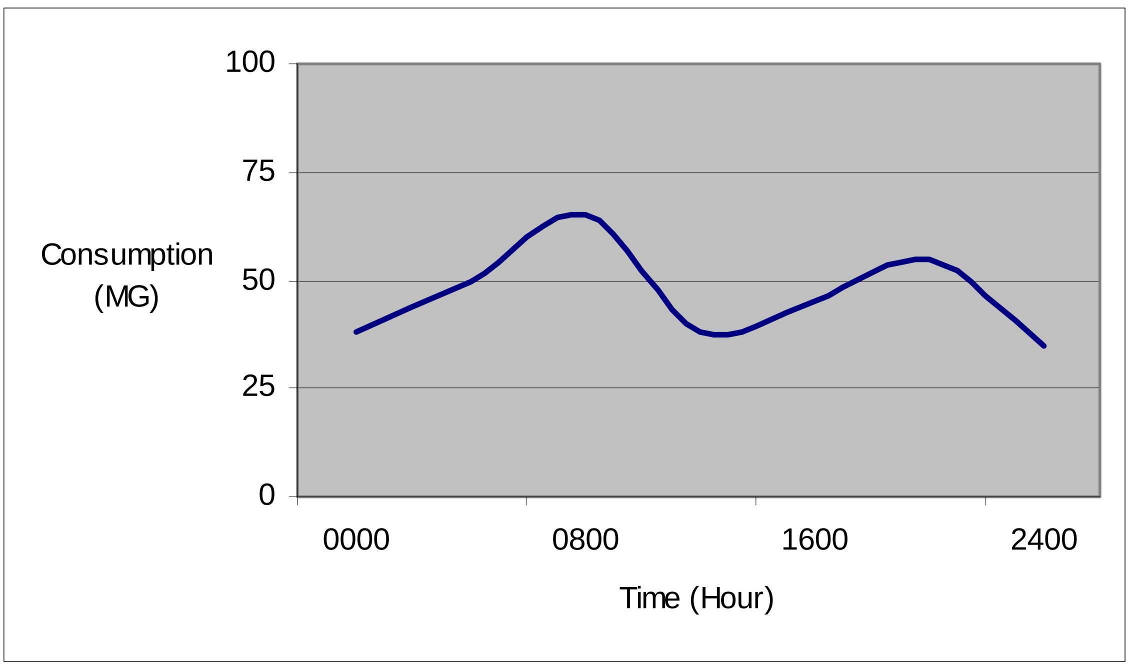

Figure 2: Hourly Consumption Profile

The objective of this research was to identify, test, and evaluate available methods and tools for making short-term water consumption forecasts necessary for optimizing pump schedules and energy use to support the implementation of an Energy and Water Quality Management System (EWQMS).

The focus of this project is on Short-term Consumption Forecasting (STCF) as it relates to energy, water quality, and water supply management in an operations environment. The best representation of integration of the STCF into water system operations is the EWQMS model (Figure 3). As shown in Figure 3, the Operations Planner and Scheduler (OPS), which is a function consisting of people and software programs (system simulator and optimizer), develops a System Operating Plan based on the Water Consumption Forecast, Maintenance Construction Schedule, Energy Cost Schedule, Water Quality, Water Supply, and "The Utility's Performance Criteria." The System Operating Plan is used by Operations to optimally control treatment plants, pumping plants, and the distribution system. The accuracy of the forecast greatly influences the cost savings and effectiveness of optimization programs. Over or under estimating consumption will result in running plants or pumps during periods of high energy and water supply costs. Assumptions made for water quality management may be invalidated if there are significant errors in the forecast.

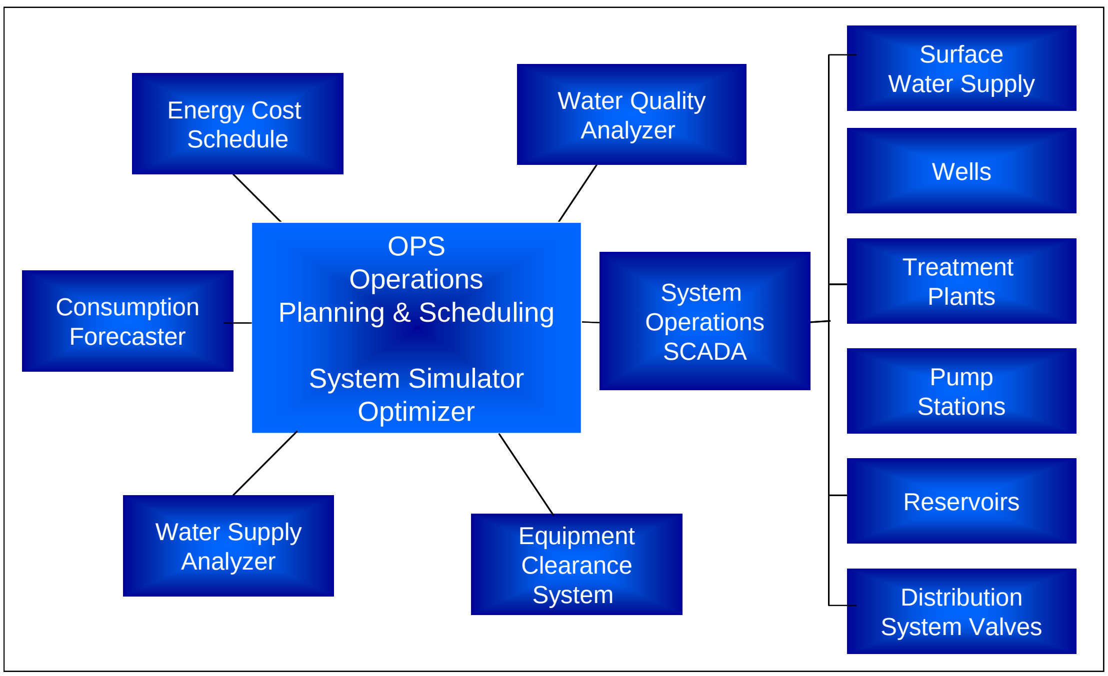

Figure 3: Energy and Water Quality System (EWQMS) Model

The decision to forecast consumption on a daily or hourly basis is significant and depends on business needs of the utility. For example, if the STCF is used for water supply and treatment only, then a daily forecast may be the only requirement. However, if the utility is minimizing time-based energy cost or demand charges, then an hourly forecast is required.

Another key requirement for the STCF is the number of service areas required for forecasting. In some cases, only a system forecast is required to satisfy optimization requirements. In other applications it is important, from a mass-balance perspective and water conveyance perspective, to forecast consumption at multiple service areas in the water system. The number of forecasted service areas depends on the size of the utility. It is not uncommon to see 40-50 services areas for larger utilities.

METHODS AND TOOLS

Research conducted for this project identified several STCF techniques used by water, gas and electric utilities. The most common are:

- Heuristic
- Linear Regression
- Similar Day
- Artificial Neural Network
- Hybrid

The classic and most utilized approach by water utilities today is **Heuristic**; that is, a system or treatment plant operator estimates today's or tomorrow's consumption based on recent consumption trends, predicted weather report, day of the week, knowledge of future events (e.g., athletic, political, cultural), and historical knowledge of utility system performance.

Linear Regression estimates consumption based on recent consumption trends, day of the week, weather, nonconforming and random consumption components. These models are also simply called regression or sometimes autoregression (AR) models. The terms are used interchangeably throughout this paper.

The **Similar Day** technique searches on a historical database for days in the past that had conditions matching the projected conditions for the upcoming day. The consumption patterns for each of these similar days are used to generate an average consumption forecast for the upcoming day.

Artificial Neural Networks (ANN) are mathematical models inspired by our understanding of biological nervous systems. They accept a large number of inputs which affect consumption and learn, from training samples, the relationships to output consumption.

Forecasting systems often use a **Hybrid** approach to consumption forecasting—that is, a blending of two or more methodologies. For example, a utility may choose to blend similar day and linear regression, or ANN with statistical techniques. Again, there is one common thread through all the approaches—the heuristics of the human forecaster. Knowledge of the utility system, consumption patterns, and influences of other factors is required to develop a consistently accurate forecast.

RESEARCH APPROACH

The approach used in the project consisted of the following phases:

Phase 1 Initiation	Researched existing STCF tools used in water, gas and electric industries
Phase 2 Analysis	Analyzed and benchmarked STCF systems in operation at several utilities. Developed and analyzed STCF prototypes at utilities not currently forecasting
Phase 3 Documentation	Documented performance of the STCF systems, defined benchmarks and develop product selection criteria

It was important that participating utilities represented multiple climatic zones (e.g., desert, tropical, oceanic, rain forest, alpine) with multiple customer demographics. Four of the utilities had STCFs in operation or experience using STCFs.

Utility	STCF Technique
Colorado Springs Utilities	Similar Day
JEA	ANN
San Diego	ANN
Las Vegas Valley Water District	ANN

Prototype STCFs were developed and tested at utilities that did not have operational STCFs:

Utility	STCF Technique
Toronto Water (TW)	Regression, ANN, Heuristic
Washington Suburban Sanitary Commission (WSSC)	Regression, ANN, Heuristic
East Bay Municipal Utility District (EBMUD)	ANN, Heuristic
Greater Vancouver Regional District (GVRD)	ANN, Heuristic
Seattle Public Utilities (SPU	ANN, Heuristic

Each prototype model used yesterday's actual consumption values to forecast tomorrow's consumption. Examples of the prototype forecasting results are shown in Figures 4-12.

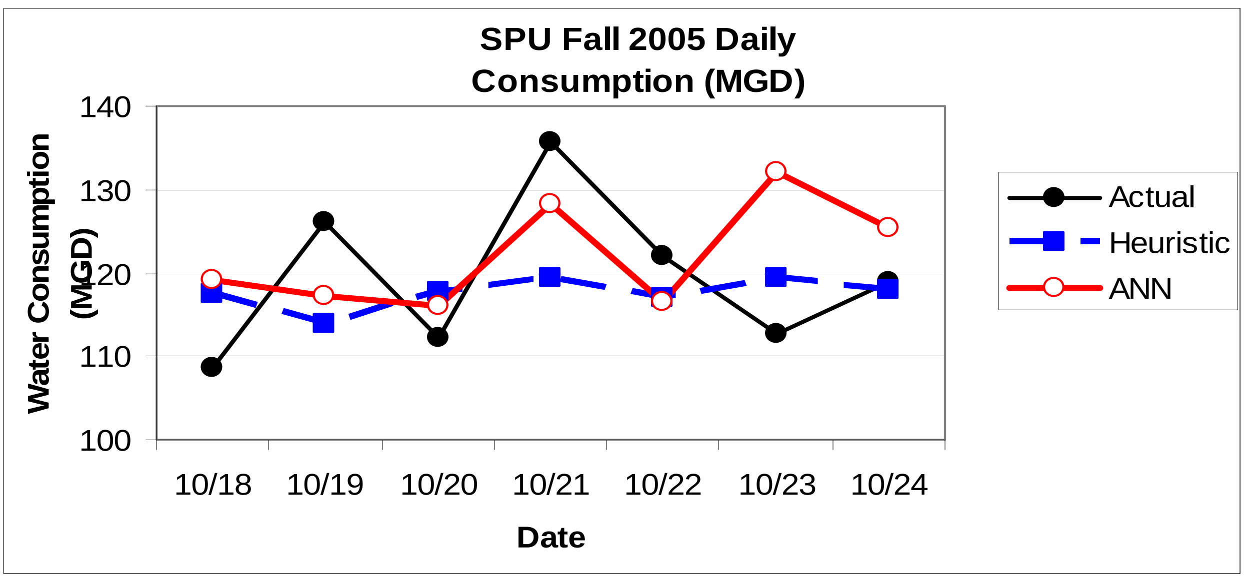

Figure 4: SPU System Daily Consumption – Fall Test Week (October 2005)

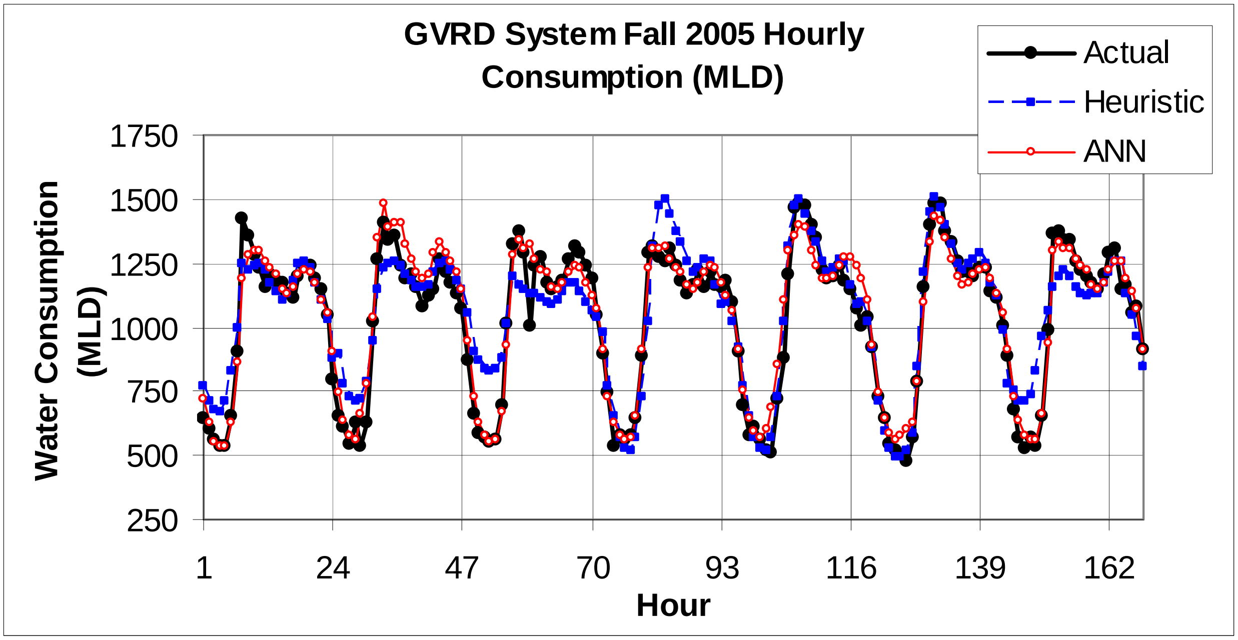

Figure 5: GVRD System Hourly Consumption – Fall Test Week (November 2005)

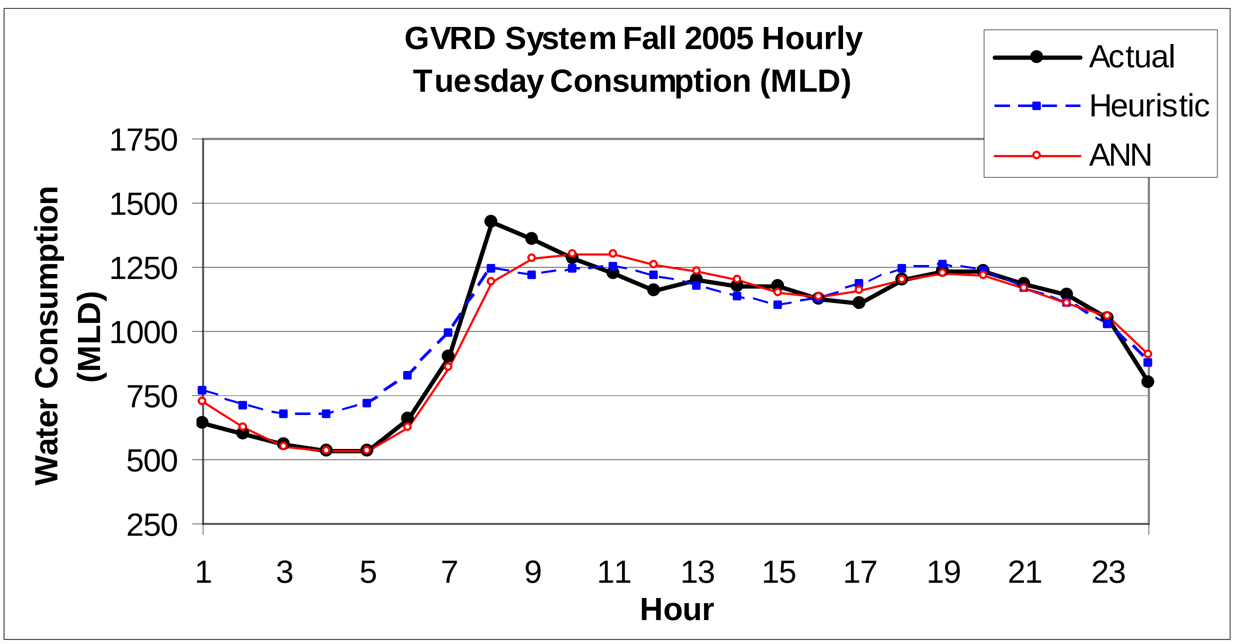

Figure 6: GVRD Test Day (Fall 2005)

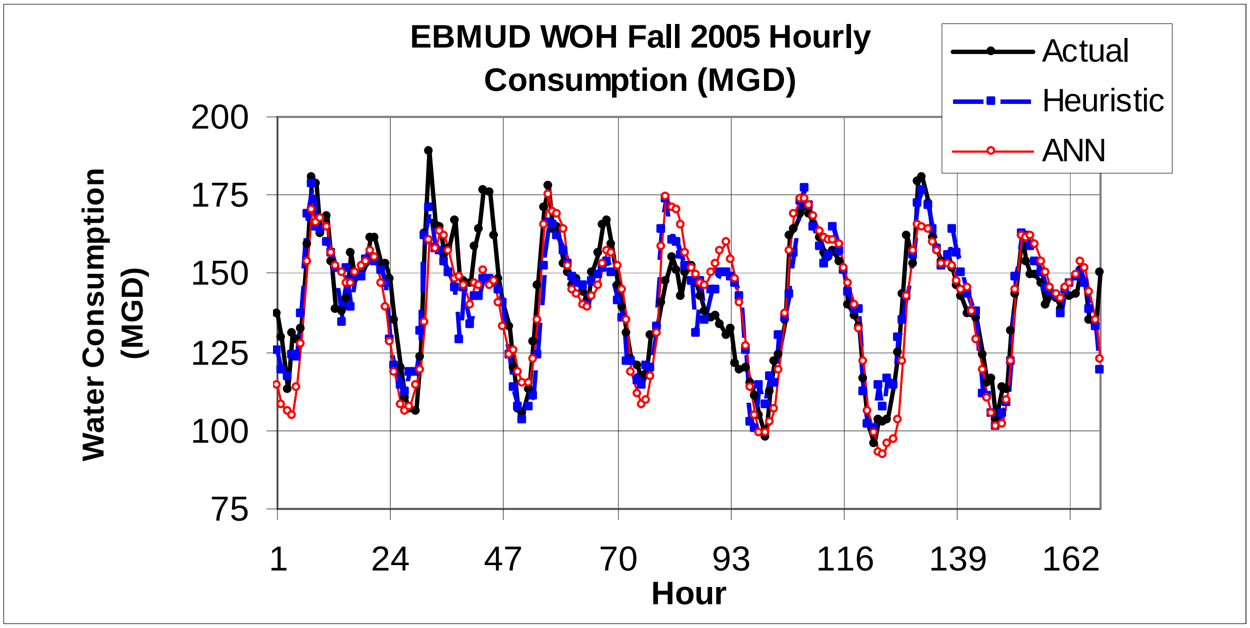

Figure 7: EBMUD System Hourly Consumption – Fall Test Week (November 2005)

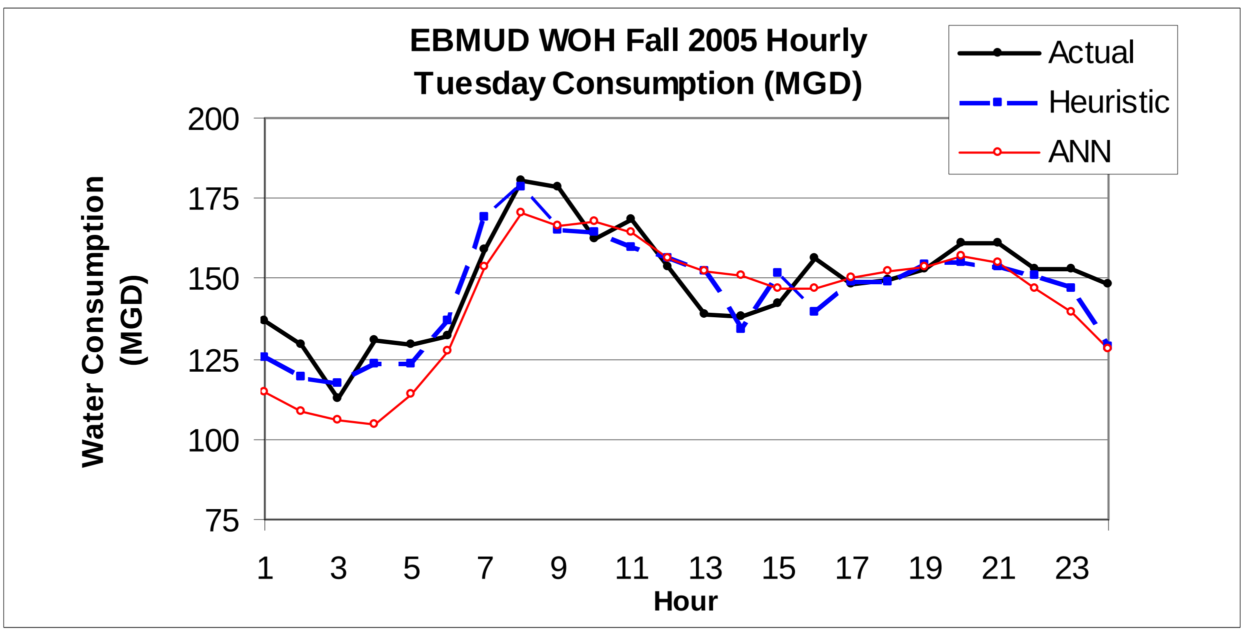

Figure 8: EBMUD Test Day – Fall (November 2005)

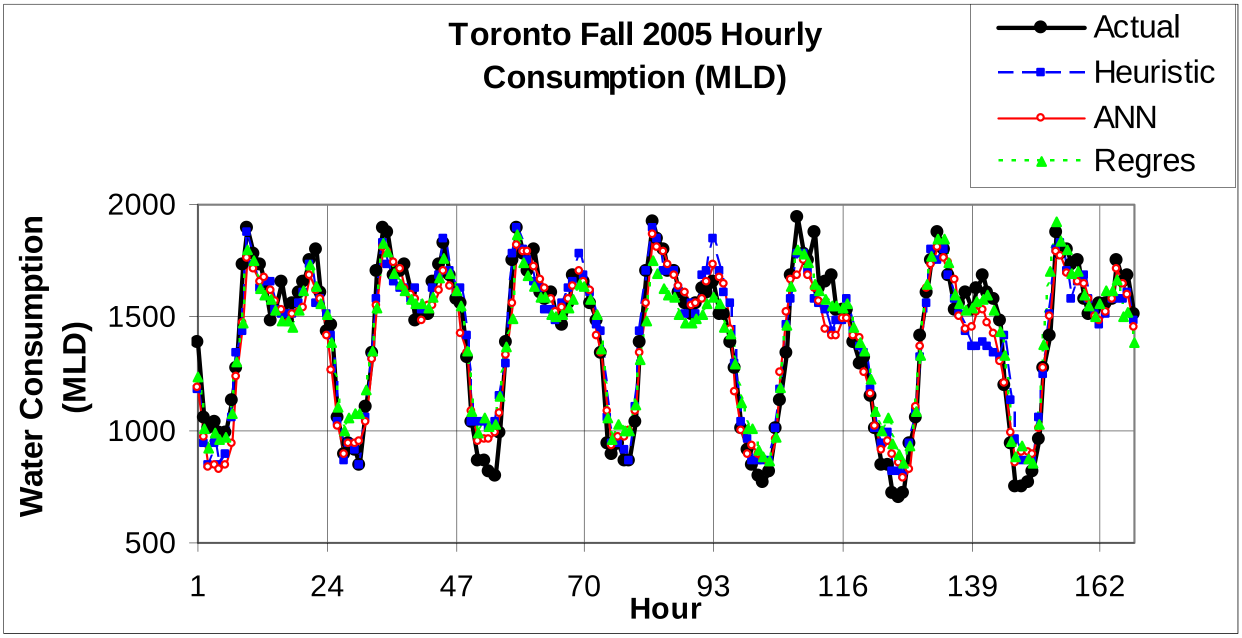

Figure 9: Toronto System Hourly Consumption – Fall Test Week (October 2005)

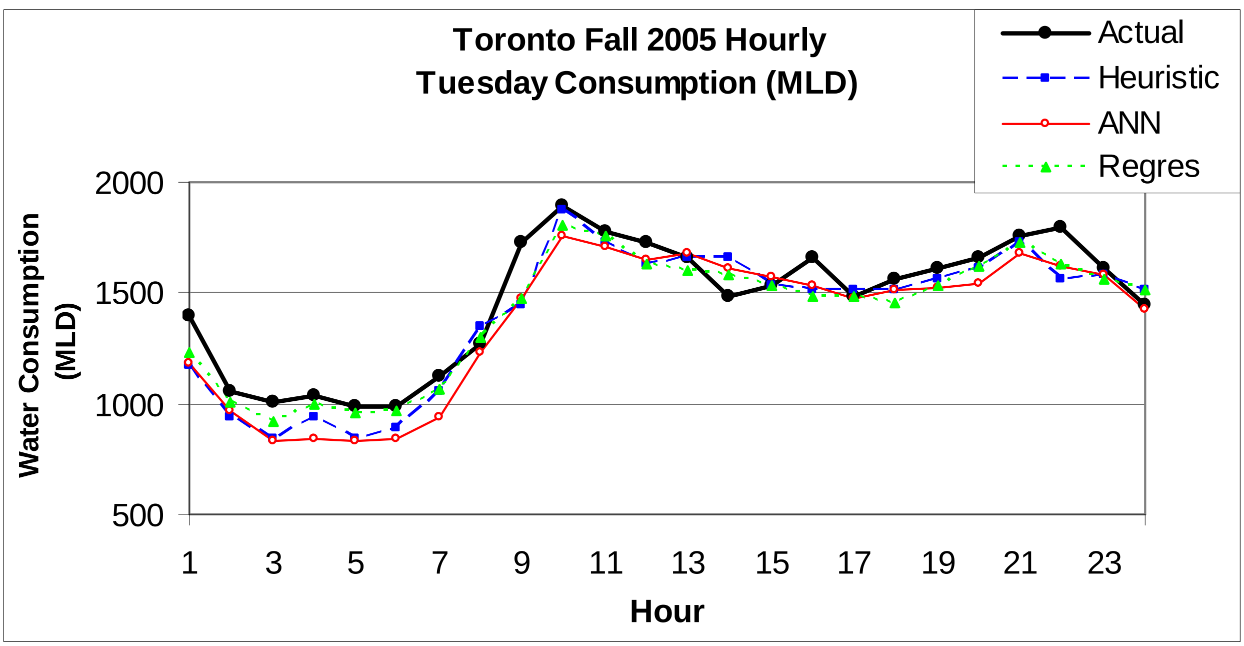

Figure 10: Toronto Test Day – Fall (October 2005)

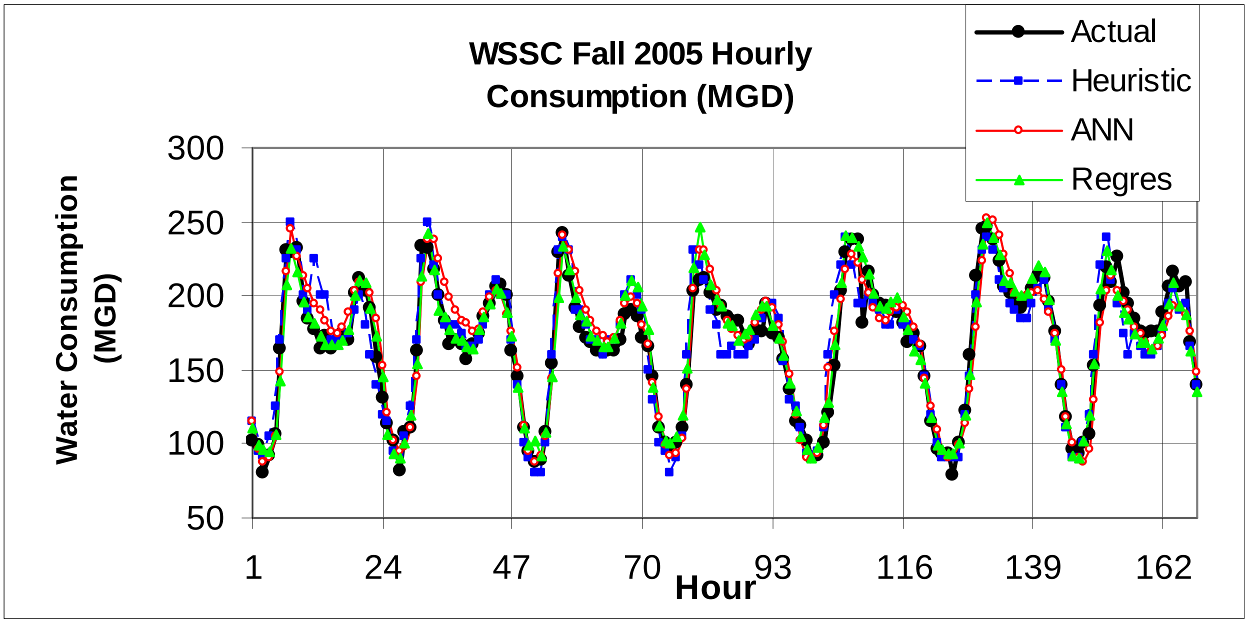

Figure 11: WSSC System Hourly Consumption – Fall Test Week (October 2005)

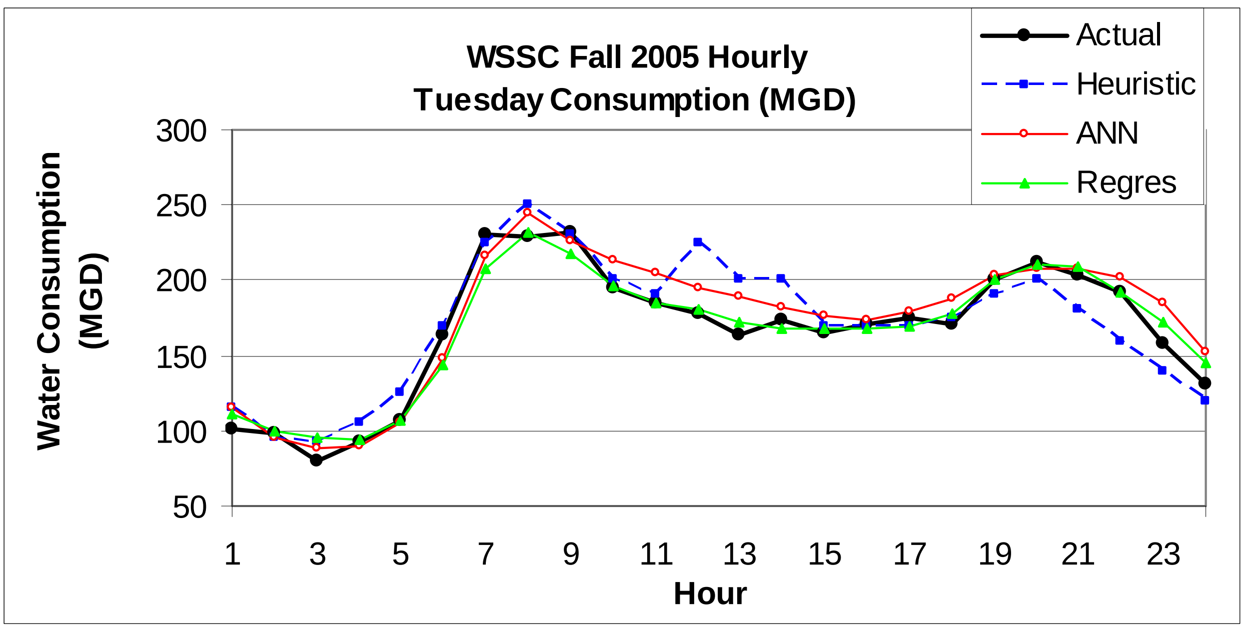

Figure 12: WSSC Test Day – Fall (October 2005)

RESULTS

Table 1 illustrates the average accuracy of the prototype models over a one week test period during each of the four seasons. Accuracy of the forecasts is based on Absolute Relative Error (ARE) which is calculated as follows:

$$ARE = 100 * \left| \frac{Forecast - Actual}{Actual} \right|$$

Hourly Average ARE (Hourly AARE) is the average of the hourly ARE values, and Daily Average ARE (Daily AARE) is the average of the daily ARE values.

The Hourly AARE ranged from 5.7 – 8.5% and the Daily AARE ranged from 2.5 – 5.1%.

Table 1: Summary of Prototype Testing

	Hourly AARE			Daily AARE		
Utilities	**Heuristic**	**ANN**	**AR**	**Heuristic**	**ANN**	**AR**
Toronto	6.0%	8.4%	5.9%	3.0%	4.8%	3.5%
WSSC	6.2%	5.8%	5.7%	2.9%	2.6%	2.5%
EBMUD	7.2%	6.7%	None	2.6%	3.2%	None
SPU	None	None	None	5.1%	4.8%	None
GVRD	8.5%	5.9%	None	4.2%	3.4%	None
Average	7.0%	6.7%	5.8%	3.5%	3.8%	3.0%

The Heuristic and ANN models had the best performance (lowest error) during the fall and winter seasons when consumption is lowest and most regular from day-to-day, and lower performance (higher error) in the spring and summer seasons when consumption is highest and subject to more day-to-day variability. The AR, however, maintained good performance during most of the seasons, except the Summer period. As a result, the AR model performed best overall on a seasonal basis, with the Heuristic and ANN models essentially tied for second place.

Table 2 contains Daily AARE arranged by day of the week.

Table 2: Summary of Pilot Utilities Results Day-by-Day

	Daily AARE		
Day of Week	**Heuristic**	**ANN**	**AR**
Tuesday	**3.2%**	**3.1%**	**3.6%**
Wednesday	**3.3%**	**4.5%**	**1.8%**
Thursday	**3.5%**	**3.5%**	**2.9%**
Friday	**3.8%**	**3.1%**	**2.7%**
Saturday	**3.0%**	**3.6%**	**3.3%**
Sunday	**4.0%**	**4.8%**	**4.1%**
Monday	**3.9%**	**3.8%**	**2.5%**
Average	**3.5%**	**3.8%**	**3.0%**

All of the forecast models demonstrated the lowest performance (highest error) on Sunday when consumption is generally lowest. In the case of the ANN model, the Sunday forecast uses the Friday Consumption as an input to the model, and the Wednesday forecast uses the Monday Consumption. This may explain the large errors on Sunday and Wednesday.

The AR model preformed best overall during the weekdays. The Heuristic models were the most consistent from day-to-day, with an AARE range between a low of 3.0% and a high of 4.0%.

Figure 13 is a plot of the average Hourly AARE for each model type displayed on an hour-by-hour basis. Each hour is an average of all four seasons of the four utilities for which hourly data were collected (GVRD, EBMUD, Toronto, and WSSC).

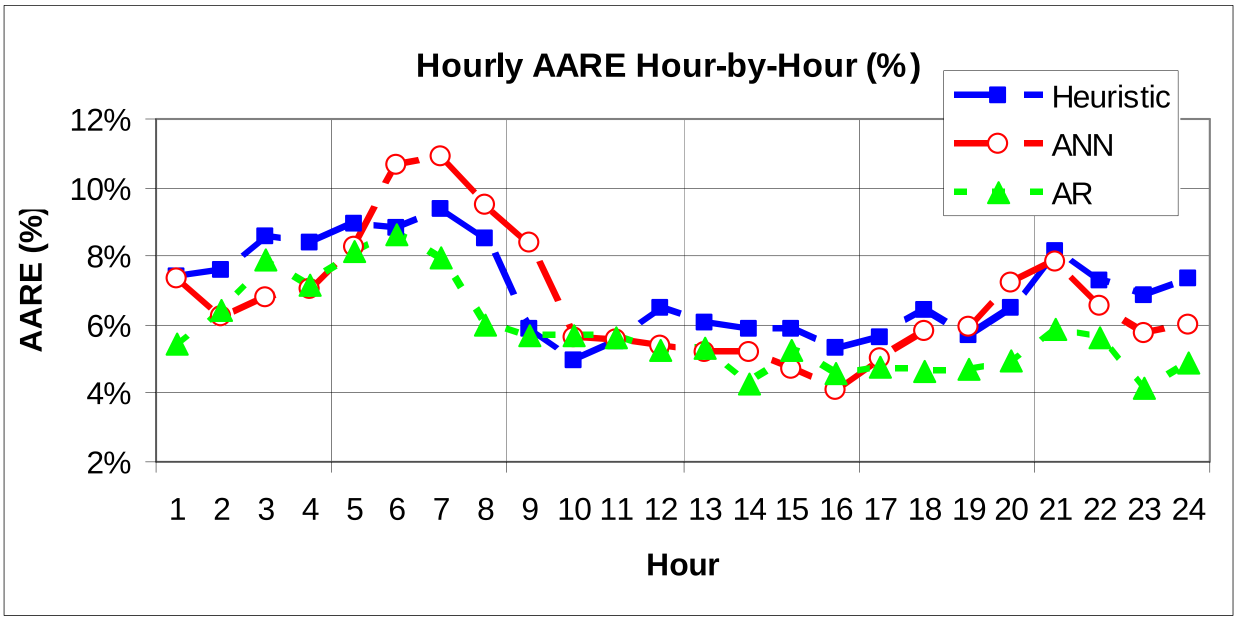

Figure 13: Hourly Average AARE For All Utilities

All of the models experienced peaks in error at the same hours as the daily peaks in consumption. Consumption during non-peak hours, between hour 10 and hour 17, tends to remain relatively constant from one hour to the next, and even one day to the next. The consumption at peak hours, between hours 4 and 10, and hours 20 and 23 tend to have greater variability, from both an hour-to-hour and day-to-day standpoint. These variations help explain the increased magnitude of errors during those hours.

The AR model performed best of the three models, but since the AR was used at only two of the utilities, this may not be definitive. When compared to the Heuristic model, the ANN model performed slightly worse during the morning peak period and better during all of the other hours of the day. Essentially, the three models performed equally well overall from hour-to-hour.

Table 3 compares the accuracy of operating forecasting systems to prototypes. The accuracy of hourly forecasts for operational systems ranged from 8.0% to 10.8% and daily averages ranged from 3.9% to 4.7%. The accuracy data for operational systems was derived from weekly sample periods for each season of a year.

Table 3: Operational Systems vs. Prototype Systems

Summary of Operational Systems		
Utilities	**Hourly AARE**	**Daily AARE**
JEA (ANN)	10.8%	3.9%
CSU (Similar day)	8.0%	4.1%
LVVWD (ANN)	None	3.2%
San Diego (ANN)	8.0%	4.7%
Average	8.9%	4.0%

Summary of Prototype Results		
Method	**Hourly AARE**	**Daily AARE**
Heuristic	7.0%	3.9%
ANN	6.7%	3.8%
AR	5.8%	3.0%

A utility's selection of a forecasting technique and its expected accuracy depend on the application(s) that are to use the forecast; initial cost and maintenance of the forecasting software/method; and complexity of the forecast (daily, hourly, system, multi-area).

Improvement in the benchmark accuracies defined in this research depends on a number of factors:

- Accuracy and repeatability of SCADA data used in the forecasts
- Sophistication and calibration of the software tools
- Continuous maintenance of the software tools

The accuracy of the forecasts at JEA, CSU, LVVWD and San Diego reflect real world conditions. This includes short-term loss of SCADA data, adverse weather conditions and other adverse/abnormal conditions. The prototypes were tested in an environment with reasonable control over the operational conditions. The prototype ANN models were developed in less than a week for each utility and the regression models were developed in a couple of days for each utility. The development costs were low.

OBSERVATIONS AND LESSONS LEARNED

The following are observations and lessons learned in the execution of this 18 month research project. Observations include:

- All methods worked well to forecast daily consumption
- Daily forecasts improved when hourly data is also forecasted
- All methods require an initial investment of labor to develop the model—ANN requires the least maintenance support
- Regression model requires new hourly data every day with no historical data gaps; ANN only requires data from the previous day
- ANN handles gaps in historical data better than regression model
- Error or noise in recent past are reflected directly in the forecasts for both ANN and regression models
- Daily consumption correlates with type of day (weekday vs. weekend)
- Not every hour of the day is sensitive to daily variations in consumption

Lessons learned include:

- Accuracy is highly dependent on quality of historical data
- Accuracy is highly dependent on quality of all input data
- Weather may not be as important as one would think
- ANN models hold up well over time
- Retraining an ANN may degrade the model—take care
- Training ANN models with a full year of data provides better results than training with only a single season of data
- Daylight savings is problematic for all forecast models
- Forecast is limited by accuracy of measurement equipment
- Multiple parameters can make models less responsive

More information concerning this project will be included in the upcoming AwwaRF Report, "*Water Consumption Forecasting to Improve Energy Efficiency of Pumping Operations.*" This report will be published by AwwaRF in the summer 2007.

BIBLIOGRAPHY

Coulbeck, B., S. Tennant and C. Orr. 1985. Development of a Demand Prediction Program for Use in Optimal Control of Water Supply. Proceedings of the 4th International Conference on Systems Engineering. Coventry Lanchester Polytechnic. Page 176.

Crommelynck, V., C. Duquesne, M. Mercier, and C. Miniussi, 1991 (Approx). Daily and Hourly Water Consumption Forecasting Tools Using Neural Networks. Syndicat des Eaux d'Ile-de-France, Paris, France.

Jain, A. and L. Ormsbee, 2002. Short-Term Water Demand Forecast Modeling Techniques. Journal AWWA. Vol. 94. No. 7.

Khotanzad, A., 1999. Natural Gas Load Forecasting With Combination of Adaptive Neural Networks. Proceedings of International Joint Conference on Neural Networks. Washington DC. July 10-16, 1999.

Lamb, P. and D. Logue, 2001. Implementation of a Gas Load Forecaster At Williams Gas Pipeline. 2001 Pipeline Simulation Interest Group Meeting. Salt Lake City Utah.

Maidment, D.R., 1986. Hourly Urban Water Use Analysis and Forecasting. Research proposal submitted to USGS by the Center for Research in Water Resources, University of Texas at Austin.

Maidment, D.R., S. Miaou and M. Crawford, 1985. Transfer Function Models of Daily Urban Water Use. Journal of Water Resources Research. Vol. 21. No. 4. Page 425.

Papalexopoulos, A.D. and T. Hesterberg, 1990. A Regression-Based Approach to Short-Term Load Forecasting. IEEE Trans. On Power Systems. Vol. 5. No. 4. Page 1535.

Papalexopoulos, A.D., S Hao and T. Peng, 1994. An Implementation of a Neural Network Based Load Forecasting Model for the EMS. IEEE Trans. On Power Systems. Vol. 9. No. 4. Page 1956.

Sugianto, L.F. and X. Lu (Date Unknown). Demand Forecasting in the Deregulated Market: A Bibliography Survey. School of Business Systems. Victoria, Australia.

Tsoukalas, L.H. and R. Uhrig, 1997. Fuzzy and Neural Approaches in Engineering. John Wiley & Sons. New York, NY.