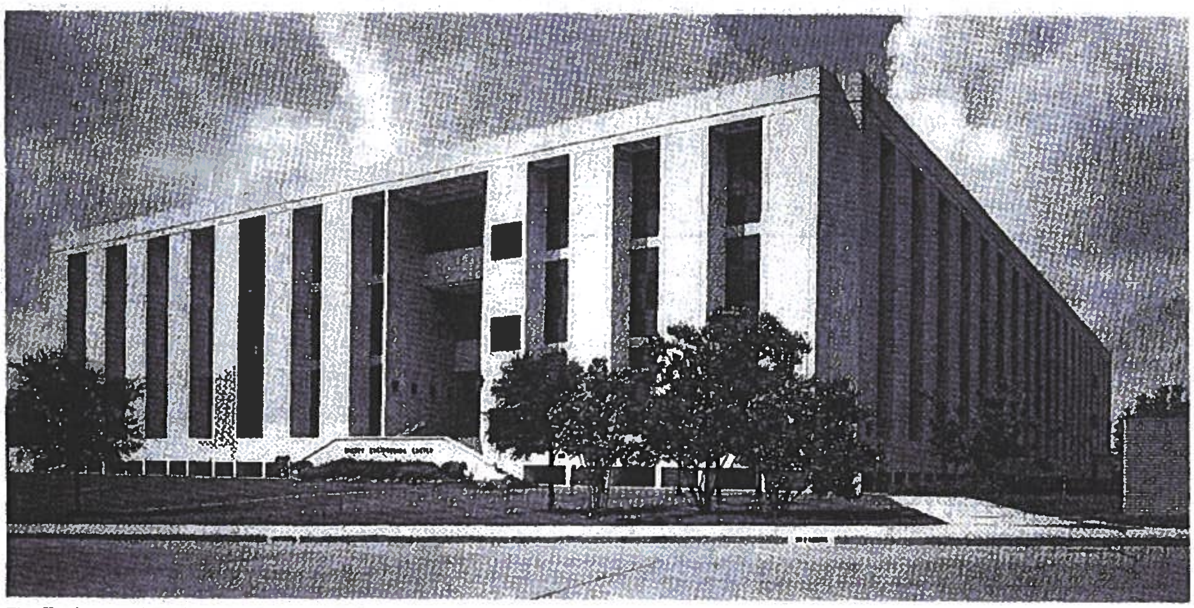
Predicting Hourly Building Energy Usage

The results of the 1993 Great Energy Predictor Shootout identify the most accurate method for making hourly energy use predictions

By Jan F. Kreider, P.E., Ph.D.; and Jeff S. Haberl, P.E., Ph.D.
Member ASHRAE Member ASHRAE



The Zachry Engineering Center at Texas A&M University.

The prediction of energy usage by HVAC systems is important for the purposes of HVAC diagnostics, system control, parameter and system identification, optimization and energy management. Many new techniques are now being applied to the analysis problems involved with predicting the future behavior of HVAC systems and deducing properties of these systems. Similar problems arise in most observational disciplines, including physics, biology and economics.

New tools (such as genetic algorithms, simulated annealing, the use of connectionist models for forecasting and tree-based classifiers or the extraction of parameters of nonlinear systems with time-delay embedding) promise to provide results that are unobtainable with more traditional techniques. Unfortunately, the realization and evaluation of this promise have been hampered by the difficulty of making rigorous comparisons between competing techniques, particularly ones that come from different disciplines.

About the authors

Jan F. Kreider is a professor and director of the Joint Center for Energy Management in the department of Civil, Environmental and Architectural Engineering at the University of Colorado, Boulder. He received his PhD in chemical engineering from the University of Colorado. Kreider is a member of ASHRAE TC 1.4 (Control Theory and Application) and chairman of the *Handbook* subcommittee of TC 4.7 (Energy Calculations).

Jeff S. Haberl is an assistant professor with the Energy Systems Laboratory in the Mechanical Engineering Department at Texas A&M University, College Station, Texas. He received his PhD in civil engineering from the University of Colorado. Haberl is chairman of the *ASHRAE Handbook* subcommittee of TC 1.5 (Computer Applications), and is a member of TC 1.7 (Operation and Maintenance Management), the *Journal/Insights* committee and SPC-140P, Standard Method of Test for the Evaluation of Building Energy Analysis Computer Programs.

To facilitate such comparisons and to foster contact among the relevant disciplines, ASHRAE TC 4.7 and TC 1.5 organized a building data analysis and prediction competition in the form of an ASHRAE seminar held in Denver in June 1993. Forecasting or prediction using *empirical models* was the goal of the competition, versus the validation of system simulation codes.

Two carefully chosen sets of energy and environmental data from real buildings were made available to over 150 contestants. Each contestant was required to prepare quantitative analyses of these data (including predictions) and submit them in a specific format to the seminar co-chairs prior to the Denver seminar. The first author evaluated all submittals using exactly the same software. The six entrants with the most accurate results each made a presentation at the seminar.

This competition was organized to help clarify the conflicting claims among many researchers who use and analyze building energy data and to foster contact among these persons and others not previously involved in building energy data analysis. The intent was not only to declare winners but also to establish a format in which rigorous and impartial evaluations of techniques could be made.

Because there are natural measures of performance accuracy, a rank-ordering was given. In all cases, the goal was to collect and analyze quantitative results to understand similarities and differences among the approaches.

The data sets

Two distinct data sets were provided to all contestants. They were given two sets of independent variables along with the corresponding values of dependent variables (energy usage or solar radiation). The accuracy of predictions of the dependent variables from values of independent variables from this data set were presented as part of each competitor's submittal. This is the classical method of testing regressions or other curve fits to data.[1]

However, a more rigorous test was also required. Some of the dependent variable values were withheld from each of the two data sets (this is explained further in Kreider and Haberl[2]) and were known only to the shootout organizers. The data sets from which the dependent variables were withheld are hereinafter called the *testing sets*, whereas the data that include a full set of both independent and dependent variables are called the *training sets*.

The independent variables in the testing set were used by each participant to make their best predictions of the corresponding dependent variables using models that were developed from the training data set. This was the essence of the competition. The contestants' predictions were then compared with the true (data) values of the dependent variables.

Two data sets were chosen to address two different sorts of building-related data analysis problems. The *A data set* (approximately 3,000 points) included time records of hourly chilled water, hot water and whole-building electricity usage for a four-month period in an institutional building. Weather data and a time stamp were also included. The hourly values of usage of these three energy forms were to be predicted for the two following months. The test-

Continued on page 74

Predicting Hourly Building Energy Usage

Continued from page 73

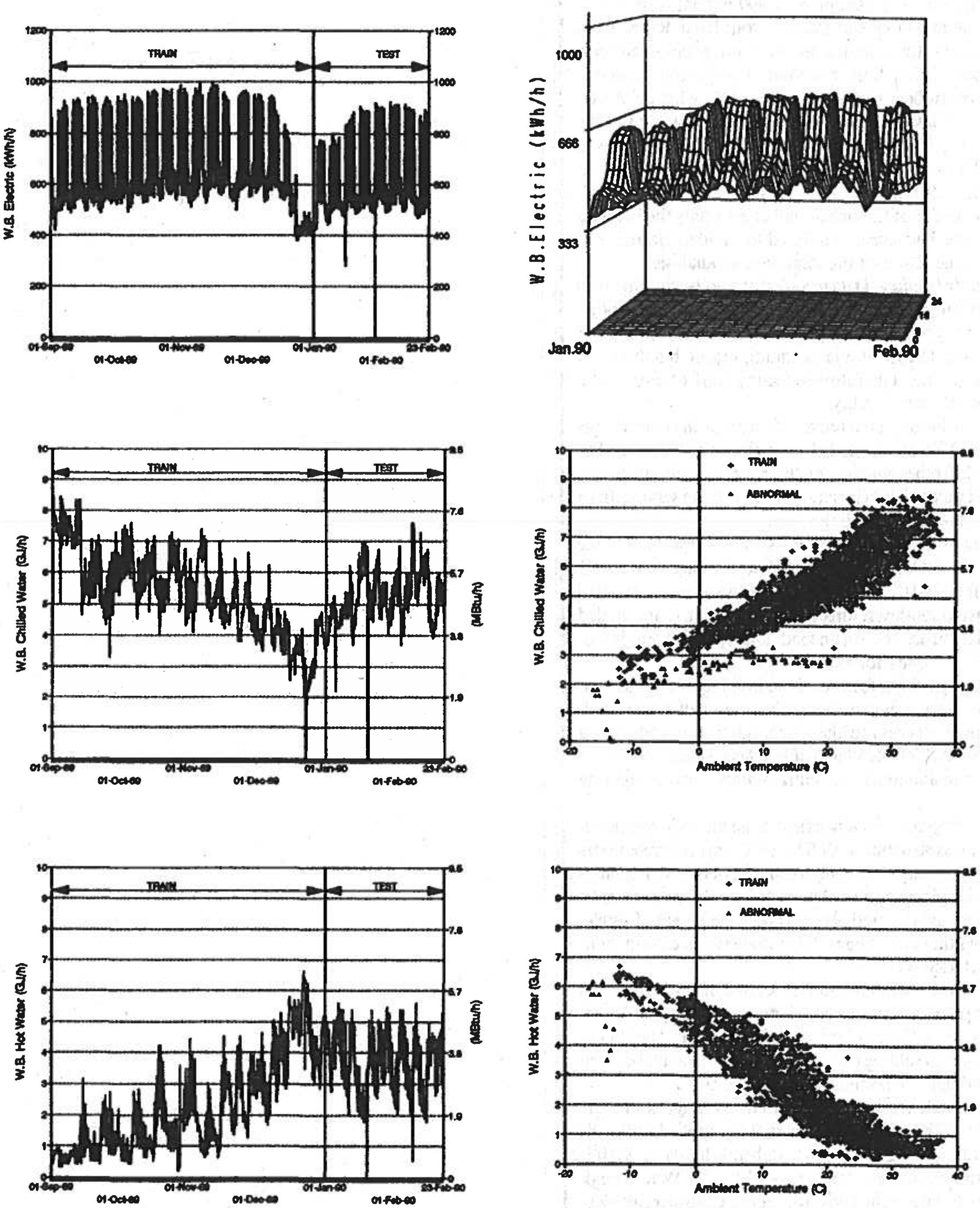


Figure 1. Training and testing data for the Engineering Center. The upper left plot is a time series trace of the whole-building lights and receptacles load; the upper right plot is a 3-D profile of electricity use for the testing period only; the center plots represent whole-building chilled water consumption; and the bottom plots represent whole-building hot water use.

ing set consisted of the two months following the four-month period.

The *B data set* (approximately 2,400 points) consisted of solar radiation measurements made by four fixed devices to be used to predict the time-varying hourly beam radiation during a six-month period. This four-pyranometer device can be used in an adaptive controller to predict building cooling loads.[3] A random sample of data from the full data set was reserved as the training set of 1,500 points. The value of beam radiation was to be predicted from data from four fixed sensors for the testing set of 900 additional points.

The remainder of this article will discuss only the building energy data set. The reader is referred to Kreider, Haberl, and Curtiss[2-4] for the results of the solar data set analysis.

Case study building. Data recorded at a university engineering center (EC) were provided to the contestants, but neither physical details, the type of building nor its location were divulged to the contestants. The EC is a large, multipurpose building that contains classrooms, laboratories, faculty/staff offices, and a large central computer facility.

The EC is located on a university campus in central Texas (30°40'N, 96°2'W) about two hours northwest of Houston. The EC and over 250 other buildings on the campus receive steam, hot water, chilled water, electricity and communication services from a centralized utility distribution system.

The four-story, 324,400 ft^2 (30 100 m^2) EC was built in the early 1970s. The building measures 339 ft (the long axis) by 221 ft and is 60 ft high (111 × 73 × 20 m). The long axis is oriented in a northeast to southwest direction; the building is not shaded by surrounding structures. An unconditioned parking area is provided under the facility for 82 cars.

One distinguishing feature of the building is a large, centralized, three-story atrium in the southwestern half of the building that provides access to the surrounding classrooms and offices. About 2,500 ft^2 (232 m^2) of northeast-facing clerestory windows help to illuminate the central staircase and computing facility.

The building can be characterized as an internal load dominated, high mass structure with 6-in. (152 mm) concrete floors and insulated concrete walls. Only about 9% of the exterior envelope is glazed. This consists of about 2,500 ft^2 (232 m^2) of single-pane clerestory lighting and about 9,000 ft^2 (836 m^2) of single-pane windows that are set back 3 ft (1 m) between exterior concrete utility chaseways.

The EC has a maximum occupancy of 2,300, which occurs during peak periods each semester. The occupancy profiles are characterized by an 8 am to 7 pm weekday schedule. Significant evening use of the building occurs during the weekdays between 7 pm and midnight. Weekend use is moderate.

The building's average air-change energy usage is strongly driven by the 10% to 20% fresh air that is supplied to the air handling units (AHUs). Internal lighting loads (2 W/ft^2; 20 W/m^2) and equipment loads (2.4 W/ft^2; 24 W/m^2) peak during the weekdays in the early afternoon. Considerable electricity is consumed in the evenings by the central computing facility.

Prior to the March 1991 VAV system retrofit, the 12 AHUs in the EC operated continuously. A detailed description of the

Continued on page 76

Continued from page 75

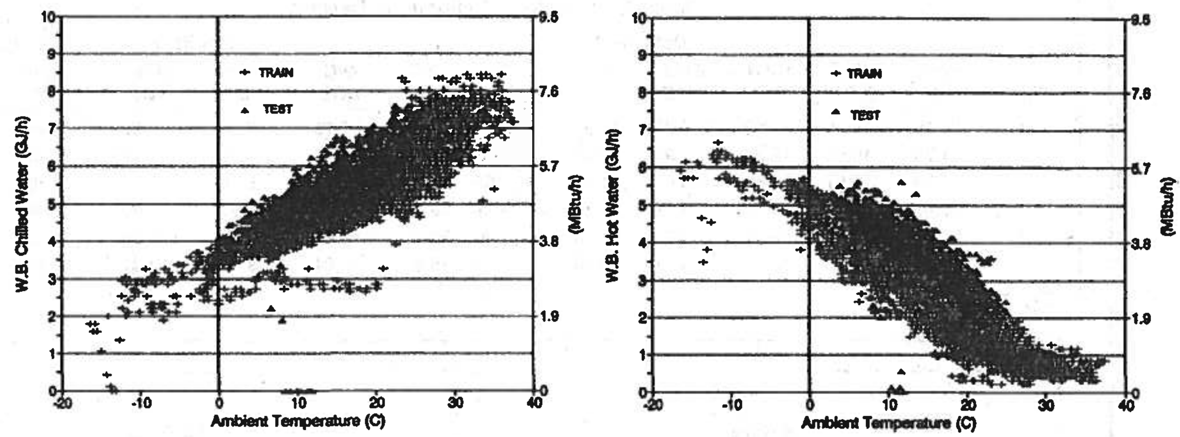


Figure 2. Plots of the chilled water and hot water data for the training and testing data sets, showing cooling and heating energy versus outdoor drybulb temperature.

HVAC system is provided by Katipamula and Claridge.[5] Additional information about the building can be found in Bronson,[6,7] Haberl,[8,9] Hinchey,[10] and Katipamula and Haberl.[11]

Time series data set for the Predictor Shootout. The data that were provided to the contest participants represent measured energy use and environmental conditions for the engineering center from September 1, 1989 to February 28, 1990, as shown in *Figure 1* and *Figure 2.*

The training set consisted of whole-building electricity, chilled water, hot water, environmental data (ambient temperature, absolute humidity ratio, wind and horizontal insolation) and an hourly and daily time stamp for the four-month period beginning September 1, 1989 and ending December 31, 1989.

The testing data set provided to contestants began January 1, 1990 and ended on February 28, 1990. This data set consisted of only the environmental data (no consumption data were provided) and time stamp. The contestants were then asked to predict the three energy end-uses for the test period using models that were developed from the training data set. Copies of these predictions were submitted for comparison with the actual energy use data for the two-month testing period.

Figure 1 shows the diurnal and weekly energy usage patterns. Because the EC is located on a university campus, it observes the same holidays as the university. These included a Thanksgiving holiday (November 23-26, 1989) and an extended Christmas holiday (December 21, 1989 through January 1, 1990). Both periods can be clearly seen in the data. The week prior to the Christmas holiday and the two weeks following the holiday in January have lower than normal consumption because this is the period between the fall and spring semesters when the building was sparsely occupied.

Figure 2 contains graphs of hourly heating and cooling energy consumption versus average outdoor dry bulb temperature. This type of plot is often used by building energy analysts because temperature is one of the most important parameters influencing building HVAC energy use.

The training data set also contained a period of abnormal consumption that was not revealed to the contestants. Such abnormal consumption periods are typical in large buildings; this represented an additional challenge to the modelers. The decision as to whether or not to include these periods in the model was left to the discretion of each analyst.

The period of abnormal consumption is December 22-23, 1989, as indicated by the triangles on the scatter plots in *Figure 1* and can be seen as the points that lie below the large cluster of chilled water data.

These data represent conditions that occurred after an extreme cold spell was encountered on the campus; many chilled water lines froze and ruptured during this period. This caused a large decrease in the amount of chilled water that was available for building cooling; hence the low chilled water data outliers. The hot water use for this period unexpectedly increased even though one might have expected the opposite because less chilled water was available to the constant volume air handlers.

The average hourly electricity use in the EC testing set is roughly 100 kW less than the electricity in the corresponding training set. This is because several academic departments (computer science and aerospace engineering) moved out of their offices in the engineering center and into newly constructed facilities in another building.

These vacant offices were then filled by faculty and staff from the other academic departments that existed within the building. This had the effect of reducing both electricity use and chilled water use, while slightly increasing the hot water consumption to make up for the lower internal heating loads, as shown in *Figure 2.* This fact was also not provided to the contestants.

Competition procedures

For the building data set, each contestant submitted hourly predictions (forecasts) for chilled water, hot water and whole-building electricity use for the two months following the four-month training set. Each contestant also prepared graphs similar to *Figure 1* and *Figure 2* for their predictions (both for training and testing sets) as well as a scatter plot of predictions versus actual values for the training set.

The seminar organizers evaluated all submissions using identical software and identical procedures for all entrants.

	Data Set A								Data Set B		Overall Results	
Winner	WBE CV	WBE MBE	CHW CV	CHW MBE	HW CV	HW MBE	AVG CV	AVG MGE	MPA CV	MPA MBE	GLOB CV	GLOB MBE
1	10.36	8.06	13.02	−6.37	15.24	−5.84	12.87	6.75	3.20	0.32	10.46	5.15
2	11.78	10.50	12.97	−5.95	30.63	−27.33	18.46	14.59	2.75	0.17	14.53	10.99
3	11.89	8.01	13.69	−6.67	31.65	−27.55	19.08	14.08	8.16	−0.15	16.35	10.59
4	16.95	6.20	14.32	−8.25	29.75	−26.19	20.34	13.55	4.91	−0.05	16.48	10.17
5	16.10	12.56	18.06	−9.79	28.08	−21.26	20.75	14.54	3.98	0.20	16.55	10.95
6	12.79	7.33	12.78	−5.31	30.98	−27.10	18.85	13.24	9.78	−0.31	16.58	10.01

Table 1. Shootout Competition Results*

*Values in percent.

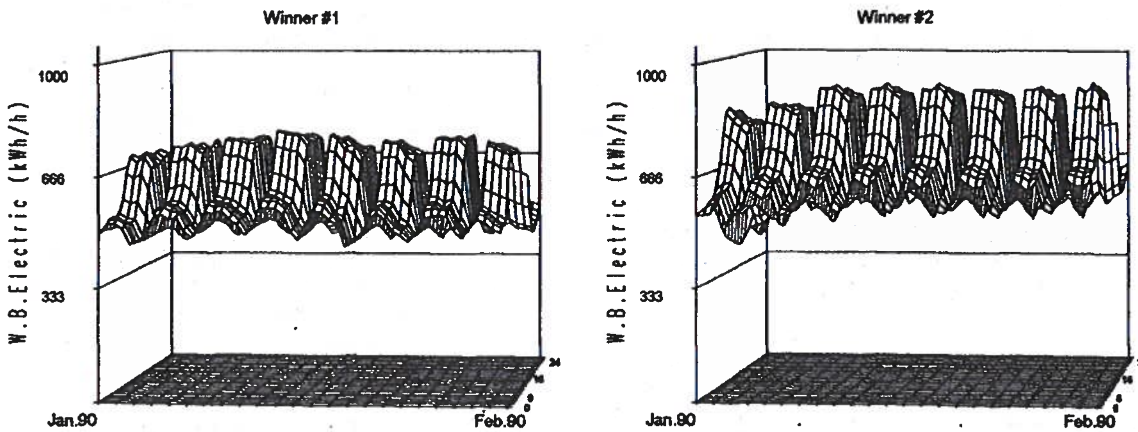


Figure 3. 3-D profiles of the predicted whole-building electricity use by winners 1 and 2. The actual data for this period are shown in the upper right plot of *Figure 1.*

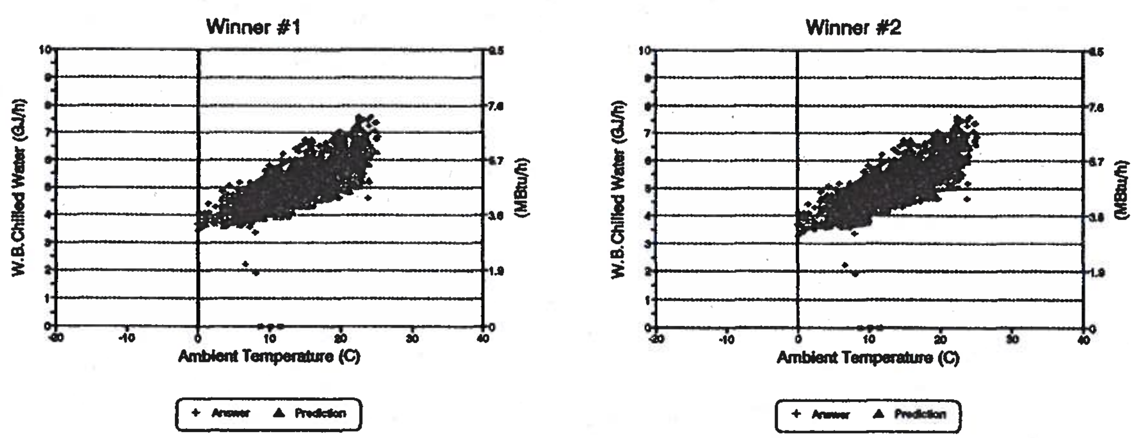


Figure 4. Chilled water predictions for winners 1 and 2. The + (plus) symbols represent the actual data for the test period. The triangle symbols are the predictions.

Continued on page 79

Predicting Hourly Building Energy Usage

Continued from page 77

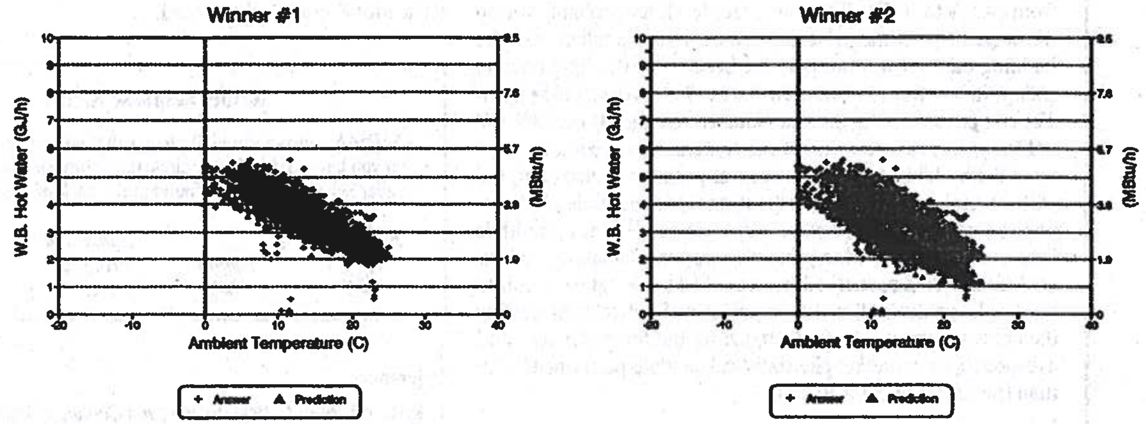


Figure 5. Hot water predictions for winners 1 and 2. The + (plus) symbols represent the actual data for the test period. The triangle symbols are the predictions.

The performance ranking was based on the standard statistical measures of coefficient of variation and mean bias error. (The mean bias error was accorded a secondary status; it would be used as a tie-breaker only if two contestants produced predictions with the same coefficient of variation, the primary ranking statistic.)

Results and summary

The results produced by the competitors were the predictions of the dependent variables for the two testing sets of independent variables. This section summarizes the performance of the top entrants.

Table 1 summarizes the results of the top six entrants. The method used by the Number 1 winner was Bayesian non-linear modeling. The Number 2 winner used a feedforward multilayer perceptron method. The fine differences in the predictions by the top two contestants can be seen by examining *Figures 1* through *5*.

In the upper right plot in *Figure 1*, a 3-D profile of the measured electricity use is presented for the EC testing set. In this 3-D profile, the x-axis contains the 59 days from January 1, 1990 through February 28, 1990. The y-axis (into the page) is the 24 hour profile of a day's electricity use; the z-axis (or the height above the x-y plane) is the electricity use. Similar 3-D profiles are provided for the two contestants' predictions in *Figure 3*.

Several features are worth noting. First, none of the contestants was able to accurately predict the first two weeks of electricity use during the testing period. This is to be expected because this represents the time between semesters when the building is sparsely occupied as the faculty, staff and students return from the holidays.

Second, there were no holidays in the EC testing data period; this was not communicated to the contestants. Possible holidays during this period could have included: Martin Luther King day, Washington's birthday, Lincoln's birthday, President's day and Ash Wednesday.

Finally, comparisons between the contestants' predictions reveal subtle differences as well. For example, the exceptional performance of the Number 1 winner (coefficient of variation, CV = 10.36%) can be clearly seen at a glance by comparing the testing data and the predictions in this 3-D presentation format. Winner Number 2 (CV = 11.78%) also had a very good whole-building electricity profile prediction.

Differences in the chilled water and hot water predictions are apparent from comparing *Figure 4* and *Figure 5* to *Figure 2*. However, they are slightly more difficult to detect. For example, in *Figure 4*, all predictions seem to have the appropriate temperature dependence in their chilled water predictions. However, differences seem to occur in the variation of the chilled water prediction for a given ambient temperature. This is to be expected in a building that has varying hourly load profiles. The contestants had visibly different predictions, although not necessarily much different CV values.

The plots of hot water predictions reveal features similar to the chilled water predictions with one exception. The Number 1 winner had a remarkably good fit for the hot water prediction (CV = 15.24%). The next best performance was the Number 5 winner (CV = 28.08%).

The remarkable accuracy of Number 1's prediction is borne out by the fact that the measured hot water for the testing period actually did increase when compared to hot water use for the training period, as shown in *Figure 2*. Again, this is thought to be due to the fact that several academic departments moved out of the building during the holiday break.

The most noteworthy feature of this event was the removal of the computer science department's computers. This had the effect of reducing the whole-building electricity, thereby increasing hot water use (to make up for the reduced heating produced by computer generated heat).

The most striking aspect of the competition (aside from the remarkable accuracy that the winners achieved) is that connectionist approaches were used in some form by all winners. Neural networks of various designs and training methods excelled at both problems (i.e., predicting the EC data and the MPA data). Traditional methods, at least as used by the entrants in this competition, were less accurate.

The overall winner had a combined CV of 10.5% for both problems. The average accuracies of the other contestants varied from 14.5% to 16.5%. These accuracy levels are probably within the range of experimental error because perfect prediction of the building data set was not possible because of the fundamental change in building use that occurred in 1990, as described earlier. The best predictions for the solar data set were slightly over 3% CV (RMSE), also near the limit of instrumentation accuracy.

The building data set for the test period was also analyzed with several traditional statistical analysis methods to determine how well these methods could perform.[2,12] Linear, multiple linear, change-point linear, principle component analysis, Fourier analysis and binning methods were applied to the data with mixed results. In general, all methods performed better at predicting the chilled water use than at predicting the hot water use, and as expected, the more complex statistical methods performed better than the simple linear methods.

Conclusion

A challenging set of building energy analysis problems was posed to the world-wide data analysis community. The top submittals to the competition were remarkably accurate given the nature of the data and the severe constraints placed on the contestants.

It appears that both decent predictions of hourly building performance and prediction of beam radiation from MPA devices, based on carefully measured past history, are now possible. However, to make the most accurate predictions, traditional statistical approaches may need to yield to novel methods such as neural networks that have not often been used by building energy analysts.

Acknowledgments

The authors express their recognition to the more than 150 people who requested the contest data from around the world. The authors would also like to express their gratitude to the six winners of the Energy Predictor Shootout:

- Winner 1, David Mackay, Cavendish Laboratory, Cambridge, United Kingdom;
- Winner 2, Mattias Ohlsson and associates, University of Lund, Lund, Sweden;
- Winner 3, Bradley P. Feuston and J.H. Thurtell, Mobil Research and Development Corp., Princeton, New Jersey;
- Winner 4, William Stevenson, Pretoria, Republic of South Africa;
- Winner 5, Makato Iijima and Ryo Takeuchi, Waseda University, Tokyo, Japan; and
- Winner 6, Minoru Kawashima, University of Wisconsin, Madison, Wisconsin.

The details of each contestant's approach are contained in their respective papers, which are listed in the Bibliography.

The authors also wish to thank the Centre d'Energetique of the Ecole des Mines de Paris, Denis Clodic, Peter Curtiss, Mike Mozer, Andreas Weigend, Marian Clark, David Claridge, Kelly Kissock, Dennis O'Neal, John Bryant, Sabaratnam Thamilseran and Mustafa Abbas.

Portions of this work were funded by the Joint Center for Energy Management of the University of Colorado, by the Texas State Energy Conservation Office (as part of Texas A&M's LoanSTAR Monitoring and Analysis contract), and by the Texas Higher Education Coordinating Board. ■

References

1. Kreider, J. Rabl, A. 1994. *Heating and Cooling of Buildings.* New York, New York: McGraw-Hill.

2. Kreider, J., Haberl, J. 1994. "Predicting hourly building energy usage: The great energy predictor shootout—Overview and discussion of results." Accepted by *ASHRAE Transactions.* Atlanta, Georgia: ASHRAE.

3. Curtiss, P. 1993. "An analysis of methods for deriving the constituent insolation components from multipyranometer array measurements." Transactions of the ASME, *Journal of Solar Energy Engineering*. New York, New York: American Society of Mechanical Engineers.

4. Curtiss, P., *et al.* 1994. "Applications of neural networks to multipyranometer array data." Submitted to Transactions of the ASME, *Journal of Solar Energy Engineering*. New York, New York: American Society of Mechanical Engineers. Vol. 115, No.1, pp.11-21.

5. Katipamula, S., Claridge, D. 1992. "Monitoring air handler performance and comparing it with a simplified system model." *ASHRAE Transactions.* Atlanta, Georgia: ASHRAE. Vol. 98, Pt. 2.

6. Bronson, D., *et al.* 1992. "A procedure for calibrating the DOE-2 simulation program to non-weather dependent loads." *ASHRAE Transactions.* Atlanta, Georgia: ASHRAE. Vol. 98, Pt. 1, pp. 636-652.

7. Bronson, D. 1992. *Calibrated Computer Simulations for the Analysis of Retrofit Energy Savings.* College Station, Texas: Energy Systems Laboratory, Texas A&M University. Master's thesis. January.

8. Haberl, J., *et al.* 1993. "Graphical tools to help calibrate the DOE-2 simulation program." *ASHRAE Journal.* Atlanta, Georgia: ASHRAE. Vol. 35, No. 1, January, pp. 27-32.

9. Haberl, J., *et al.* 1993. *A Report on the Impact of Using Measured Weather Data Versus TMY Weather Data in a DOE-2 Simulation of an Existing Building in Central Texas.* Energy Systems Laboratory report no. ESL-TR-93/09-02. College Station, Texas: Texas A&M University. September.

10. Hinchey, S. 1991. *Influence of Thermal Zone Assumptions on DOE-2 Energy Use Estimations of a Commercial Building.* College Station, Texas: Energy Systems Laboratory, Texas A&M University. Master's thesis. June.

11. Katipamula, S., Haberl, J. 1992. "A methodology to identify diurnal load shapes for non-weather dependent electric end-uses." *Solar Engineering, 1991: Proceedings of the ASME-JSES-JSME International Solar Energy Conference.* Reno, Nevada. March. pp. 457-467.

12. Kissock, K., *et al.* 1993. *Summary of Shootout Modeling Techniques*. Energy Systems Laboratory report no. ESL-PA-93/04-11. College Station, Texas: Texas A&M University. April.

Bibliography

Claridge, D., *et al.* 1991. "Improving energy conservation retrofits with measured results." *ASHRAE Journal*. Atlanta, Georgia: ASHRAE. Vol. 33, No. 10, October.

Feuston, B., Thurtell, J. 1994. "Generalized non-linear regression with ensemble of neural nets: The great energy predictor shootout." In preparation for *ASHRAE Transactions*. Atlanta, Georgia: ASHRAE.

Iijima, M., *et al.* 1994. "A piecewise-linear regression on the ASHRAE time series data." In preparation for *ASHRAE Transactions*. Atlanta, Georgia: ASHRAE.

Kawashima, M. 1994. "Artificial neural network backpropagation model with three-phase annealing developed for the building energy predictor shootout." Submitted to *ASHRAE Transactions*. Atlanta, Georgia: ASHRAE.

Mackay, D. 1994. "Bayesian non-linear modeling for the energy prediction competition." In preparation for *ASHRAE Transactions*. Atlanta, Georgia: ASHRAE.

Ohlsson, M., *et al.* 1994. "Predicting utility loads with artificial neural networks—Methods and results from the great energy predictor shootout." In preparation for *ASHRAE Transactions*. Atlanta, Georgia: ASHRAE.

Stevenson, W. 1994. "Predicting building energy parameters using artificial neural nets." In preparation for *ASHRAE Transactions*. Atlanta, Georgia: ASHRAE.