ANALYSIS OF TEXAS LOANSTAR DATA

**David E. Claridge, Jeff S. Haberl, Srinivas Katipamula, Dennis L. O'Neal, David Ruch,
Lu Chen, Tia Heneghan, Sharon Hinchey, Kelly Kissock, Jinrong Wang
Energy Systems Laboratory, Mechanical Engineering Department
Texas A&M University, College Station, Texas**

ABSTRACT

The analysis effort during the first year of the Texas LoanSTAR Monitoring and Analysis Program has emphasized selection and development of baseline analysis techniques to cover the range of buildings expected in the program. PRISM has been adopted as the baseline technique for buildings which are appropriate for treatment with one-. three- and five- parameter segmented linear, change-point models. In addition to PRISM, two- and four- parameter linear, segmented change-point models are expected to be suitable for at least preliminary analysis of monthly and daily data for all buildings in the program. Regression analysis with hourly scheduling profiles will be used for baseline analysis of hourly data.

Substantial effort has been devoted to exploratory analysis intended to refine the analysis performed with the baseline techniques. Work to date has centered on investigation of Principal Component Analysis, an improved goodness-of-fit indicator for n-parameter change-point models, and calibrated simulation modeling.

Data from five buildings is used to explain and illustrate the baseline analysis techniques and the exploratory work conducted.

OVERVIEW

The Analysis and Planning Task is responsible for selecting and developing analysis techniques, developing the necessary software and analyzing collected data to:

(1) determine the energy and dollar savings of the retrofits;
(2) reduce energy costs by identifying operational and maintenance improvements at retrofitted facilities;
(3) identify the savings of individual retrofits, as feasible, to help improve retrofit selection in future rounds of the LoanSTAR Program;
(4) initiate an end-use database of energy use for commercial/institutional buildings located in Texas.

The task is also responsible for coordinating the preparation and updates of the overall plan for the Monitoring and Analysis Program, an effort not described in this paper.

Data analyses will be performed in several phases for each monitored site. These include:

- verification/modification of audit assumptions
- pre-retrofit analysis
- preliminary post-retrofit analysis
- detailed post-retrofit analysis
- interaction and feedback to agencies and operators
- reports

The three different types of models selected and implemented during the first year are described in this paper, followed by a summary of the exploratory analysis performed to identify and develop improved analysis techniques.

ANALYSIS STRUCTURE

The procedures which are now sufficiently tested for use as baseline analysis techniques are:

PRISM analysis
Analysis with Hourly Scheduled Loads

It is expected that Extended Change-Point Regressions will soon be suitable for baseline application. Exploratory analysis is being conducted to identify and develop techniques which extend and refine the capabilities of the baseline analysis.

PRISM

The versions of PRISM (Fels, 1986) which are available or under development include one-, three- and five- parameter segmented regressions with change points as shown in Figure 1.

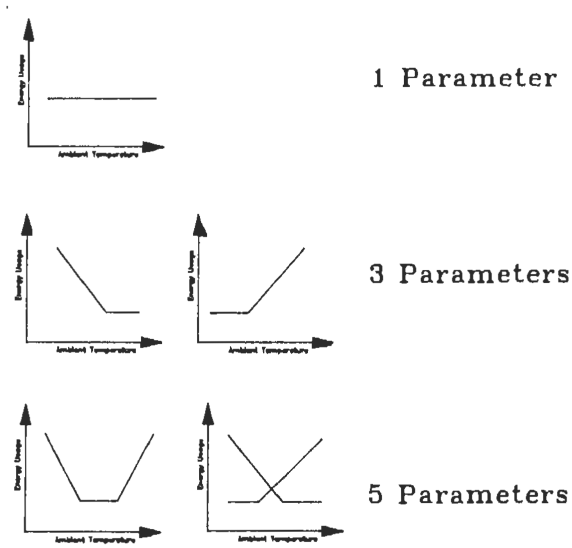

Figure 1. One-, three-, and five- parameter models of energy use as implemented in PRISM.

The one-parameter model is typical of monthly electrical use when heating and cooling influences are absent . It is also typical of sub-metered daily electrical data from many buildings expected in the program after sorting into weekday/weekend data since heating and cooling is often supplied by a central system.

The three-parameter models represent the classic PRISM Heating Only and PRISM Cooling Only models and have been used with some success on three of the five buildings for which data are currently available.

The five- parameter model represents the PRISM Heating and Cooling model (operational at Princeton) and will provide a better model for two of the current buildings.

Analysis with Hourly Scheduled Loads

Hourly data are very useful for identifying scheduling changes, equipment failures, etc. This is particularly true when energy use is predicted with a model which includes a basic schedule of the electricity use plus a temperature regression, if appropriate. Hourly data, when compared with hourly predictions, can be conveniently displayed as a residual plot. This approach will be used routinely on buildings for which hourly consumption data are available.

Extended Change-Point Regressions

The two-parameter and four-parameter change-point models shown in Figure 2 represent additional change-point types expected to be highly useful, based on examination of preliminary data.

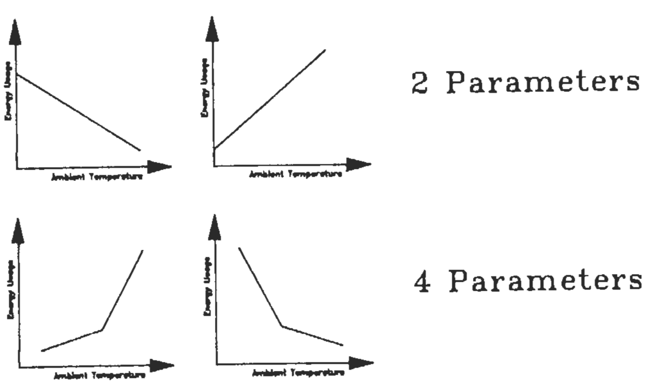

Figure 2. Two- and four- parameter models of energy use described as "extended change-point regressions" in this report.

The two-parameter model can be used to represent the hot water or chilled water consumption in buildings with conventional reheat systems, such as a large laboratory/classroom/office building to be discussed in a later section.

The four-parameter model provides a better fit to consumption than the three-parameter model for some buildings which exhibit a change point, but show a non-zero slope on both sides of the change point. This will be illustrated by electrical data from a grocery store in College Station, Texas.

Exploratory Analysis

Exploratory analysis is currently underway in three areas:

Principal Component Analysis
Improved Measures for Goodness-of-Fit
Calibrated Simulation Models

The authors have concluded that there is such strong intercorrelation between daily influencing parameters that simple multiple linear regression may be of limited value in developing predictive models because the intercorrelated parameters overdetermine the prediction. Consequently, the effectiveness of Principal Component Analysis (PCA) is being examined for this purpose. PCA is attractive because it uses the influencing parameters to create orthogonal (independent) "components" which may be interpreted to have physical significance. Other techniques which will be investigated include: 1) change-point PCA models; and 2) the use of "switching" models to isolate the influence of individual influencing parameters, or scheduling.

Goodness-of-fit has generally been evaluated primarily in terms of the correlation coefficient, R^2, when regression has been used in energy analysis. This is generally a useful measure when the model involves a single slope (e.g. PRISM CO or HO). A different method is needed when a four-parameter change-point model is used. The use of an approach which minimizes the variance for the entire data set to locate the change point for a four parameter model is currently being investigated.

Calibrated simulation models also hold promise for use in determining retrofit energy savings. A procedure which can be used to create a calibrated DOE-2 input deck has been developed and used to estimate the savings of a planned VAV retrofit.

DATA AVAILABLE

Data have been available for analysis from only one building, the Zachry Engineering Center at Texas A&M University, which has been audited and is scheduled for a retrofit under the LoanSTAR Program. Consequently, data from four other buildings have been analyzed and are presented in this paper to illustrate the techniques described.

DATA ANALYSIS AND PRE-RETROFIT MODELS

Zachry Engineering Center

The Zachry Engineering Center (ZEC) is a four-story (plus basement parking level) building on the A&M campus with approximately 324,400 gross square feet of floor area. It is a heavy structure with 6-inch concrete floors and is heated and cooled by a constant volume dual duct system. Hot water and chilled water are supplied by the central campus plant. Major uses of the building include: 1) offices, 2) class rooms, 3) computer rooms, and 4) laboratories. The building also includes hallways and a large atrium area which serves as a common space.

Electrical Consumption

The electrical consumption for ZEC from July 1989 through May 1990 is shown in Figure 3. The figure shows hours of the day from front-to-back, Julian day of the year from right-to-left and hourly electricity use on the vertical axis. The building is open seven days a week, 24 hours a day, and the HVAC systems are operated continuously. The electrical consumption shows a diurnal pattern which varies from a minimum level near 1 MW to a peak of 1.5 MW on weekdays with a slightly lower minimum and much lower peak on weekends. Some gross characteristics of the data are evident in the figure. Proceeding from right to left, consumption is seen to be lower during the break period just before Autumn Semester begins. Christmas vacation period is very evident as the "canyon" near the middle of the figure. The other "canyons" in the left half of the figure represent missing data which occurred when a technician fried an IC in the data logger.

The data were used to define an hourly schedule for weekdays and for weekends when school is in session as shown in Figure 4(b). This may be compared with the measured consumption for February 1990, shown in Fig. 4(a). The positive residuals and absolute values of the negative residuals are shown in Figs 4(c) and 4(d). The residual plots indicate that : 1) the electrical use is generally well-described by this simple model (+/- 100 KW out of 1500 KW); 2) consistent underuse of electricity can be seen on Friday afternoons (days 32,39,46 and 53 of Fig. 4(d)); and Saturday consumption is sometimes higher than expected (days 40 and 54 of Fig. 4(c)).

Chilled Water Consumption

The chilled water consumption for ZEC depends primarily on the ambient temperature as can be seen in Figure 5. There appears to be a slight difference between weekdays and weekends - physically we expect this difference to be due to the lower electrical consumption on weekends. Table 1 shows our pre-retrofit model for chilled water consumption in the ZEC. determined using SAS. (SAS 1985) Models are shown which depend on temperature (T) only and which depend on temperature, T, as well as electrical consumption for lights and equipment, LE. The second model explains slightly more of the chilled water consumption, though the difference is not statistically significant.

Note that the chilled water consumption does not show a change-point. It simply decreases as temperature decreases. It might show a change point at sufficiently low temperatures, but the available data includes some of the coldest weather ever experienced in College Station, so the two- parameter model

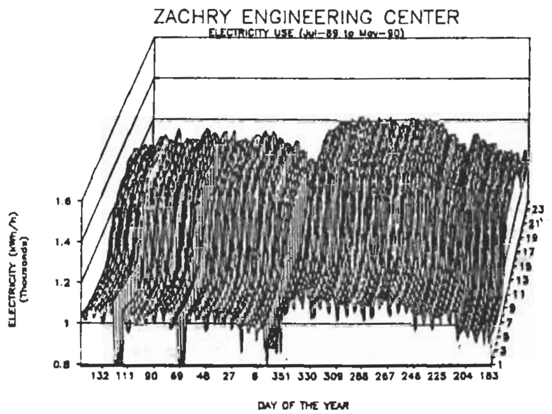

Figure 3. Hourly electricity consumption for Zachry Engineering Center from July 1989 through May 1990.

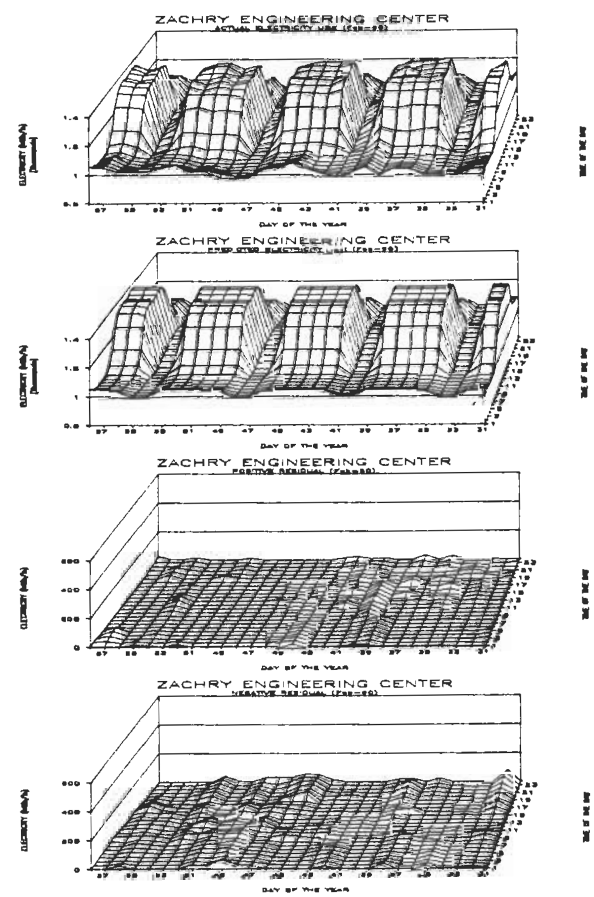

Figure 4. Zachry Engineering Center electricity use and residuals for February, 1990. (a) Measured electricity use. (b) Scheduling model of electricity use. (c) Positive residuals (measured minus predicted) of electricity use. (d) Absolute values of negative residuals of electricity use.

without a change point seems most appropriate. The relationship between chilled water consumption and electricity consumption will also be examined using a switching model to separate days when one chilled water pump is used instead of two pumps and possibly using a lagged hourly model.

Hot Water Consumption

The hot water consumption is similar to the chilled water consumption, except that the temperature dependence is negative as would be expected. This behavior is shown in Figure 6. The data again appears to exhibit a slight dependence on the electrical consumption for lights and equipment as shown in Table 2. In this case the dependence on electrical use is somewhat stronger than for chilled water, and is statistically significant, according to the SAS analysis.

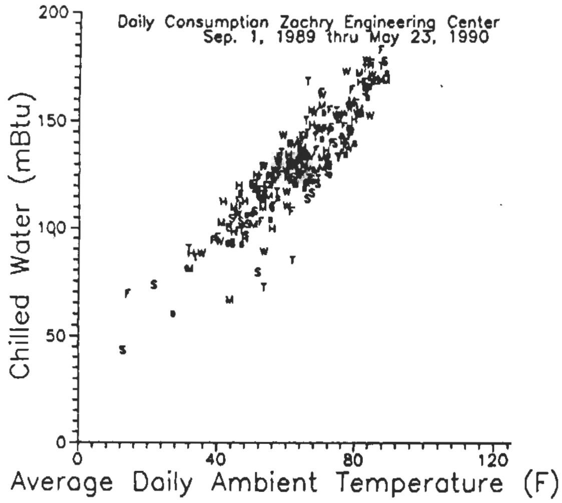

Figure 5. Chilled water consumption at Zachry Engineering Center plotted as a function of ambient temperature using data for September 1, 1989 through May 23, 1990.

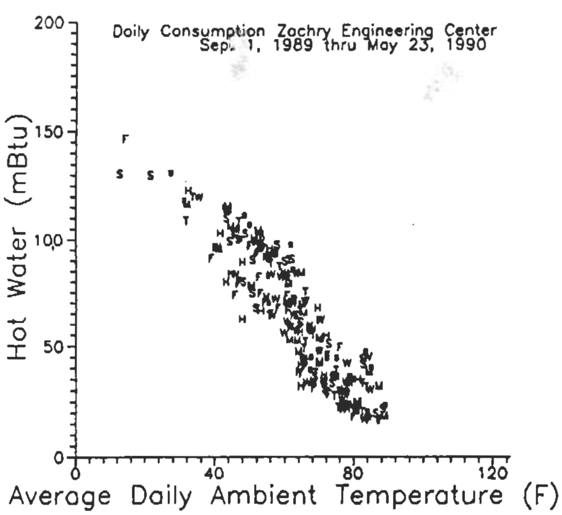

Figure 6. Hot water consumption for Zachry Engineering Center as a function of ambient temperature for September 1, 1989 through May 23, 1990.

	$CW = \alpha + \beta T$	$CW = \alpha + \beta_t + \beta_e LE$
R^2	0.86	0.87
Intercept	221.8	35.5
Slope	16.4	β_t=16.2; β_e=0.01
Statistical Significance	good	β_t = good; β_e = fair

Table 1 - Model parameters and statistics for two models of chilled water consumption at Zachry Engineering Center.

	$HW = \alpha + \beta T$	$HW = \alpha + \beta_t + \beta_e LE$
R^2	0.87	0.90
Intercept	1808.8	2178.3
Slope	-19.5	β_t = 19.2; β_e= 0.03
Statistical Significance	good	β_t = good; β_e = fair

Table 2 - Model parameters and statistics for two models of hot water consumption at Zachry Engineering Center.

A&M Consolidated High School

A&M Consolidated High School is a 209,605 sq.ft. two-story facility in College Station. Electricity dominates energy cost at the school as shown in Table 3 below which presents data for October 1988 through September 1989:

Energy Source	Amount	Cost
Electricity		$189,841
Consumption	2,853,600 Kwh	(87,945)
Demand-peak mo.	1,112 KW	(101,896)
Natural Gas	23,343 CCF	$10,439

Table 3 - Energy consumption and cost at A&M Consolidated High School for October 1988 through September 1989.

It is estimated that approximately 20 percent of the natural gas is used for water heating with the remainder used for space heating. An ASEAM (ACEC 1987) simulation of the building provides a breakdown of electricity use as shown in Table 4.

End Use	Consumption
Heating	3.1%
Cooling	36.7%
Lighting	16.5%
Fans	15.0%
Misc. Equipment	25.8%
Pumps	3.0%

Table 4 - Estimated high school electricity consumption by end use.

The electricity use of the building can best be described as erratic. The consumption is relatively constant and changes appear to be dominated more by scheduling and operation than by ambient temperature. PRISM analysis of the data is consistent with these conclusions. Analysis of monthly data for three different 12-month periods produced R^2 values of 0.35, 0.71 and 0.29. Analysis of the daily data after sorting data into weekday/weekend subsets produced R^2 values between 0.10 and 0.25.

At least part of the reason for this irregular behavior is evident in Figure 7 which plots the hourly electric consumption. Night shut-off of fans and HVAC systems has been very erratic as evidenced by early morning consumption in the 200 Kw range for much of the Autumn of 1989. Consumption during much of February shows abnormally high night-time use and low daytime use during the school year suggesting the possibility of questionable data. A dramatic drop in both nighttime and daytime consumption over the Christmas break period is clearly evident.

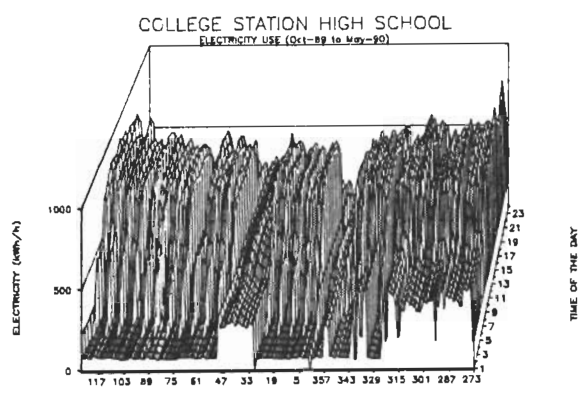

Figure 7. Hourly electricity consumption at A&M Consolidated High School for October 1989 through May 1990.

Further analysis of the data using daily and hourly totals for grouped data periods will be used in an attempt to better model and understand consumption at the high school.

Grocery Store

The grocery store analyzed is a 40,000 square-foot supermarket located in College Station, Texas. The store has the electrical energy-using systems typical of a modern supermarket. The systems and their estimated contribution to peak electrical demand are: refrigeration cases and compressors (44.3%), air conditioning (24.6%), lighting (15.8%), food preparation (12.6%), point-of-sale registers (1.2%), and miscellaneous uses (1.5%). Almost all of the heating for the store is provided by heat reclaim from the refrigeration compressors. Natural gas is used for an oven in the bakery, supplemental heating in very cold weather, and for a 40 gallon water heater. The natural gas cost is less than two percent of the electricity cost for the store, so it has been ignored in the model reported here. The store is open 24-hours per day and is closed only on Christmas and part of Thanksgiving day. Further details have been reported in Schrock and Claridge (1989).

Electrical Consumption Analysis

Electrical consumption is recorded at 15-minute intervals and read weekly via modem by the local utility. The 15-minute data has been aggregated to provide hourly and daily total consumption for analysis. The average consumption for each day, expressed in kWh/h or kW is shown in Figure 8 as a function of average daily ambient temperature. There appear to be two essentially linear regions of the data which meet at about 62 F (called the change point). Physically, the consumption appears to drop slowly with temperature (below 62 F) due to the increasing COP of the refrigeration compressors. As the temperature increases above 62 F, the COP of the compressors continues to drop, but air conditioning also becomes necessary, resulting in a sharp increase in the slope of the electrical consumption.

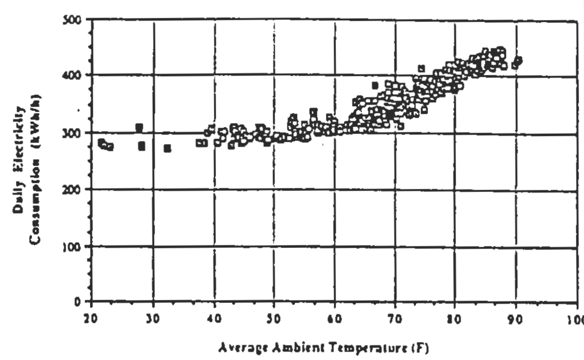

Figure 8. Average daily electricity use at the College Station Kroger Store for March 1988 through April 1989.

Consequently, the data were divided into that collected when the ambient temperature (as recorded at the local airport) was above 62 F and that collected when the temperature was at or below 62 F. Each set was then regressed against the dry bulb temperature to obtain a unique slope.

This process resulted in a four-parameter model for the daily average electric consumption: (1) a slope for the non-cooling regime, (2) a slope for the cooling regime, (3) a change-point temperature, and (4) a baseload consumption at the change point temperature. Hence the daily average electric consumption, E_d, can be expressed as:

$$E_d = E_{CP} + B_h*(T_d - T_c) \quad T_d <= 62 \text{ F}$$

$$E_d = E_{CP} + B_c*(T_d - T_c) \quad T_d > 62 \text{ F}$$

where T_d is the average daily temperature, E_{CP} is the electricity consumption at the change point of 62 F and B_h and B_c are the slope coefficients. The model parameters obtained are:

$$E_{CP} = 310 \text{ kWh}$$

$$B_h = 0.868 \text{ kWh/h/F}$$

$$B_c = 4.976 \text{ kWh/h/F}$$

The ability of the daily predictor model to estimate consumption is shown in Figure 9. The measured daily average consumption is shown for April 1 through April 29, 1989 by the asterisks near the top of the figure. The model prediction for each day is shown by the diamonds while the residual (measured minus predicted) consumption is shown by the line near the bottom of the figure. Excluding anomalies on April 4, 8 and 9, the average residual consumption is 8.6 kWh or 1.7% of the total. A more extensive discussion is provided in Schrock and Claridge (1989).

Nursing Homes

Two nursing homes in Austin and Temple are being analyzed as part of the ERAP Program. They are discussed here since electrical data from these buildings is suitable for analysis with PRISM.

The Temple facility is a 100 bed nursing home that was 80 per cent occupied in early 1990. Approximately 40 staff members are present during the day and about 20 during the night. The facility operates 24 hours per day, year round. Full food and laundry service are provided. The single story, slab on grade building was built in 1970 and has an approximate floor area of 31,000 square feet.

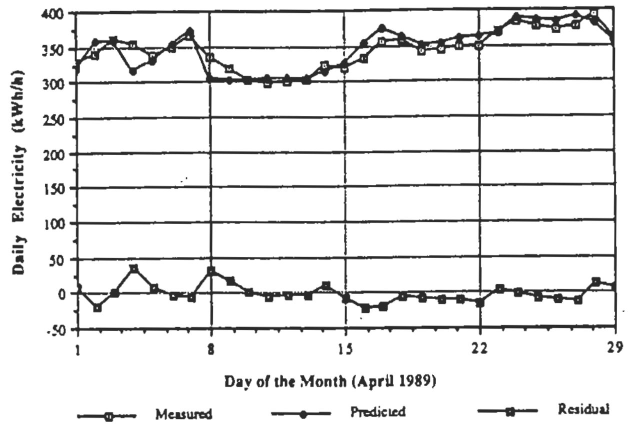

Figure 9. Measured, predicted and residual average daily electric use at the College Station Kroger Store for April 1989.

Space conditioning is provided by eight roof-top air conditioning units and six small heat pumps. Two of the roof-top units provide heat using electrical resistance heaters and six of the units use natural gas for heating. Direct expansion cooling is used by all units.

The Austin facility is nearly twice as large: it is a two-story building with approximately 58,000 square feet of floor area. Operational schedules and characteristics are similar. Space conditioning is primarily provided by 16 roof-top air conditioning units. Some small window air conditioners and heat pumps supplement the roof-top units. All roof top units provide electrical resistance heating and direct expansion cooling.

Electrical-Use Analysis

Energy use at both facilities has been analyzed using PRISM while that at Temple has also been examined using ASEAM and an equipment inventory with name plate data. The ASEAM simulation results in the following end-use estimates:

HVAC	50%
Kitchen & Laundry	24%
Lighting	26%

PRISM Analysis: The electricity billing data for both facilities are shown in Figure 10. Since some heating is present in both facilities, the PRISM cooling only model was used with winter data omitted. The Temple data provides $R^2 = 0.88$ with a cooling balance temperature of 66.8 F while the Austin data provides $R^2 = 0.96$ with a balance point of 74.2 F. Electricity consumption vs cooling degree days for these facilities is shown in Figure 11.

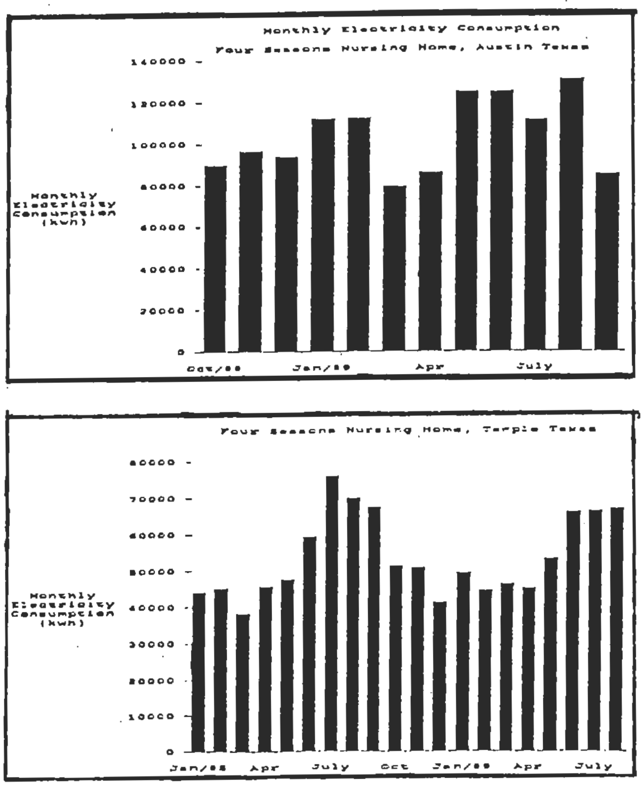

Figure 10. Monthly electrical consumption for Austin and Temple Four Seasons Nursing Homes.

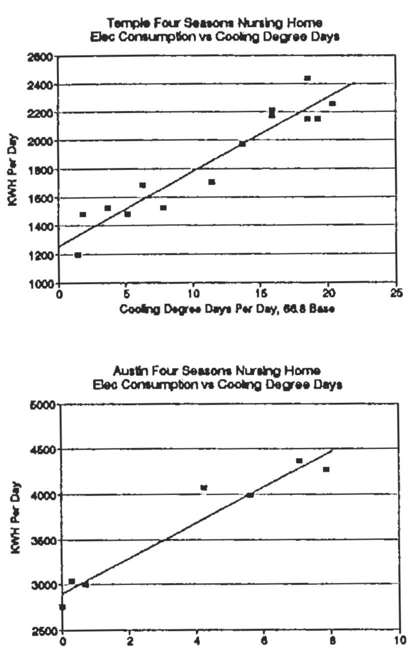

Figure 11. PRISM plots of monthly electricity use versus cooling degree-days for Austin and Temple Four Seasons Nursing Homes.

EXPLORATORY ANALYSIS

Principal Component Analysis (PCA)

Numerous investigators have attempted to use multiple linear regression analysis to develop improved models of building energy consumption. These attempts have often been frustrated by significant collinearity between the predictors used. While simulation models clearly separate the influences of temperature, solar gain, occupant gains, humidity, scheduling, etc. as inputs, the inverse problem of determining the influence of these and other factors is complicated by the level of diurnal correlation between such factors and even the annual correlation which often exists.

Principal Component Analysis has been used to tackle similar problems for some time by climatologists and more recently was used by Hadley and Tomich (1986) to examine influences on residential heating energy consumption. PCA forms orthogonal vectors as combinations of the "independent" variables which exhibit some collinearity. The data are then regressed using these orthogonal vectors or "principal components" as the new independent variables to obtain regression coefficients for each "principal component." The less significant "components" may then be dropped to obtain a set of components which provide more stable coefficients for the physical variables. This approach appears to hold promise as a way of combining physical models and insight with measured data to achieve improved models for determining retrofit savings.

The effort to date has been primarily directed toward developing software to implement PCA and beginning the exploratory analysis. The method has been applied to data for the grocery store for June 19, 1989 through June 19, 1990 considering temperature, specific humidity and global horizontal solar radiation as important influencing parameters. The daily energy use, E_p, predicted using standard regression is:

$$E_p = 2550 + 77.7*T + 19918*w + 1.56*I_h \quad T > 62\,F$$

$$E_p = 6016 + 20.5*T + 42501*w + 0.77*I_h \quad T <= 62\,F$$

where T and w are the daily average temperature and specific humidity respectively and I_h is the daily total horizontal solar radiation.

The model found using PCA and dropping the third component is:

$$E_p = 4179 + 46.7*T + 47216*w + 3.27*I_h \quad T > 62\,F$$

$$E_p = 6029 + 19.9*T + 44140*w + 0.82*I_h \quad T <= 62\,F$$

This process changed the model very little at low temperatures, but approximately doubles the apparent importance of solar and humidity at high temperatures. However, humidity and solar radiation each still account for only 8 percent of the electrical consumption on a typical June day.

Improved Goodness-of-Fit-Indicators

As noted above, the most appropriate model for the daily electric consumption for the grocery store is a four parameter, segmented linear model. The results shown in the figures are based on selection of the change-point by visual estimate. A more rigorous procedure has been developed which determines the change-point to minimize the variance with respect to the segmented model. The algorithm for calculating the optimal model is complete and a program which carries out the algorithm has been written. The user provides the data set and reasonable upper and lower bounds for the change-point temperature. The program will compute the models' parameters and various relevant statistics in less than one minute using several years of daily data.

Run on a full year of data (June 1989 - June 1990), the program yields a change-point temperature of 59.5 F. This may be compared with the visual estimate of 62 F; the difference of 2.5 F is significant and supports the need for a precise method for calculating the change-point temperature.

Calibrated Modelling

Calibrated input decks for simulation programs offer an approach for examining the expected sensitivity of individual measurement points to the retrofits installed. They also may provide improved estimates of the savings which should be expected from retrofits on individual buildings.

A procedure was defined and tested for preparing a first order calibrated input deck for DOE-2. This process identified several time-consuming software issues which had to be resolved before the procedure could be implemented. A ten-zone input deck of the ZEC was prepared for DOE-2 and important system parameters were adjusted until the measured hot water and chilled water consumption approximated the predicted consumption, using data taken with ambient temperatures between 65 F and 95 F. This input deck was then changed to incorporate a VAV system to provide an update on the expected retrofit savings. Savings indicated were 14,300 MMBtu hot water 30,500 MMBtu chilled water and 3,150,000 Kwh electricity. This model highlighted the need for better air handler data which is now being incorporated.

CONCLUSION

During the pilot year, PRISM and hourly scheduling models have been adopted as baseline analysis methods for the LoanSTAR Monitoring and Analysis Program. Extended change-point models have been identified as useful and will be adopted as baseline models as soon as adequate error diagnostics are developed. Exploratory analysis has emphasized work in three areas: (1) development of a method which locates a change-point temperature by minimizing the variance of a data set fitted by a segmented linear model; (2) use of Principal Component Analysis to provide stable regression coefficients for influencing parameters; and (3) development and investigation of calibrated DOE-2 models. These techniques have been investigated using data from five buildings for which data is currently available.

ACKNOWLEDGEMENTS

The support of this work by the Texas Governor's Energy Management Center and by the Texas Higher Education Coordinating Board under the Energy Research and Applications Program Project Number 227 is gratefully acknowledged. The GEMC is the sponsor for the LoanSTAR Monitoring and Analysis Project while the exploratory analysis of PCA and change-point methods and the monitoring and analysis of the grocery store and nursing homes are sponsored by the Coordinating Board.

REFERENCES

ACEC, 1987. *ASEAM 2.1: A Simplified Energy Analysis Method,* ACEC Research and Measurement Foundation, Washington, D. C.

Fels, M., 1986. "Special Issue Devoted to Measuring Energy Savings, The Scorekeeping Approach," *Energy and Buildings,* Vol. 9, Nos. 1 and 2.

Hadley, D. and Tomich, S., 1986. "Multivariate Statistical Assessment of Meteorological Influences on Residential Space Heating," *Proceedings of the 1986 ACEEE Summer Study,* vol. 3.

SAS, 1985. *SAS Version 5,* SAS Institute, Inc., Cary, North Carolina.

Schrock, D. and Claridge, D., 1989. "Predicting Energy Usage in a Supermarket," *Proceedings of the Sixth Annual Symposium on Improving Building Systems in Hot and Humid Climates,* Dallas Texas, October 3-4, 1989, pp.F19-F27.