# ISSUES IN LOAD SHAPE REPRESENTATION

Erik W. Pearson, Pacific Northwest Laboratory
Larry Palmiter, Ecotope

## ABSTRACT

Large volumes of high time resolution electrical end-use load data are becoming available. In raw time series form such data are cumbersome to manipulate and difficult to interpret. Thus, it is necessary to develop compact representations for such data. To be widely useful, these representations must incorporate dependence on independent variables which strongly influence load. In addition, they should provide useful visual display of the loads and be readily employed in computational work.

One approach to construction of useful representations is the generation of simple empirical models for the load date. In this paper, we investigate two techniques for generating such representations. First, we apply two-way analysis to develop a model of hourly total and refrigeration load data taken from a single residence. We then use empirical orthogonal function techniques to represent electrical consumption in a Seattle office building. The advantages and limitations of the resulting models are discussed.

# ISSUES IN LOAD SHAPE REPRESENTATION

Erik W. Pearson, Pacific Northwest Laboratory
Larry Palmiter, Ecotope

## INTRODUCTION

Electrical load data collected at high time resolution permit an analysis and interpretation of load patterns, and investigation of the dependency of loads on variables such as time of day or ambient temperature. However, a simple time series of load data is often a poor representation of the load. Such a representation is voluminous if a time series of any length has been collected and does not explicitly exhibit important dependencies. Figure 1 shows a time series of 1000 hourly total load values collected under the End-use Load and Conservation Assessment Program (ELCAP)[1] for a single residence. From this representation relatively few of the important characteristics of the load are obvious, and the predictive value of the representation is limited.

What constitutes a useful and practical representation of an electric load shape varies dramatically from application to application. However, there are a number of characteristics which a good representation should have: efficiency or compactness, facile interpretability, accuracy, and predictive power. One sensible approach to construction of such representations is to build empirical models from the data. Ideally, these will capture most of the systematic variation in load and retain the important dependencies on explanatory variables.

In this paper we illustrate this idea by generating two load shape descriptions. First, we construct a simple model based on an analysis of variance technique. The explicit dependence of the loads on independent variables, such as hour of day, day of week, and average daily temperature, are incorporated in simple functional forms. The model parameters will be few in number relative to the number of data points from which they are derived, and graphical or numerical representation of the model is facile. We apply this technique for analysis and representation of a one-year time series of residential total load data, and to the refrigeration end-use data for the same structure.

The second class of representations discussed here involves expression of the load as a sum over a set of orthogonal functions:

$$\underset{\sim}{L}(x_1,x_2,\ldots,x_n) = \sum_{i=1}^{n} w_i(x_1,x_2,\ldots,x_n)\underset{\sim}{\phi}_i \qquad (1)$$

Here the $X_j$ are explanatory variables (for instance, outside temperature or day of week), L is a vector of 24 hourly average loads, and the $\phi_i$ are a set of basic vectors each also of length 24. We obtain the basis functions for the expansion through a principal components analysis of the load data. The parameters in the functions incorporating the dependence on independent variables are then obtained through examination of the expansion of the individual

Work Supported by
the U.S. Department of Energy
under Contract DE-AC06-RLO 1830

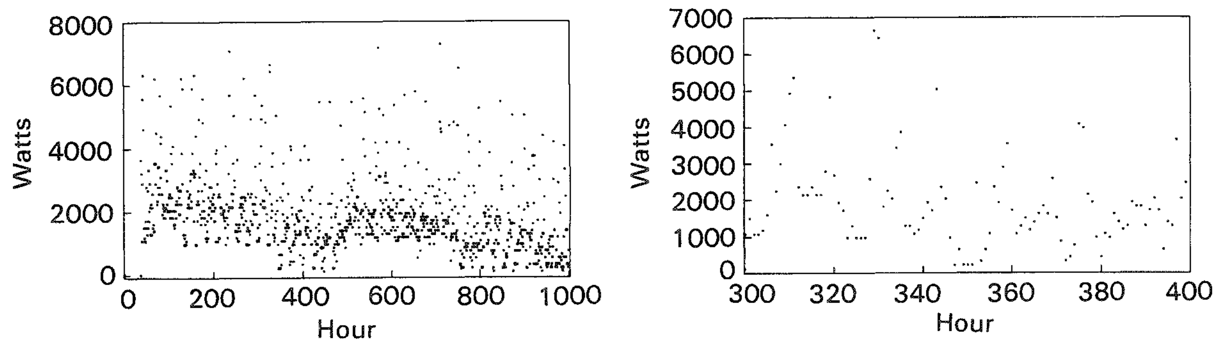


Figure 1. Hourly load data time series for a residence.

*The left panel shows the first thousand hours of total data from a year-long time series. The first few days had no data available. The right panel shows an excerpt of 100 hours from that data set. Although the diurnal pattern begins to appear on the expanded scale, this representation of the load data is clearly difficult to interpret and has little predictive power.*

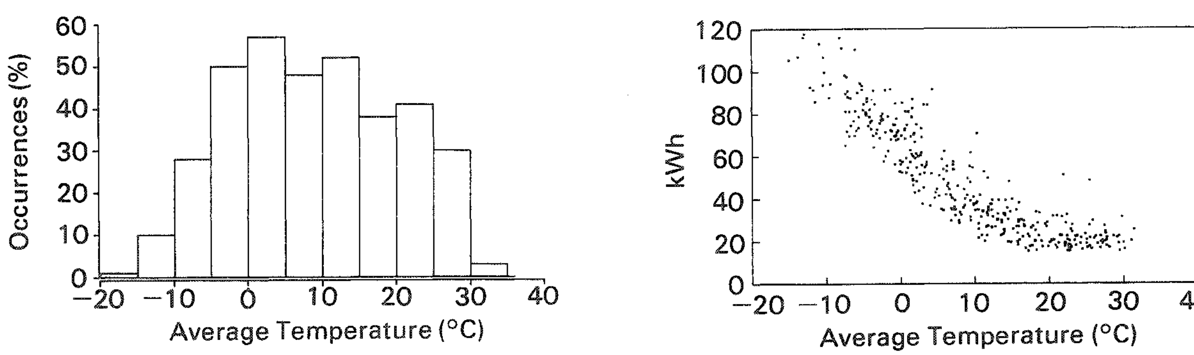


Figure 2. Distribution of temperatures and loads for a residence.

*The left panel of this figure is a histogram of daily average exterior temperature at the residence under study. The range of temperatures is fairly large, and includes both quite warm and quite cold days. The right panel is a scatter plot of daily load versus average daily exterior temperature. The pattern is typical of a residence with no air conditioning, and the data suggest that heating is required when the average exterior temperature drops below about 15° C.*

daily loads in terms of the basis functions. We illustrate this approach through an application to total load data from an office building, also collected under the ELCAP program.

The predictive power, analytic value, representational power and limitations of each method are discussed. It must be emphasized that the techniques applied here have been chosen as representative of broad classes of approach to electrical load data analysis and characterization. There is no intent to suggest that the specific tools presented here are uniquely suitable for work with electrical load data.

## MODELLING LOAD DATA WITH TWO-WAY TABLES

To illustrate the role of empirical models in load data representation, we investigate the use of two-way tables in analysis of hourly total load and refrigeration load for a single residence. We begin by discussing the construction and graphical representation of a two-way table, after which we apply the technique on an hour-by-hour basis to a year-long time series of residential data.

A two-way analysis is a procedure in which variability in the data is partitioned between two sources. Each of two independent variables is partitioned into several regions. Observations of the dependent variable or variables, in this case electrical load, are placed in bins based on the values of the independent variables associated with each observation. The data in each bin are summarized with a single value, the mean and median values being two obvious choices. Finally, the table of summary values are fit with an additive model, as shown in the following example. Summaries of this two-way analysis technique and its extension in the context of exploratory data analysis and robust statistical methods are available[2,3].

As an illustration of the construction of a two-way table, consider the dependence of total daily load on daily average external temperature and day of week. Figure 2a shows a histogram of the average daily exterior temperature for a residence, and Figure 2b is a scatter plot of the total daily load for that site against exterior temperature. We partition the days of the year into three temperature groups. We then partition each of the three temperature groups into two sub-groups, one for weekdays and one for weekends. Thus the daily load data are separated into six catagories. Table 1 is the contingency table for this partitioning. Each entry indicates how many days fell into each category. For instance, there were 66 weekdays in which the average exterior temperature was less than $0^\circ$ centigrade.

The second step in construction of the two-way table is to average the values in each of the bins. In the current case, we have taken a 10% trimmed mean (mean of the data set neglecting the highest and lowest tenths of the distribution) in each of the bins to produce the bin value. These summary values are contained in Table 2. Following the construction of the bin values, we proceed to model the data using a simple model:

$$X(i,j) = G + R(i) + C(j) + E(i,j) \quad (2)$$

Here G is the median of the full set of bin values, R(i) is a row effect associated with the ith row, C(j) is the column effect associated with the jth column, and E(i,j) contains the residuals from the fit.

TABLE I. Contingency table for simple two-way analysis of daily local data.

| | T<0 | 0<T<10 | 10<T |
|---|---|---|---|
| Weekday | 66 | 71 | 118 |
| Weekend | 23 | 34 | 46 |

*This table shows the number of days falling into each of the temperature/day type bins. The time period of the data set is 1 March 1985 to 28 February 1986; seven days data are fully or partially missing and are not included here.*

TABLE II. Average daily energy consumption (kWh).

| | T<0 | 0<T<10 | 10<T |
|---|---|---|---|
| Weekday | 78.5 | 49.6 | 23.5 |
| Weekend | 87.1 | 57.3 | 28.5 |

*Data in this table are the 10% trimmed means of total daily loads for all days falling into each bin. For instance, excluding roughly the 12 highest and lowest values, the mean of the total days for weekdays with average temperature >10°C is 23.5 kWh.*

In the current case, row corresponds to weekday or weekend, and column to temperature range. Thus, we model the daily loads as being determined by exterior temperature and day of the week, with these effects treated as being additive. For the total load data, the elements of X, R, C and E are exhibited in Tables 3 and 4.

TABLE III. Parameters at the two-way fit (kWh).

| | | | |
|---|---|---|---|
| G | 53.5 | | |
| Row | - 3.9 | 3.9 | |
| Column | 29.3 | 0 | -27.5 |

*These parameters are obtained from a two-way fit to the table by iterative extraction of medians in Table 2 (see Equation 2). The values show that temperature effects on energy consumption are larger than those of day type.*

TABLE IV. Residuals from the two-way fit.

| | T<0 | 0<T<10 | T>10 |
|---|---|---|---|
| Weekdays | -.4 | 0 | 1.4 |
| Weekends | .4 | 0 | -1.4 |

*These data are the deviations from the simple model of the load data. Thus, the grand median + row + column estimate of load for cold weekdays is high by .4 kWh while that for cold weekends is low by the same amount. Note that the residuals are much smaller than the row effects, indicating that the analysis is reasonably reliable.*

The row and column values are obtained by iterative extraction and accumulation of row and column medians, a technique referred to as "median polish"[2]. The process ceases when all row and column medians are less than some small value. Although this technique does have some disadvantages, in particular not guaranteeing that an unique result is obtained, it has the substantial advantage of being relatively insensitive to large deviations from the bulk of the data in a few of the cells.

A standard graphical representation of the two-way fit is shown in Figure 3. Each of the parallel lines sloping down from left to right represents an

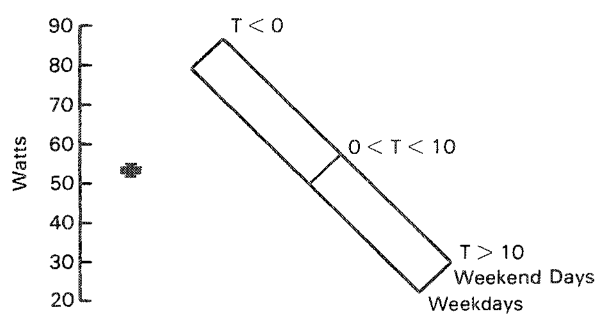


Figure 3. Illustrative plot of two-way analysis.

*This figure is the graphical representation of the simple two-way analysis of daily load data. The separation between the parallel lines with positive slope indicates the magnitude of the temperature effect on load, while the separation between the parallel lines with negative slope indicates the magnitude of the weekday/weekend effect.*

estimate of the effect of weekday or weekend. Each of the lines sloping up from left to right represent the temperature effect. Model estimates for days falling in particular bins are found at the intersection of two lines; for instance, the model value for cold weekdays is slightly less than 80 KWH. Separation between parallel lines is a visual indicator of the magnitude of row or column effects. Thus, in the present case, it appears that temperature effects are substantially more important than the weekday/weekend distinction in determining electrical consumption. The box and whiskers plot on the lefthand circle of the figure indicates the magnitude of the cell residuals, centered at the median G. That is, the elements of E are increased by G, and the resulting values then used to generate the box plot. The dispersion in the residuals relative to the row and column effects is a rough guide to the significance of these effects.

## Application to Hourly Total Load Data

We now proceed to apply this technique to a detailed analysis of the total load data for the residence. In this analysis, we construct a two-way table for total load data for each hour of the day. Thus, the derived model will consist of 24 equations of the form of (2), one for each hour of the day. The resolution of the temperature bins has been increased slightly, and we now distinguish each day of the week. The contingency table for this finer partition is given as Table 5.

TABLE V. Contingency table for hourly load analysis.

| | $T<0$ | $0<T<10$ | $10<T<20$ | $T>20$ |
|---|---|---|---|---|
| Monday | 14 | 14 | 13 | 10 |
| Tuesday | 14 | 12 | 15 | 10 |
| Wednesday | 14 | 12 | 13 | 12 |
| Thursday | 12 | 15 | 14 | 10 |
| Friday | 12 | 18 | 11 | 10 |
| Saturday | 11 | 19 | 10 | 12 |
| Sunday | 12 | 15 | 14 | 10 |

In Figure 4, we show the two-way fits for hours 4, 7, 13 and 19. Inspection of these figures immediately reveals several important facts about the dependence of energy consumption on day of week and external temperature for this site. For all four hours there is a clear tendency to consume more energy with decreasing outside temperature. The spacings between temperature bins vary substantially among the hours, thus indicating that the dependence of load on external temperature does vary over the course of the day.

It is quite evident that the day of week effects are dramatically different for the different hours. During hour 4 (3AM to 4AM) there is little day-to-day variation, as evidenced by the small spacing between the several days of the week. On the other hand, during hour 7 (6AM to 7AM) the day-of-week effect on loads is as important as the temperature dependence. The clustering of the two weekend mornings at much lower values than the weekday values is interesting, as is the evidence of a slow start on Monday morning! For hour 13 (Noon to 1PM), the day-of-week effects have reversed themselves entirely. Now it is on the weekends that the bulk of the energy is consumed, while the weekdays cluster at a lower value. Note also that the load varies little with temperature across the warmer bins, probably because internal gains are providing the

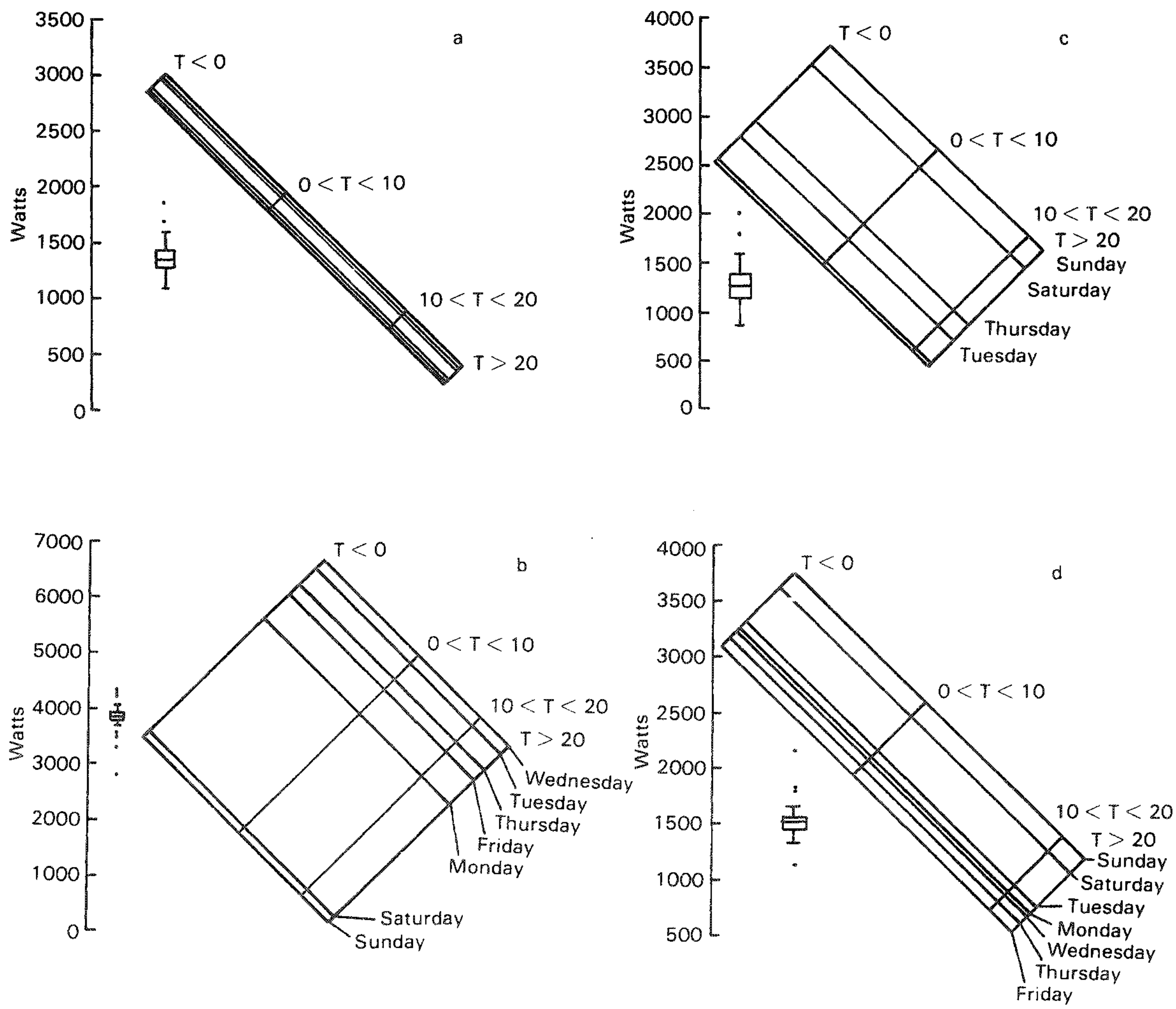


Figure 4. Two-way analysis of total load data for four hours.

*This figure shows the two-way fits for hours 4 (top left), 7 (bottom left), 13 (top right) and 19 (bottom right) for a year's worth of total load data from a single residence. The box-and-whiskers plots show the size of the residuals, with zero residual located at the median of the data. Weekday labels are omitted where the corresponding lines are superimposed.*

bulk of the required heating energy. Hour 19 (6PM to 7PM) still shows Saturday and Sunday as having the largest loads, while the weekdays are very tightly clustered at low values. Again, the distinction between the warmest temperature bins is very small.

One fault with the representation of this load data through the two-way tables is that the graphical depiction is rather difficult to absorb rapidly.

Each of the individual hourly tables is reasonably complex; if all 24 were displayed some effort to interpret them would be required. An alternate graphical representation can be obtained by constructing the 168-hour time series through the week of model estimates for each of the temperature bins. In Figure 5, three of these time series are plotted. Now the dependence of load on temperature, as well as variation across the week, is clearly depicted.

One important consideration in assessing the reliability of these simple two-way models for representing or predicting electrical loads is the difference between observed and modelled loads. Figure 6 shows the deviation between observed and mean loads, and between observed and estimated loads, for each hour of the day. Although there are several outliers, the estimates do indeed account for a substantial fraction of the variance present in the initial load data.

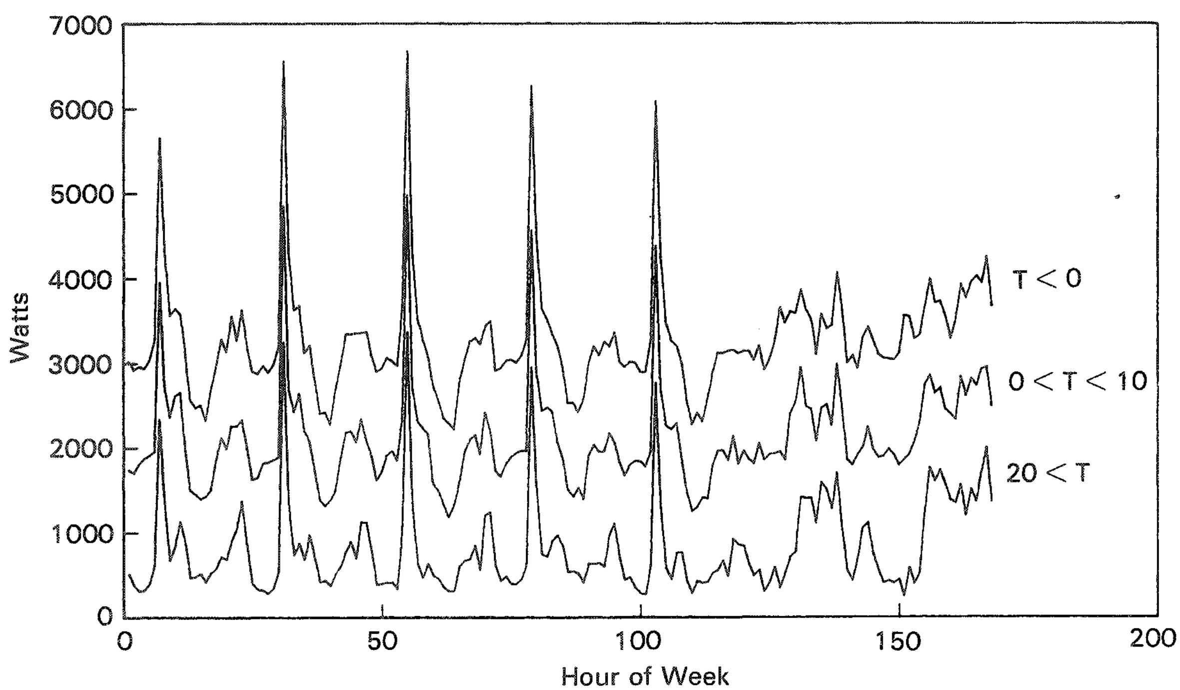


Figure 5. Weekly load profiles for three temperature bins.

*In this figure, model estimates of the weekly load profile for three of the temperature bins are shown. The model estimates are well separated, with lower temperature implying higher energy consumption for each hour of the week. Note the distinction between weekday and weekend load shapes.*

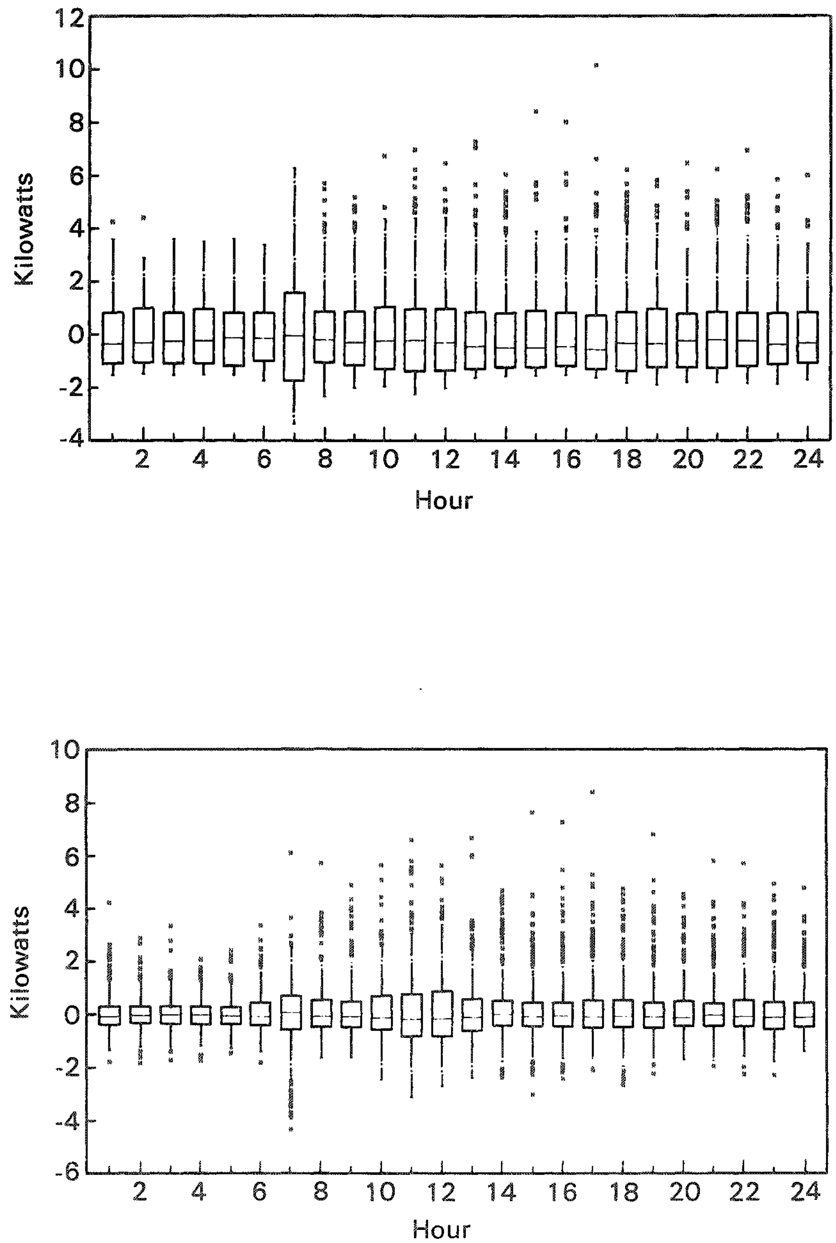


Figure 6. Residual variance in hourly loads.

*The two panels in this figure indicate the residual variance in the hourly load data, the top from mean hourly values and the bottom from model values. For all hours the model provides a clear reduction in dispersion of the data as judged by the interquartile range. As one would expect, the deviations from model values are larger during active hours than in hours where there is little occurring in the structure.*

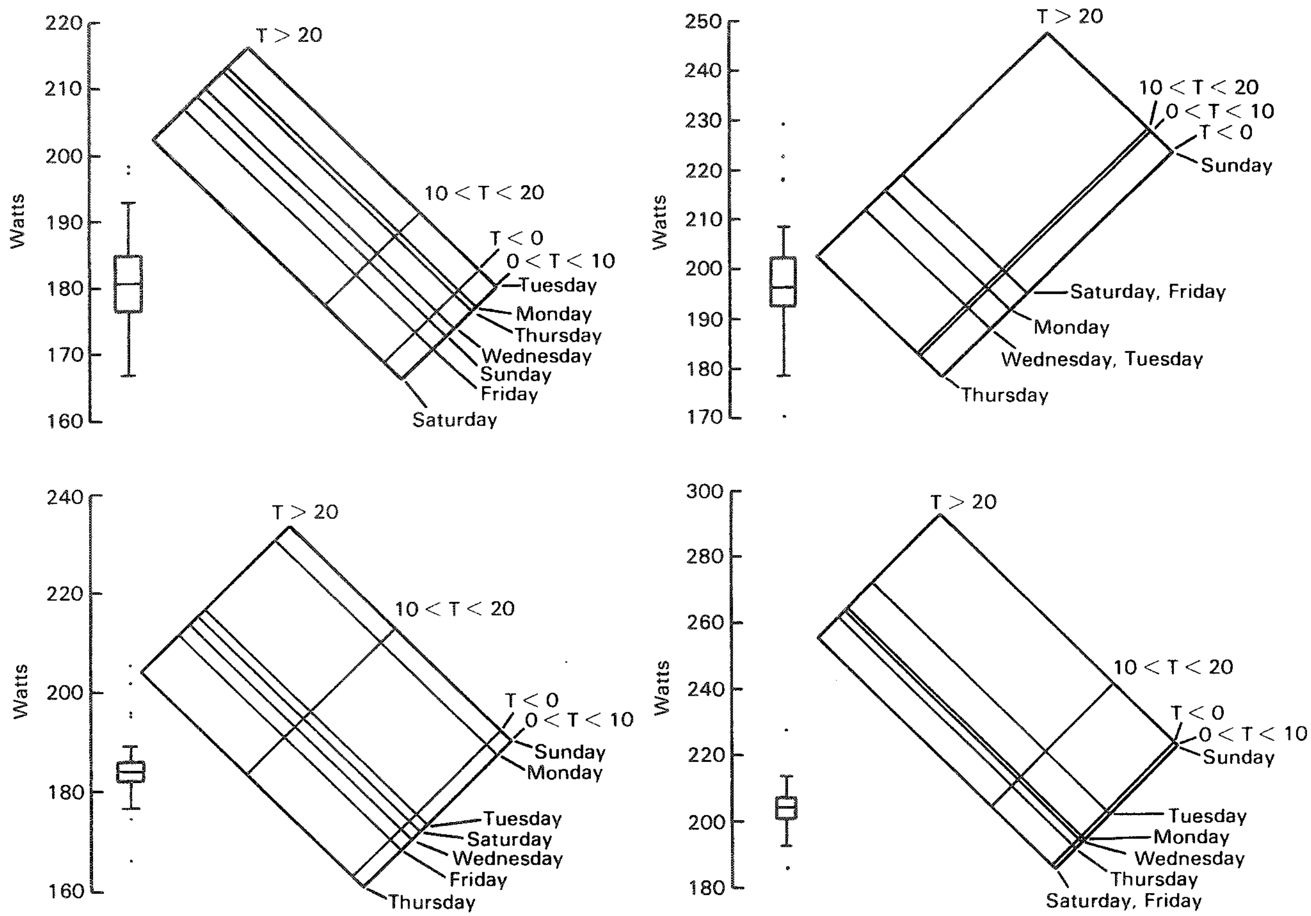


Figure 7. Two-way analysis of refrigeration load for hours 4, 7, 13 and 19.

This figure displays the two-way fits for the refrigeration load at hours 4 (bottom left), 7 (top left), 13 (top right) and 19 (bottom right). The box-and-whiskers plot on the left of each panel is generated from the residuals from the model, adjusted so that a residual of zero falls at the grand median of the data. Note the changing relative importance of outside temperature and day of week with hour of day. It is quite evident that the relation between outside temperature and refrigerator load is strong, but decidedly nonlinear. The larger day of week effects and temperature effects exceed in magnitude almost all of the residuals.

Analysis of Refrigeration Load. As a second example of the application of two-way tables, we consider the refrigeration load for the same site as analyzed above. Again, we have constructed a two-way table for each hour of the day. In Figures 7a-d, the two-way analyses for hours 4, 7, 13 and 19 are depicted. Perhaps the most important general feature to note is the increase in load with exterior temperature at all hours. The separation between temperature bins is largest for the highest temperatures, and tends to be small for the lower temperatures. This is quite reasonable; when it is cold outside, the interior temperature is reasonably constant, and so the refrigerator is operating under unchanging conditions. However, once the house begins to be warmer than the thermostat set point for extended periods of the day, the refrigerator is required to exhaust heat into a higher temperature reservoir, and load increases. There is no air conditioning in this structure, and the effect is most pronounced at the highest temperatures.

Day-of-week effects are significant at several hours. Sunday tends to be a day with generally high load on the refrigerator, particularly during the afternoon. At hour seven there appears to be little distinction between the days of the week, although Saturday does appear to have a low value. As in the case of load, the importance of day of the week in determining load varies greatly by hour of the day. Figure 8 shows the model estimates by hour of week for the highest and lowest temperature bins.

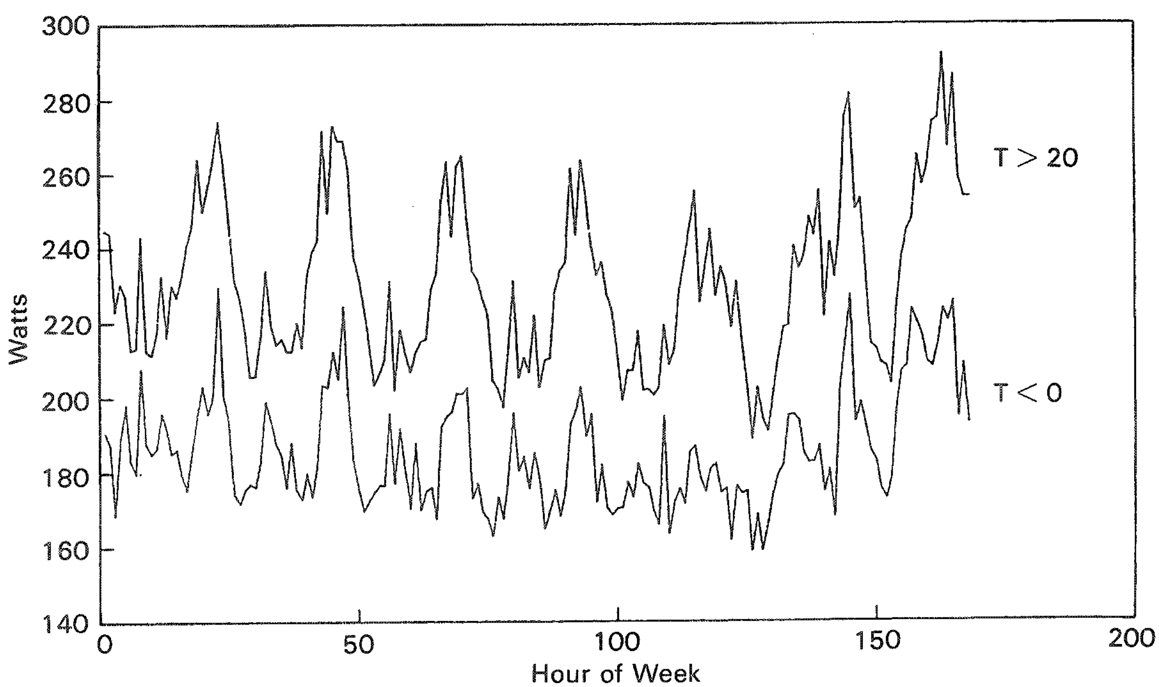


Figure 8. Model weekly load profiles for refrigeration.

*This figure shows the model estimates for weekly load profiles in the highest and lowest temperature bins. The relative importance of day-of-week, exterior temperature and hour-of-the-day may be observed in this plot.*

## Summary

It is quite evident that both as an analysis tool and as a summary or model of the data two-way tables have substantial value. The representation is compact; a total of 12 coefficients are required for each hour of the day. In the examples shown here, the analysis has sensibly determined the relative importance of various factors (hour of the day, day of week, and external temperature) in influencing load, and a substantial fraction of the variance of the load data has been captured. Simple graphical representations which permit exhibition of these dependencies are available. The residual variation among values in each bin is substantial, but this is not surprising, as random variation may be expected to be quite large in data taken from a single residential structure.

## EXPANSION IN A SET OF EMPIRICAL ORTHOGONAL FUNCTIONS

One method of providing compact representations of loadshapes, while retaining information on the impact of deterministic variables, is to express the load shape in terms of some set of functions. The coefficients in the expansion then incorporate, for instance, dependence on temperature or other variables. Thus, a typical representation incorporating dependence on temperature only would be:

$$\underset{\sim}{L}(T) = \sum_{i=1}^{n} a_i(T)\underset{\sim}{\phi}_i \qquad (3)$$

here L is a vector with 24 elements containing the electrical load by hour of the day, T is external temperature, the $a_i$ are simple functions of temperature, and the $\phi_i$ are a set of vectors of 24 elements from which the load profile is constructed.

There are many different ways in which a suitable set of basis functions could be obtained. In this paper, we investigate the use of empirical orthogonal functions (EOFs) obtained from principal component analysis in this role. The technique used contains several steps:

- Extraction of principal components from hourly load data
- Fitting of each of the daily profiles in terms of the first few components
- Determination of the relationship between the coefficients of the fit and independent variables, and
- Construction and evaluation of the model.

We illustrate this procedure by deriving a model for the total hourly electrical load for weekdays in a single office building as a function of outside temperature.

## Data Set Employed

In this work, we use ELCAP data collected from an office building located in Seattle. The data employed in this analysis are the total weekday hourly loads observed between August 1984 and July 1985, each data value representing the average demand during a particular hour. This office building has 14,920 square feet of office space located on two floors, built over a 6,000-square-foot daylight garage. It is a multitenant building and was constructed in 1976. The metering in this building permits disaggregation of the HVAC, lighting, general use (plugs, etc.), and elevator electrical consumption; here we use total load data only.

## Determination of Basis Functions

The hourly data are arranged in the form of a matrix with 24 columns each corresponding to one hour of the day and one row for each day of the year for which data are available. This matrix, denoted by X, is then multiplied on the left by its transpose to form a 24x24 symmetric crossproducts matrix. The eigenvalues and eigenvectors of this matrix are calculated, providing a representation of the crossproducts matrix as:

$$X'X = VLV' \tag{4}$$

where $X'X$ = crossproduct matrix, $V$ = eigenvector matrix where $V'V = VV' = I$, and $L$ = diagonal matrix of eigenvalues. The matrix X can then be written as:

$$X = PV' \tag{5}$$

where $P = XV$ is the principal components matrix.

The matrix X can thus be represented as the product of the principal components matrix P and the eigenvector matrix V'. Each row of V' is a set of 24 values representing an hourly load profile. Each column of P is a set of daily values. The first column of P contains a multiplier for each day for the first row of V'. Another way of looking at the matrix P is to say that its rows contain the coefficients for each day of a regression of the hourly profile for that day on the eigenprofiles given by the rows of V'. This calculation provides the desired decomposition of the data matrix X into an hour of day portion (V') and a day of year portion (P).

In Figure 9, we show the first three principal components for the office building generated from weekday data taken between 8/84 and 7/85. Visual interpretation of the components is possible, although potentially hazardous, particularly for the higher order components. In particular, it should be noted that the components are indeterminate in sign. That is, there is no guarantee that one is looking at the component "right-side up."

The first component in this building is similar to a typical load shape for the building. The second component shows either high or low afternoon load relative to the morning; it may be related to the afternoon cooling load. The third component shows a morning spike or trough, and may be related to the

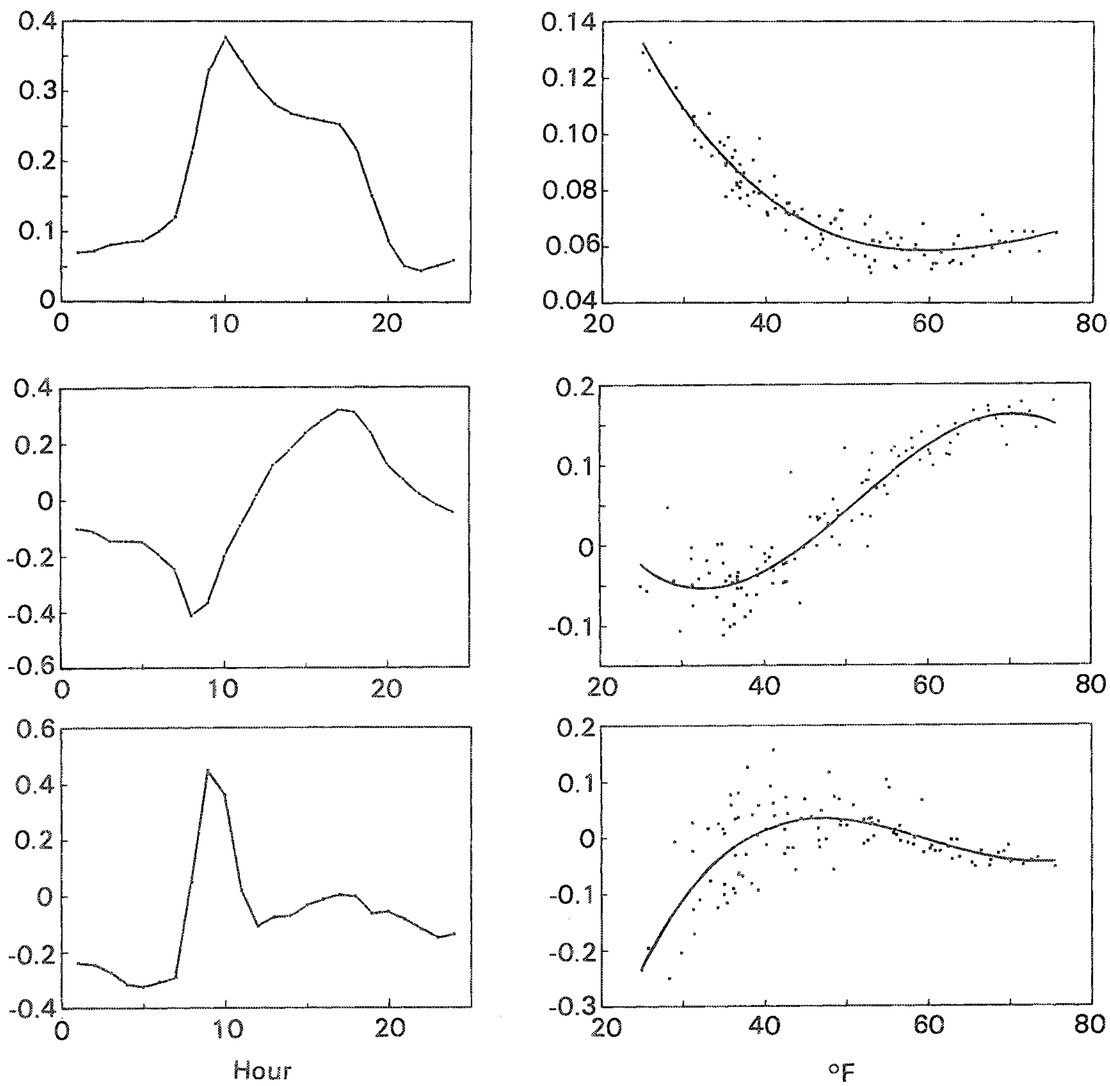


Figure 9. Principal components and temperature dependence of component weights.

*The three lefthand panels show the first three principal components of the weekday load for a Seattle office building. The first component (top) has much the same shape as the average load profile. The three righthand panels show the dependence on average daily temperature of the weight of the first three components in expansions of the individual daily load profiles. The curves are cubic polynomial fits to the scatter plots. The dependence of the coefficient for component one (top) is similar to that of total daily load on daily average temperature.*

peakiness of the morning heating load. The fourth component (not shown) is also spikelike, but, as is usually the case, interpretation is becoming ever more difficult.

## Relation of Coefficients to Ambient Temperature

In constructing the model, the second step is the determination of the relationship between coefficients of the various terms in the EOF expansion and any driving variables. To maintain reasonable simplicity, simple functional forms for these dependencies must be found. In our example, we incorporate temperature dependence only. Consequently, we expand each daily profile in terms of the first three components and examine the temperature dependence of the coefficients in the expansion.

Figure 9 also contains scatter plots of the coefficients for the first three principal components against daily average temperature. The coefficient of the first principal component has much the same dependence on temperature as does the total daily consumption. The steep increase at low temperature, the minimum at intermediate temperature and the slower increase at high temperature are typical. This observation reinforces the interpretation of principal component 1 as a mean profile.

A positive value of the coefficient for component 2 will correspond to increased load in the afternoon and decreased load in the morning, while a negative coefficient will have the opposite interpretation. Thus, the positive values of the coefficient at higher temperatures and the negative values at lower temperatures do indeed seem to correspond to a component reflecting the intensity of afternoon cooling. The scatter in the coefficients for the daily fits seems to be higher at low temperatures than at high temperatures.

The coefficient of the third function shows somewhat more complicated, and, as indicated by the scatter, perhaps less well determined behavior as a function of temperature. If one takes the fits to the data seriously, there is a maximum in the coefficient at some intermediate value, with a sharp decrease to minimum values at the lowest temperatures. The negative values indicate a sharp enhancement of morning load between midnight and 7 AM against the period between 8 AM and 11 AM. There is relatively little clear evidence of temperature dependence in the coefficient for the fourth function.

To construct our model of this building, we have fit cubic polynomials in temperature to the scatter plots. Our predicted energy consumption profile for each day is then obtained by summing over the basis functions, with each weighted by a temperature-dependent coefficient calculated from a simple polynomial fit to the data.

## Evaluation of Model

In Figures 10a and b, we show the fit to a typical profile and to an exceptional profile. These figures provide a good basis for discussion of the advantages and limitations of EOF models of the type developed here. Figure 10a shows the metered consumption and model estimates for an early summer day. The deviation in load from the model shown here is typical of most weekdays. The model captures the bulk of the systematic variance both from building schedule (that is, by hour of the day) and external temperature. Figure 10b makes the same comparison of metered and estimated data for the following day. Note that a large morning spike in the metered data is not fitted by the model. This is not surprising; the average external temperature for the two days is almost identical, so the model predictions must be similar. However, the two metered load profiles are quite dissimilar.

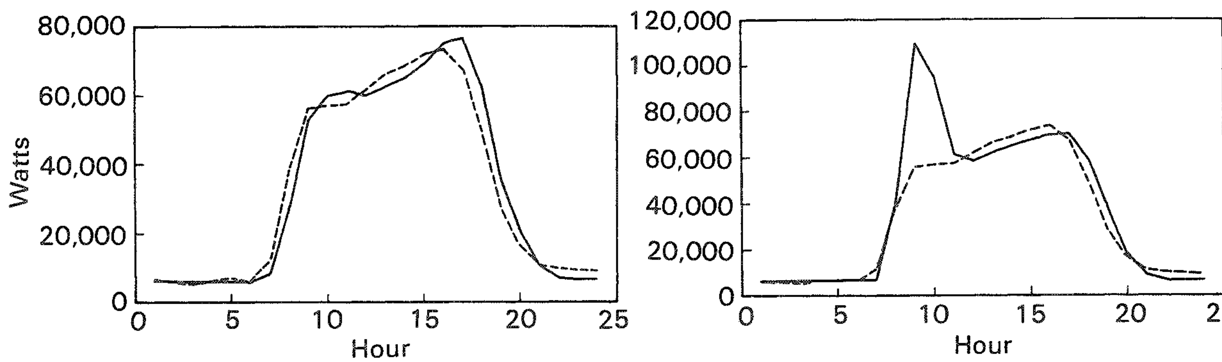


Figure 10. Comparison of model and observed loads.

*In the left figure, the model is compared to the actual daily profile for June 20, 1985. The solid curve shows the measured load, while the dashed line shows the model predictions. This comparison gives a typical picture of deviations between measured data and the model.*

*The right figure shows the comparison between model results (dotted line) and actual energy consumption for June 21, 1985. Note the failure to predict the spike between 8 and 10 am. Such spikes occur in very few single day metered load profiles for this building, and are particularly atypical at this time of year. Consequently, they are not reflected in the basis functions. The deviation is reflective of the fact that models of the type applied here will not characterize large random fluctuations.*

It turns out that on a small number of days during this summer, the metered data show these spikes, which are associated with unusual operation of the HVAC equipment. For the purposes of discussion, the spikes may be viewed as large random excursions from typical building operation. As rare events, these spikes are not reflected in the low-order principal components derived from the load data. In any event, because the incidence of such spikes during warm weather does not appear to be related to temperature, no model of the type developed here will reflect them. Indeed, if it was desired to develop a model in which these spikes would appear with a frequency similar to that with which they actually appear, a stochastic model would be required.

## Summary

On balance, the model described here is successful at capturing the bulk of the variance in the load data for the office building in question in compact form. Although unable to capture occasional large random variations in the load, it does provide quite reasonable predictive power for determining weekday load by hour of the day as a function of external temperature. It is, however, important to understand one additional limitation of this approach before its utility can be completely assessed.

The office building studied here is heavily scheduled, and the schedule varied little over the course of the year from which the metered data have been taken. Had there been major shifts in building schedule, unrelated to temperature or other independent variables, low-order components reflecting the schedule shift would have been generated. Although the data could have been modelled with these functions, its predictive power would be substantially diminished. To put it another way, in order to predict a load for a scheduled building, it is necessary to know the schedule. This limitation is, of course, much more important for single buildings than for large sets of structures. In addition, interpretation of the principal components in terms of the building schedule, something clearly feasible in the case of at least the first component, should permit estimation of load profile as a function of schedule.

## CONCLUSION

In this paper we have argued that compact representations of load shape which contain the principal systematic determinants of load shape are required. We have then illustrated two approaches--two-way analysis and expansions in terms of empirical orthogonal functions--to generation of simple models for electrical load. Each of these approaches has been applied to metered data from single structures and has been shown to capture the major part of the systematic variance in the data.

We conclude with a number of observations. First, it does seem likely that the two approaches here, as well as any of a large variety of other techniques, can produce reasonably reliable empirical models of electrical load data. These models have substantial value both in analysis and interpretation of load data, and presumably in summarizing or representing load data for conservation or load forecasting applications. There is no limitation to incorporating dependencies on day of week or external temperature; any variable systematically related to load may be included.

Second, models of this type are most useful for studying loads where the schedule exhibits changes unrelated to the independent variables only rarely. Unless more elaborate models which take explicit account of schedule are developed, this limitation can be significant with regard to the predictive power of empirical models for single structures. However, presumably schedules for loads aggregated across large numbers of buildings change relatively slowly and infrequently, so that this limitation may not be of great importance.

Third, there is a place for a very different type of electrical load model. A stochastic model which incorporates information about the variability in loads, and in particular about the statistics of the fluctuations about the systematically varying load, clearly has substantial areas of applicability. Such a model could easily be generated starting with one of the empirical models as a basis. One would then add a random component, with some parameter dependency on hour of day and other variables.

This work was supported by the U.S. Department of Energy under Contract DE-AC06-76RLO 1830.

REFERENCES

1. ELCAP is a Bonneville Power Administration project being carried out by Pacific Northwest Laboratories. It involves the detailed end-use electrical metering of roughly eight hundred residential and commercial structures in the Pacific Northwest.

2. Tukey, J. W. 1977. Exploratory Data Analysis. Adison-Wesley Publishing Co. (Chapters 10-12).

3. Hoaglin, D. C., F. Mosteller, and J. W. Tukey, (editors). 1985. Exploring Data Tables, Trends and Shapes. John Wiley and Sons. (Chapters 2-4).