EXPLORATORY ANALYSIS OF RESIDENTIAL END-USE CONSUMPTION

G. M. Stokes and N. E. Miller
Pacific Northwest Laboratory

ABSTRACT

The data being collected as part of the Bonneville Power Administration's End-use Load and Conservation Assessment Program (ELCAP) include highly disaggregated residential data taken with hourly integration periods. The detailed end-use patterns seen in this data offer the opportunity to explore occupancy effects in residential energy consumption to an unprecedented degree. Preliminary analyses have suggested that aggregations of data according to similar consumption profiles, an activity-based end-use aggregation scheme, provide an interesting complement to analyses based on simple equipment-based analyses.

An earlier analysis of this data was based on a series of multivariate statistical techniques including multidimensional scaling and principal component analysis. In this paper the analyses are continued and a series of waveform classification schemes are used to attempt to establish a firm basis for comparing energy consumption profiles between residences.

EXPLORATORY ANALYSIS OF RESIDENTIAL END-USE CONSUMPTION

Gerald M. Stokes and Nancy E. Miller
Pacific Northwest Laboratory

INTRODUCTION

One of the most difficult aspects of conservation assessment and load forecasting in the residential sector is the effective treatment of occupancy effects. While the term "occupancy" has been used to cover a variety of issues in the residential sector, it primarily refers to the way residential utility customers operate their homes and use electricity.

Many of the difficulties arise from the multiple use nature of residential structures. The simplest way to understand the problem is to look at residences in comparison with commercial buildings. From this commercial perspective, the "economic activities" in a residence can resemble restaurant, motel, warehouse, office building, laundry, and health facilities. However, unlike a commercial structure, very few of these activities have either regular business hours or a standard set of products. Each of these activities has the potential to consume large amounts of energy, and, therefore, understanding the frequency and intensity of each of the activities is important. Furthermore, as the design of homes continues to take advantage of the interaction of the various uses of energy in a home, such as using waste heat from refrigeration for space heating, understanding the interrelatedness of the various residential activities will become just as important as understanding individual schedules.

This paper describes an exploratory research program that has been undertaken in an attempt to characterize residential occupancy effects through the analysis of highly disaggregated hourly load data. The goal of this program is to develop characterizations of residential load shapes that can be combined with basic demographic data to gain insight into the problem of residential occupancy effects on energy consumption. The data set is from the End-Use Load and Conservation Assessment Program (ELCAP), which is sponsored by the Bonneville Power Administration (BPA) and managed by Pacific Northwest Laboratory (PNL) in Richland, Washington.

APPROACH

The characterization of residential end-use loads has been widely discussed. The ELCAP data set is so large that an effective method of characterizing the data is essential. Ideally, the method of characterization will compress the size of the data set without reducing the information content in an unacceptable fashion.

The foregoing discussion of the distinction between residential and commercial buildings strikes at the heart of the peculiar difficulties of characterizing highly disaggregated end-use data in the residential sector. In this context it is useful to consider the concept of a typical day's energy consumption. In the commercial sector, one can imagine a typical weekday lighting profile for an office building. The uniformity of business hours and the standard operational mode of the structure makes the load profile very similar from day to day. As a result of this similarity, a profile constructed from the average hourly energy consumption on weekdays or from the median for the same set would result in a lighting load profile that typifies the lighting load. The resulting profile would be typical in that one would reasonably expect the profile to be fairly similar to some day's worth of data on lighting energy consumption.

In the residential sector, however, the situation is much more difficult. If a profile was constructed from means or medians of the laundry end-use in a residence, it would not likely resemble any single day's consumption of energy. In particular, the profile constructed from the medians would most likely be zero; the one based on the means would be a very complicated convolution of a typical use profile, with the joint probability distribution of the likelihood of any laundry activity taking place on the day and the likelihood of that activity beginning at a given time.

The ideal residential end-use characterization should incorporate information concerning the amplitude, the phasing, the frequency of occurrence of a particular end-use activity, and the general characteristics of the waveform.

THE DATA AND INITIAL APPROACH

The data being analyzed come from the ELCAP metering system, which has been described in several places (Schuster and Tomich, 1985; Parker, Pearson, and Sandusky, 1985). For the present work, the most important aspect of the metering system is that up to 16 channels of residential electric consumption are collected with hourly resolution. The 16 channels are selected following a protocol that attempts to achieve the maximum end-use disaggregation possible at the distribution panel level. The level of disaggregation is affected by some quality control constraints, but it is usually possible to get a reasonable number of the end-uses described in Table I.

In attempting to structure this investigation it was noted that hourly end-use data has some interesting properties. One of the most interesting is that each hourly record can be viewed as a testing instrument applied to the residents of a house. As conditions vary, we can therefore study the answers to the same set of "questions" as a function of time. While the testing metaphor may be stretched somewhat, it does suggest that a whole range of statistical tools developed to interpret sociological and psychological testing results may be appropriate to the study of end-use data as well.

Table I. ELCAP residential end-uses.

BASIC (ALL RESIDENCES)	EXPAND (AS PRACTICAL)
HVAC	Refrigerator
Hot Water Heat	Freezer
Other (expanded at right)	Food Preparation
	Dishwasher
	Clothes Washer
	Clothes Dryer
	Dehumidifier
	Disposal
	A/A Heat Exchanger
	Lights and Convenience Outlets*
	Specialty Appliances

---

*Location of lights and convenience outlets are noted as possible.

ACTIVITY-BASED END-USES

There are two classes of end-use interaction that can be seen in a residence. The first is the obvious physical interaction such as the displacement of heating load by waste heat from refrigerators. The second is that individual appliances are many times used in conjunction with each other. In the former case, the interaction is physical while in the latter case, the interaction is a function of the occupants.

Stokes and Owen (1985) have described two views of residential end-uses which they have called functional end-uses and activity-based end-uses. The functional end-use set is the conventional view of residential end-use consumption and is, in fact, the basis of the end-use approach taken for ELCAP and summarized in Table I. The approach is engineering and equipment oriented and has several important advantages. As an equipment-based approach, it is well suited to the study of the performance and overall importance of specific appliances and energy-consuming equipment in a residence. Next, it has the valuable experimental advantage in that the end-use set can be specified prior to the installation of metering equipment.

An activity end-use set focuses on the way people use electricity rather than the equipment used to consume the energy. The end-use patterns in an activity end-use approach can be viewed as the schedules for the basic residential activities. An activity is therefore defined as a collection of energy consuming actions with a common schedule. The experimental disadvantages of this approach are obvious. Knowledge of the residents and their energy consumption schedules is needed before the end-use set can be specified. Therefore, the classification of data into activity end-uses is necessarily an *a posteriori* activity, which would be done to compare of the data with demographic and structural information.

Stokes and Owen applied several multivariate statistical techniques to the study of the ELCAP data to assess the appropriateness of the activity end-use approach. All three of the techniques used in the analysis, principal component analysis, cluster analysis and multidimensional scaling, confirmed that there are several basic schedules in the residences studied. The results are illustrated in Figure 1, which shows the results of the cluster analysis and the multidimensional scaling, performed on 14 energy channels for a single residence. The results for both analyses were considered in making the final classification of the individual channels into activity end-uses.

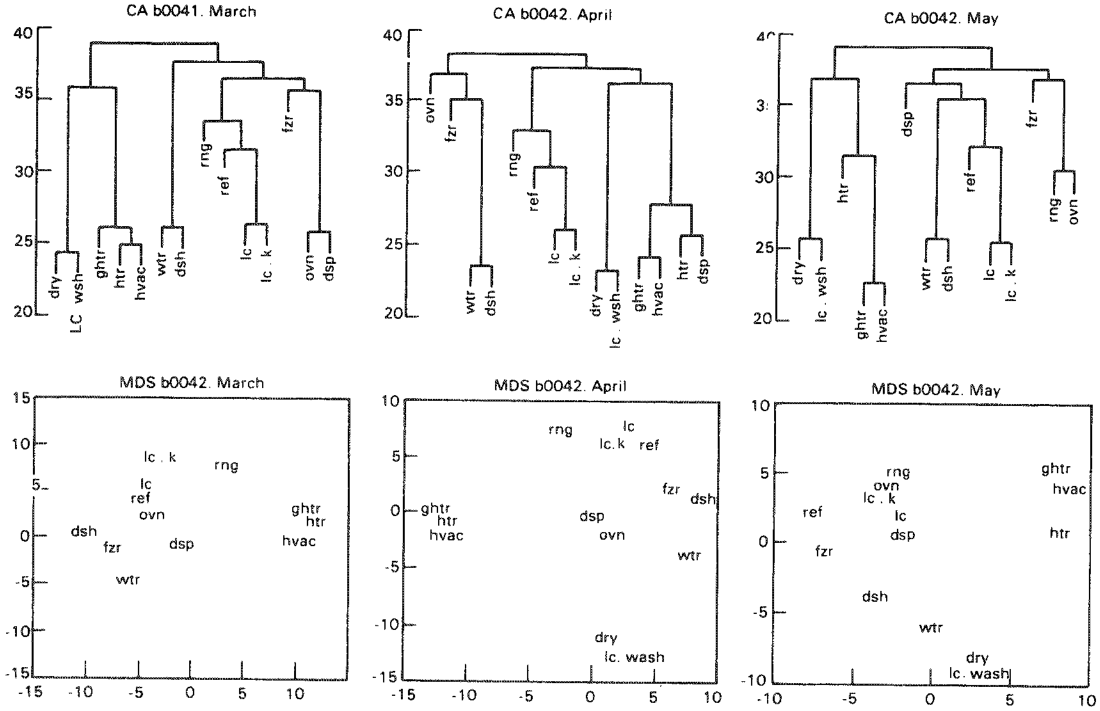


Figure 1. Identifying activity-based end-uses for a single residence, March - May 1985.

The results in Figure 1 are based on the application of cluster analysis and multidimensional scaling to a Euclidean distance matrix constructed from the channel level data for a single residence from March, April, and May 1985. The matrix was constructed by viewing a four week segment of hourly data (672 data points) for each channel of data as a vector. The distance was then calculated between each of the pairs of end-use vectors. The analogy would be the distance matrix found in a road atlas showing the distance between cities. The scale in the multidimensional scaling results represents a crude measure of the distance apart of the vectors in a 2016-dimensional energy space. For the road map example, the multidimensional scaling result should reproduce the relative distribution of the cities in the matrix as they would appear on the map.

The most interesting data in Figure 1 is from the multidimensional scaling. In this type of analysis, the proximity of one data point to another indicates similar schedules for the use of electricity on the respective channels. The channels shown in this set of figures are: wtr (hot water heater), rng (kitchen range), lc (lights and convenience outlets), lc.k (kitchen lights and convenience), lc.wsh (lights and convenience in the laundry; the washer is connected to one outlet on the circuit), dry (clothes dryer), htr (baseboard heater), hvac (central hvac system), ghtr (heater for garage), dsh (dishwasher), ovn (oven), ref (refrigerator), fzr (freezer), and dsp (garbage disposal).

In Figure 1, several features should be noted. First, the stability of the clusters from month to month seems to be quite strong. The clusters that display this most readily are the clothes washing cluster (lc.wsh, dry); hvac cluster (hvac, htr, ghtr); and a kitchen-related cluster (ovn, ref, lc.k, lc, dsp, rng, fzr). In addition, it is interesting to note the relative relationship of the hot water heater (wtr) to the dishwashing (dsh) and laundry cluster (lc.wsh, dry). (The process of classification is, of course, somewhat arbitrary and one of the difficulties more recent work is attempting to remedy.) The hot water channel (wtr) is quite interesting in its relationship between two hot water consuming activities, dishwashing and laundry. The criteria loosely applied during the classification procedure were that the cluster be fairly stable in time and that the relative relationship, primarily "distance" (in the sense of energy) among the various channels, be preserved.

In seven residences that were studied in detail, the number of activity clusters varied from three to six. While most of the inferred activities were fairly obvious and intuitive others were not. The results of the classification of channels into activity end-uses is shown in Table II. In Table II, we show the occurrence of clusters associated with different activities for each of the seven residences. The Xs denote the presence of a cluster with those basic characteristics. The number 2 indicates two clusters, e.g., site 42 (kitchen). The presence of any other letter indicates that the activity has been subsumed into another activity (e.g., the dishwashing activity for Sites 42 and 26). The occurrence of multiple kitchen activity end-uses is fairly common. Each of these different kitchen activity clusters represents real differences in schedules.

The multivariate analyses lead us to several conclusions about residential energy consumption. Specifically, the concept of an activity end-use appears to be a valid one. Not surprisingly, electricity is consumed following several basic schedules that involve several functional end-uses. The activity end-uses appears to be stable with time. In other words, the cluster of electrical consuming activities revealed through the analysis appear to remain clustered from month to month. The relationship between the various activity clusters seems to vary in reasonable ways revealing, for example, that some activities such as heating water are related closely to other activities that have very different schedules.

Unfortunately, in spite of the interesting elements of the foregoing conclusions and the interesting aspects of the analyses that produced them, there are several important limitations to the applicability of the analysis. For

Table II. Activity cluster matrix.

ACTIVITY	SITE						
	42	88	67	26	39	69	94
HVAC (H)	X	X	X	K	X	X	X
HOT WATER (W)	X	X	X		X	X	L
LAUNDRY (L)	X	X	X	X	X	X	
KITCHEN (K)	2	2	X	2	2	X	2
DISHWASHER (D)	K			K		X	
CONVENIENCE (C)		X	X				
A-A HEAT EX (A)							H
REFRIGERATION (R)	K	K	K		X	X	

example, the activity end-uses, particularly the complicated ones, are not replicated in a simple fashion from household to household. Next, while the activity end-uses appear to be fewer in number than the functional end-uses, the resulting average profiles are, on a residence-by-residence basis, afflicted by the same problems as functional end-uses. The average profile generally represents more of a probability distribution of an activity than a typical daily use pattern.

With these concerns and limitations in mind, it appeared more appropriate to focus on approaches to characterizing residential energy profiles themselves. Any characterization of the load profiles should allow the comparison of the profiles and facilitate a re-creation of the foregoing analysis. We do, however, conclude that in the residential case, this multivariate analysis is instructive but limited in applicability.

APPLICATION OF TREE-BASED WAVEFORM CHARACTERIZATIONS

The data, used for waveform construction, consists of residence-specific averaged hourly power observations aggregated to the end-uses listed in Table I. In addition, most sites have averaged hourly indoor and outdoor temperature data. The daily end-use load shape is constructed by connecting consecutive hourly data points. Hours of the day are represented on the horizontal axis; averaged watts per hour are represented on the vertical axis. The construction process is analogous for the daily meteorological plots, except that the vertical axis is averaged degrees. The resulting curve can then be viewed as a waveform.

The peaks and valleys of the wave detail the changes in electrical consumption for a given end-use or the changes in inside or outside air temperature.

The use of trees to describe waveforms began with the work of Ehrich and Foith (1976), who developed the concept of a relational tree. The extension of this work that we have followed is from Cheng and Lu (1985) and Lu (1983). In particular, we have focused on a tree representation that Cheng and Lu refer to as a skeletal tree. The relationship between a load shape and a skeletal tree is shown in Figure 2.

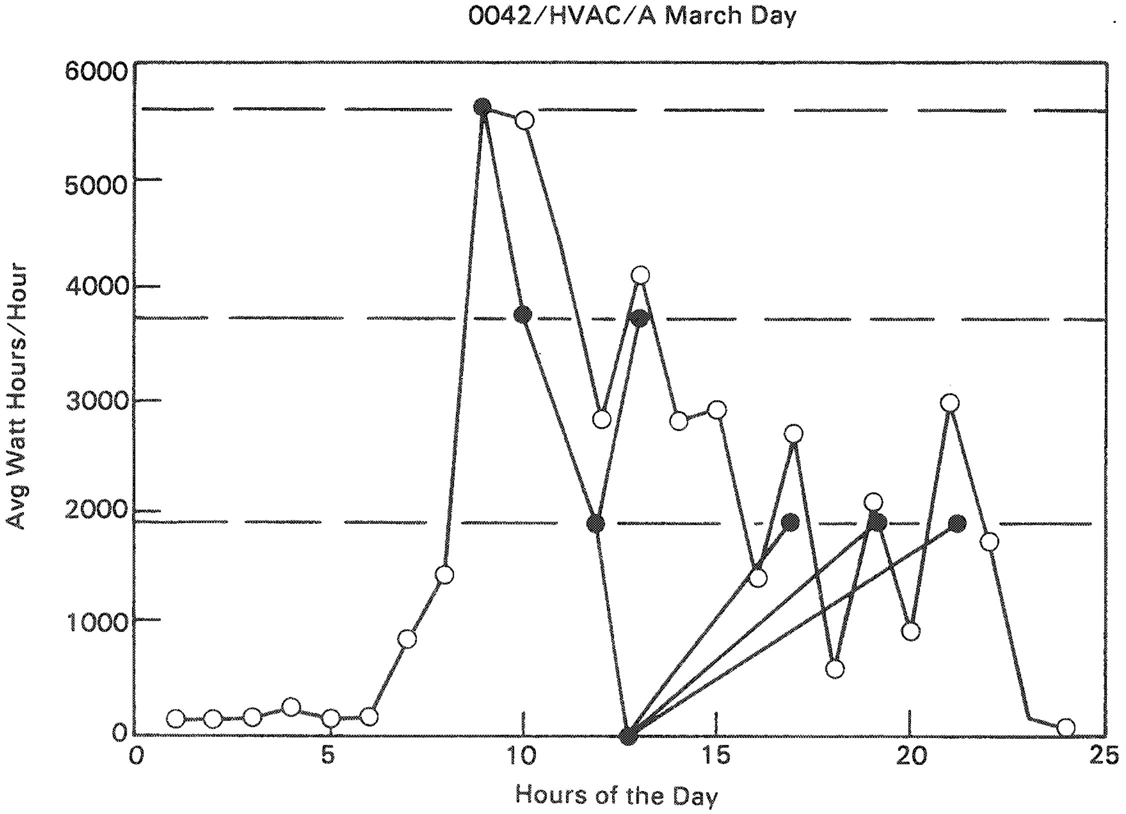


Figure 2. The creation of a skeletal tree.

In practice, we have elected to normalize the daily load curves to the highest demand before constructing the tree. We also calculate the total power for the load curve as well, which we hold along with the amplitude for further analysis.

To go from the daily end-use waveform to a skeletal tree requires that the curve's maximum value be quantized. The selected number of quantization levels form partitions of equal width, beginning at the horizontal axis or at the wave's baseline, and extending upward. The intersection of the waveform with each line of quantization generate pairs of intersection points that define a node generating interval on the quantization line. The midpoint of this intercepted quantization line segment will be taken as a tree node. After a node is selected on level k, the midpoint of the intercepted quantization segment directly above on level (k+1) will generate the son of the node on level k. This process continues for every node until no further nodes can be found.

A tree consists of the nodes and the relationship between those nodes. Two nodes are connected in the tree if, and only if, the node on level (k+1) is the son of the node on level k. Peaks in the waveform, discernible at the chosen level of quantization, are given by nodes with no sons (at some terminal point) and valleys show up when a single node has multiple sons. The single node on the first quantization level is called the root of the tree.

A chain is a sequence of nodes that starts at a peak node and goes, level by level, down to the root. The individual nodes on the chain form the links for a given chain. In Figure 3, the 3 discernible peaks generate three chains. Chain 1 contains links 6,4, 2, and 1. There are two valleys detected by nodes with multiple sons. Node 2 has two sons, nodes 4 and 5; and the root (node 1) has two sons; nodes 2 and 3.

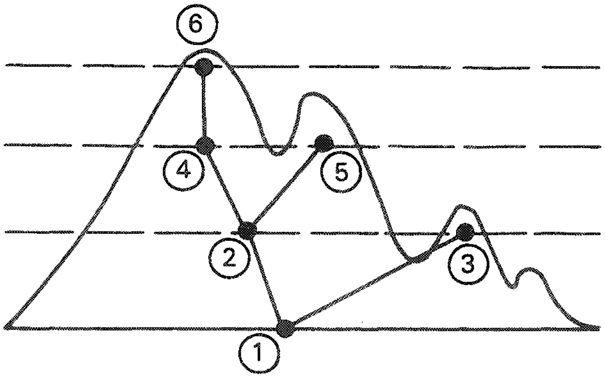


Figure 3. Geo-order node numbering scheme.

The number of partitions selected in the quantization process does impact both nodes and chains but in different ways. That difference helps to capture the quantity and quality of the peaks and valleys in the waveform. Our analysis requires a single skeletal tree from each daily waveform. This could mean adding an artificial root, and noting it as such, should a valley intersect the first quantization level.

We are able to parameterize the tree decomposition by selecting from options such as normalization, baseline load removal, or alteration of the number of partitions desired in the quantizing step. The following information is kept for each daily waveform:

- parameterizing options
- day of the year
- day of the week
- maximum energy value
- minimum energy value

- total energy used for the day
- energy represented for the partition elements.

Very detailed information for each tree is also kept, including:

- the total number of chains for the tree
- the length of each chain (this equals the number of link nodes)
- the exact link sequence for each chain.

The following functions (explained below) are computed for each node:

- q_level
- ancestor
- time of day
- node load
- geo-order number
- postfix number.

The q_level is the quantization level that the node falls on. The root is always on the first q_level. The ancestor of a node j is that node for which node j is a son. There is always a unique ancestor. The time of day for a node is the projection of the node onto the horizontal (time) axis. The node load is the amount of time represented by each generating interval for the node. The geo-order node numbering scheme details the order in which the nodes are encountered beginning at the first q_level and moving up the waveform from left to right. The postfix node numbering scheme allows any node with all its generations of sons to be viewed as a subtree with each subtree having complete tree properties. The postfix numbers decrease from the root to the peak of any chain. Figure 4 illustrates a legitimate postfix ordering scheme. This numbering scheme allows us to count the number of transformations required to change one tree into another using matching tables (Lu 1984).

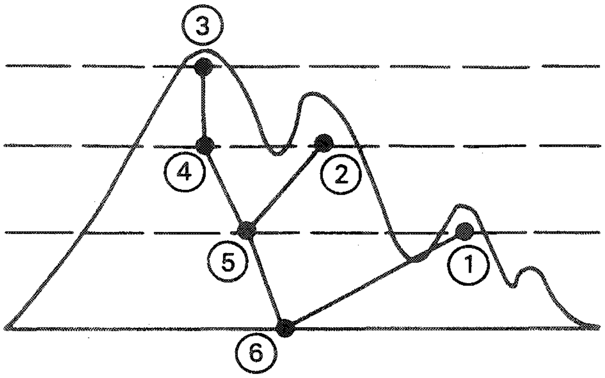


Figure 4. Postfix node numbering scheme.

Tree-like characterizations have several properties that lend them to analysis of load profiles. Among these parameters are:

1. It is possible to compare waveforms in a systematic fashion.

2. The level of quantization, and therefore the complexity of the characterization, is adjustable.

3. The characterization emphasizes the relative magnitude of successive features in the waveform and separates out the relative timing of those features to another data structure.

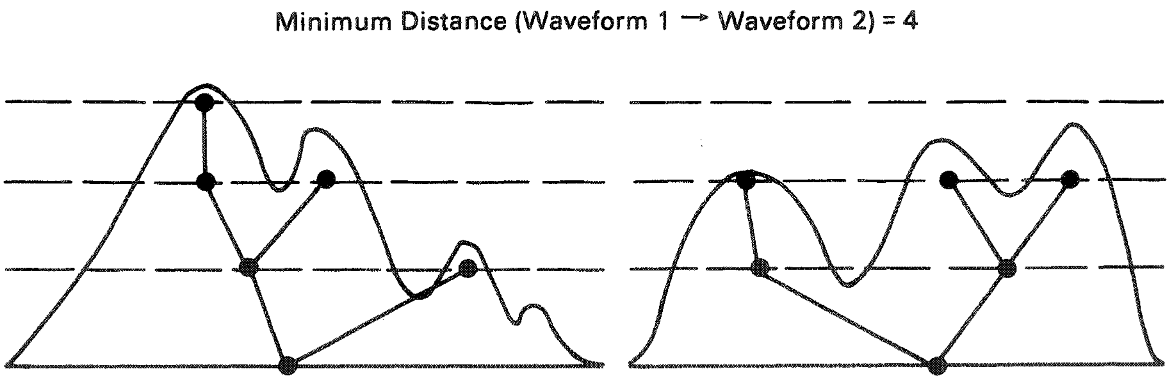
The most important of the three properties noted above is the ability to compare trees and therefore properties. The process of comparison is an attractive one. A measure of dissimilarity can be computed by calculating the number of transformations that are necessary to change one tree into another. The four allowed transformations as they have been described by Cheng and Lu (1985) are as follows:

1. The father-son splitting.
2. The father-son merging.
3. The brother-brother splitting.
4. The brother-brother merging.

This minimum distance computation and the four types of operations on tree nodes are illustrated in Figure 5. Waveform 1 is deformed to waveform 2 in a total of four steps. The smallest number of required transformations to deform one wave to another may not, however, require a unique sequence of the steps (1-4) above. In Figure 5c, one node has been added onto the node associated with the lowest quantized peak from Figure 5a--this a father-son split. One father-son merge reduces the chain of length four in Figure 5c to a chain of length three in Figure 5d. In a brother-brother split, one node is added at the same level, carrying all, some, or no nodes carried previously by the brother node. This split corresponds to a deepening of a valley in moving from Figure 5d to 5e. Conversely, the brother-brother merge corresponds to the shallowing of a valley denoted in the final step, Figure 5f.

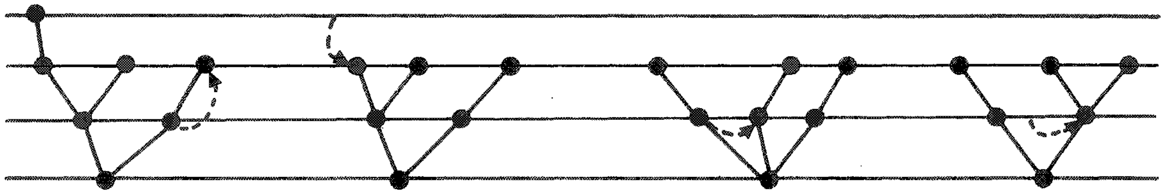
The dissimilarity index can be viewed as a distance between the two waveforms. The distance obtained in this fashion is quite different from the Euclidean distance used to compare to the channel level data in the multidimensional scaling analysis and cluster analysis described earlier. In spite of this difference, the techniques such as multidimensional scaling and cluster analysis can be applied to the resulting distance matrix in a way that is quite analogous to the earlier application of the methods.

Besides using the minimal distance computation as a measure of waveform similarity, consideration of the associated skeletal tree's total nodes, number of chains and length of the chains (with a normalizing option) allows for waveforms comparison independent of the amplitude and phasing considerations. Amplitude may be incorporated in one of several ways: either by fixing the amount of energy each quantization partition represents or by using a classification scheme that factors in the amplitude of each daily load shape in some way. Scheduling and on-time issues can be addressed with time of node and node load functions for specific chain elements. Contributing end-uses can be



Waveform 1
(a)

Waveform 2
(b)



Father-Son Split
(c)

Father-Son Merge
(d)

Brother-Brother Split
(e)

Brother-Brother Merge
(f)

Figure 5. Deformation using the four allowed transformations.

deconvolved from the more general end-use categories by fixing partitions to a particular energy level and incorporating the time of node functions into the skeletal trees.

Since waveform analysis is essentially pattern matching, the grouping of waveforms may be an extremely useful tool in solving the following types of problems:

- detection of the presence or absence of thermostat control behavior for a set of residences
- detection of behavior patterns that describe how specific end-uses are used, such as whether one or many loads of laundry are done on wash days
- detection of "day types," such as when the residence is unoccupied for several days.

Each of these problems can be reformulated in terms of particular waveforms to search for, or changes in overall waveform patterns.

Let us illustrate possible grouping strategies. In general, the exact strategy is driven by the problem one is attempting to solve. As a first pass at a grouping strategy, cluster analysis was applied to a subset of tree properties computed for a three month set of daily hvac waveforms for a single residence. The data are from March to May, 1985. During the 92-day period, five days were excluded due to at least on missing observation, and 14 days exhibited zero power for the hvac end-use under study. The level of detail was kept low as only four quantization levels were selected for the tree analysis.

The three tree properties selected for each day's tree were: total nodes, total chains, and the sum of the lengths of the secondary chains. The total nodes gives an overall gauge for the level of wave complexity. The total chains is exactly the number of quantized peaks, while the sum of the secondary chains is the sum of the nodes on all but the first chain. (The first chain's length is static due to the daily peak load normalization.) On the basis of these three characteristics, only 28 distinct waveforms emerge for the 73 days. The majority of the trees are relatively simple. The median node total for the 73 days is 7. The three parameters chosen do not uniquely determine the skeletal tree, but do provide much information concerning each tree's structure. The frequency of occurrence for each wave type and the characteristic properties are summarized in Table III. The last column, entitled group, gives the classification rating from the cluster analysis to be described below.

A matrix was then constructed with the columns from Table III for nodes, chains, and the sum of secondary chain lengths (ΣL_i). Each column was normalized prior to the construction of a Euclidian distance matrix. A cluster analysis was then performed on this distance matrix.

Four major groups emerge from the cluster analysis. These groupings are marked A-D on Figure 6. Table IV also shows the means and standard deviations for the three parameters used in the cluster analysis for all the waveforms in each of the categories A-D. There is a temptation to construct the "generic" waveform for each of the four classifications. For example, consider category C with the averaged parameters of nine nodes, five chains and 11 as the sum of the lengths of the secondary chains. Since each chain has a minimum of two nodes there are three additional nodes available for distribution in addition to the minimum two nodes for each of the four secondary peaks. These three additional nodes govern the relative heights of the secondary peaks in the "generic" waveform.

Due to the smaller range of values taken on by the number-of-chains parameter, it tends to dominate the cluster analysis. In Table III, when the last column is compared with the chains column, it shows a strong grouping tendency toward the number of quantized peaks. Furthermore, the three parameters selected do not include any ordering of the height of the peaks as would be desired in the ideal classification scheme. As different parameters, such as time of day for the first highest peak in the waveform, are to be incorporated in the analysis, another column is added to the matrix that drives the cluster analysis. Although the example presented is strictly for technique illustration and consists of only a single residence, a single end-use, and a single quarter of the

Table III. Skeletal decomposition for hvac data March - May 1986, single residence.

Waveform	Nodes	Chains	ΣL_i	# of Days	Group
1	4	1	0	13	A
2	5	2	2	8	A
3	5	2	3	1	A
4	6	2	2	9	A
5	6	2	3	1	A
6	6	3	4	3	B
7	7	2	3	7	B
8	7	3	5	3	B
9	8	2	4	1	B
10	8	3	5	2	B
11	8	3	6	2	B
12	8	4	6	3	D
13	8	4	7	2	D
14	8	5	9	1	C
15	9	3	7	1	D
16	9	4	8	2	D
17	9	4	9	1	D
18	9	5	9	1	C
19	9	5	10	2	C
20	9	5	11	2	C
21	9	5	12	1	C
22	10	4	8	1	D
23	10	5	11	1	C
24	10	6	12	1	C
25	11	3	8	1	D
26	11	4	9	1	D
27	11	4	10	1	D
28	11	5	13	1	C

Days not used------------------------------------ 5

Days with no power-------------------------------14

92

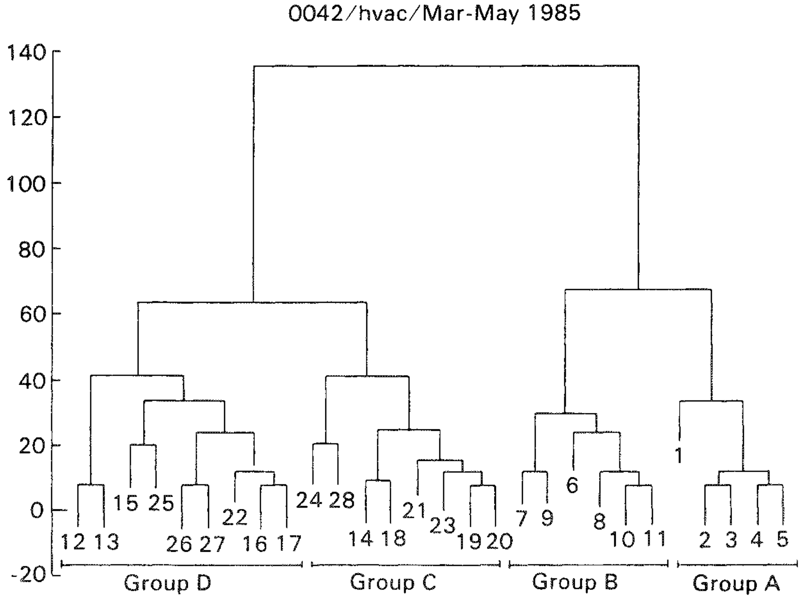


Figure 6. Cluster analysis for waveforms.

Table IV. Average parameters for the cluster analysis in Figure 6.

	Category A	Category B	Category C	Category D
Total Nodes	4.9 ± .9	7.1 ± .7	9.3 ± .8	9.2 ± 1.2
Total Chains	1.6 ± .5	2.6 ± .5	5.1 ± .3	3.8 ± .4
Sum of Secondary Chain Lengths	1.3 ± 1.1	4.1 ± 1.1	10.8 ± 1.3	7.6 ± 1.3

year, the technique need not be so limited in application. The approach is easily extended to go across end-uses, season, and residences.

One might wish to consider how the cluster analysis based on the three tree properties would compare to a grouping of waveforms based upon minimizing the dissimilarity indices of Cheng and Lu. In this scheme, waves are clustered when a very few number of steps are required to deform one waveform into another. The advantage of using the distance between waveforms as a method of picking the classification sets is that the ordering of the peaks and their heights is incorporated automatically. The cost for this is the failure to include phasing information, such as the time of day certain peaks take place. Also variance in the number of peaks is much less significant in this ordering method. In fact, if the four classes, A-D, that emerged from the previous cluster analysis are compared in terms of tree distance within their respective classes, we find the greatest distance to be four to six steps and the smallest distance within class to be one to two steps. The minimum distances outside of class for the 28 distinct waveforms varies from one to three steps. The maximum distance outside of class is three between wave four from class A and wave 18 from class C.

To further contrast the two grouping schemes, consider that the maximum distance of six, noted above, occurs between two class C waves that are 180 degree rotations, whereas for the three parameter cluster analysis these waves are not even distinct. The grouping of the waveforms is quite different clustering on the basis of the dissimilarity indices of Cheng and Lu rather than the three parameter approach. Of course, the best grouping strategy depends upon the problem one is attempting to address with the waveform analysis.

Our research has only begun to utilize the potential of skeletal trees as an aid to waveform classification. An interesting grouping scheme appears to be one that will incorporate both the distance information and the other parameters previously discussed. Current work is focused this direction. As we look to the future, tree analysis appears to be a powerful tool to both capture occupancy-driven differences in end-use waveforms and, most importantly, to quantify this occupancy effect that will facilitate cross-residence comparisons.

REFERENCES

Cheng, Y-C, and S-Y Lu. 1985. "Waveform Correlation by Tree Matching." IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-7, pp.-299-305.

Ehrich, R.W., and J.P. Foith. 1976. "Representation of Random Waveforms by Relational Trees." IEEE Transactions on Computers, Vol. C-25, pp. 725-736.

Lu, S-Y. 1984. "A Tree-Matching Algorithm Based on Node Splitting and Merging." IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-6, pp. 249-256.

Parker, G.B., E.W. Pearson, and W.F. Sandusky. 1985. "ELCAP: The Residential Pilot Study." DOE/BPA-13795-7, Bonneville Power Administration, Portland, Oregon.

Schuster, G.J., and S.D. Tomich. 1985. "ELCAP: Description of Field Data Acquisition System for Building Energy Monitoring." DOE/BPA/13795-4, Bonneville Power Administration, Portland, Oregon.

Stokes, G.M., and D.R. Owen. 1985. "Electrical Energy Use and Residential Activity Clusters." Presented at ELCAP Workshop II, Portland, Oregon, November 7-8, 1985.