Developing Load Shapes: Leveraging Existing Load Research Data, Visualization Techniques, and DOE-2.1E Modeling

Jedd L. Parker, RLW Analytics, Inc.

Developing statistically accurate end-use load shapes for marketing, evaluation, forecasting, and customer information has traditionally been an expensive, multi-year undertaking. A new method uses DOE-2.1E simulations calibrated with existing, transferred, or new metered data to develop prototype building models. Simulation results are then ‘‘expanded’’ to a population of program participants or market segment, to provide a utility with accurate 8760 load shapes. These load shapes are stored as models so that the utility can adjust the model inputs to investigate different scenarios or determine the impacts of many types of programs. The results are proving to be more accurate (when expanded to a population), much less expensive, much less intrusive, and quite a bit quicker than metering.

The methodology is being used in a number of projects, from impact analysis of residential new construction programs to developing of market segment loadshapes. One of the exciting things about this methodology is that much existing data can be used with little new field work required.

This paper focuses on the challenges, successes, and techniques that were developed to apply this method.

DATA LEVERAGING FOR LOAD SHAPE DEVELOPMENT

Traditionally, end-use load information has been generated using either energy simulation models of prototype buildings or end-use metering of a small sample of actual buildings. RLW Analytics, working with the Electric Power Research Institutes Center for Electric End-Use Data (EPRI/CEED) and a number of Utilities in Tailored Collaboration Projects has created a new methodology for developing 8760 total load and end use load shapes. The methodology integrates statistical sampling, whole-premise and end-use metering, site-specific DOE-2.1E modeling, and visual data analysis—supported by survey data and limited end-use metering.

The data leveraging methodology benefits the utility because the load shapes developed are:

- Developed more quickly than metered data—months instead of years.
- Much less expensive that metered data—$7,000 per site instead of $20,000.
- More accurate than metered data (more on this surprising statement later).
- Much more flexible—data is stored as models for easy What-If analysis.
- Developed using existing data with very little new data collection required.

This paper gives an over view of the data leveraging and presents the benefits to the utility of deploying this methodology. Examples are used throughout from a pilot project targeting the Commercial Office sector of British Columbia Hydro (BCH), but the underlying method is applicable to any evaluation that can benefit from accurate load shapes. Since this method relies on DOE-2.1E models, it should be valuable for base case compared to as built evaluations.

An Overview of the Data Leveraging Methodology

This methodology leverages existing billing, metered, building characteristic survey, and audit data and a sample of calibrated DOE-2.1E models to generate 8760-hour end-use load shapes with associated error ratios. Data leveraging is an application of the Engineering Calibration Approach (ECA™) (Townsley & Wright). In the data leveraging method (DLM), very accurate total load and end-use energy use information (typically hourly demand) for a sample of buildings is ‘‘expanded’’ to a target population using supporting audit, characteristics, and billing information.

Effective sampling and statistical analysis techniques are necessary for reliable results. The DLM uses:

- Statistical sampling—to minimize selection bias and provide measurable precision,

- Stratification—to control the size and distribution of buildings in each sample,

- Ratio or regression estimation—to link the results of each level of the sample design to supporting information from lower levels, and

- Optimal design—to allocate a suitable fraction of total resources to each level of the sample design.

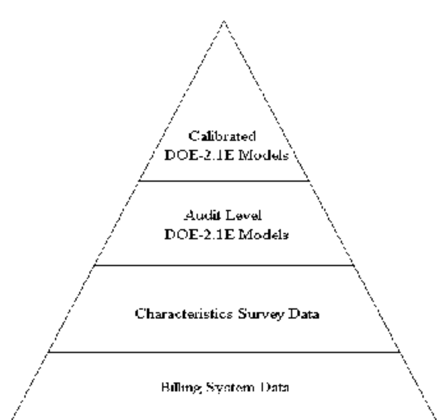
The Layers of Data. Figure 1 illustrates four tiers of data that are utilized in the DLM. The data structure is pictured as a pyramid since the sample at each level is regarded as nested within the lower samples.

The base of the pyramid is the billing data available for all customers. The next level is the characteristics sample, comprised of data that provides basic information about building operation, fuel types and equipment stock. The third level of data is the DOE-2.1E models built based on audit information. This is the first level of DOE-2.1E models. The fourth level of data is the subset of DOE-2.1E models that are calibrated to total load. There may be a fifth level, if a subset of the total load calibrated models can be calibrated to end use metered data.

The Calibrated DOE-2.1E sample provides the peak of the pyramid. This level provides the best practical results for a relatively small sample of buildings. Here, higher unit costs of the more detailed simulation and calibration are offset by smaller sample sizes.

The Expansion. A strategy is required for combining and leveraging the information from the various layers of the data hierarchy. Figure 2 illustrates the analysis methodology. The billing data is used to develop information about kWh sales by SIC-coded market segment, and the survey data is used to develop square footage information. The analysis adjusts for potential bias arising from the fact that the billing data is at the account level whereas the survey and modeling data are at the premise level. The analysis also adjusts for SIC-coding misclassification. Together the billing and survey data provide an estimate of the total square footage of each market segment, together with information about the distribution of square footage among premises in each market segment.

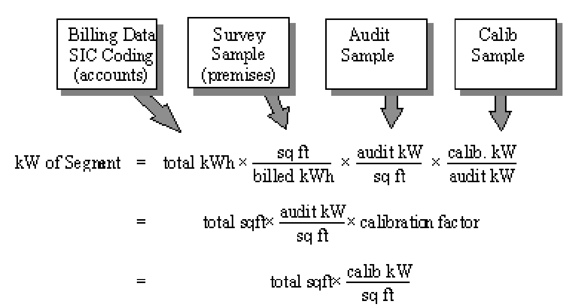
Figure 1. Hierarchy of data



Next the audit-level DOE-2.1E models developed for each site in the audit sample are used to estimate the 8760-hour end use load profiles per square foot for each targeted end use. Each audit-sample model is used to generate site-specific 8760-hour end use loads which are extrapolated to the target market segments using the survey data. The square footage of the audit-sample sites is also extrapolated to the target market segments and stratified ratio or difference estimation is used to calculate the end use wattage per square foot.

Finally, the calibration-level DOE-2.1E models developed for each site in the calibration sample are used to estimate 8760-hour calibration factors for each targeted end use. The calibration factors are used to correct the audit-sample results for any systematic bias identified from the metered data. The calibration factors are developed by using another application of ratio or difference estimation. In this case, the calibration-level and audit-level end use profiles are both expanded from the calibration sample to the target market segments using the survey data, and the calibration factors are calculated as the ratio between the end use demand from the calibration-level models divided by the end-use demand from the audit-level models. All results are developed for 8760 hours, for each targeted end use.

Figure 2. Analysis methodology



Essentially the DLM allows one, by leveraging the nested samples and using rigorous statistical sampling and ratio or difference estimation techniques, to develop load shapes for market segments, or other well defined populations, with greater accuracy than possible with a typical metering project.

DOE-2.1E Calibration, Data Visualization and Goodness of Fit Statistics

The most detailed data is developed using DOE-2.1E models calibrated to total load and/or end use data.

Traditionally DOE-2.1E models have been calibrated to monthly billing data—peak demand and energy usage. The problem with the traditional method is that serious modeling errors may be mutually offsetting and not apparent at the monthly energy level. An EPRI report (EPRI, 1992 *Engineering Methods for Estimating the Impacts of Demand-Side Management Programs*) expresses concern that under-predictions for one end use may cancel out over-predictions for another end use, resulting in simulations that closely match monthly energy use but incorrectly describe actual hourly end-use demand.

A recent ASHRAE Journal article (Kreider & Haberl) suggested that graphical VDA techniques together with standard statistical measures of goodness of fit can be used to calibrate model predictions to whole-premise load and end-use metered data. Our experience confirms this suggestion.

Goodness of Fit Statistics. The following is a brief description of the goodness of fit statistics used. Mean Bias Error (MBE) takes the Mean of the residual load (residual load = metered—DOE-2.1E for each interval) and divides it by the Mean of the metered data. Root Mean Square Error (RMSE) is the square root of the mean of the square of each interval of the residual load. Coefficient of Variation of the Root Mean Square Error (CV(RMSE)) is the RMSE divided by the mean of the metered data. In all cases intervals from both data sets where there is missing metered data are excluded from the calculations. These metrics are developed on a monthly and annual basis. Table 1 shows a sample of the goodness of fit statistics generated.

In this example, the models fits the metered data well as measured by the RMSE which indicates the average difference between the two is ±5.22%. In general, a CV(RMSE) of 0.2 or below is very good. The CV(RMSE) here is a little high, indicating a little more variation from hour to hour than is desirable. In this case the modeler would probably try one more round of calibration to attempt a better CV(RMSE), as long as the RMSE did not increase.

Table 1

	MBE	RMSE	CV(RMSE)
Annual	0.0269124	5.22 %	0.313301
Jan.	−0.0402185	2.88 %	0.238287
Feb.	0.0262515	5.14 %	0.383433
Mar.	−0.0446691	3.27 %	0.244788
Apr.	0.0244103	5.86 %	0.36444
May	0.0269487	5.40 %	0.305624
Jun.	0.0205838	5.37 %	0.252675
Jul.	−0.00816728	5.69 %	0.290635
Aug.	0.107781	6.34 %	0.270721
Sep.	0.100282	7.03 %	0.339305
Oct.	0.0859239	5.24 %	0.319142
Nov.	0.0115682	4.48 %	0.319562
Dec.	−0.102058	4.45 %	0.383976

Model Calibration. We have found that several iterations are necessary to calibrate each individual model. The strategy is to avoid making a single new model incorporating several changes. Instead, the model is changed incrementally as suggested by the following guidelines:

Base Case: Select or create prototype model.
Iteration 1: Apply the ‘‘obvious and easy’’ modifications.
Iteration 2: Apply the ‘‘obvious but not so easy’’ modifications.
Iteration 3: Apply the ‘‘not so obvious’’ modifications.

After each iteration, the modeler compares the DOE-2.1E 8760 output with the available metered data and notes areas where the model is over or under predicting the metered data. The goodness of fit statistics are recorded for each iteration to know when a model has been calibrated to an acceptable level. There is a point of diminishing returns in model calibration and it appears to be when the CV(RMSE) falls below 0.2.

Model calibration requires experience and skill in working with DOE-2.1E. It is necessary to avoid or work around known quirks in DOE-2.1E. A certain level of familiarity and expertise is also needed to represent certain building characteristics when areas of a building and/or system are inaccessible—as is often the case. Moreover, it is necessary to avoid ‘‘over-modeling’’ a site by making time-intensive changes which have a relatively small effect.

Visual Data Analysis. This methodology uses VDA techniques as an integral part of the calibration process. VDA techniques can range from line graphs comparing monthly peak demands for model data vs. metered data to load duration curves, to color renditions of 8760 load shapes. These techniques have been described by Parker and McCray, Bailey, Gillman and Parker. We found that the following capabilities are very helpful when comparing modeled data to metered data in a VDA tool:

- Comparing monthly total usage and peak demand.
- Compare average load shapes for analyst selected periods.
- Average week end and week day load shapes for each month.
- Interactive exploration of single day load shapes.
- An 8760 hour residual plot of the difference between the metered data and the modeled data.

Using VDA as part of the model calibration process maximizes the value of readily available whole building total load data. Visual data analysis also provides rapid, readily understandable feedback to the building modeler and allows for interactive exploration of the modeling results.

Figure 3 shows a screen capture from a VDA tool that shows a number of things:

- Load shapes for the peak day in the upper left corner.
- Goodness of fit statistics in the upper right corner for the total load.
- 8760 profiles for the metered total load, DOE-2.1E total load, residual load, and many of the DOE-2.1E end uses (including temperature). The lighter areas indicate higher demand, the x-axis is hours, and the y-axis is days.

THE EXAMPLE CASE

The Data Leveraging Methodology has been successfully employed to develop market segment or program level load shapes for a number utilities. The results of the pilot project is presented here as an example of the data leveraging methodology. It was a tailored collaboration project with CEED and BC-Hydro that featured some end use metered data from BCH and some end use data ‘‘loaned’’ from CEED. The specific task was to develop 8760-hour profiles and confidence intervals for BCH’s office sector for the following ten end uses:

- Interior Lighting,
- Exterior lighting,
- Electric Cooling,
- Electric Heating,
- Auxiliary (pumps and ventilation),
- Miscellaneous Equipment (plug loads),
- Transportation (elevators),
- Domestic Hot Water,
- Refrigeration, and
- Cooking.

This project sought to answer whether these end-use load profiles can be effectively developed using end-use-metered data borrowed from another service area. Specifically, the project sought to address the following methodological questions:

(1) **DOE-2.1E Modeling**—Can reliable estimates of end-use loads be developed using DOE-2.1E modeling informed by on-site audits whole premise metering, and end use metering?

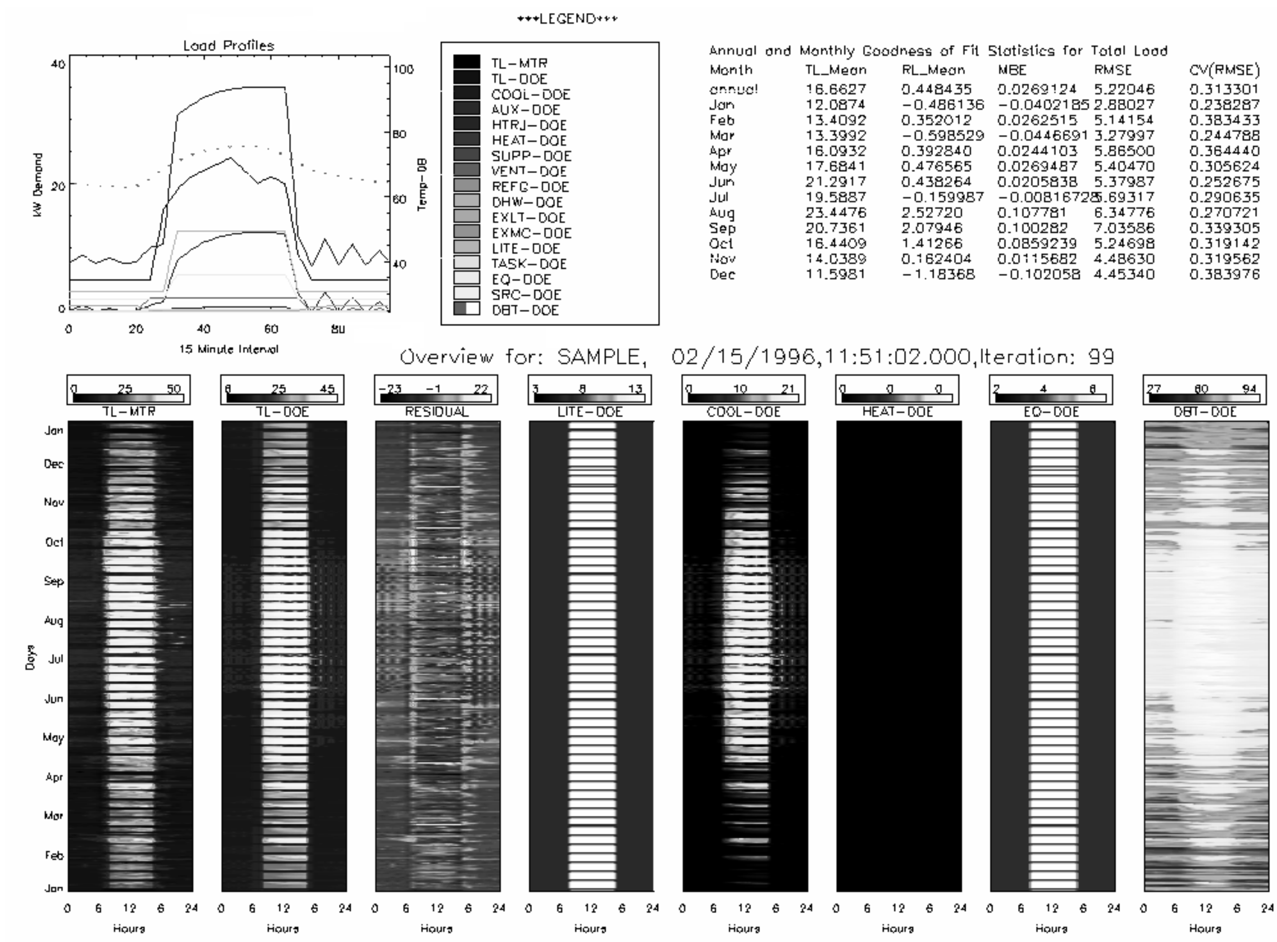
(2) **End-Use Metering**—How useful is end-use metering in refining these models? How useful is borrowed end use metering data compared to a utility’s own end use metering data?

(3) **Data Leveraging**—Can billing, survey, and load research data be integrated to improve end-use estimation?

(4) **Transferring Models versus Data**—Do DOE-2.1E models provide a more useful product than the end use data itself?

The broader objectives of this project included the evaluation of the likely effectiveness of the DLM for BCH’s additional commercial segments.

Figure 3. *VDA screen capture*



What Data Was Available?

Data about the Commercial Office Segment was used from BCH, as well as ‘‘donated’’ end-use data from CEED. Table 2 lists the data and sources.

In this example, the audit level models were informed by audits, and calibrated to total load metered data for the Load Research Department. The calibration level models were calibrated to metered end-use data. In the case of the 9 sites donated from CEED, the information from the buildings (billing, characteristics, audit, and total load) in their original location (Seattle, WA) were nested within the sample. The CEED sites were selected from all of the available CEED sites with similar weather characteristics, using the same sampling methodology as that used to pick the sites from BCH. The samples were drawn from 5 strata, based on Mwh.

Table 2.

Data	Amount	Source
Billing	27,000 accounts	BCH
Characteristics Survey	213 accounts	BCH
Audit Level Models	29 sites	20 BCH, 9 from CEED
Calibration Level Models	13 sites	4 from BCH, 9 from CEED

What did the Utility Get?

In this project, 8760-hour load profiles were developed for each of ten end uses. Four day types were used to summarize the results:

(1) Average January weekday

(2) Average January weekend

(3) Average August weekday

(4) Average August weekend

Figure 4 shows the 24-hour end use profiles for the average January weekday. This profile is of particular interest to BCH since BCH is a winter peaking utility. The results were projected to the BCH office population and reported in units of watts per square foot. The figure shows, for example, that the peak lighting load is almost 2 watts per square foot while the plug load is about 1 watt per square foot. The electric heating load, averaged over all offices, is also almost 1 watt per square foot, reaching its peak at 7 am. For each of the end uses, the total demand in the office segment can be calculated by multiplying the watts per square foot by the total square footage in the office segment, estimated from the survey and billing data to be 107 million square feet. All results are based on typical meteorological year (TMY) weather for calendar year 1994.

Confidence Intervals. 90% confidence intervals were calculated for the total load and for each of the major end uses for each of the four day types. Figure 5 shows the January weekday profile for the total office load.

The hourly confidence intervals vary across the day, averaging about $\pm 30\%$ during the night time hours and approximately $\pm 17\%$ during the day light hours. This is due to the fact that all of the metered data calibration sites had relatively large loads during the day, but many had very little load at night, thus effectively reducing the sample size for the night time hours, and widening the confidence intervals.

__Figure 4.__ January weekday end-use loads

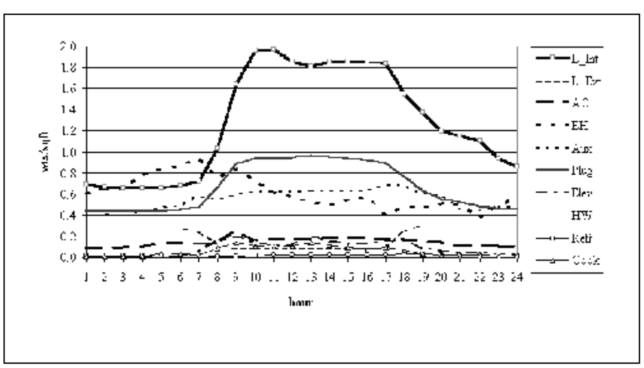
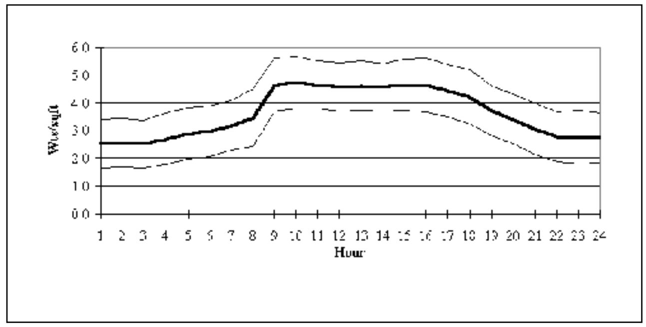


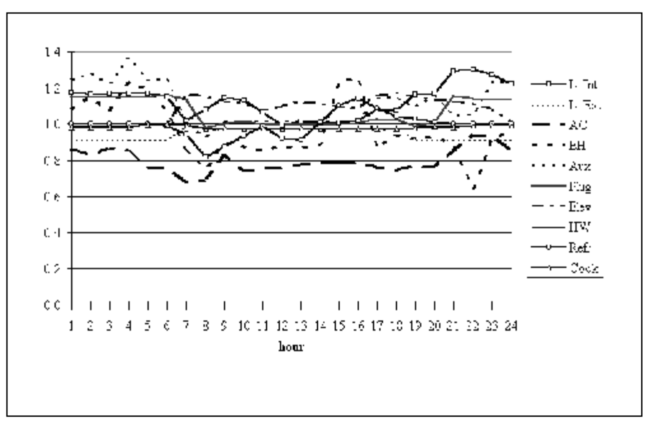
__Figure 5.__ Confidence interval for total load, January weekday



Calibration Factors. Calibration factors were calculated to anchor the DOE-2.1E models to the end-use-metered data available from CEED and BCH. These factors indicate the potential bias in the end-use results developed from the audit-level DOE-2.1E models developed without end use metering. A factor of 1.1, for example, would indicate that the end use load estimated from the calibration-level DOE-2.1E models is 10% higher than the end use load estimated from the audit-level DOE-2.1E models. The final results were obtained by adjusting the audit level results by the calibration factors.

Figure 6 shows the calibration factors for each end use for the average January weekday. During the business hours, 8–18, the calibration factors were generally in the range 0.8 to 1.2, indicating very little if any bias. During the night hours, the calibration factors were somewhat larger and more variable, reflecting the greater difficulty in identifying the level of actual use during these hours without end-use metering.

__Figure 6.__ Calibration factors for January weekday



Calibration Sample Error Ratios. Besides the calibration factors, additional statistics called calibration-sample error ratios were calculated to summarize the relationship between the audit-level and calibration-level DOE-2.1E models. Whereas the calibration factors measured bias in the kW loads from the audit-level models relative to the calibration-level models, the calibration-sample error ratios measured the strength of the association between these results. A small calibration-sample error ratio would indicate that the audit-level end-use load for a given hour was an accurate predictor of the calibration-level end-use load for the same hour within the sample of calibration sites, after applying the calibration factor to correct for any systematic bias.

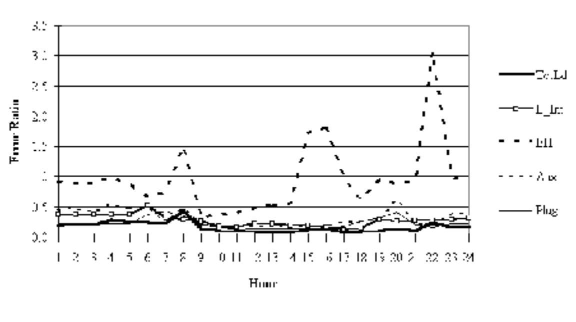
Figure 7 shows the calibration-sample error ratios for total load and the four major end uses. With the exception of electric heating, the error ratios were very small throughout the average January weekday typically in the range 0.1 to 0.2. The error ratios for electric heating were much larger, typically in the range 0.4 to 1 but reaching 3 at hour 22. This indicated that, except for the electric heating component, the loads from the audit-level models were quite accurate predictors of the loads from the calibration-level models, especially during the hours of high demand. The electric heating component error ratios are so high due to the large variability from site to site in electric heating scheduling.

Findings

In the introduction, a number of claims were made regarding the benefits of the data leveraging methodology. Specifically they were that the load shapes were:

- Developed more quickly than metered data—months instead of years.
- Much less expensive that metered data—$7,000 per site instead of $20,000.
- More accurate than metered data (more on this surprising statement later).
- Much more flexible—data is stored as models for easy What-If analysis.
- Developed using existing data with very little new data collection required.

Figure 7. *Calibration-Sample error ratios, January weekday*



Cheaper, More Accurate Load Shapes. Per unit costs and error ratios are at a minimum using the data leveraging methodology. Table 3 shows a conservative example. Assume error ratios at the upper end of the observed range, i.e., 0.6 for the audit-level sample and 0.3 for the calibration-level sample. The indicated sample sizes would be 42 sites in the audit-level sample and 15 sites in the calibration-level sub-sample. The 42/15 site sample design using DOE-2.1E modeling would have a data development cost of $180,000. By contrast, traditional end use metering would require a sample of 30 sites for a cost of $600,000.

These cost and error ratio comparisons are only hypothetical and will vary depending on the circumstances of each utility. Nevertheless, these examples indicate that the approach demonstrated here may yield substantial savings, generally 50% or more.

Reduced Bias. Conventional end-use metering may be exposed to potentially serious bias. The sample can be selected to favor customers that are thought to be receptive and sites that are expected to be relatively easy or valuable to monitor. Circuit and equipment layouts can make it impractical to monitor end uses separately, completely, and consistently from one site to another. The danger of bias

Table 3.

Approach	Error	Unit	Sample	Total
Sample	Ratio	Cost	Size	Cost
DOE-2.1E Modeling				
Audit	0.60	$2,500	42	$105,000
Calibration	0.30	$5,000	15	$75,000
Total				$180,000
Conventional EUM	0.67	$20,000	30	$600,000
Savings				70%

from these and other causes can be reduced through the techniques demonstrated in this methodology, especially VDA for model-calibration and double ratio estimation to link the EUM to the larger supporting samples.

Better Statistical Precision. With lower unit costs and less customer intrusiveness, samples can be large enough to provide statistically reliable results. This method integrates information from nested samples to achieve statistically reliable results.

Flexibility. Since the data is delivered as models as well as numerically and with the right tools and training to expand the modeling results, there is a tremendous amount of flexibility that is not normally found in load research data. Typical load research data is a snapshot of a buildings performance, that is out of date by the time the data is cleaned, assessed and made available in a usable format. With the data leveraged data the utility can change key parameters in the models as the appliance stock and saturation changes, or change for current year weather conditions, or any number of ''What If'' analysis.

Test Scenarios—Play ''What If''. ''What if'' analysis can be conducted using the DOE-2.1E models to assess the impact of technological changes, fuel-switching, DSM measures, etc. The profiles can be easily weather normalized by rerunning the DOE-2.1E models using typical meteorological year (TMY) weather files. The models can also be updated over time and transferred to other service territories.

Quick Turnaround. A typical load research project to develop end use data takes at least 1 year of metering. This does not include the time it takes to clean the data, plan the project, or convert the data into usable information. With the data leveraging approach, the end use data can be generated in less than one year.

LESSONS LEARNED

We have learned that it is possible to accurately model a site's end-use demand during hours of high demand. By contrast, during periods of light demand, it is more difficult to model the end uses accurately without information from some form of end-use metering, especially for HVAC.

This approach is most effective if a load research sample is available that was designed and selected with energy modeling in mind. To facilitate DOE-2.1E modeling, load data should be collected at the site level rather than at the account level. In other words, all meters serving a site should be monitored whenever possible. In addition, the sample should be suitably stratified by annual use so that large sites are oversampled relative to small sites. It may also be useful to stratify the load research sample by market segment to ensure that each segment is adequately represented.

Visual data analysis techniques prove to be a very powerful way to examine all 8760 hours of data simultaneously and interactively, allowing the modeler to recognize the characteristic signatures of various end-use loads and schedules and to refine the models quickly and appropriately.

End-Use Metering

Both transferred and original end-use metering can be used to calibrate the DOE-2.1E models.

The use of transferred end-use metering is worthwhile and provides the basis for the preceding conclusions, but it is not a free ride. Problems that may be encountered include:

- The available audit information may not provide adequate information about the operation and control of equipment because it may not be intended to support modeling but rather to provide a general description of the site and its relevant end uses.

- Available audit information may not be timely. Due to the time lag between original audits and metered data, discrepancies may appear which can only be explained by changes in the building(tenants, equipment, schedules, etc. This underscores the importance of allowing the modeler access (directly or indirectly) to the site being modeled.

In addition, both transferred and original end-use metering has proved to be of less value than expected. Because of the arrangement of circuits and other practical considerations, a significant fraction of the whole-premise load may not be end-use monitored. In addition, end-use metered channels may included mixed loads. Thus, the end-use metering is often more informative about end use schedules and control strategies than actual kW levels. The VDA techniques can provide an effective way of extracting the information from the end use metering and creating a consistent decomposition of the whole-premise load into the component end uses of each site.

Based on this experience, model calibration seems to require an innovative approach to end-use monitoring. The conventional approach has been to try to end-use meter all significant loads (e.g., greater that 5%) in each building in a selected sample. Once the monitoring equipment is installed, data are usually collected for several years. Instead, the experience of this project suggests that future end-use monitoring be undertaken only after the initial DOE-2.1E modeling and comparison to whole-premise load data. Monitoring should be primarily used to reconcile problems between the model and whole-premise load data, or to validate key features of the model. For most of this work, spot measurements or short-term monitoring would be adequate. The whole-premise load data should be relied on for most longer-term information. The method of double sampling should still be used to control the cost of this type of monitoring.

The Commercial Survey

An excellent commercial survey is a joy to have available. Together with billing data, the survey is used to characterize the population square footage(minimizing multiple account bias, correcting for SIC misclassification and providing case-weights for the audit and calibration samples. A survey of this type must be considered an integral component of a comprehensive strategy for understanding the target population and for developing detailed end-use information through monitoring and modeling.

Some difficulties may be encountered in obtaining current billing information for the survey respondents, either because of occupant changes or problems in account matching. This may reduce the size of the final survey sample quite substantially. A better approach may be to use billing data matched to each site at the time of the survey. In addition, energy intensity (kWh per square foot) should be used as a consistency check on both the billing data and the reported square footage of each site.

It has generally worked well to use a single measure of the square footage for the site for all end uses but consideration should be given to employing a separate survey variable for each of the primary end uses, e.g., the square footage for interior lighting, the square footage of air conditioned space, etc. In addition, the survey instrument should be designed to avoid double counting square footage when a building contains two or more premises.

CONCLUSION

As the utility industry is deregulated the value of flexible, precise, inexpensive market segment information will increase. We believe that the data leveraging method presented here is a valuable new customer information tool for the utility industry. The data leveraging method provides accurate, cheap, quality data in a fraction of the time of a traditional load research project.

REFERENCES

Townsley & Wright 1990. Measuring DSM Impacts: End-Use Metering and the Engineering Calibration Approach. Paper presented at the End-Use Information and its Role in DSM Conference.

EPRI, 1992 *Engineering Methods for Estimating the Impacts of Demand-Side Management Programs*, Volume 1, Palo Alto, CA., Electric Power Research Institute.

Kreider & Haberl 1994, ‘‘Predicting Hourly Building Energy Usage’’, *ASHRAE Journal*: 72–81.

Parker 1994 ‘‘Using Data Visualization to Better Understand Electric Load and End-Use Data’’, *In Proceedings of the ACEEE 1994 Summer Study on Energy Efficiency in Buildings*, 2:225–228. Washington D.C.: American Council for an Energy-Efficient Economy.

McCray, Bailey, Gillman and Parker. 1995, ‘‘Using Data Visualization Tools for the Calibration of Hourly DOE-2.1E.1 Simulations,’’ *In Proceedings of Building Simulation ‘95, Fourth International Conference*. 461–466. Madison Wisconsin: International Building Performance Simulation Association.