DEVELOPING LOADSHAPES: NEW TECHNIQUES USING EXISTING DATA TO DEVELOP ACCURATE MARKET SEGMENT LOADSHAPES

Jedd L. Parker, RLW Analytics, Inc., Sonoma, CA
James McCray, RLW Analytics, Inc., Sonoma, CA
Roger L. Wright, RLW Analytics, Inc., Sonoma, CA
Ilene Obstfeld, EPRI-CEED, Portland, OR

Data Leveraging For Load Shape Development

Total load and end-use loadshapes for entire market segments will become increasingly important as the utility industry de-regulates. Market segment loadshapes can be used to develop pricing structures that are market segment specific, or investigate the profitability or cost to serve a particular market segment. Traditionally, total and end-use loadshapes have been generated using either energy simulation models of prototype buildings or end-use metering of actual buildings then assuming those loadshapes represented an entire market. RLW Analytics, working with the Electric Power Research Institute's Center for Electric End-Use Data (EPRI/CEED) and a number of utilities in Tailored Collaboration Projects has created a new methodology for developing full year hourly (8760) total load and end-use load shapes *for market segments.* These are known as ReShape projects. This paper presents this new methodology which integrates statistical sampling, whole-premise and end-use metering, site-specific DOE-2.1E modeling, and visual data analysis.

The data leveraging methodology benefits the utility because the market segment level load shapes are:

- Developed more quickly than metered data - months or weeks instead of years.
- Much less expensive than metered data - $7,000 or less per site instead of $20,000.
- More flexible - data is stored as models for easy What-If analysis.
- Usually developed using existing data with very little new data collection required.

Market segment level loadshapes are loadshapes that describe how a group of customers use energy. The group can be based on SIC code and annual kWh (groceries with less that 500,000 annual kWh), ownership (all the "Super-Save" groceries in a territory), or another variable like program participation. The market segment level loadshape provides valuable information on program impacts, peak demands, and profitability for groups of customers.

This paper gives an over view of the data leveraging process and presents the benefits to the utility of deploying this methodology. Examples are used throughout from projects with BCHydro, Salt River Project, and a group of southeastern utilities that made up the Southeast Data Exchange.

An Overview of the Data Leveraging Methodology (DLM)

This methodology leverages existing billing, metered, building characteristic survey, and audit data and a sample of calibrated DOE-2.1E models to load shapes with associated error bounds (error bounds show, on an interval basis, how closely a different stratified statistical sample would project to the same population). Data leveraging is an application of the Engineering Calibration Approach (ECA™) (Townsley & Wright.) In the data 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 often utilized in the DLM. The data structure is pictured as a pyramid since the sample at each level is nested within the lower samples. The base of the pyramid is the billing data available for all customers in the group of interest. 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 are the DOE-2.1E models 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.

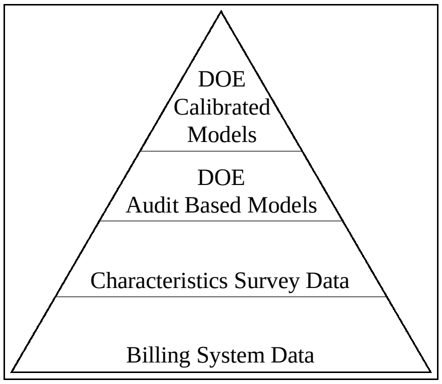


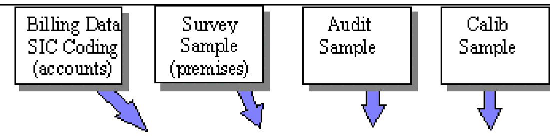
Figure 1. Hierarchy of Data

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.

<u>*The Expansion.*</u> 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. For example, we usually aggregate billing data based on a location identifier in the billing data. The analyst must also be aware of the potential for SIC-coding misclassification. The assignment of the market segment that a DOE model belongs to should be based on the segment the site is assigned in the billing data. 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.

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 for each targeted end use. Each audit-sample model is used to generate site-specific 8760-hour end use loads that 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 segment 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. Using an application of ratio or difference estimation develops the calibration factors. 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.



$$\text{kW of Segment} = \text{total kWh} \times \frac{\text{sq ft}}{\text{billed kWh}} \times \frac{\text{audit kW}}{\text{sq ft}} \times \frac{\text{calib. kW}}{\text{audit kW}}$$

$$= \text{total sqft} \times \frac{\text{audit kW}}{\text{sq ft}} \times \text{calibration factor}$$

$$= \text{total sqft} \times \frac{\text{calib kW}}{\text{sq ft}}$$

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 Visual Data Analysis (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 the residual load for each interval. 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.

Table 1. Goodness of fit statistics

	MBE	CV(RMSE)
Annual	2.69%	31%
Jan.	-04.02%	23%
Feb.	02.62%	38%
Mar.	-04.46%	24%
Apr.	02.44%	36%
May	02.69%	30%
Jun.	02.05%	25%
Jul.	-00.81%	29%
Aug.	10.77%	27%
Sep.	10.02%	33%
Oct.	08.59%	31%
Nov.	01.15%	31%
Dec.	-10.20%	38%

In this example, the model fits the metered data well on average as measured by the MBE which indicates the average difference between the two is ± 2.7% for the year and 11% or less for each month. In general, a CV(RMSE) of 20% 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 another round of calibration to attempt a better CV(RMSE), as long as the MBE did not increase.

Model Calibration

We have found that several iterations are necessary to calibrate each individual DOE2.1E 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 show 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 20%.

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 that have a relatively small effect.

Visual Data Analysis

Our 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. Parker and McCray, and Bailey, Gillman and Parker have described these techniques. 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 summer, winter, and shoulder periods.
- Average weekend and weekday 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.

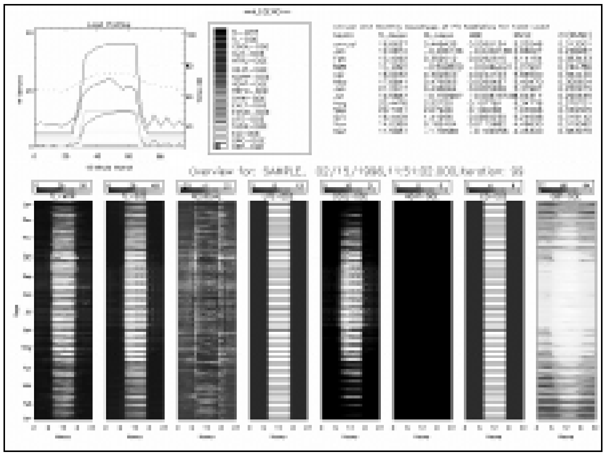


Figure 3. VDA screen capture

The Examples

We have used this methodology to develop segment level loadshapes in a number of projects. The projects have ranged from the pilot project where we investigated the use of “donated” end-use metering, to projects where large numbers of audit level models have been computer generated. This section discusses some of the lessons we have learned in those various projects.

BC-Hydro

For this project we developed loadshapes for BC-Hydro’s (BCH) office segment. The challenge was that BCH had very little applicable end-use load research data. Metered data provided by CEED from a northwestern utility was used to bring the sample size to a suitable level. To develop prototype models, we used audits and whole premise metering provided by BCH to develop audit level models. We then combined the audit level models with end-use level models developed and calibrated using the donated end-use data. Then the models were re-run with the correct weather data. We were then able to say that the twenty nine office models, when properly stratified, accurately represented the 26,000 offices in the target BCH service territory. There is an EPRI report on this project titled “Leveraging Limited Data Resources: Developing Commercial End-Use Information”.

Figure 4 shows the loadshapes for a number of end-uses for a typical August weekday. Figure 5 shows the 90% confidence interval for interior lighting (the largest end-use loadshape shown in Figure 4) for the same time period.

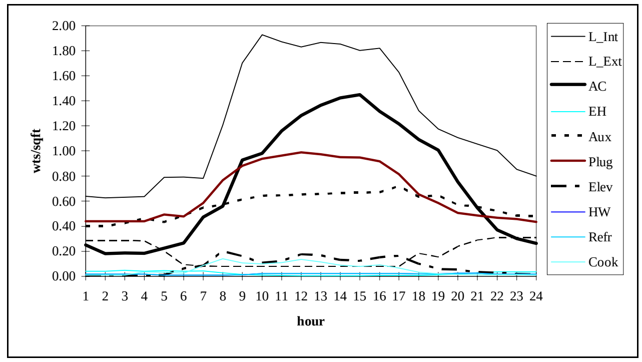


Figure 4. Office Load Shapes for Typical August Weekday

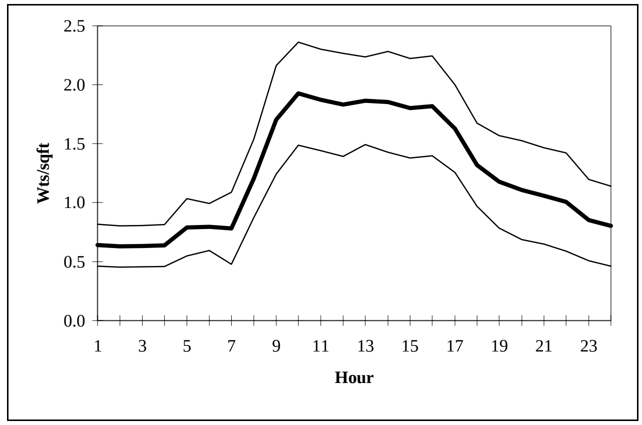


Figure 5. Office Interior Lighting Load Shape for Typical August Weekday with Confidence Interval.

The BC-Hydro project successfully demonstrated the validity of the data leveraging methodology.

<u>*Salt River Project*</u>. For Salt River Project (SRP) one of the goals was to provide the ability to rapidly test “What-If” scenarios. We used the data leveraging methodology to develop loadshapes for the Grocery, Hotel, School, Retail, Restaurant and Hospital segments. We then met with their analysts and investigated a number of What-If scenarios. Over the course of two days of training, we investigated three scenarios.

A scenario investigated was to determine both the energy and financial impact of a program targeted at the largest groceries that would result in a 30% drop in the energy use of lighting and motors in refrigerated cases. The results are shown below in Table 2.

Table 2. What-If Results

Item	Reduction	Relative Precision
Avoided kW	3.35 MW	11%
Avoided kWh	29.4 GWh	14%
Cost to Utility	$ 701,000	9%

We determined that the program would save 29.4 GWh in annual energy, and 3.35 MW in demand (winter) at a cost to SRP of $ 701,000 (this is the margin that would be lost, not the cost of the program).

South East Data Exchange

Duke Power had completed a large metered end-use study. They agreed to share the results with a group of eleven southeastern Utilities. In this project we developed models based on the Duke end-use data then modified a sample of the models to reflect utility specific characteristics data, then re-ran each model with the appropriate weather data for each utility. The results were then expanded to eleven market segments for each utility. A unique thing about this project was the absence of square footage data that until now had been used as the "expansion" variable. In the absence of square footage data, we used annual kWh as the expansion (or explanatory) variable. This prompted a small internal investigation where we found that using kWh as the expansion variable instead of the square footage led to slightly better relative precision for total load and some end-uses (like refrigeration), but for end-uses like lighting the relative precision suffered a little. Given the problem of measuring square footage accurately at the survey level, we have continued to explore the use of annual energy use (kWh) as the explanatory variable.

PG&E / SCE Commercial New Construction

In this implementation of the methodology we projected many iterations of a sample of DOE2 models to both the participant population and a similar population of non-participants, and compared the results to determine the savings due to the program as a whole and for various end-uses within the program. In this project we also demonstrated the effectiveness of an automated procedure for converting audit data into DOE2.1E models.

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 than metered data - $7,000 or less per site instead of $20,000.
- Much more flexible - data is stored as models for easy What-If analysis.
- Developed using existing data with very little new data collection required.

Quick Turnaround

A typical load research project to develop end use data takes at least one year of metering, and often another year or two 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 months rather than year. The longest project listed here took less than one year to develop the load shapes.

Cheaper, Accurate Load Information

Per unit costs and error bounds are at a minimum using the data leveraging methodology. Table 3 shows a conservative example. Assume error bounds 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 forty-two sites in the audit-level sample and fifteen 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 thirty sites for a cost of $600,000.

These cost comparisons are only hypothetical (but based on experience) 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.

Table 3. Sample sizes and costs assuming high error ratios

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

<u>*Reduced Bias*</u>**.** 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 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 end-use metered data to the larger supporting samples.

<u>*Better Statistical Precision*</u>**.** 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

The data is delivered as models as well as numerically. 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 model generated 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, energy service 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.

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 low 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 end-use metering (local or transferred) 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 include 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 and/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 extremely important. Together with billing data, the survey is used to minimize multiple account bias, correct for SIC misclassification and provide 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.

If the survey was done some time ago, 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

1. 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.
2. EPRI, (1992) *Engineering Methods for Estimating the Impacts of Demand-Side Management Programs*, Volume 1, Palo Alto, CA., Electric Power Research Institute.
3. Kreider & Haberl (1994), "Predicting Hourly Building Energy Usage", *ASHRAE Journal*: pp. 72-81.
4. Parker, Jedd L., (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*, pp. 2:225-228. Washington D.C.: American Council for an Energy-Efficient Economy.
5. 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*. pp. 461-466. Madison Wisconsin: International Building Performance Simulation Association.