An Approach to Calibrating DOE-2 Simulation Models Using Hourly End-Use Data and Data Visualization Software

David A. Cohen, Architectural Energy Corporation, Boulder, CO
Pat Bailey, RLW Analytics, Inc., Sonoma, CA

ABSTRACT

Existing hourly end-use data are increasingly being considered as a resource for use in calibrating building energy simulation models. This paper describes the results and the approach used to calibrate 100 commercial building simulation models using hourly end-use data, data visualization software, and newly developed modeling and decision-making protocols. Simulation models were generated using a major utility's commercial building survey information and hourly end-use metered data for a three-year period. Models from nine commercial building segments were spawned from well-defined prototypes using existing building characteristics survey data, technical literature, and actual weather data for the simulation year. The buildings were modeled using the DOE-2.1E building energy analysis program and calibrated using a data visualization program[1]. The building simulation models were created with a high level of parametric capability for calibration purposes and for future analysis. A calibration approach was then developed based on experiential knowledge of building simulation experts[2]. The calibration approach gave the modelers a general set of rules to follow with respect to the types of parameters to change during the calibration process and the order in which these changes should be made. The primary reason for developing the approach was to maintain consistency in the calibration results and to provide historical documentation of the process. The data visualization software was used to provide graphical and statistical feedback to the modeler during the calibration process. This feedback was essential for determining the hourly variation in building operation. The program assists the modeler in determining qualitative insights and overviews of the large data sets used in the study. This paper will report on the calibration approach used, the results of the calibrated models, and the usefulness and limitations of the approach.

Introduction

Calibrated building energy simulation models can be used to produce valuable information for utility load research, DSM evaluation and assessment, and benchmarking studies. This paper describes

[1] DOE-2.1E, developed by Lawrence Berkeley National Laboratory and JJH and Associates, is an hourly building energy use simulation program. DOE-2.1E release 116 from JJH and Associates was used for this project. An "in house" version of the VisualizeIT™ data visualization program developed by RLW Analytics was used for all calibration efforts.

[2] Part of this approach included administering informal verbal and written surveys to approximately 20 simulation practitioners. The practitioners included staff at Architectural Energy Corporation, RLW Analytics and the Joint Center for Energy Management at the University of Colorado. The purposes of this process was to determine which building simulation variables should be formatted as 'include' parameters. These parameters were then used for calibration purposes, or after the project was completed, for "what-if" analyses by the project's clients.

the data used to develop calibrated models, how the models were created, how they were made "user friendly" for use by non DOE-2 engineers, the calibration process, and an evaluation of the quality of the calibration. A significant number of utilities and building owners have, for various reasons, developed simulation models of their buildings in the past. These models can be useful as energy efficiency evaluation tools, but only if they accurately represent the buildings' demand and energy profiles.

Others (Kreider & Haberl 1994) have done previous work on the use of data visualization to facilitate the calibration of building simulation models. These rigorous prior efforts are valuable because they began the process of developing methodologies and setting calibration targets. Our continuation of those efforts advances the understanding and application of calibrated simulation models as follows:

1. This project described in the paper used existing databases of building characteristics data to develop the simulation models,
2. The models described here all contain robust code which will allow less experienced simulation engineers to modify key calibration and building parameters in the future.
3. The calibration and building variables available for easy modification by future users were chosen after a survey of approximately 20 simulation practitioners,
4. The project represents our attempt at broadening the application of simulations to a larger group.

The authors have found that it is possible to spawn DOE-2 simulation models from a well designed prototype, customize them using available data sources, and calibrate them to within reasonable standards. The mean bias error and coefficient of variation statistics used to evaluate the models' goodness of fit show that it is reasonable to expect calibrated models for medium to large buildings to be within 5-10% of the actual building's annual energy consumption and to have an annual CV(RMSE) of under 0.3–0.4. Smaller office, retail, and education buildings exhibited more variability in their operation, were more difficult to calibrate, and did not meet these targets. Data visualization tools and statistical metrics were used to assess the match between monitored and modeled data. In general, the statistical metrics determined whether the model was complete, while the visual assessment of the results told the modeler where to make changes if the model did not meet the calibration targets.

Summary of Characteristics Data

The building characteristics database used to create the DOE-2 models was available from a commercial buildings load research survey conducted by a major east coast utility. The survey included the following general building information: a short description of the building type; whether the building stood alone or was attached to other structures; building operating hours; variation in building operation throughout the year; building vintage; and whether the building had one, two, or more than two floors.

More specific building information was available in the categories of: building envelope; lighting, equipment, and other internal loads; HVAC system; cooking; and refrigeration. Information relating to the building envelope consisted of a qualitative assessment of the insulation level of the building, the percent of exterior wall area covered by glazing, construction assemblies, and building total square footage. Internal loads information included lighting and equipment connected loads and schedules, number of people at full occupancy and corresponding schedules. The Heating, Ventilating, and Air Conditioning (HVAC) information included heating and cooling thermostat setpoints and control strategies; heating and cooling equipment type and fuel source, efficiency, and fan efficiency and size.

Other building end-use information included water heating type, fuel, usage code, and capacity; cooking equipment type, fuel, capacity, and age; and refrigeration equipment type, capacity, cubic feet, and age.

Missing Information

Although the above database of characteristics was useful for defining general building features, it was too sparse to be the only source of data for detailed building simulation models. Several building characteristics, whose values can have significant energy impacts, were missing or incomplete in the database: secondary HVAC system type, number of floors, and building geometry and shading. The database included information such as the cooling type (evaporative, chiller, or direct-expansion); heating type (heat-pump, electric resistance, or gas); and whether the system was unitary roof-top or a built-up system. Engineering judgment and modeling experience was required when choosing the specific type of HVAC system installed in the building. It was necessary to specify the type of secondary system such as dual-duct vs. fan coil, and specify control features such as constant volume vs. variable-air-volume, or to specify other control features such as economizers, reset, and a host of other system related features which have a significant impact on the building energy consumption. No information was available on the building geometry, number of floors, exterior wall and roof construction, and orientation. Engineering judgment was used for creating models based on the best available information for a particular building type.

End-use-metered Data

In addition to the building characteristics database, end-use-metered data were available for each of the buildings. Many end-use channels were available, but not all end-use channels were available for each building type. A set of standard end-use outputs was created for analysis and calibration purposes. Table 1 lists the channels which were created by aggregating and mapping the available end-use-metered (EUM) data channels. These end-uses were output by the DOE-2 program and, if available, were created from the EUM data.

Table 1. End-uses for Model Calibration and Output

Number	Unit	Variable Name
1	kW	Whole building electric
2	kW	Cooling electric
3	kW	Auxiliary electric pumps
4	kW	Heat rejection electric
5	kW	Heating electric
6	kW	Heat-pump supplemental heating electric
7	kW	Ventilation electric
8	kW	Refrigeration electric
9	kW	Domestic Hot-water electric
10	kW	Exterior lighting electric
11	kW	Exterior misc. electric
12	kW	Interior lighting electric
13	kW	Task lighting electric
14	kW	Equipment electric
15	kW	Source electric
16	^oF	Outside dry-bulb temperature

Summary of DOE-2 Models

A set of 100 DOE-2 models was created representing nine market segments and corresponding building occupancy types. The building models were meant to represent most of the commercial building sector. The building types were: large and small office; large and small retail; restaurants; grocery stores; hotels; nursing homes; and school buildings. These models were based on a library of well-developed prototypical DOE-2 models derived from in-house studies and other building prototype research projects. For each market segment, a set of building prototype models were selected to represent that segment. The best information from each was then combined into a single prototype for that segment. The 100 individual site-specific building models were spawned from these prototypical building models.

The prototypical models were updated using a combination of (1) a utility supplied building characteristics database, (2) the CBECS database of commercial building characteristics (EIA, 1992 & 1995), (3) in-house research, and (4) engineering judgment. The goal was to create detailed, site-specific models representing actual buildings from the above data sources.

Model Creation

The first step used to create the models was to write a program to query the building characteristics database. These characteristics were then mapped into a set of two building description files used to update the building models. A building characteristics file and a calibration parameters file were written from the program in the form of a DOE-2 macro "include" files. These macro files were designed to allow the user to incorporate external files containing pieces of BDL into the main BDL input stream, selectively skip portions of input, perform arithmetic and logical operations, debug input code, and perform future "what if" analysis. These "include" files were composed of DOE-2 building parameters that could be updated and changed over time and were initially used for model creation, and used later for model calibration. Figure 1 shows part of a calibration "include" file.

STATEMENT/MACRO	VALUE
$ This File Name:	175cal
$ Calibration Include File:	
$ siteID:	9cal
$ description:	LARGE INSURANCE OFFICE
$ INTERNAL LOADS	
##set1 ext_light_kw	9.5
##set1 lighting_w_sqft	1.8
##set1 lites_occ	0.85
##set1 lites_unocc	0.11

Figure 1. Sample Calibration "include" File

In future work, these files may be used to update the models with more detailed building information, to re-calibrate the models, and to perform "what-if" analysis. The calibration include files consisted of all the anticipated building parameters and variables deemed necessary for calibrating the buildings using the hourly end-use-metered data. These variables were put into a consistent, concise file format to ensure that future modifications of the models could be performed by client engineering and non-engineering staff.

For each of the building prototype models, any end-use attributes particular to that building type but not specified in the characteristics database, or that were considered to vary according to location and site, were programmed into the models based on engineering judgment and the previously mentioned technical literature. An example of this occurred in the grocery store models. The refrigeration end-use information did not give enough information to calculate refrigeration loads accurately. Therefore, engineering knowledge and prior modeling experience were used to program refrigeration loads based on available refrigeration characteristics data. Another feature which was included in the prototype models was scalability. The characteristics database included the total building area but did not specify the geometry or number of floors. The models were designed to be scaleable by including logic to resize each of the zones and number of floors based on total square footage. Spawning a site-specific model was performed using the following steps:

- Choose the building prototype that best represents the described building
- Run program to extract characteristics and calibration parameters from database
- Import these characteristics and calibration parameters as include files
- Review spawned models for reasonableness with literature and engineering judgment
- Debug input files
- Models ready for pre-calibration phase

Figure 2 summarizes the steps necessary for model creation.

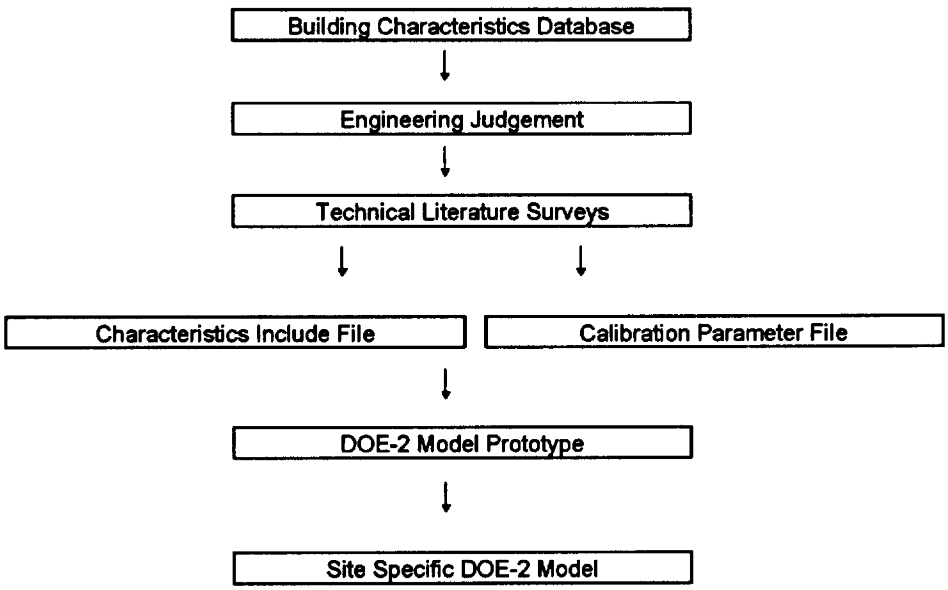

Figure 2. Steps in the Model Creation Process

At this point, the updated site-specific models are ready for the pre-calibration phase. This consisted of including available non-weather-dependent end-use-metered data such as lighting and equipment data in the models.

The Calibration Approach

Creation of Calibration Variables

All models created for the project were designed to utilize an identical set of parametric variables. These variables were listed in the external files DOE-2 "include" files described above. Table 2 lists the variables available in the external calibration and characteristics "include" files for calibration and modification by future model users.

No attempt was made to include every simulation variable that might be varied. For example, the grocery store models often had numerous variables associated with the distribution of refrigerated cases. These refrigeration variables can be changed inside the model, but were thought to be beyond the ability of the average user and were not included in the easily modified, externally referenced parameter list presented in Table 2.

Table 2. Calibration Parameters Used for the One Hundred Simulations

Parametric Calibration Variables

INTERNAL LOADS PARAMETERS
- Internal lighting power density
- % of lighting on (occupied)
- % of lighting on during unoccupied
- Lighting heat to space
- % of receptacle loads on (occupied)
- % of receptacle loads on (unoccupied)
- Receptacle load power density (W/m2)
- Receptacle load latent %
- Receptacle load sensible %
- Miscellaneous sensible load
- Miscellaneous sensible load

BUILDING SHELL PARAMETERS
- Building orientation
- Window wall ratio
- Glass shading coefficient
- Window U value
- Ceiling insulation
- Wall insulation
- Floor Insulation
- Infiltration
- Building Mass
- Miscellaneous exterior electrical loads
- External lighting kW
- Exterior lighting % on (day)
- Exterior lighting % on (night)

OCCUPANCY SCHEDULES
- 24 hour weekday occupancy flag
- Weekday begin occupancy hour
- Weekday end occupancy hour
- Saturday closed flag
- 24 hour Saturday occupancy flag
- Saturday begin occupancy hour
- Saturday end occupancy hour
- Sunday closed flag
- 24 hour Sunday occupancy flag
- Sunday begin occupancy hour
- Sunday end occupancy hour

SYSTEM CALIBRATION
- Heating setpoint (occupied)
- Heating setback (unoccupied)
- Cooling setpoint (occupied)
- Cooling setback (unoccupied)
- Fan flag during unoccupied periods

ZONE PARAMETERS
- VAV turndown ratio
- Electric baseboard heating capacity

SYSTEM LEVEL PARAMETERS
- System sizing ratio

SYSTEM-TERMINAL
- Degrees of reheat available

SYSTEM-CONTROL
- Supply air temperature control
- Maximum supply air temperature
- Minimum supply air temperature

SYSTEM-AIR
- Minimum outside air %
- Airside economizer control
- Integral economizer flag

SYSTEM-FANS
- Fan power
- Fan speed control
- Fan cycling control

SYSTEM-EQUIP
- Cooling electrical efficiency
- Heating electrical efficiency
- Heating fuel efficiency

HEAT PUMPS
- Minimum HP operating temperature
- Maximum supplemental heating setpoint
- Supplemental heating capacity
- Supplemental heating fuel type
- Defrost type
- Defrost control
- Supplemental heating sizing ratio

REFRIGERATION
- Sensible schedule (load to space)
- Latent schedule (load to space)
- Minimum condensing temperature
- Compressor sizing ratio
- Condenser fan kW

DOMESTIC HOT WATER
- Peak DHW usage rate
- DHW water heater efficiency
- DHW tank and pipe loss
- DHW usage (occupied)
- DHW usage (unoccupied)

CENTRAL PLANT
- Boiler fuel efficiency
- Cooling tower fan efficiency
- Cooling tower pump head
- Cooling tower setpoint
- Cooling tower setpoint control
- Hot water loop pump control
- Hot water loop head
- Hot water loop pump efficiency
- Cold water loop pump control
- Cold water loop head
- Cold water loop pump efficiency

Assessment of Available Calibration Data Sources

After the simulation model was built using a database of building characteristics data, the model was calibrated to existing whole premise metered (WPM) and end-use metered (EUM) electrical data. Prior to this calibration, an initial validation exercise was carried out to ensure that the monitored data was reasonable. In general, complete 15-minute interval data sets were available for whole premise electric, cooling electric, heating electric and sometimes lighting end-uses. Partial data for other end uses was available, but typically only a portion of the end use had been monitored. In cases where partial end-use data sets were available, the monitored data was used to determine operating schedules, but could not in general be used to carry out power density calibrations. For analysis purposes, all hourly simulation data was converted to 15-minute interval data to facilitate interval comparisons. This process did not change the values for any hour interval, rather the hour integration of energy to produce demand was changed to a 15-minute integration with the demand remaining constant for all four 15-minute intervals

Validation of the WPM and EUM data was carried out using data visualization techniques which allowed the modeler to determine whether the data was reasonable. An automated batch process was used to produce graphical presentations of the monitored end-use metered data for all 100 sites. Because of the wide range of building types involved in the study, no attempt was made to create strict decision rules, rather this step was considered a validation or reasonableness check.

Figure 3 shows an example of the simplest level of data validation using monthly demand and energy plots. The monthly HVAC energy and demand for this large office was labeled as total HVAC, but showed no temperature dependence and represented less than 10% of the site's total demand. It was hypothesized that the HVAC monitored data for this site was incomplete and represented only fan energy. Because the HVAC data was incomplete for this site, the data was only used to calibrate to the fan operating schedules. No attempt was made to force the modeled chiller to match this profile.

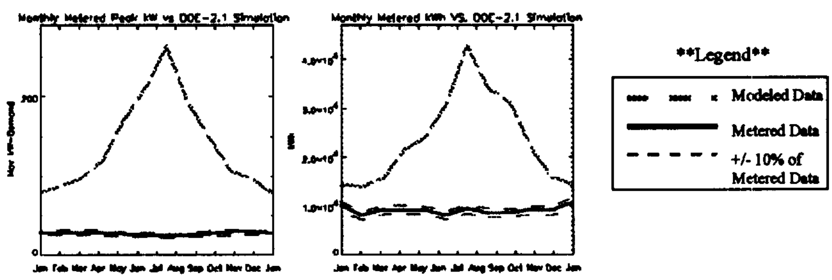

Figure 3. Data Visualization Showing the HVAC Metered Data

Calibration Steps: Developing a Process for Calibration

During this project, a set of consistent steps were taken to calibrate models. The funding level of project required that the calibration process be completed in approximately two to six hours of an engineer's time rather than the more long-term research and evaluation efforts typical at academic institutions. Previous calibration work at consulting and research organizations was examined prior to the kickoff of this project in an effort to capitalize on what had been done by others (Bou-Saada 1995, Kreider & Haberl 1994, Weaver, Hepting & Jones 1995). The prior work was particularly useful in determining initial goodness of fit calibration targets.

All model calibration efforts were carried out using actual year weather data for a period matched to the monitored electrical data. A DOE-2 compatible weather file was developed from NCDC Surface Airways Observations files containing dry bulb temperature, dew point temperature, atmospheric pressure, wind direction and wind speed, and cloud cover data. Ground temperature data was taken from values in TMY weather tapes for the same sites. The DOE-2 weather processor was used to convert the available cloud cover data to solar data.

Regardless of whether the model represented a grocery store or a large office building, the general calibration process was the same. The steps were:

1. Modify power density of modeled, non-temperature dependent loads to match peak monitored data values,
2. Adjust scheduling for modeled non-temperature dependent loads to match metered data during off peak periods,
3. Modify temperature dependent (typically HVAC) variables to match peak monitored data,
4. Modify temperature dependent (typically HVAC) schedules to match monitored data.

It was our practice to have benchmarks available that the modeling engineers could follow to validate their models. This helped ensure that no member of the modeling staff used unrealistic model inputs to force agreement between monitored data and model output. The benchmarks were loose ranges of values that corresponded to the *expected* range for model parameters. For example, one would not expect to see lighting levels of 6 W/ft^2 in an office building.

Table 3 shows validation ranges used during this project. It was possible for a modeling parameter to be outside of this range however, any outliers were flagged for further examination to ensure that modeling engineers maintained consistency.

Table 3. Typical Ranges for Common Model Parameters

Category	Typical Value
EUI by Building Type	
Education	8.7 kWh/ft^2
Food Service	39.6 kWh/ft^2
Health Care	23.4 kWh/ft^2
Lodging	12.4 kWh/ft^2
Retail	12.1 kWh/ft^2
Office	20.9 kWh/ft^2
Warehouse	5.4 kWh/ft^2
Cooling Area/ton	100 - 1000 ft^2/ton (300 - 600 typ.)
Lighting W/ft^2	0.8 - 2.8 W/sf
Equipment W/ft^2	0.1 - 1.5 W/sf
Supply Air Temp.	50 - 60 F (55 deg. F typ.)

EUI values from ASHRAE. 1995

The modeling engineer almost always had a choice of several calibration parameters that would cause the desired change in model output. For example, if the modeled peak cooling electrical demand was too low, the modelers had several options. Among other things, they could have lowered the thermostat set point or lowered the cooling system's efficiency. In general, enough detailed audit information existed so that the modeler could make an informed estimate regarding which calibration parameter to modify.

Within the broad framework of the four calibration steps listed above was a second iterative process which consisted of evaluating the changes made to the models to ascertain whether changes to the model had the expected effects. By combining data visualization and statistical metrics, an assessment of the progress toward calibration targets was made. The assessment was carried out at both the end-use and total load level for each model. Figure 4 shows the iterative data visualization/ model modification process that was followed in the creation of calibrated models.

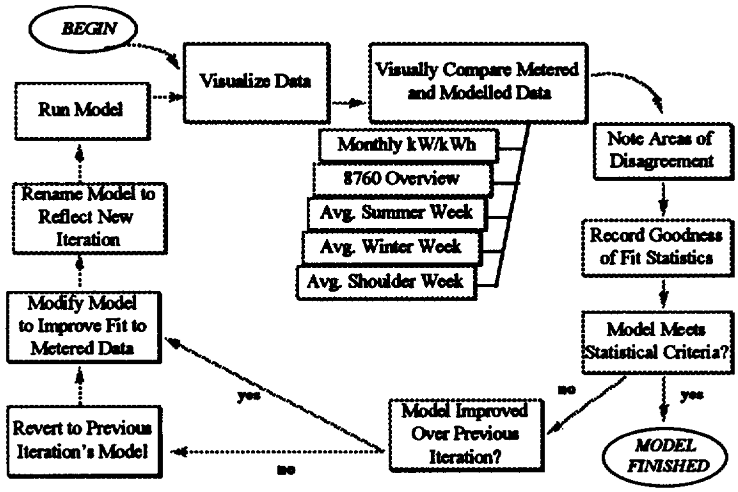

Figure 4. Steps in the Calibration Process

As a practical matter, it is rare to have detailed end-use monitored data available for use in calibration exercises. However, it is increasingly possible to acquire and utilize 15 minute total load electrical data for existing large buildings. This total load data, in conjunction with data visualization software, a building characteristics database or audit, and short term or spot end-use electrical measurements, can be used where the budget or capability of gathering longer term end-use data does not exist. Although this paper concentrates on calibration of sites with end-use data, the same general process is followed regardless of the type of data being used for calibration. The signatures of cooling and heating end-uses can be seen in the total load using data visualization, and these signatures used as proxies for end-use information.

Two final notes on the data available for calibration:

1. Actual year weather data for the nearest National Weather Service location was used for all calibrations. No attempt was made to quantify the differences in building performance resulting from the variation between site level weather and the NWS airport data. Comparisons between airport and building level weather data have been done by others. (Haberl, Bronson & O'Neal 1994).
2. Many large buildings have sophisticated Energy Management Systems which acquire premise level electrical demand data but this data is often overwritten after a period of seven to thirty days because of limited storage capacity. It should be possible to utilize this data in the near

future to calibrate simulation models. In principle, the same potential exists for end-use data, however building owners do not generally specify EMS electrical measurement points at the end-use level.

Data Visualization and Statistical Goodness of Fit - Evaluating the Model

Model calibration as practiced by the authors for this and other projects is an iterative process which relies on successive rounds of data visualization and immediate evaluation of results. A model is built, its fit to billing or monitored data is summarized quantitatively and qualitatively, decisions are made regarding whether the model meets pre established standards for goodness of fit, and the model is improved or considered complete depending on how well the model meets the standards. Data visualization allows the modeler to carry out the iterative process efficiently and receive immediate feedback on progress. This section will attempt to answer two questions relating to data visualization and evaluation of model results.

1. How are data visualization techniques used to summarize model results/calibration?
2. How good is good enough with respect to the model's fit to monitored data?

Data Visualization. The use of data visualization software allows the modeler to quickly assess the results of a model iteration, and identify areas requiring additional work. It has been our experience that the ability to look *qualitatively* at model results has improved the quality and cost effectiveness of the simulation models. In the past, many modeling projects have relied on gross graphical measures of model goodness of fit. These graphical measures were typically customized spreadsheet applications which plotted loadshape averages by season or daytype as well as monthly simulated vs. billed electrical energy and demand plots like the one shown in . While this type of analysis is useful, it does not go far enough. It is our opinion that the graphical tools greatly assist simulation efforts. It was possible, using desktop PC computers, data visualization software and modeling engineers, to complete a calibration iteration in well under an hour. An iteration consisted of running the model, using data visualization software to examine the output, assessing goodness of fit visually and statistically, either modifying parameters or considering the model complete, documenting the changes made and re-running the model. Sometimes the iteration only took 15 minutes if the changes to the model were simple and the result of an obvious error. It is important to realize that errors or mistakes in a model that are obvious when using data visualization may not have been readily observed in a traditional calibration effort.

Figure 6 shows an example of the EnergyPrint screen used in the model calibration effort[3]. This screen allowed the modeling engineer to assess an entire year of modeled or metered data at the premise or end-use level interval. The EnergyPrint is a three dimensional plot of either 15 minute (metered) or one hour (simulation) interval data in which color is the z axis, hour of the day is on the y axis and day or month of the year is on the x axis. The demand for each interval is represented by a color gradient in which 0 kW appears black, low kW blue, medium kW red and orange, and high kW yellow to white. When used for calibration, this graphic is viewed with a high resolution color monitor. The mouse is moved around to areas of interest. A second image presents dynamic load shapes for the day chosen by the cursor.

[3] The EnergyPrint is similar to images used by Craig Christensen of NREL in the mid 1980s.

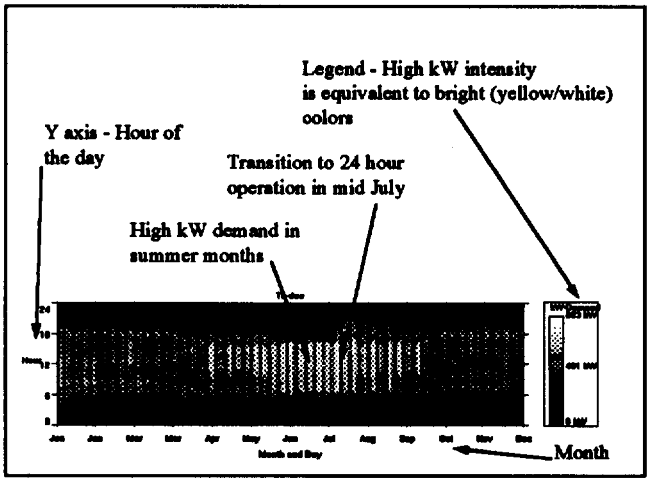

Figure 5. Example EnergyPrint for a Large Office

Figure 6 shows the three most commonly used screens of the data visualization software. The modeling engineers typically started an iteration by used monthly kW and kWh comparison screens to do a billing data comparison between the model and the actual building. 8760 hour EnergyPrint screens were then used to easily determine seasonal variations in equipment scheduling and to ensure that the simulation model did not contain obvious errors. Finally, the average week screens were an easy way to compare average loads, both total and end-use for different seasons. In addition to the graphics shown here, the software also produced statistical summaries of the goodness of fit between the metered and modeled data sets.

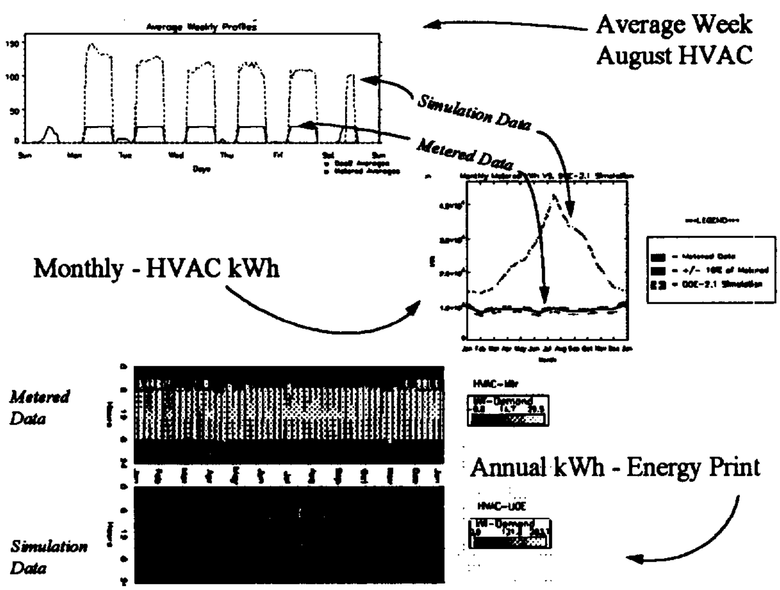

Figure 6. Some of the Tools used in the Model Calibration Effort

How Good is Good Enough? Statistical metrics were used to assess the goodness of fit between modeled and metered data. As reported in previous works, the residual load coefficient of variation, CV(RMSE) and mean bias error, MBE were the statistical measures used to quantify goodness of fit between modeled and metered data (Krieder & Haberl 1994, McCray et.al. 1995). The

CV(RMSE)[4] can be thought of as an estimate of the model's consistency in predicting building demand on an hour by hour basis. The MBE describes how the model predicts monthly or annual energy use and is simply the percentage difference between the modeled and metered energy use.

Figure 7 shows a *schematic* example of why both the MBE and CV(RMSE) calibration statistics must be used together when evaluating the level of model calibration. The example shows a worst case scenario in which the annual energy would be equal for the simulation and the monitored data, but the goodness of fit would be terrible because the loads are 180 degrees out of phase. This comparison is exaggerated, but illustrates the point that a simple comparison of energy consumption between the model and utility bills is not, by itself, an adequate measure of how well a model tracks a building's energy use.

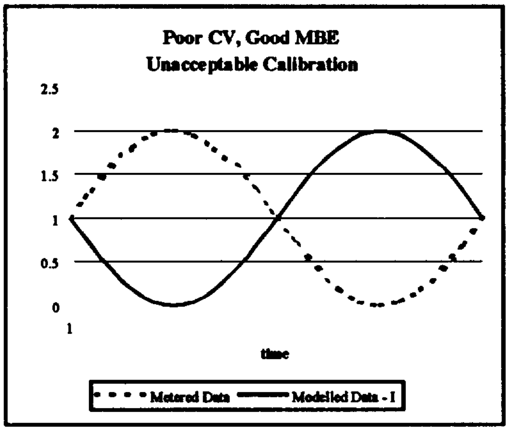

Figure 7. Goodness of Fit Schematic Example

Both the CV(RMSE) and MBE statistics must be acceptable for a model to be considered calibrated. What are acceptable values for these statistics? The answer depends on how the simulation results will be used. Based on our experience with electric utility clients, annual and monthly MBE's in the 5% to 10% range provide a model that is of high enough quality to be useful to clients and is possible in a reasonable time frame. Acceptable values for CV(RMSE) are less well understood by most clients and modelers, dependent on building and occupancy categories and tend to vary widely. However, CV(RMSE) values in the 25 – 40% range are acceptable. The calibration statistical targets that were proposed at the beginning of this project are summarized in Table 4.

---

[4] $CV_{(RMSE)}$ is the coefficient of variation of the root mean square error. It is defined as

$$\sqrt{\frac{\sum_n (y-\hat{y})^2}{n}} \Bigg/ \frac{\sum_n y}{n}$$

where y is the hourly value for the monitored data, $\hat{y}$ is the hourly value for the modeled data, and n is the number of interval data points in the monitored and modeled data sets. The CV(RMSE) is a measure of the monitored vs. modeled error relative to the standard deviation of the error. The CV(RMSE) is especially useful for identifying situations where the absolute errors are self canceling.

Table 4. Goodness of Fit Statistical Targets

	CV(RMSE) (%)	MBE (%)
Annual Total Load	30	10
Monthly Total Load	40	15
Annual Peak kW	n/a	10
End-Use Total Load	40	15

The initial goodness of fit statistical targets proved to be overly pessimistic in some cases and overly optimistic in others. In general, the MBE targets were met or exceeded. However, the CV(RMSE) targets varied over a large range. Large building with relatively stable loads had CV(RMSE)s ranging from 10% - 30% whereas schools, small office and small retail sites had CV(RMSE) statistics which were worse, ranging from 30% to 70%. Based on this and other work, we've found that building tenants at smaller facilities do not run always operate their buildings consistently. Cooling systems that are said to be turned off at 6 pm were often found to be on at 10 pm, lighting that is supposedly 100 percent off on weekends was found to be 20% on. Seemingly random variation was common in small buildings. For this reason, small building models proved to be difficult to calibrate. There is some agreement that CV(RMSE)'s below 20% are very difficult to obtain for any building type because of behavioral drivers associated with energy usage. An exception to this .20 CV(RMSE) floor appears to be large grocery stores which are dominated by consistent lighting and refrigeration loads. Although the nursing home models in this work also tracked metered data very closely and had good calibration statistics, no general statement about the ease of calibrating that segment will be made because there were only two sites in the sample.

Figure 8 summarizes the mean and range of calibration statistics by segment for the 100 models in this project.

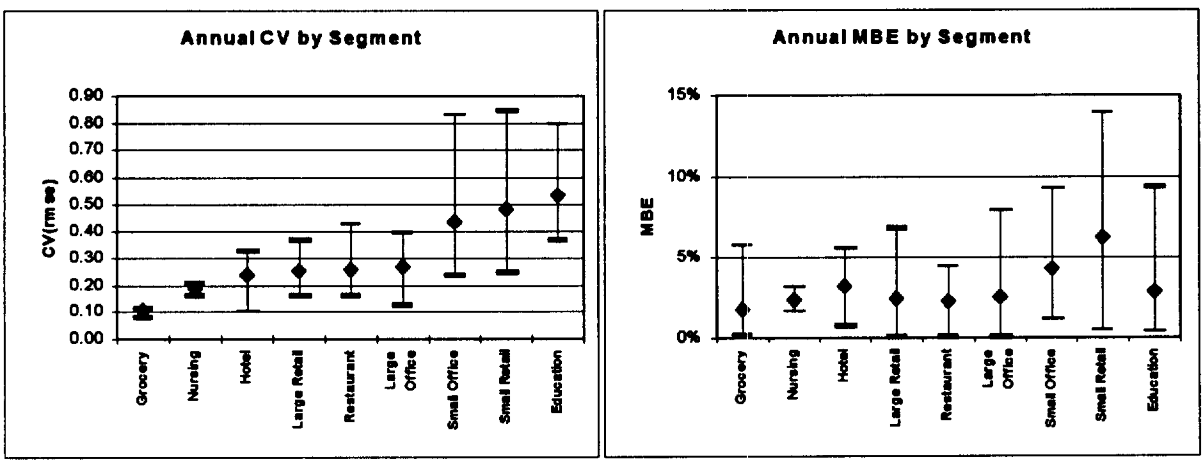

Figure 8. Calibration Results for 100 Models

One final note on model goodness of fit and overall quality. The quality of simulation models is clearly related to the quality of the data and the amount of staff time available for generating the models. As described at the beginning of the paper, these models were generated based on prototype models which were modified using pre-existing database information, not on audits carried out by the modelers. In addition, budgetary constraints limited the average amount of time that could be spent developing and calibrating each model to approximately one or two days. With additional resources, primarily the ability to audit the modeled facility or contact facility staff, further refinement of troublesome sites would have been possible.

Conclusions

Approximately 100 calibrated building simulation models were developed using measured hourly end-use data, data visualization tools, and an ordered calibration approach. Having hourly end-use data was found to be extremely useful in both understanding the building operation, and calibrating the models. The visualization tools provided a much needed function in the calibration process, allowing the modeler to see relationships among the data and to make good judgments based on the time and weather dependent factors evident in the data. The models developed for larger buildings demonstrated a high level of accuracy in terms of the hourly match of the whole building load as well as for any measured channels that were available during calibration. Smaller buildings and buildings such as some schools with a high degree of operational variability did not fare as well.

Calibrated simulation models are useful for assessing the impacts for various energy efficiency measures, identifying building systems which are not functioning properly, optimizing building control systems, and for load research purposes. The models created in this study were used in a data leveraging study in which the building load profiles were transferred to other utility territories. In the future, as electricity markets deregulate, there will likely be additional uses for having calibrated computer models. There will still be a need for the traditional uses of energy efficiency assessment in the context of energy services, as well as new uses such as rate design, profitability analysis, customer segmentation, and others from power marketers and aggregators.

Future work is planned which will further the standardization of the calibration process including a refined step-by-step approach for calibration, a suite of tools for changing calibration parameters, performing data visualization, assessing calibration statistics, and for reading in non-weather dependent end-uses into the model directly.

References

Kreider, J and J. Haberl 1994. "Predicting Hourly Building Energy Use: The Great Energy Predictor Shootout – Overview and Discussion of Results. " ASHRAE Transactions. 100(2).

EIA. U.S. Department of Energy, Energy Information Administration, 1992 and 1995 Commercial Buildings Energy Consumption Survey (CBECS).

Weaver, Norm, Curt Hepting and Chris Jones 1995. "Office Segment Load Research Assessment – Development of Audit/Model Calibration Protocols." *Report to BC Hydro, Market and Load Research Department*. ERG International. 1995.

Haberl, Jeff and Tarek Bou-Saada 1995 "Procedures for Calibrating Hourly Simulation Models to Measured Building and Environmental Data." College Station, TX.: Texas A&M University System Press.

ASHRAE. 1995 Handbook - HVAC Applications. Table 4, 32.6.

Haberl, J., J. Bronson and D. O'Neal 1994. "An Evaluation of the Impact of Using Measured Weather Data versus TMY Weather Data in a DOE-2 Simulation of an Existing Building in Central Texas." Draft submission of article to ASHRAE Journal.

McCray, J., P. Bailey, J. Parker and R. Gillman 1995, "Using Data Visualization Tools for the Calibration of Hourly DOE-2.1E Simulations." *IBPSA Proceedings*, 461- 466.

Haberl, J., Oral Presentation on Model Calibration, ASHRAE 1997 Winter Meeting, Philadelphia, PA. January 28, 1997.