SERI/TP-254-1953
UC Category: 59c
DE83009402

Measured Versus Predicted Performance of the SERI Test House: A Validation Study

R. Judkoff
D. Wortman
J. Burch

May 1983

To be presented at the
National Heat Transfer Conference
Seattle, Washington
July 1983

Prepared under Task No. 1053.00
WPA No. 304

Solar Energy Research Institute
A Division of Midwest Research Institute

1617 Cole Boulevard
Golden, Colorado 80401

Prepared for the
U.S. Department of Energy
Contract No. EG-77-C-01-4042

Printed in the United States of America
Available from:
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161
Price:
Microfiche $4.50
Printed Copy $7.00

MEASURED VERSUS PREDICTED PERFORMANCE OF THE SERI TEST HOUSE: A VALIDATION STUDY

R. Judkoff
D. Wortman
J. Burch

ABSTRACT

For the past several years the United States Department of Energy (DOE) Passive and Hybrid Solar Division has sponsored work to improve the reliability of computerized building energy analysis simulations. Under the auspices of what has come to be called the Class A Monitoring and Validation program, the Solar Energy Research Institute (SERI) has engaged in several areas of research that includes: (1) developing a validation methodology; (2) developing a performance monitoring methodology designed to meet the specific data needs for validating analysis/design tools; (3) constructing and monitoring a 1000-ft^2, multizone, skin-load-dominated test building; (4) constructing and monitoring a two-zone test cell; and (5) making sample validation studies using the DOE-2.1, BLAST-3.0, and SERIRES-1.0 computer programs. This paper reports the results obtained in comparing the measured thermal performance of the building to the performance calculated by the building energy analysis simulations. It also describes the validation methodology and the Class A data acquisition capabilities at SERI.

The Class A, B, and C performance monitoring programs were initiated in 1979 because of the demand from researchers and industry for passive and hybrid building performance data at various levels of detail (1). Class A monitoring provides detailed data (approximately 200 channels per building) under controlled conditions at a few sites for algorithm development and validation of building energy analysis simulation programs. Class B provides limited detail (about 20 channels per building) in approximately 100-200 occupied buildings for field testing passive and hybrid designs and statistically evaluating simplified design tools. Class C provides utility bill data and a survey of occupant reactions.

SERI's involvement in validating building energy analysis simulations (BEAS) resulted from two comparative studies conducted in 1980 and 1981 (2,3). These studies showed significant disagreement between four state-of-the-art simulations: DOE-2.1, BLAST-3.0, DEROB-4.0, and SUNCAT-2.4 when given equivalent input for a simple, direct-gain building with a high and low mass parametric option (Figure 1). The studies also indicated the need for high quality, controlled validation data and a validation methodology. SERI assumed responsibility for defining the data acquisition criteria for validation, developing a validation methodology, and constructing a Class A data collection facility. Class A facilities were also constructed at the National Bureau of Standards (NBS) and several universities.

VALIDATION METHODOLOGY

The overall validation methodology uses three different kinds of tests (4): (1) analytical verification (5), (2) empirical validation, and (3) code-to-code comparisons. The advantages and disadvantages of these three techniques are shown in Table 1.

Each comparison between measured and calculated performance represents a single data point in an immense N-dimensional parameter space. We are constrained to establishing very few data points within this space, yet, we must somehow be assured that the results at these points are not coincidental and do represent the validity of the simulation elsewhere in the parameter space. The analytical and comparative techniques minimize the uncertainty of the extrapolations we must make around the limited number of Class A empirical data points it is possible to sample. These extrapolations are classified in Table 2.

Figure 2 shows the process by which we use the analytical empirical and comparative techniques together. The first step is to run the code against the analytical test

---

Solar Energy Research Institute;
1617 Cole Blvd.; Golden, CO 80401

cases. This checks the numerical solution of major heat transfer models in the code. If a discrepancy occurs, the source of the difference must be corrected before any further validation is done.

The next step is to run the code against Class A empirical validation data and to correct discrepancies. A quantified definition of these discrepancies has been proposed by Los Alamos National Laboratory (LANL) (6). SERI and several other Class A sites are currently collecting these data.

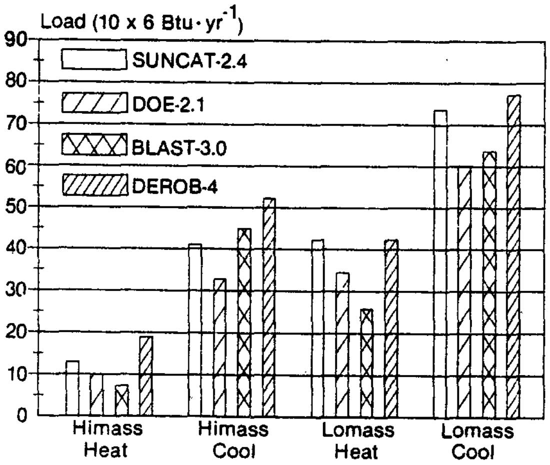

Figure 1. Phase II Comparative Study: Albuquerque

The third step involves checking the code against several prevalidated building energy analysis simulations (BEAS) in a number of comparative studies. If the code passes all three steps, it can be considered validated for the range of climates and building types represented by these studies. The prevalidated BEAS will have successfully passed steps one and two and will have shown substantial agreement for all the comparative study cases. These comparative study cases will use Class B data where possible. SERI is currently prevalidating the DOE, BLAST, and SERIRES programs as part of its Class A empirical validation project.

DATA COLLECTION METHODOLOGY

There are many levels of validation depending on the degree of control exercised over the possible sources of error in a simulation. These error sources consist of seven types divided into two groups:

External Error Types

1. Differences between the actual weather surrounding the building and the statistical weather input used with BEAS.
2. Differences between the actual effect of occupant behavior and those effects assumed by the user.
3. User error in deriving building input files.

Table 1. Validation Techniques

Technique	Advantages	Disadvantages
Comparative Relative test of model and solution process	No input uncertainty Any level of complexity Inexpensive Quick, many comparisons possible	No truth standard
Analytical Test of numerical solution	No input uncertainty Exact truth standard given the simplicity of the model Inexpensive	No test of model Limited to cases for which analytical solutions can be derived
Empirical Test of model and solution process	Approximate truth standard within accuracy of data acquisition system Any level of complexity	Measurement involves some degree of input uncertainty Detailed measurements of high quality are expensive and time consuming A limited number of data sites are economically practical

4. Differences between the actual thermal and physical properties of the building and those input by the user (generally from ASHRAE handbook values).

Internal Error Types

5. Differences between the actual thermal transfer mechanisms taking place in the real building and the simplified model of those mechanisms in the simulation.
6. Errors or inaccuracies in the numerical solution of the models.
7. Coding errors.

At the most basic level, the actual long-term energy usage of a building is compared to that calculated by the computer program with no attempt to eliminate sources of discrepancy. This level is similar to how the BEAS would actually be used in practice and, therefore, is favored by many representatives of the building industry. However, it is difficult to interpret the results of this kind of validation exercise because all possible error sources are simultaneously operative. Even if good agreement is obtained between measured and calculated performance, the possibility of offsetting errors prevents drawing conclusions about the accuracy of the method of calculation. More informative levels of validation are achieved by controlling or eliminating various combinations of error types. At the most detailed level, all known sources of error are controlled to identify and quantify unknown error sources. This is the approach taken in Class A data acquisition for validation.

Detailed meteorological and microclimate measurements are taken at the site to eliminate error 1. The buildings are kept unoccupied to eliminate error 2. Input files are derived independently by several experienced users and then cross-checked until collective agreement is reached to control error 3. Thermophysical properties are directly measured through destructive and nondestructive testing to control error 4. Once all

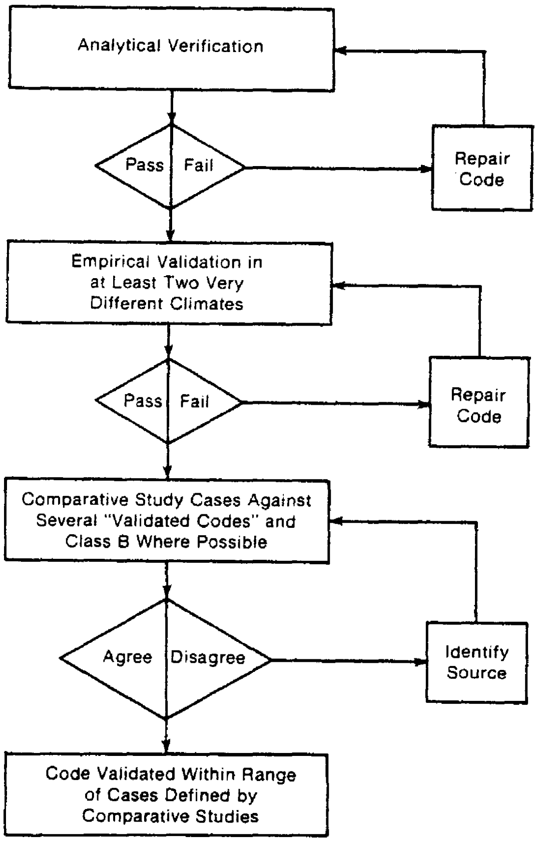

Figure 2. Validation Method

Table 2. Types of Extrapolation

Obtainable Data Points	Extrapolation
A few climates	Many climates
Short-term (e.g., monthly) total energy usage	Long-term (e.g., yearly) total energy usage
Short-term (hourly) temperatures and/or flux	Long-term (yearly) total energy usage
A few buildings representing a few sets of variable mixes	Many buildings representing many sets of variable mixes
Small-scale, simple test cells and buildings	Large-scale complex buildings

external error types have been controlled, it is possible to isolate internal errors.

To validate the key thermodynamic models, which comprise errors 5 and 6, two different kinds of data are needed. First, data must be taken to define the overall building energy performance. This overall system level includes zone air and globe temperature data and (if temperature controlled) auxiliary energy measurements. These data summarize building energy performance. Second, data must be taken at the energy transport mechanism level. Energy transport mechanisms are summarized in Table 3. Where this is not possible because of state-of-the-art measurement limitations or where no acceptable models exist for a mechanism, the mechanism may be physically suppressed as was done in our test cell for ground coupling. This two-level approach allows us to identify those mechanism inaccuracies that lead to system level errors.

To ensure that all major transport mechanisms are monitored, we provide for internal consistency checks. Failure to achieve closure on the measured heat balance $Q_{in} = Q_{out} + Q_{stored}$ can be attributed only to faulty data or to important mechanisms not represented in the measurements.

Table 3. Energy Transport Mechanisms

- CONDUCTION: Measure Temperatures and Conduction Fluxes
  - Structural elements
    - Skin and interzonal opaque walls
    - Glazings
  - Ground coupling
- CONVECTION: Tracer Gas, Special Experiments
  - Film coefficients
    - Inside surfaces: free convection
    - Outside surfaces: forced convection
  - Air Motion
    - Infiltration
    - Zone to zone
      - Natural convection through doorways
      - Natural convection through cracks
    - Stratification
- RADIATION: Measure Radiant Fluxes
  - Infrared surface coupling
    - Internal surfaces
    - External surfaces (sky temperature)
  - Solar
    - External absorption
    - Glazing transmission and absorption
    - Internal absorption

SERI CLASS A DATA FACILITY

The SERI Class A validation facility consists of two structures: a 1000-ft^2 residence and a 120-ft^2 two-zone test cell. These two structures are instrumented with approximately 250 sensors each to achieve the degree of experimental control previously discussed. The sensors include type J thermocouples, heat flux transducers, Hall effect watt-hour meters, Kip & Zonen and Eppley pyranometers, and an Eppley pyrheliometer. Wind speed, direction, and humidity are also measured.

Details of the house and the test cell are provided in two handbooks (7,8). Figure 3 shows the plan and south elevation of the house. The cell and the house were designed to complement each other and other Class A facilities. The approach in the cell was to suppress all difficult mechanisms. These included ground coupling, interzonal and cavity convection, stratification, and infiltration. The house was operated in a more realistic fashion, and attempts were made to measure such difficult transport paths as ground coupling via a crawlspace and multizone infiltration. The crawlspace configuration was chosen to complement the floor slab configuration at NBS. For multizone infiltration, we initiated a project to develop an apparatus capable of continuous multizone infiltration monitoring (9). A prototype of this apparatus has been collecting data since April 1982. Table 4 shows the measurement approach taken for various mechanisms in the house and the cell.

We monitored the house and cell through a number of configurational changes in the winter and spring of 1982. In the case of the house, this consisted of several conservation and solar retrofits including: (1) insulation blown into walls and attic, (2) batt insulation on foundation walls in crawlspace, (3) storm windows, (4) caulking and weatherstripping, (5) orientation of largest glazed areas to south, and (6) addition of thermal mass to south-facing rooms. These retrofits reduced the effective crack area as measured by a blower door from approximately 200 to 50 $in.^2$ (see Figure 4).

We will continue to collect and analyze data from the house and cell. Complete results from the fiscal year 1982 work are in Wortman et al. (10).

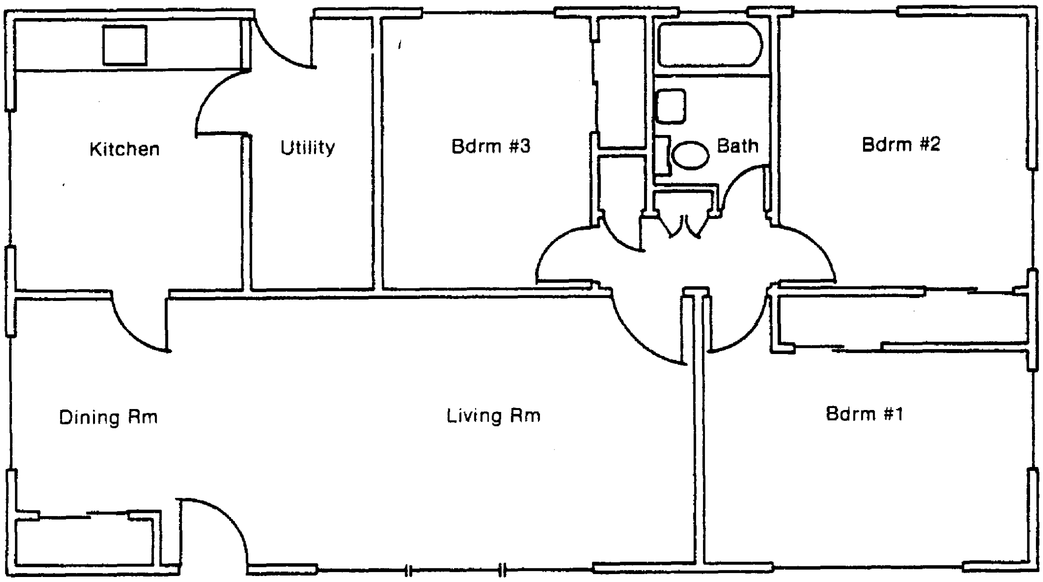

Figure 3a. Validation Test Residence: Floor Plan

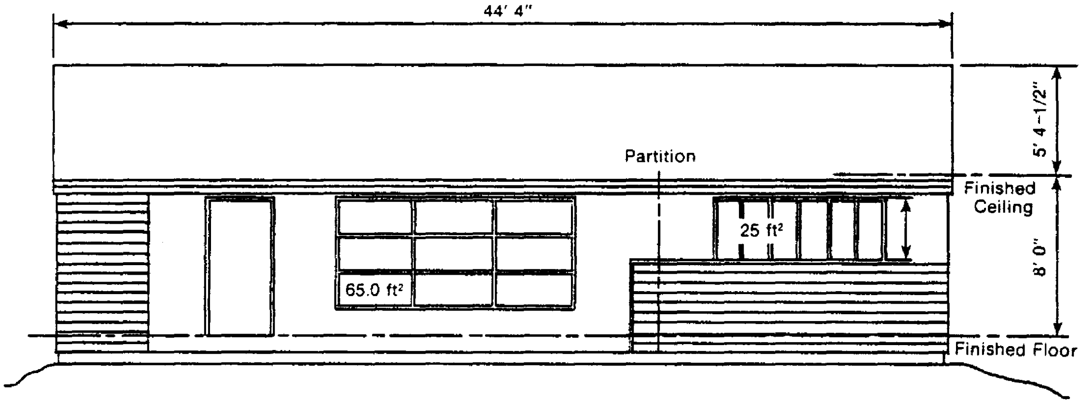

Figure 3b. Validation Test Residence: South Elevation

Validation Studies

Figures 5 through 10 show preliminary results from the validation study on the DOE-2.1A, BLAST-3.0, and SERIRES computer programs. These results are presented as nine cases, each representing a different level of input and variable accuracy. The definition for each of these cases is given below:

1. **Base Case:** Handbook or assumed values are used for all thermophysical inputs. Meteorological and geometric inputs are measured.

2. **Infiltration:** Same as base case except hourly zonal infiltration rates were measured and used to generate the infiltration input for the computer codes.

3. **Ground Temperature:** Same as base case except measured ground temperature was used as input to the ground coupling subroutines in the codes.

Table 4. Measurement Approaches

Mechanism	Code Approach	Measurement Approach	
		Test Cell	House
Building-Related Processes			
Wall Conduction			
Basic assumption	One-dimensional flow	Insulate edges where possible to ensure one-dimensional flow	One-dimensional flow assumed
Wall conductivities	Inputs, constants	Measurements to directly determine U_{wall}, U_{layer}	Same as cell
Mass Storage	Not directly available; can be computed from temperature	Compute from temperature data, ~10 rakes in mass	Compute from temperature data, ~2 locations per zone
Ground coupling	One-dimensional flow to ground temperature, neglecting edge effects	Eliminate entirely.	Study in detail, flux and temperature at ~10 locations
Boundary Conditions			
Interior surfaces	Varied approaches from h_{tot} = const. (e.g., SUNCAT) to explicit IR + convection correlations (e.g., DEROB)	Measure h_{conv} separately; define effective interior temperature, and compute infrared flux ($\dot{Q}_{IR}$)	Measure films only on glazing, same techniques as for cell
Exterior surfaces	Varied approaches, from constant (SUNCAT) or wind-driven only or wind and sky infrared	Measure $\dot{Q}_{IR}^{sky}$, T_{ground}; deduce h_{conv} on average	Same as cell
Zone-Related Effects			
Zone mixing	Always isothermal	Destratify to force zone to be isothermal	Destratify continually (FY 1982); study destratification in FY 1983
Interzonal advection and conduction	Uncertain, approximate algorithms for advection; wall conduction included	Measure conduction directly; advection minimized by careful caulking	Measure conduction; closed doors between cells (FY 1982). Study natural advection in FY 1983
Occupancy effects	Schedules input, major uncertainty	None	None
Furnishings	Neglected or approximate	None	Unfurnished
Internal humidity	Latent heat usually included	Not measured	Not measured
System Effects			
Heating systems	Set points; ramp	Measure $\dot{Q}_{heater}$ with electrical inputs of known efficiency, $\eta = 1.0$; small deadband	Electric heaters, to be computer-controlled for night setback at night
Night ventilation	Schedule or constraint for $\dot{V}_{night}$; volume flow $\dot{V}$ is input	None	Measure $\dot{V}_{once}$ by tracer decay
Environment-Related Processes			
Solar Radiation			
Descriptive inputs	Need I_{beam}, G_H	Measure I_{beam}, G_H directly	Same as cell
Tilted surface irradiance	Various models, mostly isotropic or anisotropic	Exterior: measure south irradiance broken into south sky and ground diffuse components	External: same as cell
		Internal: floor, north wall, east wall	Internal: measure vertical transmitted, each orientation; and floor and mid-wall irradiance in living room

Table 4. **Measurement Approaches** (Concluded)

Mechanism	Code Approach	Measurement Approach	
		Test Cell .	House
Environment-Related Processes (Continued)			
Glazing transmissions	Beam transmission calculated from input index of refraction and extinction coefficient, diffuse transmission = some input or default constant	Measure beam and diffuse transmission directly; extract best fit index of refraction and extinction coefficient from data. Done only occasionally.	Same as cell, for the south glass only, before and after storm glazings
Ground reflections	Input α_{GR}	Measure α_{GR} continuously; $\alpha_{GR}^{\mathrm{Eff.}}$ once	α_{GR} is some as for cell, use cell data
Solar glazing back losses	Calculatable from various models, or input constant (SUNCAT)	Measure cell albedo directly for clear, cloudy conditions	No albedo measurements
Wind	Input velocity, direction; assume same value for film calculation and infiltration model, very uncertain	Measure at two heights at ~100 yards from cell; uncertain microscale problems - Average $h_{convection}$ to be calculated - Reduce effects by tight construction	Same as cell
Other: humidity, pressure	Inputs used for air heat capacity, latent loads	Adequate direct measure	Same as cell
Precipitation	No impact on thermal models	Field site observation, plus α_{GR} data effects	Same as cell

4. **Ground Albedo**: Same as base case except measured ground albedo was used in the calculation of radiation incident upon glazed surfaces.
5. **Set Point**: Same as base case except a correction was made to the thermostat set point based on the average temperature of air in the zone when the heater actually turned on.
6. **Wall and Roof Conductance**: Same as base case except measured wall and ceiling conductances were used.
7. **Window Conductance**: This case was not run because measured window conductances were the same as those given by the ASHRAE *Handbook of Fundamentals*.
8. **Absorptivity**: This case was not run because the measured solar spectrum absorptivity on opaque surfaces was not significantly different than assumed values.
9. **Measured**: All of the measured values in cases 2 through 6 were used. This case represents the highest degree of control over external error sources and should presumably yield results closest to the measured temperature and energy performance of the building.

Results

Figure 5 shows the whole-house heating load (in kWh) during the week of April 20-26, 1982. The loads predicted by the DOE-2.1A, BLAST-3.0, and SERIRES computer programs are shown along with the measured load for cases 1 through 9. In case 1, where handbook input values were used, the code predictions were high by 59%-66% compared with measured loads. In case 5, where the correction was made for the actual thermostat set point, the code predictions were high by 47%-52%. In case 9, where all known measured input values were used, the code predictions were low by 10%-17%. In general, the predictions were most accurate for case 9.

Figure 6 shows the root mean square (RMS) difference between measured and predicted temperatures in zone 2 of the house for all 9 cases. Zone 2 is the southern living room and has a massive floor surface. In general the results from zone 2 are typical of the results from the whole building. Case 1 has RMS errors of between

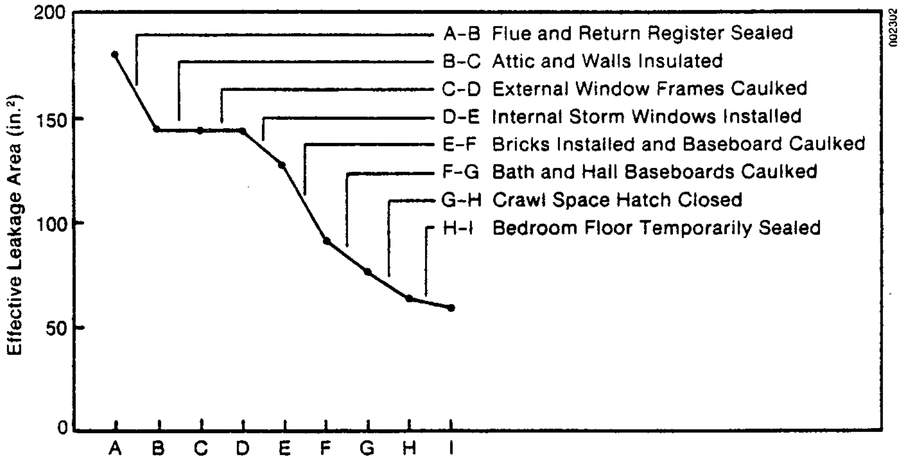

Figure 4. Blower-Door results for SERI Retrofit House

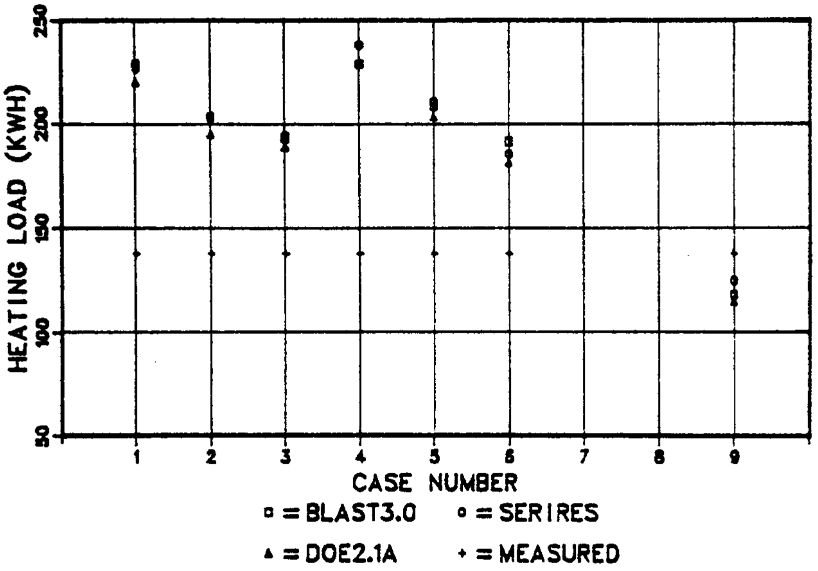

Figure 5. Weekly Whole-House Heating Loads

.9° and 1.2°C. Case 5 has RMS errors of from .6° to .8°C. Case 9 has the largest RMS errors of from .4° to 1.6°C.

Figure 7 shows the zone 2 measured peak heating load and the peak heating loads predicted by the three computer codes in cases 1 through 9. The case 1 predictions of peak load are high by 36%-49%. The case 5 predictions are high by 31% to 43%. The case 9 predictions are the most accurate and fall within ±5% of the measured peak load.

Figure 8 shows the peak load for the whole house. The pattern is similar to that observed for zone 2 with case 1 predictions being least accurate and case 9 predictions being most accurate.

Figure 9 shows the hourly temperature profile predicted by the DOE-2.1A code in

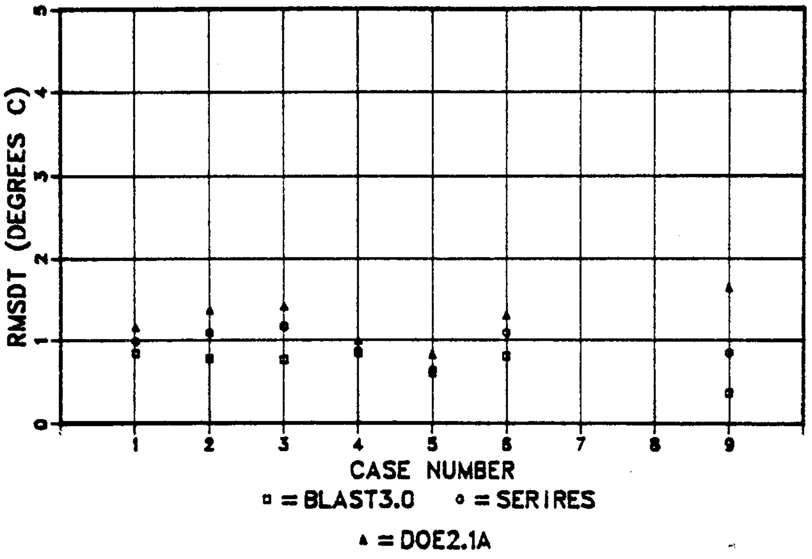

Figure 6. Zone 2 Root Mean Square Temperature Difference

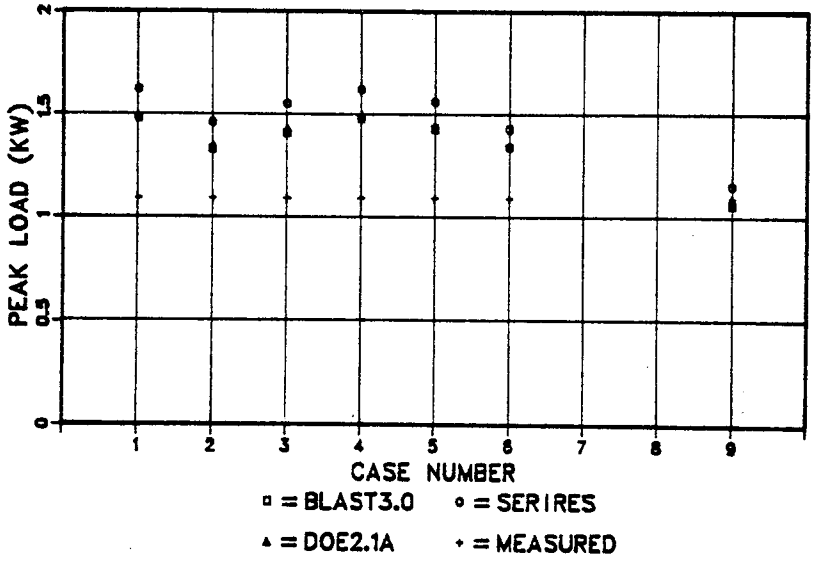

Figure 7. Zone 2 Peak Load

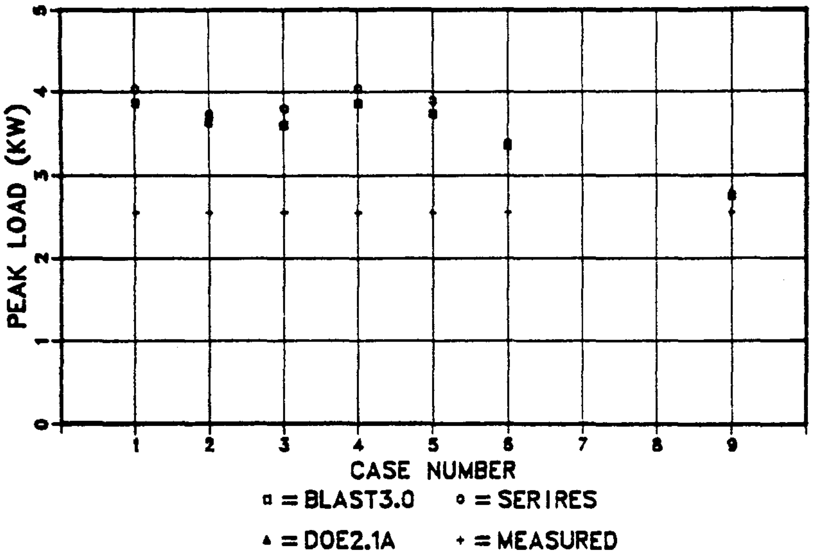

Figure 8. Whole-House Peak Load

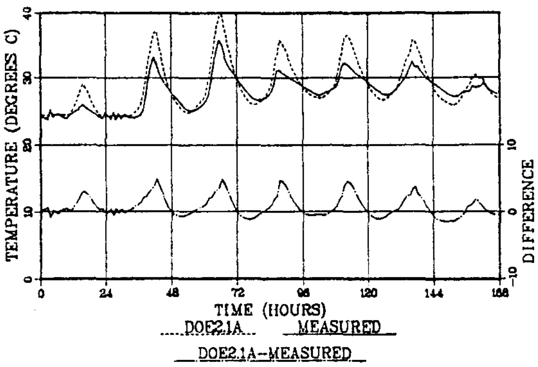

Figure 9. DOE-2.1A vs. Measured Temperatures, Case 1: Zone 1

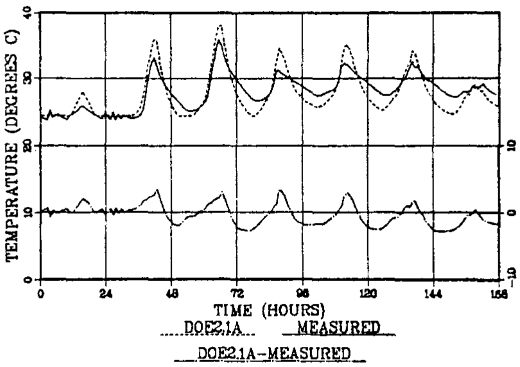

Figure 10. DOE-2.1A vs. Measured Temperatures, Case 9: Zone 1

relation to the measured temperature profile for case 1, zone 1. Zone 1 was primarily a free-floating zone during this time period because temperatures remained above the thermostat set point from hour 36 to hour 168. Figure 9 shows that the predicted temperature tends to overshoot measured temperature during the day and undershoot measured temperature at night.

Figure 10 shows the same information for case 9 as was shown for case 1 in Figure 9. In case 9 we see that predicted temperatures overshoot by even more during the day than in case 1 and that they undershoot by less at night than in case 1.

INTERPRETATION OF DATA

There is some apparent inconsistency in the data with respect to the presumption that case 9 would always yield the most accurate predictions. The most obvious is seen in Figure 6 where case 1 and case 5 had smaller RMS temperature errors than case 9. This trend is the reverse of that seen in Figure 5, where, as expected, the most accurate load prediction was obtained in case 9. The most likely hypothesis now is that (1) the amount of solar energy absorbed in the building is being overpredicted in all cases, and (2) the conductive losses through walls and roof are being overpredicted in cases 1 and 5.

This conjecture is partially supported by the large (approximately a factor of two) difference found between measured and assumed wall and roof resistances as shown in Table 5.

Table 5. Measured vs. Assumed Wall and Ceiling Resistances

	R-Value $W/m^2\ ^oC$ ($Btu/h\ ft^2\ ^oF$)
Average measured wall resistance	3.05 (17.3)
Assumed wall resistance from ASHRAE	1.56 (8.83)
Average measured ceiling resistance	13.19 (75.03)
Assumed ceiling resistance from ASHRAE	7.04 (40.00)

The smaller RMS temperature errors in cases 1 and 5 could be explained, therefore, by the offsetting effects of too high an envelope conductance and the calculation of too much solar radiation absorbed in the building. This explanation is consistent with the hourly temperature profiles seen in Figures 9 and 10 where the case 1 predicted temperature was high in the day and low at night, while the case 9 predicted temperature was even higher during the day but not so low at night. This also explains how the heating loads in case 9 could be most accurate while the RMS temperature errors in case 9 were greatest. The large RMS temperature errors were caused primarily by the daytime overprediction of temperature. The greater accuracy in load prediction was still possible because at night the performance of the building was primarily governed by the conductive skin losses. The effect of the overprediction of solar energy adsorbed resulted in the 10% to 17% underprediction of loads in Figure 1, case 9. Finally, the high degree of accuracy in the case 9 peak load predictions is also consistent with this explanation. The code predictions were most accurate when stored solar energy was most depleted and envelope conduction was most dominant.

CONCLUSIONS

This work is part of a multiyear, multilaboratory effort on the part of DOE to improve calculational methods for building energy analyses by collecting high quality detailed data and applying rigorous validation techniques. Although this work is far from complete, several conclusions can be drawn that should help guide future activities.

- Input assumptions based on standard engineering references can cause prediction errors of approximately 60% even when using measured meteorological data.
- Accurate temperature prediction does not guarantee accurate load prediction, nor does it guarantee an accurate temperature prediction on the next building studied. There is evidence of compensating errors giving a false sense of confidence. Any validation methodology must account for the possibility of hidden compensating errors.
- The heating load predictions for the three codes for all cases were within about 7% of each other.
- Even when most input inaccuracies are eliminated using measured thermophysical input data, prediction errors ranging from 10% to 17% have still been found. This can have a large impact on building and HVAC system design options.
- A more detailed level of analysis and experimentation will be necessary to determine if these inaccuracies are caused by unknown remaining external or internal error sources. This additional work should include:
  - Corroborating the conductances measured in the walls and ceiling with an ASTM standard large section clamp-on guarded hot-box.
  - Installing a simpler window assembly in the test house.

  - Developing a measurement technique to determine the amount of solar radiation absorbed in the building.
  - Determining the sensitivity of output accuracy to isotropic versus anisotropic sky models.
- The methodological approach used in this work for skin load dominated buildings should be expanded to include the mechanical systems in commercial buildings.

ACKNOWLEDGMENTS

This work was sponsored by the U.S. Department of Energy, Systems Research and Development Branch, Passive and Hybrid Solar Heat Technologies.

LITERATURE CITED

1. Holtz, M., Program Area Plan for Performance Evaluation of Passive/Hybrid Solar Heating and Cooling Systems, SERI/PR-721-788, Solar Energy Research Institute, Golden, Colo. (1980).

2. Judkoff, R. et al., A Comparative Study of Four Building Energy Simulations: DOE, BLAST, SUNCAT-2.4, DEROB-III, SERI/TP-721-837, Solar Energy Research Institute, Golden, Colo. (1980).

3. Judkoff, R. et al., A Comparative Study of Four Building Energy Simulations, Phase II: DOE-2.1, BLAST-3.0, SUNCAT--2.4, and DEROB-4, SERI/TP-721-1326, Solar Energy Research Institute, Golden, Colo. (1981).

4. Judkoff, R. et al., Validation of Building Energy Simulations, SERI/TR-254--1508, Solar Energy Research Institute, Golden, Colo. (forthcoming).

5. Wortman, D. et al., The Implementation of an Analytical Verification Technique on Three Building Energy Analysis Codes: SUNCAT-2.4, DOE-2.1, and DEROB-III, SERI/TP-721-1008, Solar Energy Research Institute, Golden, Colo. (1981).

6. Hunn et al., Validation of Passive Solar Analysis/Design Tools Using Class A Performance Evaluation Data, LA-UR-82--1732, Los Alamos National Laboratory, New Mexico (1982).

7. Burch, J. et al., Site Handbook for the SERI Validation Test Residence, Solar Energy Research Institute, Golden, Colo. (forthcoming).

8. Burch, J. et al., Site Handbook for the SERI Two-Zone Validation Test Cell, Solar Energy Research Institute, Golden, Colo. (forthcoming).

9. Wortman, D. et al., A Multizone Infiltration Monitoring System, SERI/TP-254--1638, Solar Energy Research Institute, Golden, Colo. (1982).

10. Wortman, D. et al., Empirical Validation of Building Energy Analysis Simulations Using Class A Data from the SERI Validation Test House, SERI/TR-254-1840, Solar Energy Research Institute, Golden, Colo. (forthcoming).