Title
A SUMMARY REPORT OF BUILDING ENERGY COMPILATION AND ANALYSIS (BECA) PART V: VALIDATION OF ENERGY ANALYSIS COMPUTER PROGRAMS

Authors
Wagner, B.S.
Rosenfeld, A.H.

Publication Date
1982-08-01

Lawrence Berkeley Laboratory

UNIVERSITY OF CALIFORNIA

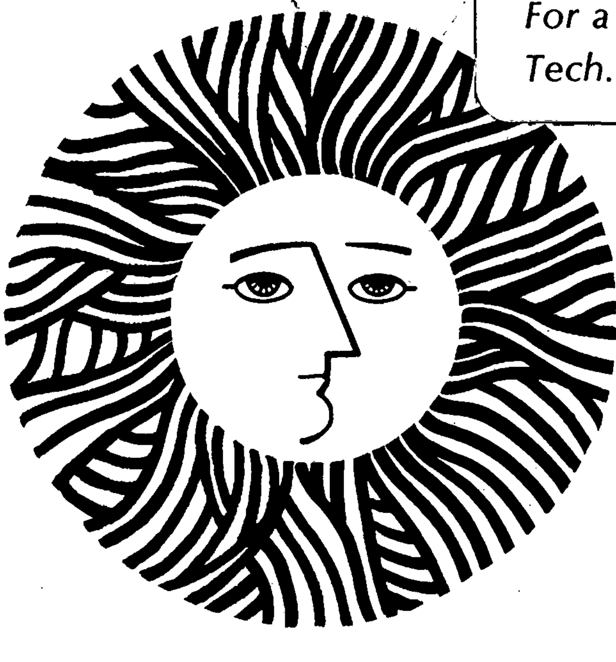
ENERGY & ENVIRONMENT DIVISION

Presented at the Second ACEEE Summer Study on Energy Efficient Buildings, Santa Cruz, CA, August 22-28, 1982

A SUMMARY REPORT OF BUILDING ENERGY COMPILATION AND ANALYSIS (BECA) PART V: VALIDATION OF ENERGY ANALYSIS COMPUTER PROGRAMS

Barbara Shohl Wagner and Arthur H. Rosenfeld

August 1982



Prepared for the U.S. Department of Energy under Contract DE-AC03-76SF00098

DISCLAIMER

LBL-14838
EEB-BED 82-10

Paper to be presented at the Second ACEEE Summer Study on Energy Efficient Buildings, Santa Cruz CA, August 22-28, 1982.

A SUMMARY REPORT OF BUILDING ENERGY COMPILATION AND ANALYSIS (BECA) PART V: VALIDATION OF ENERGY ANAYLSIS COMPUTER PROGRAMS

Barbara Shohl Wagner and Arthur H. Rosenfeld

Energy Efficient Buildings Program
Lawrence Berkeley Laboratory
University of California
Berkeley CA 94720

August 1982

This work was supported by the Assistant Secretary for Conservation and Renewable Energy, Office of Buildings and Community Systems, Buildings Division of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

A SUMMARY REPORT OF
BUILDING ENERGY COMPILATION AND ANALYSIS (BECA)
PART V: VALIDATION OF ENERGY ANALYSIS COMPUTER PROGRAMS

Barbara Shohl Wagner and Arthur H. Rosenfeld
Energy Efficient Buildings Program
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720
August 2, 1982

ABSTRACT

BECA-V assesses the accuracy of computer programs in predictiong measured building energy use. This paper summarizes preliminary results for the 12 studies reviewed to date. For commercial buildings, detailed computer programs were accurate to within about 10% when correct i input data were available. For residential buildings, accuracy of analysis programs is generally better than 10% when the buildings analyzed have been intensively instrumented and monitored to eliminate errors in input. Simplified programs suitable as building energy labelling tools were accurate to within 15% both for submetered and for less intensively monitored houses. Average accuracy for groups of non-submetered houses on a yearly basis was within 15%.

The work described in this report was funded by the Assistant Secretary for Conservation and Renewable Energy, Office of Buildings and Community Systems, Buildings Division of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

1. INTRODUCTION

In calculating building energy consumption for a label, the accuracy of the output of course depends on the accuracy of the following three inputs and algorithms:

1. Weather,

2. Schedules for thermostats, appliances, window management, and venting,

3. Input data that describe a house: U-values, dimensions, infiltration, etc.,

4. Algorithms used in energy analysis, and microclimate corrections to weather data.

In order to make meaningful comparisons between houses, items 1 and 2 can be standardized -- i.e., all programs can be run on standard reference weather, and assumed to have standard schedules and occupancies. The remaining sources of error are then INPUT DATA, and ALGORITHMS. This suggests a two-step process to validate the accuracy of a labelling system:

1. Validation of the program algorithms, by researchers using carefully measured building data and known or simulated occupancy and weather. Occupancy and weather used in validation of algorithms should be in the range in which includes the "standard reference" conditions for the energy label.

2. Certification of users, by testing the professional auditor or labeller's ability to provide the correct inputs.

This paper will focus primarily on the first step. We review studies of the ability of energy analysis programs, given correct occupancy, weather, and construction data, to predict measured energy consumption. In practice, errors in input are not always easily detected, even under research conditions. However, we indicate the source and quality of the input for the studies cited, and point out examples in which known input errors affected energy use predictions. We divide the remainder of the paper into three parts: some brief examples of validation of commercial building analysis; a review of residential building program validations; and conclusions. Studies in the residential building section are divided according to the detail of input data and degree to which effects due to occupancy were controlled or monitored. Such a categorization helps to indicate whether a given comparison was likely to reflect errors from sources other than program algorithms (i.e. measurement of weather, occupant schedules, or building characteristics).

2. COMMERCIAL BUILDINGS

Commercial building energy analysis presents an obvious application of computer programs, because the complexity and expense of the building and heating, ventilating, and air-conditioning (HVAC) systems often require engineering analysis already, and because increasing energy costs also increase the attractiveness of investment in design of more efficient systems. Presumably, then, there has been a strong incentive to develop accurate analysis programs for commercial buildings and the results may indicate the potential for residential energy analysis. Figure 1 summarizes results of three studies of predicted vs. measured energy use in commercial buildings.[1-3] The eleven buildings represent a wide variety of building types, locations, and HVAC systems. Most were simulated as part of a research project, and underwent detailed audits. The results are quite encouraging: Figure 1 shows that the predictions all fell within about 10% of the measured use.

3. RESIDENTIAL BUILDINGS

3.1 Comparisons in Instrumented Buildings, Unoccupied and Monitored Occupancy

An early program with both the capability to model residential buildings and extensive experimental verification of algorithms is the National Bureau of Standards Load Determination (NBSLD) program. NBS has undertaken verification of NBSLD in their environmental test chamber using full-scale buildings with known thermal characteristics. The environmental test chamber offers the advantage of precisely controllable temperatures and humidities, which can be set to average sol-air temperatures for the monitored building. Effects of wind, rain, and visible solar gain are excluded. Narrowing the range of weather variables, while it restricts the scope of verification, simplifies comparison of predicted and measured data. Future comparisons under natural weather conditions have been planned.

Peavy et. al.[4] compared predicted vs. measured electric heating in a 20'x20'x10' concrete building inside the test cell (shown by "+"s in Fig. 2). Intensive monitoring provided heat flow and temperature distribution data. Temperatures in the test cell simulated a 24 hour diurnal cycle, which was repeated several times to allow transient heat flow patterns to die out before final measurements were made. Five of the ten test runs (one 24 hour cycle per test run) were made with no heating, in order to see the effect on indoor temperature swings of variations in window and internal mass configurations. In the remaining five runs, a thermostat regulated electric heaters to keep indoor temperature constant. The predicted heating load for one 24 hour cycle averaged 4.6% lower than the measured heating load, with a standard deviation (SD) of 3.4%. In these five tests, the predicted maximum heat flow rates (of interest when sizing HVAC equipment for buildings) were within 8% of measured rates, while maximum rates predicted by a conventional steady state method ranged from 29% to 69% higher than measured values. Steady state methods did, however, predict total daily heat load to within an average of 7.2% (see Figure 2), with an SD of 3.6%. The discrepancy in predictive accuracy for maximum heating loads is not surprising in the high mass building, since NBSLD takes thermal storage (thermal lag) into consideration, whereas the steady state model does not.

In a second experiment, Peavy et. al. investigated the thermal performance of a four-bedroom, wood-frame townhouse.[5] Temperature cycles simulated recorded weather conditions for winter, summer, and fall (one day each) in Macon, Georgia and Kalamazoo, Michigan. Both gas and electric furnace systems were tested in seven "winter day" test runs; the eighth run was a "pull-down" test and runs 9 and 10 were "summer" and "fall" day tests, respectively. Lightbulbs and appliances simulated occupancy (including metabolic heat from occupants) by a family of six during five of the tests. For the six tests of interest for simulation on NBSLD, error in predicted 24-hour energy use averaged -0.67%, with an SD of 3.1% (represented by "X"s in Fig. 2).

In conjunction with the development of a detailed building simulation program for the Electric Power Research Institute (EPRI), Sepsy et. al. of the Ohio State University (OSU)[6] instrumented and monitored 57 residences. Data from the most intensively submetered and audited residences (6 single-family dwelling and 3 townhouse apartments, all located in Columbus) were used for modification and verification of the simulation program. Energy consumption predicted by the Residential Energy Analysis Program (REAP) was compared to metered data on an hourly, daily, and total (varying from 7 to 27 days) basis. Three of the houses, operated under a variety of weather and occupancy conditions (including unoccupied) and, in one case, with a variety of heating systems, provided data for a total of ten comparison periods. Average percentage errors for furnace and baseboard heating, weighted by monitoring duration, were under 5%. Predicted heatpump heating was low by 16%, a result attributed by the researchers to inaccurate modelling of defrost cycles. Air conditioning average error was -8.2% (see circles in Fig. 2).

In further validation work on REAP, Talbert et. al. of Battelle, Columbus Laboratories selected for simulation six of the 48 buildings monitored by OSU outside Columbus. [7] Monitored data were less complete than in the OSU study, necessitating engineering assumptions to complete input decks. Missing data ranged from such items as thickness of wall panelling and presence of curtains to critical information, for a few houses, such as indoor temperature and type of air-conditioner. In contrast to the OSU study which was designed to use audit and monitoring data to improve program accuracy, the goal of the Battelle study was to make "blind" assumptions, i.e. with no subsequent revisions of input or algorithms, in order to test the current state of model accuracy.* Of the six heating and four cooling simulations, two cooling simulations were eventually rejected because of critical missing information. Results of the remaining runs are shown in Fig 3 (represented by "+"s and "x"s). Average error was 5.1%, with a SD of 36%. The large SD reflects, in part, one run for which the error was +60%. Weather data for the run were questionable, however, discrepancies were inadequate to explain the total error.

3.2 Comparisons in Submetered, Unmonitored Buildings

Submetering of heating/cooling equipment allows separation of a large fraction of the building-dependent energy use from occupant-dependent energy use (e.g. stove use). Although submetering only the space-conditioning use may be inadequate for detailed verification of complex programs, it can be quite useful, particularly if indoor temperatures are also measured, in overall verification of simple or complex heating models. Dickinson and Sonderegger compared predictions by Lawrence Berkeley Laboratory's (LBL) Computerized, Instrumented Residential Audit (CIRA) to submetered heating data for 14 electrically heated houses near Hanford, Washington.[8] The houses were divided into 3 cells of 4-5 houses each, according to type of retrofits received. Average

*With one exception, a house with nighttime thermostat setback patterns where the simulation was rerun to include the setback.

heating consumption of each cell for each test year was normalized to 1979 weather and compared to CIRA predictions for an "average" house description of that cell. Average error was -3.1%, with a SD of 6.9% (circles in Fig. 3).

3.3 Comparisons in Uninstrumented, Unmonitored Houses

Comparisons of predicted vs. metered energy use in occupied, uninstrumented houses actually tests the accuracy of input (auditor skill) as well as program accuracy. However, in the following three comparisons, the audit and input were made by researchers, and in the second two comparisons some houses were instrumented and submetered for at least a short time. Further, input requirements for the three programs tested were less complex than for the detailed comparisons discussed elsewhere in this paper. For these reasons, comparison results should reflect reasonably accurate input in all cases.

Wagner[9] analyzed total yearly gas consumption of eight tract houses in Walnut Creek, California, comparing CIRA predictions to utility bills. Because weather files for Walnut Creek was unavailable at that time, she used data from Sacramento. Comparison of average monthly temperatures for the two locations showed agreement to within a few °C. The houses had been selected for another experiment in part because of a good degree-day vs. gas use correlation, which may have biased the results favorably. However, it also quite possible that the houses whose gas use correlated poorly with standard degree days either 1) had heated pools, spas, or hot tubs, which would not have affected CIRA's ability to accurately predict the building energy use; or 2) practiced thermostat setbacks which CIRA can model, but which do not necessarily yield good correlations. Average error in total annual gas use was -0.26%, with a SD of 14% (circles in Fig. 4).

Using a simplified computer program designed for analysis of space heating in residential buildings (HOTCAN), Dumont et. al.[10] analyzed heating consumption of 9 detached houses in Saskatoon. Most of the buildings used substantial passive solar heating (fuel requirements for space heating ranged from 36 to 80 GJ/year) and all but one used air-to-air heat exchangers. Three of the houses were submetered; heating use for the others was estimated by subtracting summer base use and adjusting for number of occupants. Interior temperature estimates were based on one week hygrothermograph measurements. Average error was -1.6%, with an SD of 9.7% ("+"s in Fig. 4).

Most of the studies cited in this paper investigated heating use. Cooling energy use, in general, presents a much more difficult analytical problem, since there are a larger number of important driving forces, temperature differentials are smaller, and use of an air-conditioner often appears more discretionary than furnace use. Comparison to non-submetered utility data by separation of cooling and estimated base appliance load is also difficult. In a study to carry out such an analysis, Vine et. al. of LBL[11] compared total electric and cooling use predicted by DOE-2.1A to metered total and estimated cooling use in 74 single family dwellings in Davis, California. Audit data were obtained by utility auditors and verified from city records and discussions with builders, under the constraint that the detail be no greater than that obtainable in a relatively low-cost audit program. Thermographs recorded indoor temperatures in twenty-two houses. Predicted total electricity use for this subset of houses, using thermograph data for inside temperatures, was 26% higher than the measured average. Substituting occupant-reported indoor temperatures in the same group reduced the overprediction to 14% The prediction for all 74 houses was 18% high (triangles in Fig. 4.). Regression analysis of predicted vs. actual cooling loads (the latter estimated from billing and audit data) yielded an r^2 of only 0.12. The r^2 increased to 0.30 when actual consumption in April was used to estimate electric baseload, rather than audit data. Problems with audit and weather data are cited to explain the degree of error in individual estimates of electric use.

In a sufficiently large sample of houses, variations in occupant effects should average out, so that a prediction of average use may be accurately based on average inputs. Colborne et. al. of the University of Windsor analyzed electric heating use of 75 houses in Windsor, Canada.[12] The houses were all of the same design, built by the same contractor, and in the same subdivision. Indoor temperature, infiltration, and exact internal gain data were unavailable. Standard assumptions were made for all unknown data and were not subsequently changed to improve agreement of predicted and measured consumption. The average heating use of the 75 houses was estimated by subtracting the summer baseload from each billing period. DOE-2 simulations of the house with and without basement heating came within 9% of measured use; weighting the simulated results by the percentage of heated and unheated basements (determined by audit) yields an average predicted consumption which is within 3% of the measured average. Given the large number of assumptions and the lack of submetered data it is not possible to tell how much of the agreement is attributable to 1)cancelling of errors in inputs and algorithms; 2) cancellation of variations in occupant use; and 3) accuracy of algorithms. A modified degree day calculation using the same input assumptions yields a weighted average error of 9%, which suggests that the DOE-2 algorithms increased the accuracy of calculations compared to degree day analysis (see Fig 4).

4. CONCLUSIONS

Tables 1 through 4 summarize the results from the 12 studies reviewed. Accuracy of energy analysis programs was generally better than 10% for buildings with accurate construction data and carefully monitored or controlled operation. Accuracy tended to decrease as quality of input data decrease, but for buildings with submetered data or reasonably detailed audit data, 2 simplified programs, suitable for audit/labelling programs, predicted heating and total gas consumption within 10-15%. These results are encouraging, if the goal of a labelling system is in the range of $\pm 15\%$ accuracy in predictions of "standard house" consumption. However, we note that:

1. The number of comparisons cited is still small, as is the total building sample size.

2. Comparisons of predicted vs. measured cooling consumption are scarce, and, to date, not encouraging on an individual building basis

3. We know of no experimental results from tests of the accuracy of auditor inputs.

4. Published studies tend to reflect comparisons in which accuracy was relatively good. It would be useful to know when and why computer predictions fail.

This paper is a preliminary version of a more complete review of energy analysis program validations. We invite contributions to our continuing data collection and analysis effort.

REFERENCES

1. S.C. Diamond, B.D. Hunn, C.C. Cappiello; DOE-2 Verification Project, Phase I, Interim Report; Los Alamos National Laboratory, LA-8295-MS; 1981. (DOE-2)

2. Yuill, GK, EG Phillips, "Comparison of BLAST Program Predictions with the Energy Consumptions of Two Buildings"; ASHRAE Transactions, Vol 87, Pt. 1, CH-81-18, No. 2, 1981.

3. D. Herron, L. Windingland, D. Hittle; Comparison of Building Loads Analysis and System Thermodynamics (BLAST) Computer Program Simulations and Measured Energy Use for Army Buildings; Construction Engineering Research Laboratory, US Army Corps of Engineers, Interim Report CERL-IR-E-161; Champaign, IL; 1980.

4. Peavy, BA, FJ Powell, DM Burch; Dynamic Thermal Performance of an Experimental Masonry Building; US National Bureau of Standards, Building Science Series 45; Washington, DC; 1973.

5. Peavy, BA, DM Burch, CM Hunt, FJ Powell; Comparison of Measured and Computer-Predicted Thermal Performance of a Four-Bedroom Wood-Frame Townhouse; US National Bureau of Standards, NBS Building Science Series 57, Washington, DC; 1975.

6. Sepsy, C, MF McBride, RS Blancett, and CD Jones; Fuel Utilization in Residences; report by The Ohio State University for the Electric Power Research Institute; EA-894, Research Project 137-1; Palo Alto, CA; 1978.

7. Talbert, SG, FE Jakob, RD Fischer, Analysis of Field Test Data on

Residential Heating and Cooling; prepared by Battelle, Columbus Laboratory for the Electric Power Research Institute; EA-1649; Palo Alto, CA; 1980.

8. Bruce Dickinson, LBL, personal communication.

9. BS Wagner, Lawrence Berkeley Laboratory, personal communication.

10. Robert Dumont, Division of Building Research, National Research Council Canada, personal communication.

11. E. Vine, P. Craig, J. Cramer, T Dietz, B Hackett, D Kowalczyk, M Levine; The Applicability of Energy Models to Occupied Houses: Summer Electric Use in Davis; forthcoming in Energy.

12. Colborne, WG, NW Wilson, and JD Hall, "A Comparative Study of Methods to Determine Annual Energy Consumption for Space Heating", University of Windsor, presented at Energex 82, Regina, Saskatchewan.

TABLE 1. Commercial Buildings: Comparisons of Predicted vs. Metered Energy Use

A	B	C	D	E	F	G
Program Name	Ref. #	# of Bldgs	Onsite Weather	Systematic Error %[1]	SD[2] %	Duration of Monitoring, Comments
DOE-2.0A	1	7		-3.4	7.7	1 year. Monthly SD~17%.
BLAST	3	1	Y	+12.1	NA	1 month. Dental clinic. Erroneous occupancy gave -7.2%, but correction improved individual predictions of cooling and appliance use
BLAST	3	1	Y	-5.2	NA	1 month. Battalion HQ, total electric use. Erroneous occupancy gave +49.6%
BLAST	2	1		-13 E -8 S	NA NA	Office building. E= electricity, S= steam
BLAST	2	1		-1 E +0.3 G	NA NA	Office building. E= electricity, G= gas

[1]Systematic error is the average of ((Predicted - Measured)/Measured).

[2]SD= standard deviation of percentage error for several runs or buildings NA= not available or not applicable

TABLE 2. Residential Buildings, Intensively Instrumented and Monitored Predicted vs. Measured Energy Use

A	B	C	D	E	F	G
Program Name	Ref. #	# of Bldgs	Onsite Weather	System Error %[1]	SD[2] %	Duration of Monitoring, Comments
NBSLD	4	1	Y	-4.6	3.4	1 day, 5 simulation runs. Massive house in environmental test cell.
NBSLD	5	1	Y	-0.67	3.1	1 day, 6 simulation runs. Townhouse in environmental test cell.
REAP	6	3	Y	-3.4*		1-4 week. *Time weighted average error.

[1]Systematic error is the average of ((Predicted - Measured)/Measured).

[2]SD= standard deviation of percentage error for several runs or buildings NA= not available or not applicable

TABLE 3. Residential Buildings; Submetered, no Monitoring. Predicted vs. Measured Energy Use

A	B	C	D	E	F	G
Program Name	Ref. #	# of Bldgs	Onsite Weather	Systematic Error %[1]	SD[2] %	Duration of Monitoring, Comments
REAP	7	6		5.1	36	1 week.
CIRA	8	14		-3.1	6.9	1 year. 3 groups of 4-5 houses each.

[1]Systematic error is the average of ((Predicted - Measured)/Measured).

[2]SD= standard deviation of percentage error for several runs or buildings NA= not available or not applicable

TABLE 4. Residential Buildings; No Submeters or Monitoring. Predicted vs. Measured Energy Use

A	B	C	D	E	F	G
Program Name	Ref. #	# of Bldgs	Onsite Weather	Systematic Error %[1]	SD[2] %	Duration of Monitoring, Comments
CIRA	9	8	N	-0.26	14	1 year. Tract houses.
HOTCAN	10	9		- 1.6	9.7	1 year. Detached houses in Saskatoon.
DOE2.1A	11	74	Y	+18	NA	3 months. Total electricity use in detached houses in Davis, CA.
"	"	22	Y	+26	NA	
DOE-2	12	75		3	NA	1 year. Space heating average of identical houses in Windsor.

[1]Systematic error is the average of ((Predicted - Measured)/Measured).

[2]SD= standard deviation of percentage error for several runs or buildings

NA= not available or not applicable

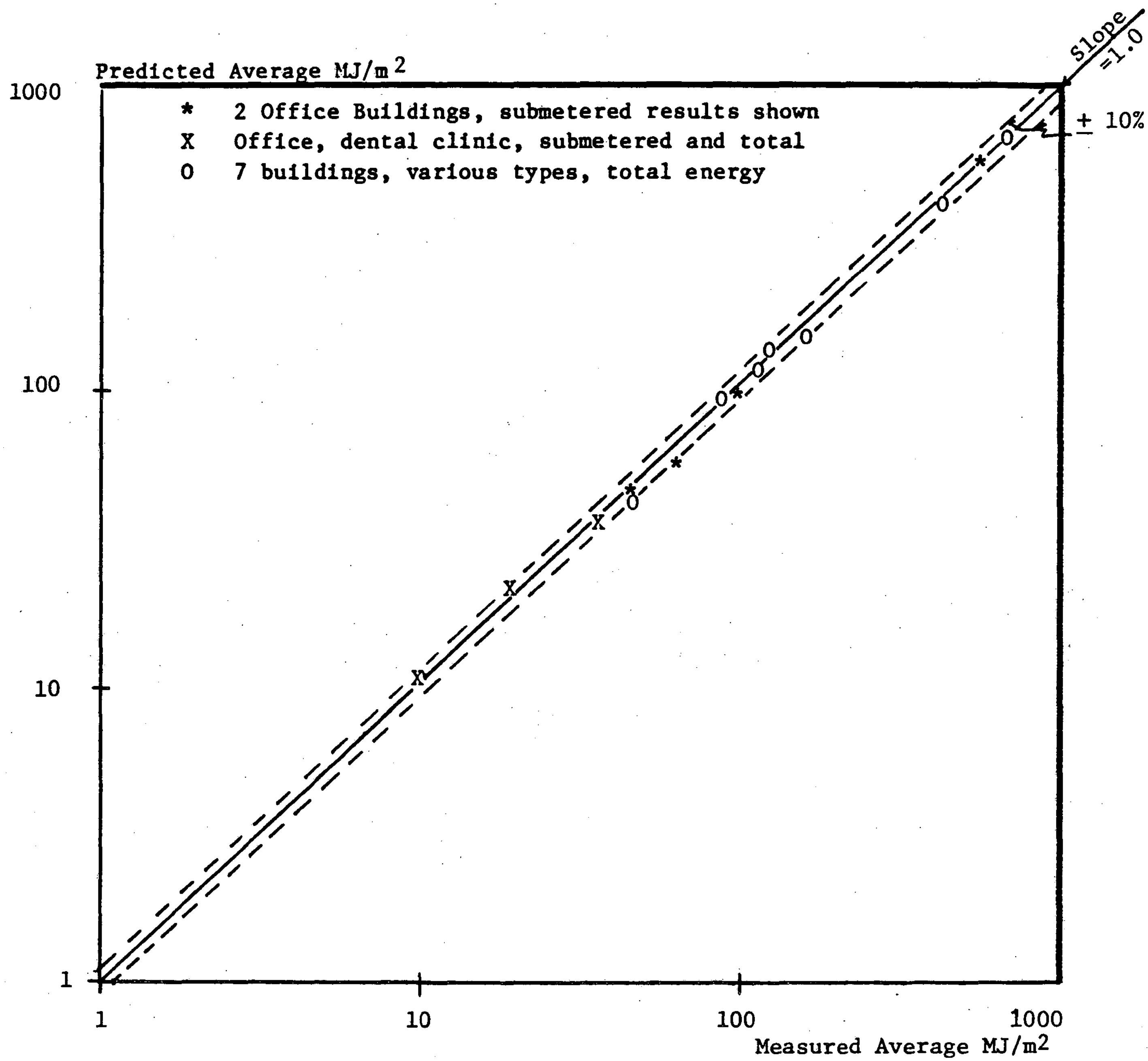


Fig. 1. Commercial Buildings; Predicted (DOE-2 and BLAST) vs. Metered Site Energy Use, averaged over Metering Period (1 month to 1 year).

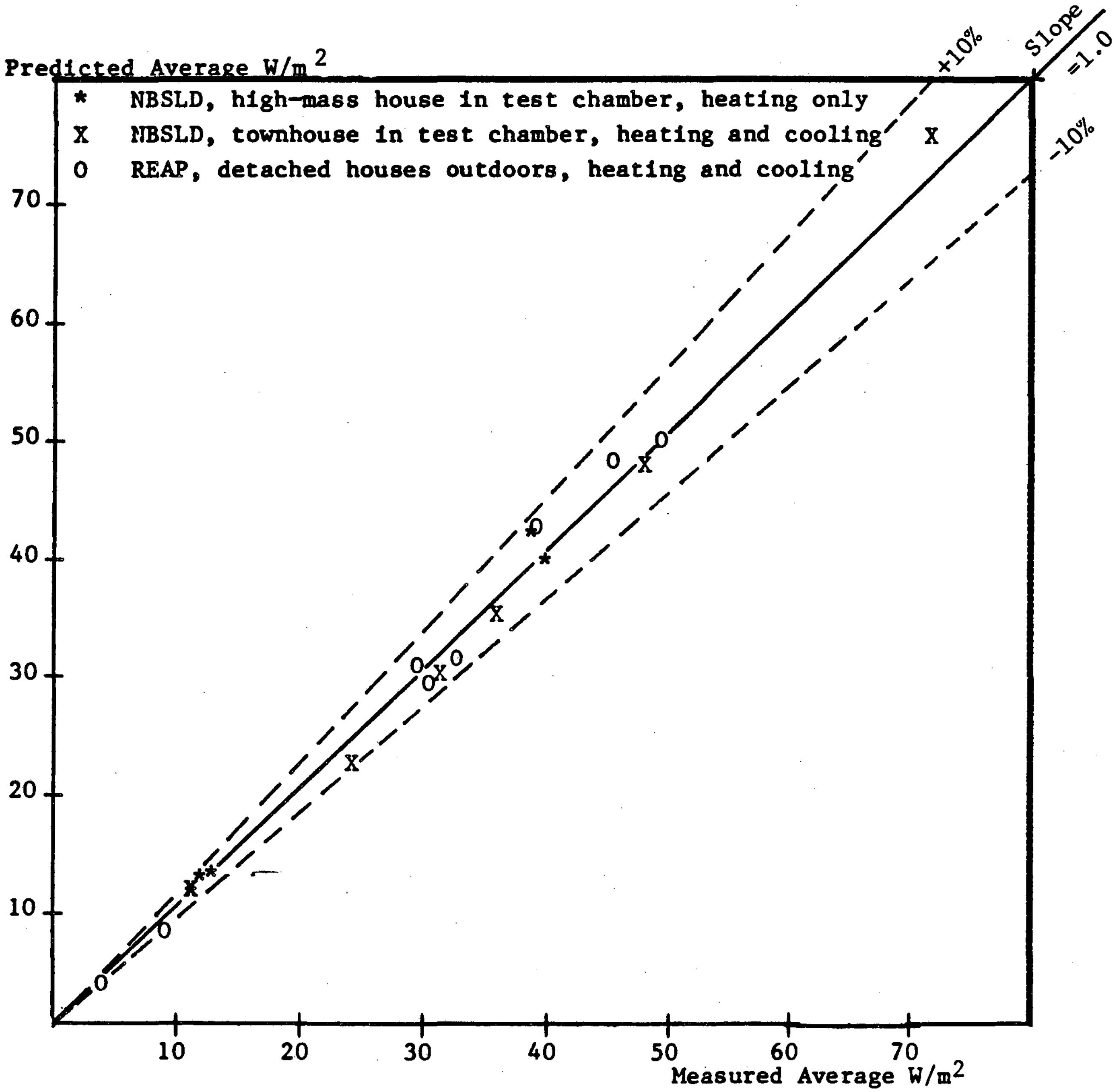


Fig. 2. Residential Buildings; Intensively Instrumented and Monitored. Predicted vs. Metered Site Energy Use, Averaged over Monitoring Period (1 day to 1 year).

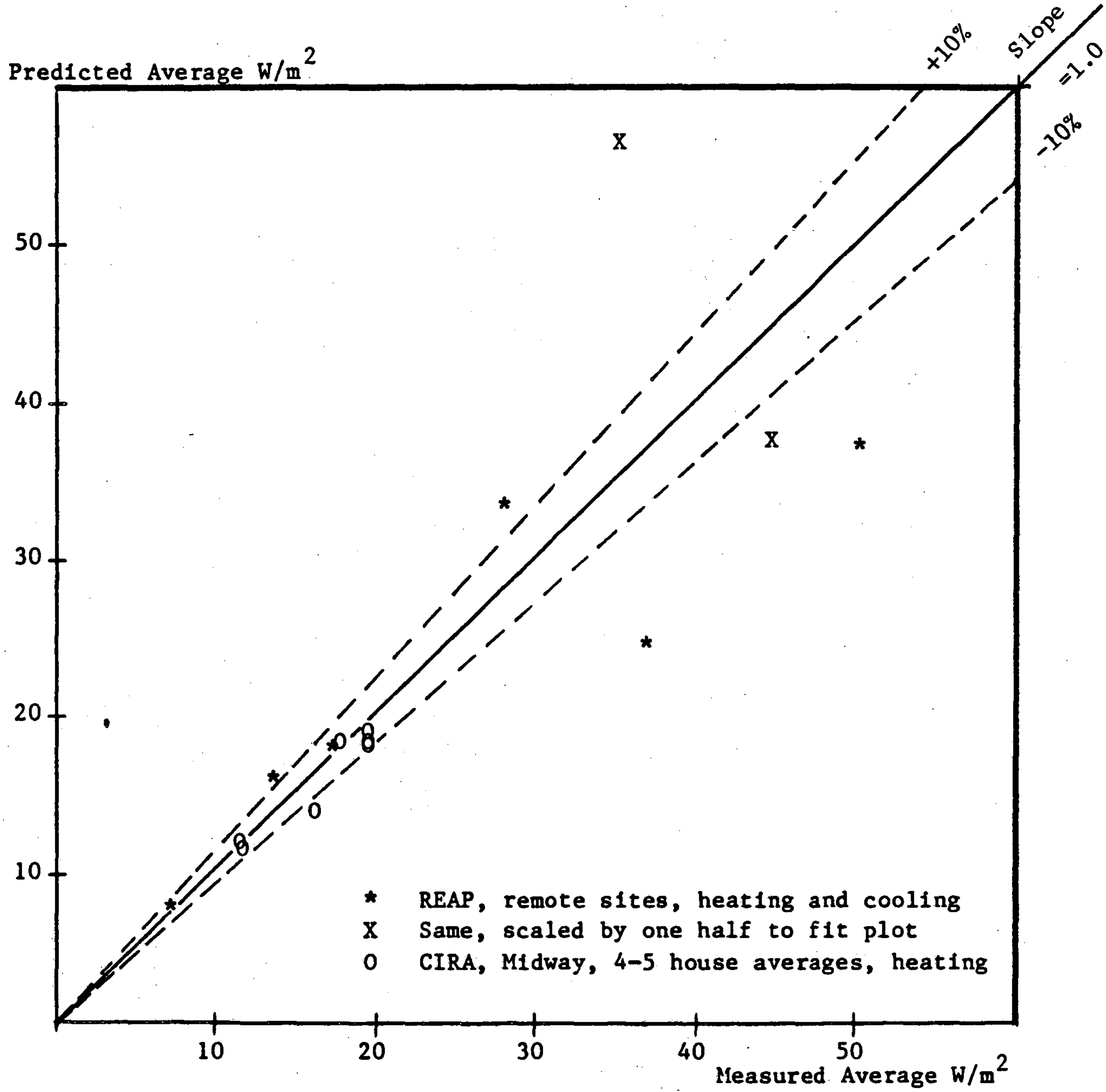


Fig. 3. Residential Buildings; Submetered, no Intensive Monitoring. Predicted vs. Measured Site Energy Use, Averaged over Monitoring Period (4 day to 1 year).

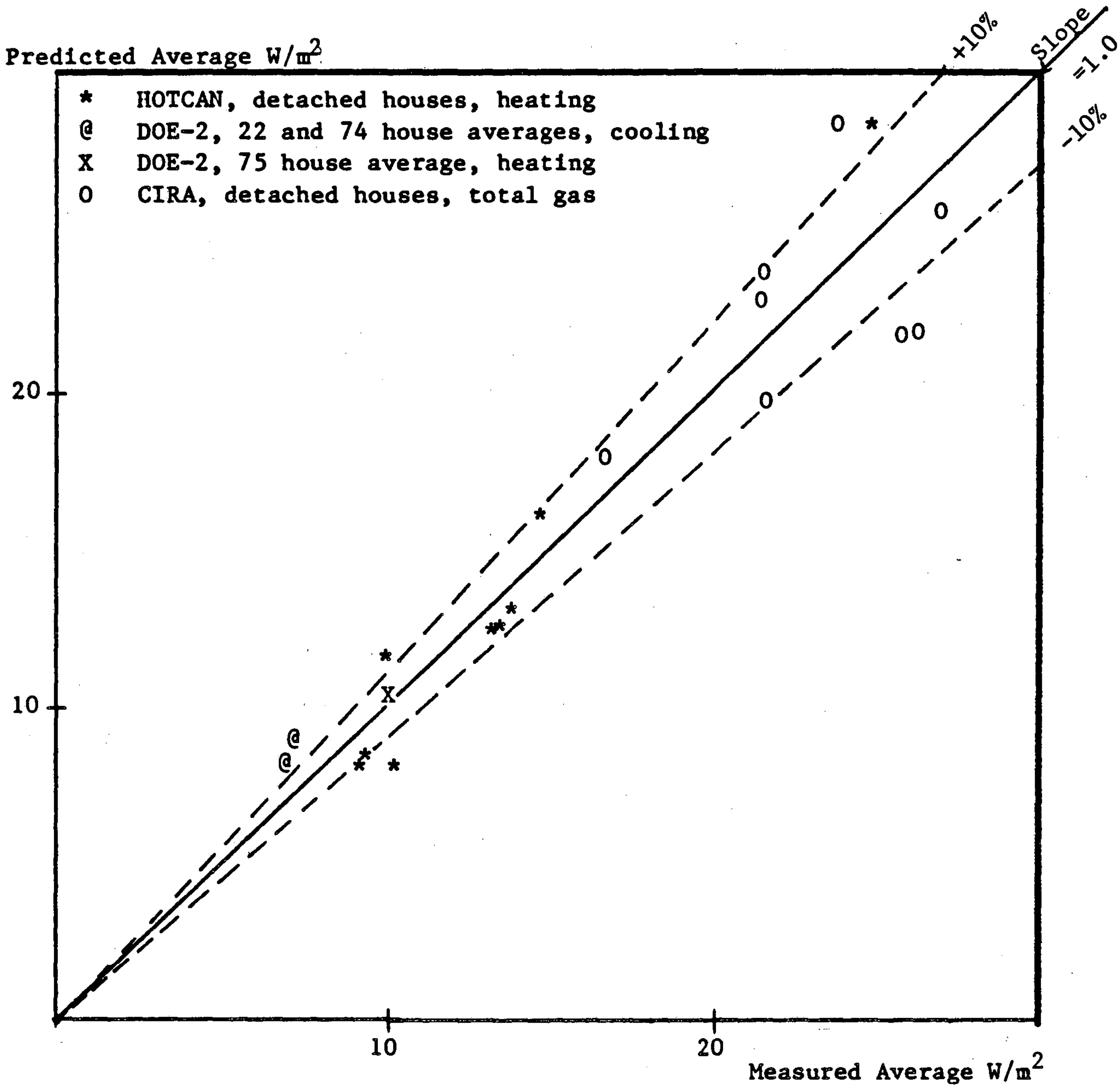


FIG. 4. RESIDENTIAL BUILDINGS; NO SUBMETERS OR MONITORING. PREDICTED VS. MEASURED SITE ENERGY USE, AVERAGED OVER MONITORING PERIOD (3 MONTHS TO 1 YEAR).

TECHNICAL INFORMATION DEPARTMENT
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720