SERI/TP-254-3371
UC Category: 232
DE88001173

International Energy Agency Design Tool Evaluation Procedure

R. Judkoff (SERI)

S. Barakat (National Research Council, Canada)
D. Bloomfield (Building Research Establishment, United Kingdom)
B. Poel (Bouwcentrum, Holland)
R. Stricker (Fraunhofer Institute of Building Physics, W. Germany)
P. van Haaster (Bouwcentrum, Holland)
D. Wortman (Architectural Energy Corporation, USA)

July 1988

Prepared for
Annual Meeting of ASES
Cambridge, Mass.
June 20-24, 1988

Solar Energy Research Institute
A Division of Midwest Research Institute

1617 Cole Boulevard
Golden, Colorado 80401-3393

Prepared for the
U.S. Department of Energy
Contract No. DE-AC02-83CH10093

NOTICE

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 A01
Printed Copy A02

INTERNATIONAL ENERGY AGENCY DESIGN TOOL EVALUATION PROCEDURE

R. Judkoff, Solar Energy Research Institute, USA
S. Barakat, National Research Council, Canada
D. Bloomfield, Building Research Establishment, United Kingdom
B. Poel, Bouwcentrum, Holland
R. Stricker, Fraunhofer Institute of Building Physics, W. Germany
P. van Haaster, Bouwcentrum, Holland
D. Wortman, Architectural Energy Corporation, USA

ABSTRACT

Detailed state-of-the-art building energy simulation models from nations participating in International Energy Agency (IEA) Task VIII are used to develop a quantitative procedure to evaluate more simplified design tools. Simulations are performed with the detailed models on a series of cases that progress systematically from the extremely simple to the relatively realistic. Output values for the cases, such as annual loads, annual maximum and minimum temperatures, and peak loads, are used to set target ranges with which the results from more simplified design tools can be compared. The more realistic cases, although geometrically simple, test the ability of the design tools to model such combined effects as thermal mass, direct gain windows, overhangs, internally generated heat, and dead-band and set-back thermostat control strategies.

INTRODUCTION

With the increased acceptance of personal computers has come a proliferation of building energy design tool software. A survey among IEA participating countries listed 215 such design tools, 156 of which were developed in the United States alone (1). There is little if any objective quality control of this software. It is essentially a buyer beware situation. An evaluation of a number of design tools carried out in IEA Task VIII showed large unexplained differences between these tools, even when run by relatively "sophisticated" users (2). It is important that the design industry not become disillusioned with these tools, since the potential for energy savings and comfort improvements are great through their use.

In recognition of this problem, a joint effort was begun in 1986 under IEA Task VIII to investigate the possibility of developing a quantitative design tool evaluation procedure and to provide a technical basis for doing so. In this paper, we summarize the results of that effort. A full-length IEA Technical Report will also be available on this topic by the time this paper is published. Based on this work, several of the participants have indicated the intention to develop evaluation procedures "customized" for conditions in their own countries. Recently, concern about this issue in the United States prompted the formation of a Standards Subcommittee within ASHRAE TC-4.7. The goal of that subcommittee is to formulate standard test procedures for building energy design and analysis tools.

APPROACH

It was not the intention of this project to create procedures for "validating" design tools. This would have been a large, if not impossible, undertaking, and was beyond the scope of Task VIII. Instead, the strategy was to use a number of detailed computer programs to generate "reference" data against which more simplified tools could be compared. The reference programs were selected by experts in each of the participating nations as representing the current state of the art in building energy simulation. These reference programs had been subjected to a number of validation studies both within the context of IEA Task VIII and by individual countries (3,4). It is well known from these studies and others that even the reference programs frequently differ depending on the climate and building-type modelled (5). However, in this project we chose to accept legitimate internal modelling differences between the reference codes to establish useful ranges of target output values for the simplified design tools. Legitimate modelling differences are those not due to input errors or code bugs.

The objective of the project was to see if a series of cases could be developed for which agreement between the reference programs was adequate to help in determining A) the overall credibility of a design tool and B) the appropriateness of a design tool for a given application. Additionally, it was hoped that the reference cases would be useful for diagnostic purposes when a design tool disagreed significantly with the reference simulations. While such disagreement would not necessarily mean that the design tool was faulty, it would indicate that the design tool disagreed with the current state of the art in building energy simulation as defined by the project team.

SPECIFICATION OF THE CASES

A series of buildings were specified that proceeded from the thermally simple to the realistic one parameter at a time. The cases were defined so that thermal properties, geometric proportions, and thermal responses were meaningful in terms of actual residential buildings. Since this was an international effort, the building specification represented a compromise between American and European construction. In general, the buildings contained more thermal mass than common in the United States.

The cases were jointly defined by the project team to test the ability of design tools to calculate shell loads in residential-type buildings. Such features as thermal mass, direct gain windows, window overhangs, internally generated heat and dead-band and set-back thermostat control strategies were included to emulate those found in a variety of residential buildings. We did not define cases that required simulating mechanical equipment operation. The mechanical equipment was assumed to be 100% efficient and adequately sized to meet the peak load.

Figure 1 shows the basic building geometry. The building geometry remains constant for all cases except for size and orientation of windows and the presence or absence of overhangs. Table 1 lists the key thermophysical properties of the building. These properties either define a heavyweight or lightweight case. The floor was assumed to be thermally isolated from the ground because the coupling between the building and the ground is not well modelled even in the detailed simulation programs. Material thicknesses were adjusted so that the conductance of the walls, roof, and floor were equal in the light and heavy cases. Table 2 lists the characteristics of each case.

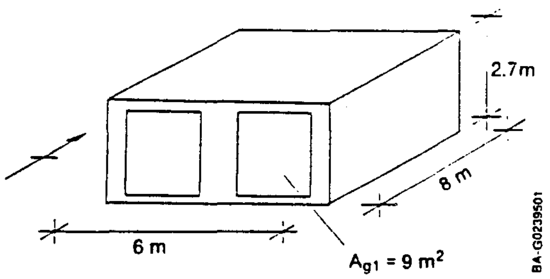

Fig. 1 The Basic Building Geometry

INPUT EQUIVALENCY AND REFERENCE CODE DIFFERENCES

All the project participants were experienced in building energy simulation. They were aware of the need for maintaining strict equivalence between the input files of the various programs. We jointly developed the building specifications from which the input files would be derived. We attempted to document the specifications as unambiguously as possible. Even so, several iterations were required until we could reach mutual agreement on the equivalence of our input files. Input equivalence is not always a straightforward concept, especially where the modelling approach is very different between codes. Even where no input errors exist, legitimate differences in interpretation can lead to significant differences in simulation results. In some instances, we ran parametric studies to better understand the consequences of different modelling approaches. Some of these parametrics are discussed later in the paper.

During the project, we sometimes observed relatively large output differences between the "reference" codes. A serious investigation of those differences would have been a large project and was beyond the primary mission of the working group. As a group, we tried to weed out input-induced differences. Internal modelling differences were considered part of the legitimate output scatter. Where the reference codes differed markedly, a large target range would exist for the simplified design tools. Where the reference codes showed close agreement, a tight target band would exist for the design tools. In several instances, participating countries withdrew a reference code for debugging when an internal problem was suspected. Thus, the procedure proved useful for identifying problems in the reference codes as well as the simplified design tools.

The reference codes originally selected were:

ESP and HTB2	England
SERIRES	England and Holland
DEROB	West Germany
BLAST-3.0 and DOE2.1C	USA

The design tools were run voluntarily by groups interested in the procedure. These were:

EASI and ENERPASS	Canada
BREDEM	England
EBIWAN	Austria

TABLE 1. THERMO-PHYSICAL PROPERTIES

LIGHTWEIGHT			HEAVYWEIGHT		
Element	Thickness (MM)	R-Value	Element (m^2K/W)	Thickness (MM)	R-Value (m^2K/W)
EXT. WALL		1.92	EXT. WALL		1.92
Plasterboard	12		Plaster	16	
Batt	47		Conc. Block	100	
Cavity	29		Foam	50	
Plywood	9		Cavity	50	
Cavity	50		Brick	102	
Brick	105				
INT. WALL			INT. WALL		
Plasterboard	12		Plaster	16	
Cavity	50		Conc. Block	100	
Plasterboard	12		Plaster	16	
FLOOR		25	FLOOR	25	
Timber	25		Screed	50	
Insulation	1003		Conc. Slab	150	
			Insulation	1000	
ROOF		3.13	ROOF		3.13
Plasterboard	10				
Batt	100				
Cavity	25		(Same as lightweight)		
Fiberboard	13				
Asphalt	19				
WINDOW GLASS		.37	WINDOW GLASS		.37
Double Pane			Double Pane		

TABLE 2. CHARACTERISTICS OF THE CASES

Case #	Set-Points Heat (°C)	Cool	Mass	Glass (m^2)		Infilt (ach)	Intgen (W)	Base Case	Other
0	20	20	LW	0	S	0	0		
1	20	20	LW	0	S	1	0		
2	20	20	HW	0	S	1	0		
3	20	20	LW	9	S	1	0		
4	20	20	HW	9	S	1	0		
7	20	27	LW	0	S	1	0		
8	20	27	HW	0	S	1	0		
9	20	27	LW	9	S	1	0		
10	20	27	HW	9	S	1	0		
11	FLOATING		LW	9	S	1	0		
12	FLOATING		HW	9	S	1	0		
13	20	27	LW	4	S	1	200	9	
14	20	27	HW	4	S	1	200	10	
15	20	27	LW	4	E	1	200	13	
16	20	27	HW	4	E	1	200	14	
17	20	27	HW	9	S	1	200	10	overhang Denver only
18	FLOATING		HW	9	S	1	200	10	overhang Denver only
19	SET-BACK		LW	4	S	1	200	13	
20	SET-BACK		HW	4	S	1	200	14	
21	20	27	HW	9	S	1	200	10	adiabatic E,W,N walls
22	20	27	LW	9	S	1	200	9,13	
23	20	27	HW	9	S	1	200	10,14	
24	SET-BACK		LW	9	S	1	200	19	
25	SET-BACK		HW	9	S	1	200	20	
26	FLOATING		LW	9	S	1	200	22	Denver only
27	FLOATING		HW	9	S	1	200	23	Denver only

Where: LW = Lightweight
HW = Heavyweight
Infilt = Infiltration in air changes per hour
Intgen = Internally generated heat from lights, appliances, etc.
E,W,N,S = East, West, North, South
20 20 = Heat on if temp < 20°C
Cool on if temp > 20°C
20 27 = Heat on if temp < 20°C
Cool on if temp > 27°C
FLOATING = Temperatures in building allowed to float freely
SET-BACK = Set-back thermostat control from 2300-0700 hrs.
Set-back temperature = 10°C
(A blank cell contains the previous value in the column.)

RESULTS

Simulations were performed for each case using annual hourly statistical weather data from Denver, Colo., and Copenhagen, Denmark. The winter Denver climate is generally clear and cold with about 3600 degree days (°CDD), while that in Copenhagen is cloudy and cold with about 3800 °CDD. It is not possible to show all the results in the context of this necessarily short paper. Here, we show selected results concentrating on heating in the Denver climate. For complete results, please refer to the previously mentioned IEA report.

Figure 2 shows annual Denver heating loads obtained with five detailed simulation programs for cases 0 through 10. This sequence of cases starts with an opaque box in which solid conduction is the dominant mode of heat transfer and finishes with a passive solar example (19% south glass-to-floor area ratio) in which many dynamic heat transfer mechanisms are simultaneously operative. For most of the cases, agreement is quite close. However, in case 4, SERIRES and ESP represent high and low outliers, respectively, with a disagreement of about 40%. The difference between the maximum and minimum result obtained for each case establishes the target band for that case.

Figure 3 shows annual Denver heating loads for cases 13 through 21. These cases are variants of a relatively realistic building with such features as overhangs, east window, and night-setback thermostat added. The "overhang" cases contain a 9-m^2 window, while all other cases in this series contain a 4-m^2 window. Most of the codes agree quite closely in these cases.

Figure 4 shows annual Denver heating loads for cases 22

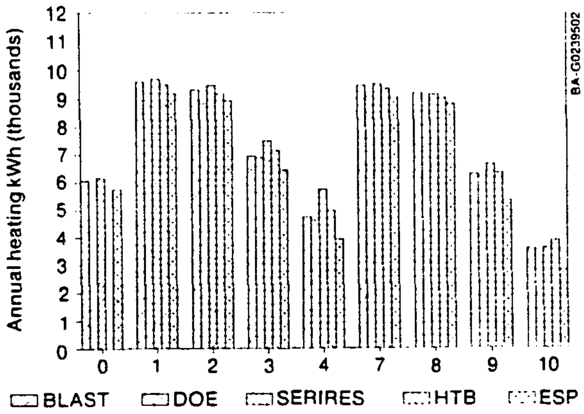

Fig. 2. IEA Reference Codes Cases 0-10 (Denver TMY)

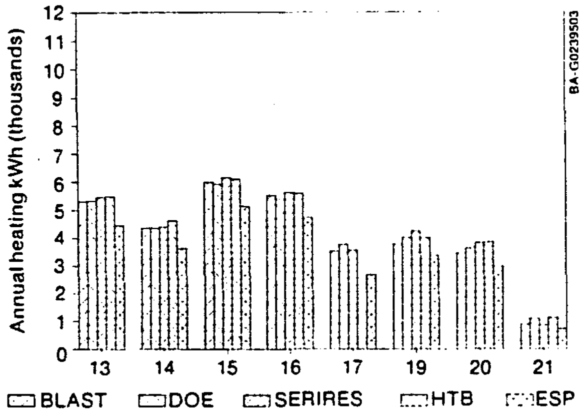

Fig. 3. IEA Reference Codes Cases 13-21 (Denver TMY)

through 25. These cases are similar to those in the previous series except that the window is 9 m^2 instead of 4 m^2.

Figure 5 shows the annual solar energy transmitted through the 9-m^2 south-facing window and absorbed in the building. Both the DOE and ESP codes contain anisotropic sky algorithms that tend to place more solar energy on the south than the isotropic algorithms in BLAST or SERIRES. However, the anisotropic model in ESP appears to place significantly more solar radiation on the south than does the anisotropic model in the DOE code. This would explain the consistently lower ESP heating load predictions in figures 2, 3, and 4. This does not necessarily indicate a fault in ESP. No sky modelling algorithm generally accepted as "correct" has yet been developed.

Figure 6 shows annual heating load "target" ranges in horizontal black lines for cases 0 through 10 in the Denver climate as established from Figure 2. The target ranges were taken from the maximum and minimum reference code predictions for each case. Also displayed are the outputs from three design tools. The most obvious out-of-range prediction is from BREDEM in case 10. Cases 9 and 10 are the light and heavy variants of a fairly realistic case with a south-facing 9-m^2 window and a 20°-27°C banded thermostat control strategy. BREDEM shows no change in heating load prediction from the lightweight to the heavyweight case.

Figure 7 shows identical cases as those in figure 6 except that the climate is Copenhagen. Here, none of the design tools are far out of range. The significance to the designer is

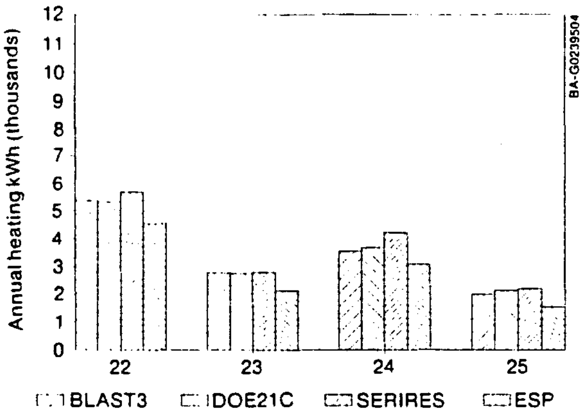

Fig. 4. IEA Reference Codes Cases 22-25 (Denver TMY)

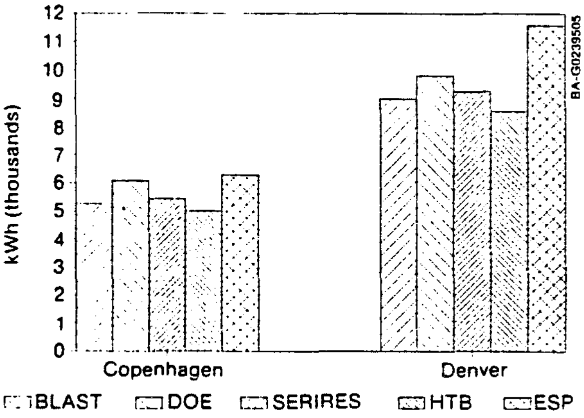

Fig. 5. IEA Reference Codes Solar Gain (Annual gain through 9-m^2 window)

obvious. Some tools will not be appropriate for certain applications. In this example, BREDEM is not appropriate for passive buildings in sunny climates, but it is quite adequate for cloudy climates. The significance to design tool developers is also quite clear. Know the limitations of the method. Document major limitations prominently so as not to mislead designers.

Figure 8 shows the differences in heating loads predicted by the reference programs for the light and heavy buildings. These can be thought of as load savings due to the addition of thermal mass. In some ways, the responsiveness of a design tool to parametric changes is more important than the magnitude of the load prediction for any one case. The designer wants to know what effect design changes will have on the thermal performance of the building. Does the design tool show the right trend by approximately the right amount? If we added BREDEM to figure 8, it would be immediately apparent that this tool shows no response to changes in building capacitance. This may not be a serious drawback for buildings with limited window area or in predominantly cloudy climates. However, in many other instances, use of a tool that is insensitive to thermal mass could mislead the designer.

Note in figure 8 that ESP, which has been somewhat lower than the other reference programs in the prediction of annual heating loads, agrees quite well with the other programs in its response to thermal mass. This was also true for many of the other sensitivities studied, including set-back thermostat control, shading, and window size.

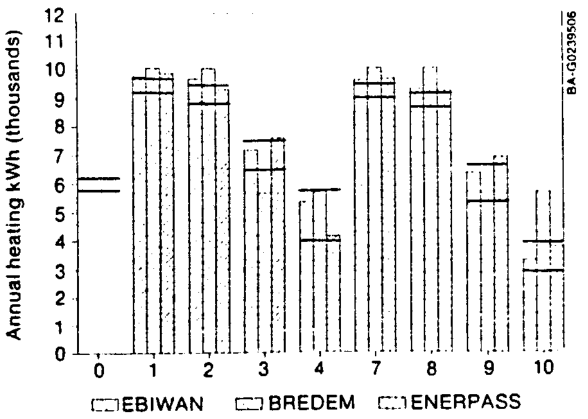

Fig. 6. IEA Design Tool Examples (Denver TMY cases 0-10)

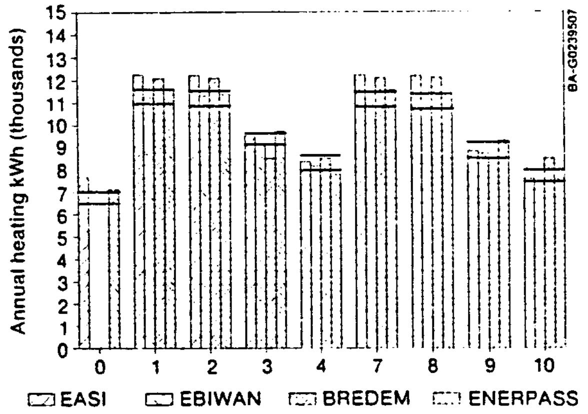

Fig. 7. IEA Design Tool Examples (Copenhagen TMY cases 0-10)

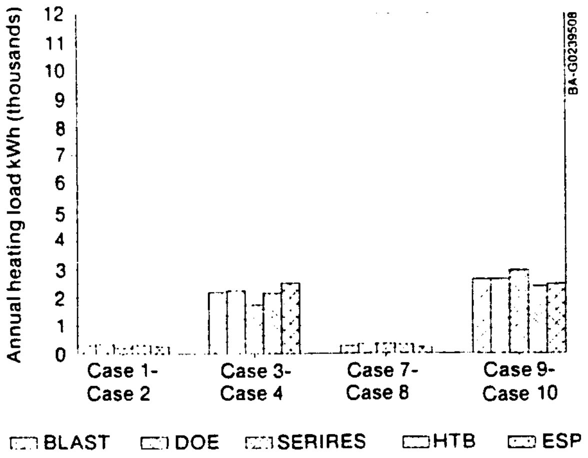

Fig. 8. IEA Reference Codes Mass Response (Denver savings due to added mass)

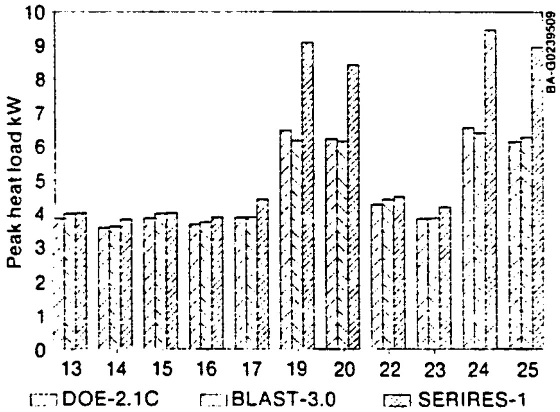

Fig. 9. IEA Reference Codes Peak Heat (Denver annual peak cases 13-25)

SENSITIVITY STUDIES

Figure 9 shows annual peak heating loads in the Denver climate calculated using three of the reference codes. In all cases with a night set-back thermostat control strategy (19,20 and 24,25), SERIRES predicted substantially greater peak loads than the other programs. Since the development of SERIRES was sponsored by the Solar Energy Research Institute (SERI), the SERI participant decided to investigate the disagreements pertaining to cases 4, 19, 20, 24, and 25.

The SERIRES simulations had been performed by researchers in England and Holland. A careful review of the input files for those cases revealed no obvious input problems. However, an interpretive input issue was revealed concerning the interior surface coefficients. The SERIRES Manual advises using the constant combined radiative and convective surface coefficients as defined in the *ASHRAE Handbook of Fundamentals*. The European researchers used the values from the CIBSE guide, which are essentially identical to those in the ASHRAE Handbook. The specifications had called for the thermostat control to be based on the zone air temperature. The control temperature in SERIRES is the conductance weighted average of the interior surface temperatures. Thus, when the normal ASHRAE combined coefficients are used, the control temperature in SERIRES is more like a radiant temperature. This would be similar in physical reality to painting the thermostat black versus silver. In actuality, real thermostats behave somewhere between these extremes, responding to both the radiant and convective environment.

In BLAST and DOE, the thermostat control temperature is modelled more like an air temperature. Both BLAST and DOE calculate the interior surface radiation hourly based on algorithms of varying sophistication. In most instances, these somewhat subtle distinctions have very little effect on thermal performance predictions. However, under certain conditions, the effect can be significant.

For the set-back cases, the standard combined coefficient used in SERIRES couples the building mass to the thermostat control node more closely than in DOE or BLAST. This becomes very noticeable in the calculation of peak loads, which occur at the "set-up" hour. The more closely coupled the building mass to the control node, the harder the heating system has to work to bring the control temperature up to the new thermostat setting. Since the heating systems were sized to always meet peak load, this shows up as a large difference in the peak load.

To test this hypothesis, we reduced the interior surface coefficient input values in SERIRES by the amount normally associated with the radiative portion of the coefficient (about 2/3 radiative to 1/3 convective). We increased the wall conductance by the same amount to keep the overall building load coefficient identical.

Table 3 shows the influence of these changes on annual and peak loads for case 25. As anticipated, the peak loads calculated with SERIRES are now close to those calculated with DOE and BLAST. The effect on annual loads is negligible. They remain close to those for DOE and BLAST, as in the original run.

TABLE 3. RADIANT VS. CONVECTIVE THERMOSTAT CONTROL
Case 25 Denver

	Annual (kWh)	Annual (kWh)	Peak (kW)	Peak (kW)
SERIRES original	2212	1549	8.95	3.48
SERIRES modified	2156	1416	6.45	2.67
BLAST3.0	2001	1291	6.27	2.59
DOE2.1C	2149	1463	6.14	2.69
ESP	1565	1435		

TABLE 4. RADIANT VS. CONVECTIVE THERMOSTAT CONTROL
Case 4 Denver

	Annual Heat	Annual Cool
SERIRES original	5617	5839
SERIRES modified	4729	4611
BLAST3.0	4738	4242
DOE2.1C	4609	4754
HTB2	4956	4416
ESP	3915	4779

For case 4, the source of the disagreement is the same, but the effect is somewhat different. Here, it is the ability to store transmitted solar energy that is affected by the value chosen for the interior surface coefficient. The large effect this has on the annual heating and cooling loads is an artifact of the unrealistically tight thermostat control strategy for that case. Since the control temperature cannot drift, the heating and cooling system response is very sensitive to the coupling between the control temperature and the building mass. When the standard combined interior surface coefficient is used in SERIRES, the coupling is effectively tighter than in BLAST or DOE. Thus, less solar energy is stored in the mass increasing both heating and cooling loads. For more realistic thermostat control strategies, this effect would be much less apparent. To test this theory, we ran case 4 with the same input changes used in case 25.

Table 4 shows the influence of these changes on the case 4 results. As expected, the heating and cooling load predictions from SERIRES now resemble those from BLAST and DOE.

A similar sensitivity study was conducted by Bloomfield in England using the ESP model. ESP allows the user to specify the degree to which the thermostat responds to mean radiant temperature versus "air" temperature. For cases 1 and 9 in the Copenhagen climate, annual heating load differences of 10% and 28% were found, respectively, when thermostat response was based on the air temperature versus a mixed radiant and convective temperature (2/3 radiant, 1/3 air).

CONCLUSIONS

The work conducted to date demonstrates the feasibility of using detailed building energy analysis simulation programs to quantitatively evaluate more simplified design tools. Agreement among the detailed "reference" programs was sufficient to establish reasonable target ranges against which to compare the output from simplified design tools. Cases for which agreement was not close indicate heat transfer phenomena that are somewhat beyond the current modelling state of the art. Target bands were consequently wider, presenting a less severe test for simplified tools. This is legitimate, since the simplified tools should not be expected to surpass detailed tools in modelling capability. The prototype design tool evaluation procedure was successful in uncovering limitations in the sample design tools tested. The procedure was also successful in illustrating the significance of algorithmic differences in the reference programs.

The exact nature of the tests and the definition of what constitutes passing a test remain to be defined. This will probably be done within the context of individual countries, engineering societies, or trade associations. Several of the participants in this international effort have expressed the intention of their respective governments to pursue domestic versions of the design tool evaluation procedure. In the United States, ASHRAE has formed a subcommittee to formulate standard procedures. It is hoped that this work will be useful as an initial basis for those efforts.

It should be reiterated that failing a test does not necessarily indicate a faulty tool. However, the tool developer should understand why the tool disagrees and should warn users of important limitations in the tool. Conversely, passing all the tests does not completely validate a tool. The tests represent a fairly course filter. Simplified design tools should be passed through that filter as a minimum quality assurance procedure. The target ranges and types of tests will evolve as the state of the art in building energy analysis simulation progresses, and our understanding of building physics improves. The outputs of the "reference" programs do not represent truth. They do, however, collectively represent our best current objective knowledge on the calculation of building thermal behavior.

Design tool evaluation procedures will provide practitioners with a rational basis for selecting the appropriate tool for a given application.

REFERENCES

(1) Design Tool Survey, IEA Task VIII. P.R. Rittlemann, S.F. Ahmed. Burt Hill Kosar Rittlemann Assoc., Butler, PA., USA. May 1985.

(2) Design Tool Comparison IEA Task VIII Working Document. P.R. Rittlemann, S.F. Ahmed. Burt Hill Kosar Rittlemann Assoc., Butler, PA., USA.

(3) Simulation Model Validation Using Test Cell Data, IEA Task VIII. Edited by Ove Mork. Technical University of Denmark. June 1986.

(4) Validation of Building Energy Analysis Simulation Programs at the Solar Energy Research Institute, R. Judkoff. ENERGY AND BUILDINGS, Elsevier Sequoia S.A., Lausanne, Switzerland. Volume 10, 1988.

(5) International Energy Agency Building Simulation Comparison and Validation Study, R. Judkoff. Proceedings International Climatic Architecture Congress, Louvain-La-Neuve, Belgium. July 1-3, 1986.

ACKNOWLEDGMENTS

This work was supported by the United States Department of Energy Solar Buildings Program. It was cosponsored under Task VIII of the International Energy Agency. We also would like to thank Manfred Bruck, Austria; Stephen Carpenter, Canada; and Hans Erhorn, West Germany; and Don Alexander and Peter Lewis, United Kingdom, for their contributions to this project.