BESTEST: Comparing Predictions of Building Energy Simulation Software

Ron Judkoff[1]*, Joel Neymark*[2]

CONFIDENCE IN COMPUTER-SIMULATION PREDICTIONS

Numerous software programs simulate energy performance in buildings. But these programs—even if considering identical structural designs, energy-related equipment, and energy usage patterns—often produce widely divergent results when calculating overall energy performance. Consequently, architects and engineers have not trusted the programs, and instead, have continued to design buildings without focusing on energy use.

To improve the accuracy of energy software and help designers gain confidence in computer predictions, the National Renewable Energy Laboratory (NREL), in cooperation with International Energy Agency (IEA) Task 12 and Annex 21, developed IEA BESTEST—the International Energy Agency Building Energy Simulation Test and Diagnostic Method.[1] This procedure systematically compares whole-building energy software packages and determines the algorithms—or computer-coded computational routines—responsible for prediction differences.

BESTEST BASICS

The energy, comfort, and lighting performance of buildings depends on many complex interactions. And computer simulation is the only practical way to bring such a large-scale systems integration problem within the grasp of building designers.

To study these simulation models, the BESTEST technique applies a series of carefully specified test-case buildings that progress systematically from the extremely simple to the relatively realistic. Output values for the cases—such as annual loads, temperature ranges, and peak loads—are compared and used with diagnostic logic to pinpoint the routines responsible for prediction differences.

In IEA BESTEST's 36 cases, *qualification* cases represent a set of buildings with relatively realistic thermal characteristics. These cases test a program's ability to model such features as windows at different orientations, external shading devices, setback thermostats, night ventilation economizer cooling, a passive solar sunspace, and ground coupling. *Diagnostic* cases attempt to isolate the effects of individual algorithms by varying a single parameter from case to case, to minimize confusion caused by interacting heat-transfer phenomena.

BESTEST example results are compared using eight programs run by the various IEA participants as described in Table 1.

HELPING TO DEVELOP ENERGY SOFTWARE

BESTEST helps software developers in several ways. Predictions from a building-energy program of interest can be compared to the reference results from detailed programs already studied, or the algorithm-based differences in predictions observed between several simulation programs can be diagnosed. A previous version of a program can be checked against itself after a programmer has modi[illegible] to ensure that only the intended changes actually resulted. And the sensitivity of an algorithm to changes may be investigated by checking the modified version against the original.

By itself, the BESTEST procedure is not a complete validation method, which would also include analytical and empirical techniques.[2] Instead, it compares a given program with other state-of-the-art programs that have been analytically verified and field-validated with actual buildings. Anomalous results or "failing" a test do not necessarily indicate a faulty program, but rather, differences to be studied and understood.

In practice, the diagnostic procedures have revealed bugs, faulty algorithms, and modeling limitations in every one of the world-class building-energy computer programs studied by NREL and IEA researchers. Figure 1 (next page) indicates results from six simulation programs that fell within the shaded envelope; results from a version of TRNSYS[3] then being used in Europe (TRNSYS 12.2v1) shown by the dotted curve deviate drastically from the other programs. When computer code "bugs" revealed by BESTEST diagnostics were fixed, the results fell within the range of the other programs.[4]

As another example, DOE2[5] is one of the U.S. Department of Energy's most advanced building energy simulation programs. One series of diagnostic tests on DOE2.1D detected problems with the treatment of solar absorptivity on exterior surfaces. Use of IEA BESTEST traced the problem to a bug in the solar absorptance

Table 1 Participating Organizations and Computer Programs

Computer program	Authoring organization	Implemented by
BLAST-3.0 level 193 v.1	CERL[a], U.S.	NREL[b], U.S. Politecnico Torino, Italy
DOE2.1D 14	LANL/LBL[c], U.S.	NREL, U.S.
ESP-RV8	Strathclyde University, U.K.	De Montfort University, U.K.
SERIRES/SUNCODE 5.7	NREL/Ecotope, U.S.	NREL, U.S.
SERIRES 1.2	NREL/BRE[d], U.S./U.K.	BRE, U.K.
S3PAS	University of Sevilla, Spain	University of Sevilla, Spain
TASE	Tampere University, Finland	Tampere University, Finland
TRNSYS 13.1	University of Wisconsin, U.S.	BRE, U.K. Vrije Universiteit (VUB) Brussels, Belgium

[a]CERL—Civil Engineering Research Laboratory
[b]NREL—National Renewable Energy Laboratory
[c]LANL/LBL—Los Alamos National Laboratory/Lawrence Berkeley Laboratory
[d]BRE—Building Research Establishment

[1] *National Renewable Energy Laboratory, 1617 Cole Boulevard, Golden, Colorado 80401-3393, USA*
[2] *independent simulation consultant, Golden, Colorado, USA*

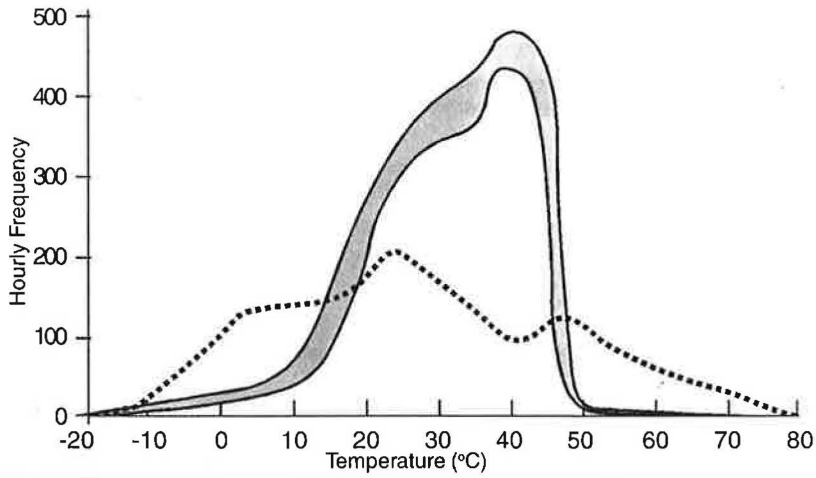

Figure 1 BESTEST Case 900FF: Annual Hourly Temperature Frequency

algorithm associated with surfaces defined as doors. Once the algorithm was corrected, the problem disappeared.[4] Similarly, BESTEST helped isolate a bug that caused improper calculation of weighting factors in a beta-test version of PowerDOE; this bug, which has now been corrected, did not exist in the preceding version of DOE2 (DOE2.1E).[6]

USING VALIDATED ENERGY SOFTWARE

BESTEST is designed to help develop reliable energy software. But the ultimate goal is to assure potential software users that a particular simulation program gives reasonable results or that a program is appropriate for their particular application. NREL has also developed a version of BESTEST that is the basis for testing software in the US National Home Energy Rating System initiative (see box below).[7]

The list of BESTEST users continues to grow. The American Society of Heating, Refrigerating, and Air-Conditioning Engineers is adopting BESTEST as a "standard method of test." Additionally, California's State Energy Commission approved the method for evaluating its building-energy analysis software. Other prominent BESTEST users include:

- New Zealand (for certifying software used for energy code compliance)
- Canada Department of Natural Resources
- National HERS Program (Cited by NOPR 10 CFR 437 and HERS Council Guidelines)
- National Weatherization Program (US)
- Energy Bureau, Iowa State Department of Natural Resources

BESTEST will improve building-energy software and will increase confidence in its predictions among architects and engineers, enabling them to design increasingly energy-efficient buildings.

To obtain a copy of IEA BESTEST[1] and/or HERS BESTEST[7] please contact:

Ron Judkoff
National Renewable Energy Laboratory
1617 Cole Boulevard
Golden, Colorado 80401-3393, USA
Phone: *1 303 384 7538*
e-mail: *ron_judkoff@nrel.gov*

CERTIFYING SOFTWARE FOR THE HOME ENERGY RATING SYSTEM (HERS)

In the US, the National Energy Policy Act of 1992 calls for certifying the technical accuracy of analysis tools that determine building energy-efficiency ratings. NREL has developed software testing protocols with the HERS council, a U.S. group representing utilities, government, environmental and consumer groups, and members of the real estate, finance, and building industries. The protocols are based on NREL's HERS BESTEST—an adaptation of BESTEST specific to residential building models.

Mortage companies willing to make larger loans to cover energy efficiency improvements are especially interested in the accuracy of home energy rating systems. Lenders want some confidence that the energy cost savings will more than offset the increased loan amount. Use of HERS BESTEST increases the credibility of home ratings by screening out inaccurate HERS software.

Software used by NREL in developing example results for HERS BESTEST includes:
DOE2.1E - W54
BLAST 3.0 Level 215
SERIRES/SUNCODE 5.7.

REFERENCES

1 Judkoff, R., and J. Neymark, 1995, *International Energy Agency Building Energy Simulation Test (BESTEST) and Diagnostic Method.* NREL/TP-472-6231. Golden, Colorado, USA: National Renewable Energy Laboratory.

2 Judkoff R., 1988, "Validation of Building Energy Analysis Simulation Programs at the Solar Energy Research Institute (now NREL)." *Energy and Buildings*, Vol. 10, No. 3; pp. 221-239. Lausanne, Switzerland: Elsevier Sequoia.

3 Klein S. et al., Sept. 1990, *TRNSYS A Transient System Simulation Program,* Solar Energy Laboratory, University of Wisconsin, Madison, Wisconsin, USA.

4 Judkoff R.D. and J.S. Neymark., 1995, "A Procedure for Testing the Ability of Whole Building Energy Simulation Programs to Thermally Model the Building Fabric." *Journal of Solar Energy Engineering*, Vol 117, No. 1, pp. 10-12, American Society of Mechanical Engineers, New York, New York, USA.

5 Winkelmann F.C., B.E. Birdsall, W.F. Buhl, K.L. Ellington, A.E. Erdem, J.J. Hirsch and S. Gates., 1994, *DOE-2 Supplement Version 2.1E.* Berkeley, CA: Lawrence Berkeley Laboratory.

6 Hirsch J., James J. Hirsch Associates, Camirillo, California, USA, personal communications with J. Neymark, December 1996 and January 1997.

7 Judkoff, R., and J. Neymark., 1995, *Home Energy Rating System Building Energy Simulation Test (HERS BESTEST).* NREL/TP-472-7332. Golden, Colorado, USA: National Renewable Energy Laboratory.