NREL National Renewable Energy Laboratory

Innovation for Our Energy Future

Model Validation and Testing: The Methodological Foundation of ASHRAE Standard 140

Preprint

R. Judkoff
National Renewable Energy Laboratory

J. Neymark
J. Neymark & Associates

NOTICE

Available electronically at http://www.osti.gov/bridge

Available for a processing fee to U.S. Department of Energy
and its contractors, in paper, from:
      U.S. Department of Energy
      Office of Scientific and Technical Information
      P.O. Box 62
      Oak Ridge, TN 37831-0062
      phone: 865.576.8401
      fax: 865.576.5728
      email: mailto:reports@adonis.osti.gov

Available for sale to the public, in paper, from:
      U.S. Department of Commerce
      National Technical Information Service
      5285 Port Royal Road
      Springfield, VA 22161
      phone: 800.553.6847
      fax: 703.605.6900
      email: orders@ntis.fedworld.gov
      online ordering: http://www.ntis.gov/ordering.htm

Model Validation and Testing: The Methodological Foundation of ASHRAE Standard 140[1]

ABSTRACT

Ideally, whole-building energy simulation programs model all aspects of a building that influence energy use and thermal and visual comfort for the occupants. An essential component of the development of such computer simulation models is a rigorous program of validation and testing. This paper describes a methodology to evaluate the accuracy of whole-building energy simulation programs. The methodology is also used to identify and diagnose differences in simulation predictions that may be caused by algorithmic differences, modeling limitations, coding errors, or input errors. The methodology has been adopted by ANSI/ASHRAE Standard 140 (ANSI/ASHRAE 2001, 2004), Method of Test for the Evaluation of Building Energy Analysis Computer Programs. *A summary of the method is included in the* ASHRAE Handbook of Fundamentals *(ASHRAE 2005). This paper describes the ANSI/ASHRAE Standard 140 method of test and its methodological basis. Also discussed are possible future enhancements to Standard 140 and related research recommendations.*

INTRODUCTION

Ideally, whole-building energy simulation programs model all aspects of energy use and thermal and visual comfort in buildings. Such programs may contain tens of thousands to hundreds of thousands of lines of code. An error in even one character of one line of code can lead to seriously flawed results. Therefore, an essential component of the development of such computer simulation models is a rigorous program of validation, testing, and diagnostics. The United States Department of Energy's National Renewable Energy Laboratory (NREL) has developed an overall methodology to evaluate the accuracy of whole-building energy simulation programs, and to identify and diagnose differences in simulation predictions that may be caused by algorithmic differences, modeling limitations, coding errors, or input errors. This method has been adopted by ANSI/ASHRAE Standard 140 (ANSI/ASHRAE 2001, 2004), Method of Test for the Evaluation of Building Energy Analysis Computer Programs. The method has recently been summarized for inclusion in the ASHRAE Handbook of Fundamentals (ASHRAE 2005).

ANSI/ASHRAE STANDARD 140

ANSI/ASHRAE Standard 140 (ANSI/ASHRAE 2001, 2004), *Method of Test for the Evaluation of Building Energy Analysis Computer Programs*, was the first codified method of test for building energy software in the world, and has recently been referenced by ASHRAE Standard 90.1 (ANSI/ASHRAE/IESNA 2004, 2005) for approval of software used to show performance path compliance. Standard 140 is structured to allow for the addition of all elements of a complete validation approach as they become available. This structure corresponds to the validation methodology introduced above with subdivisions creating a matrix of six areas for testing including:

---

a) Comparative tests – building envelope
b) Comparative tests – mechanical equipment and on-site energy generation equipment
c) Analytical verification – building envelope
d) Analytical verification – mechanical equipment and on-site energy generation equipment
e) Empirical validation – building envelope
f) Empirical validation – mechanical equipment and on-site energy generation equipment.

This is a highly abbreviated way of representing the overall parameter space in which building energy simulation programs operate, and each cell in the matrix represents a very large region in the space. The current set of tests are classified under categories "a" and "d". These tests are based on procedures developed by the National Renewable Energy Laboratory and field-tested by the International Energy Agency (IEA) over three IEA research tasks. (Judkoff and Neymark 1995a; Neymark and Judkoff 2002) The category "a" procedures (ANSI/ASHRAE 2001) were refined and converted to standards language and format by the ASHRAE SPC-140 Standards Project Committee. The category "d" procedures were refined and converted to standards language and format by ASHRAE Standing Standards Project Committee (SSPC) 140 and are included along with the category "a" procedures in the most recent revision of Standard 140 (ANSI/ASHRAE 2004). Continuing oversight of Standard 140 by SSPC 140 allows validation procedures used within Standard 140 to evolve as the state of the art in whole-building energy simulation evolves. Additional tests have been or are being developed under ASHRAE research projects (Spitler, Rees, and Xiao 2001; Yuill and Haberl 2002) and under a joint IEA Solar Heating and Cooling Programme and Energy Conservation in Buildings and Community Systems Task 34/Annex 43 (Judkoff and Neymark 2004) that are intended to fill in other categories of the above validation matrix.

METHODOLOGICAL BASIS FOR ANSI/ASHRAE STANDARD 140

There are only a few ways to evaluate the accuracy of a whole-building energy simulation program (Judkoff et al. 1983):

- *Empirical Validation*—in which calculated results from a program, subroutine, algorithm, or software object are compared to monitored data from a real building, test cell, or laboratory experiment.
- *Analytical Verification*—in which outputs from a program, subroutine, algorithm, or software object are compared to results from a known analytical solution or a generally accepted numerical method for isolated heat transfer under very simple, highly constrained boundary conditions.
- *Comparative Testing*—in which a program is compared to itself or to other programs.

Table 1 compares these techniques (Judkoff 1988). In this table the term "model" is the representation of reality for a given physical behavior. For example, heat transfer may be simulated with one-, two-, or three-dimensional thermal conduction models. The term "solution process" encompasses the mathematics and computer coding to solve a given model. The solution process for a model can be perfect, while the model remains inappropriate for a given physical situation, such as using a one-dimensional conduction model where two-dimensional conduction dominates. The term "truth standard" represents the standard of accuracy for predicting real behavior. An analytical solution is a "mathematical truth standard," but only tests the solution process for a model, not the appropriateness of the model. An approximate truth standard from an experiment tests both the solution process and appropriateness of the model within experimental uncertainty. The ultimate (or "absolute") validation truth standard would be comparison of simulation results with a perfectly performed empirical experiment, with all simulation inputs perfectly defined.

Establishing an absolute truth standard for evaluating a program's ability to analyze physical behavior requires empirical validation, but this is only possible within the range of measurement uncertainty, including that related to instruments, spatial and temporal discretization, and the overall experimental design. Test cells and buildings are large, relatively complex experimental objects. The exact design details, material properties, and construction in the field cannot be perfectly known, so there is some uncertainty about the simulation model inputs that accurately represent the experimental object. Meticulous care is required to describe the experimental apparatus as clearly as possible to modelers to minimize this uncertainty. This includes experimental determination of as many material properties and other simulation model *inputs* as possible, including overall building parameters such as overall steady-state heat transmission coefficient, infiltration rate, and thermal capacitance. Also, required are detailed meteorological measurements. For example, many

experiments measure global horizontal solar radiation, but very few experiments measure the splits between direct, diffuse, and ground reflected radiation, all of which are inputs to many whole building energy simulation programs.

Table 1. Validation Techniques

Technique	Advantages	Disadvantages

Empirical
Test of model and solution process
	• Approximate truth standard within experimental accuracy
	• Any level of complexity
		• Experimental uncertainties:
		  - Instrument calibration, spatial/ temporal discretization
		  - Imperfect knowledge/ specification of experimental object (building) being simulated
		• High quality detailed measurements are expensive and time consuming
		• Only a limited number of test conditions are practical

Analytical Test of solution process	• No input uncertainty • Exact mathematical truth standard for the given model • Inexpensive	• No test of model validity • Limited to highly constrained cases for which analytical solutions can be derived
Comparative Relative test of model and solution process	• No input uncertainty • Any level of complexity • Many diagnostic comparisons possible • Inexpensive and quick	• No absolute truth standard (only statistically based acceptance ranges are possible)

The National Renewable Energy Laboratory (NREL) methodology divides empirical validation into different levels, because many past validation studies produced inconclusive results. The levels of validation depend on the degree of control over possible sources of error in a *simulation*. These error sources consist of seven types divided into two groups:

External Error Types

- Differences between the actual microclimate that affects the building versus weather input used by the program
- Differences between actual schedules, control strategies, effects of occupant behavior, and other effects from the real building versus those assumed by the program user
- User error in deriving building input files
- Differences between the actual physical properties of the building (including HVAC systems), versus those input by the user.

Internal Error Types

- Differences between the actual thermal transfer mechanisms in the real building and its HVAC systems versus the simplified model of those processes in the simulation (all models, no matter how detailed, are simplifications of reality)
- Errors or inaccuracies in the mathematical solution of the models
- Coding errors.

The simplest level of empirical validation compares a building's actual long-term energy use to that calculated by a computer program, with no attempt to eliminate sources of discrepancy. Because this is similar to how a simulation tool is used in practice, it is favored by many in the building industry. However, it is difficult to interpret the results because all possible error sources are acting simultaneously. Even if there is

good agreement between measured and calculated performance, possible offsetting errors prevent a definitive conclusion about the model's accuracy. More informative levels of validation involve controlling or eliminating various combinations of error types and increasing the density of output-to-data comparisons (e.g., comparing temperature and energy results at various time scales ranging from subhourly to annual). At the most detailed level, all known sources of error are controlled to identify and quantify unknown error sources and to reveal causal relationships associated with the error sources.

This principle also applies to intermodel comparative testing and analytical verification. The more realistic the test case, the more difficult it is to establish causality and diagnose problems; the simpler and more controlled the test case, the easier it is to pinpoint sources of error or inaccuracy. Methodically building up to realistic cases is useful for testing interactions between algorithms modeling linked mechanisms.

A comparison between measured and calculated performance represents a small region in an immense N-dimensional parameter space. Investigators are constrained to exploring relatively few regions in this space, yet would like to be assured that the results are not coincidental (e.g., not a result of offsetting errors) and do represent the validity of the simulation elsewhere in the parameter space. Analytical and comparative techniques minimize the uncertainty of the extrapolations around the limited number of sampled empirical domains. Table 2 classifies these extrapolations. Use of the term "vice versa" in Table 2 is intended to mean that the extrapolation can go both ways, (e.g., from short-term to long-term data and from long-term to short-term data). This does not mean that such extrapolations are correct, but only that researchers and practitioners have either explicitly or implicitly made such inferences in the past.

Table 2. Types of Extrapolation

Obtainable Data Points	Extrapolation
A few climates	Many climates
Short-term total energy usage	Long-term total energy usage, or vice versa
Short-term (hourly) temperatures and/or fluxes	Long-term total energy usage, or vice versa
A few equipment performance points	Many equipment performance points
A few buildings representing a few sets of variable and parameter combinations	Many buildings representing many sets of variable and parameter combinations, or vice versa
Small-scale: simple test cells, buildings, and mechanical systems; and laboratory experiments	Large-scale complex buildings with complex HVAC systems, or vice versa

Figure 1 shows one process to combine the analytical, empirical, and comparative techniques. These three techniques may also be used together in other ways, for example, intermodel comparisons may be done before an empirical validation exercise, to better define the experiment and to help estimate experimental uncertainty by propagating all known error sources through one or more whole-building energy simulation programs (Hunn et al. 1982, Lomas et al. 1994).

For the path shown in Figure 1, the first step is running the code against analytical verification test cases to check its mathematical solution. Discrepancies must be corrected before proceeding further.

Second, the code is run against high-quality empirical validation data, and errors are corrected. Diagnosing error sources can be quite difficult and is an area of research in itself. Comparative techniques can be used to create diagnostic procedures (Achermann and Zweifel 2003; Judkoff 1988; Judkoff and Neymark 1995a, 1995b; Judkoff, Wortman, and Burch 1983; Judkoff et al. 1980; Morck 1986; Neymark and Judkoff 2002, 2004; Purdy and Beausoleil-Morrison 2003; Spitler, Rees, and Dongyi 2001; Yuill and Haberl 2002) and to better define empirical experiments.

The third step is to check agreement of several different programs with different thermal solution and modeling approaches (that have passed through steps 1 and 2) in a variety of representative cases. This uses the comparative technique as an extrapolation tool. Deviations in the program predictions indicate areas for further investigation.
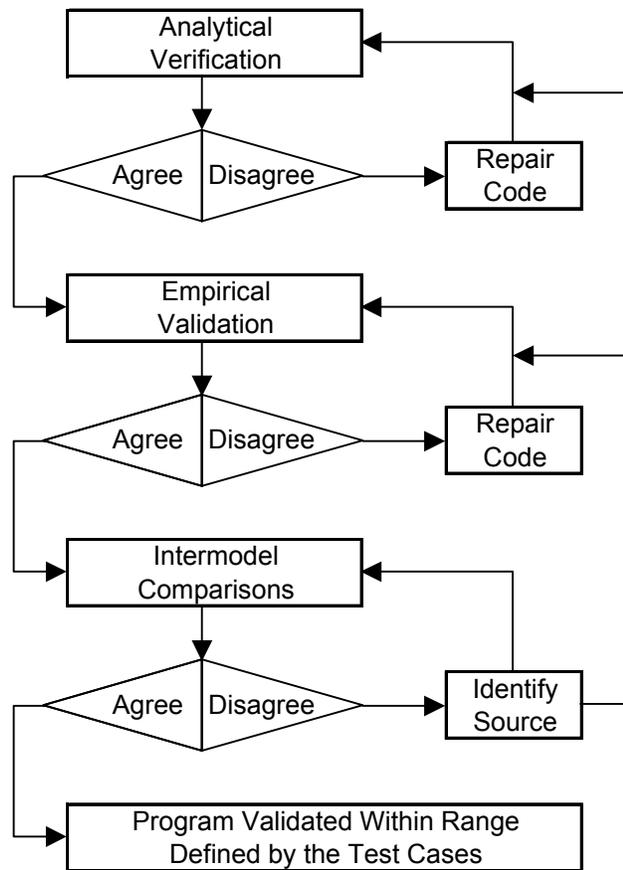
Figure 1. Validation method

When programs successfully complete these three stages, they are considered validated for cases where acceptable agreement was achieved (i.e., for the range of building, climate, and mechanical system types represented by the test cases). Once several detailed simulation programs have satisfactorily completed the procedure, other programs and simplified design tools can be tested against them. A validated code does not necessarily represent truth. It does represent a set of algorithms that have been shown, through a repeatable procedure, to perform according to the current state of the art.

The NREL methodology for validating building energy simulation programs has been generally accepted by the International Energy Agency (Irving 1988), ASHRAE (ANSI/ASHRAE 2001, 2004; ANSI/ASHRAE/IESNA 2004, 2005), and elsewhere with refinements suggested by other researchers (Bland 1992; Bloomfield 1988, 1999; Guyon and Palomo 1999; Irving 1988; Lomas 1991; Lomas and Bowman 1987; Lomas and Eppel 1992). Additionally, the Commission of European Communities has conducted considerable work under the PASSYS program (Jensen 1989; Jensen and van de Perre 1991).

EXAMPLES OF APPLYING THE NREL VALIDATION/TESTING METHODOLOGY

Herein is an example applying a portion of the above methodology to show how comparative tests can be used to extrapolate beyond analytical verification test cases. The process is illustrated for testing the ability of simulation programs to model a simple split-system air conditioner using the steady-state analytical verification tests of HVAC BESTEST Volume 1 (Neymark and Judkoff 2002a) as the basis for further, more realistic, comparative tests of HVAC BESTEST Volume 2 (Neymark and Judkoff 2004). Initially, Volume 1 analytical verification test cases were developed and distributed to IEA field-trial participants, *before* analytical solutions were developed. The initial results are illustrated in Figure 2; these are "blind" results, meaning that the field-trial participants had no prior knowledge of each others' results. Development of steady-state analytical solutions after this initial round of simulations yielded a mathematical truth standard for

the field-trial participants to compare to their results. Using this mathematical truth standard the field-trial participants discovered and corrected a number of errors in their simulation programs. Such improvements to simulation programs or simulation inputs made by participants must have a mathematical and physical basis, and must be applied consistently across tests. Also, all improvements were required to be documented in modeler reports. Arbitrary modification of a simulation program's input or internal code just for the purpose of more closely matching a given set of results is not allowed. The participants' corrections resulted in improved agreement among the simulation programs, as shown in Figure 3.
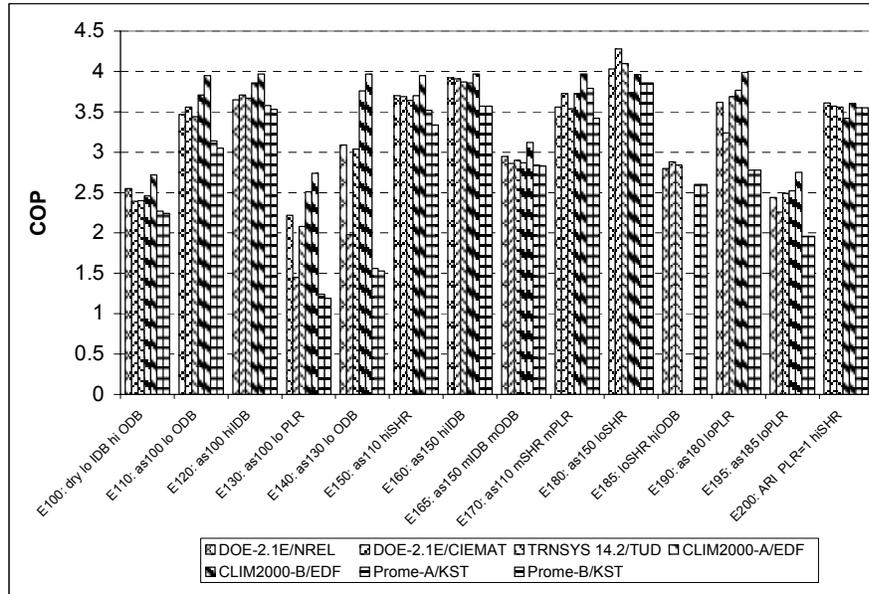

Figure 2. HVAC BESTEST Volume 1: mean COP, initial simulation results, before analytical solutions developed (abbreviations along the x-axis are shorthand for the case descriptions; see Neymark and Judkoff [2002a] for full case descriptions)
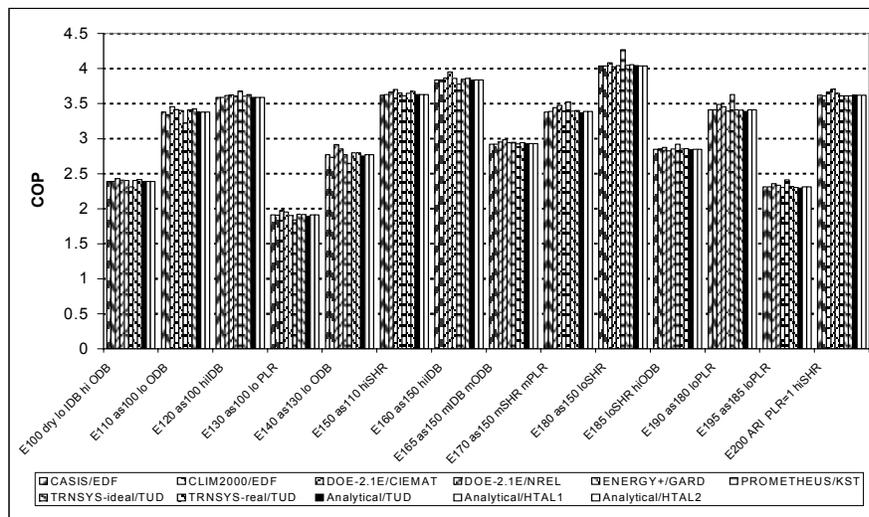

Figure 3. HVAC BESTEST Volume 1: mean COP, final analytical solutions and simulation results (abbreviations along the x-axis are shorthand for the case descriptions; see Neymark and Judkoff [2002a] for full case descriptions)

The HVAC BESTEST Volume 2 comparative test cases, also applying a split-system air conditioner, were then developed. These cases include more realism (e.g., hourly varying internal gains and weather data, outside air mixing, etc.) such that obtaining analytical solutions outside the environment of a whole-building energy simulation program was not possible. Because these cases are an extension of the HVAC BESTEST Volume 1 analytical verification test cases, it was possible to use results from the programs that were run and improved in the Volume 1 test cases, and then further tested with Volume 2, as approximate benchmark results for comparing with programs that did not previously run the Volume 1 test cases. Figure 4 illustrates "blind" initial results for the Volume 2 comparative tests, while Figure 5 illustrates final Volume 2 comparative test results after technically justifiable corrections to the simulation programs were applied according to the rules previously stated. These figures show that – because good agreement among simulations was achieved in the Volume 1 tests, and because the Volume 2 tests are an extension of the the Volume 1 tests – there was some propropagation of the level of agreement among simulations of the Volume 1 analytical verification tests through to the Volume 2 comparative test simulation results. There was also some further mprovement to the simulations achieved because of the Volume 2 comparative tests.
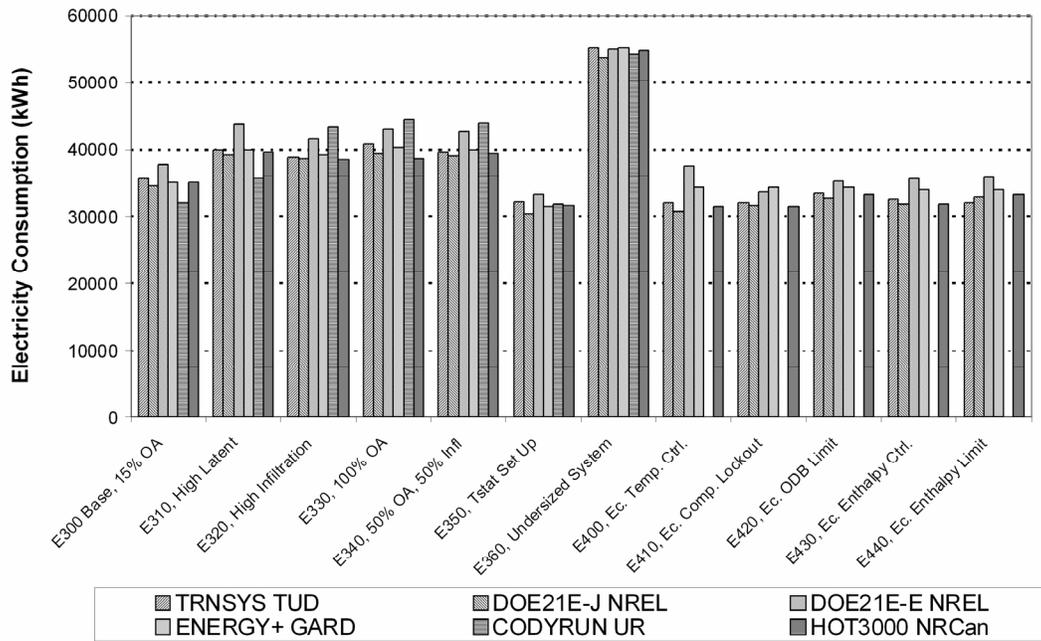

Figure 4. HVAC BESTEST Volume 2: total electricity consumption, initial simulation results (abbreviations along the x-axis are shorthand for the case descriptions; see Neymark and Judkoff [2004] for full case descriptions)
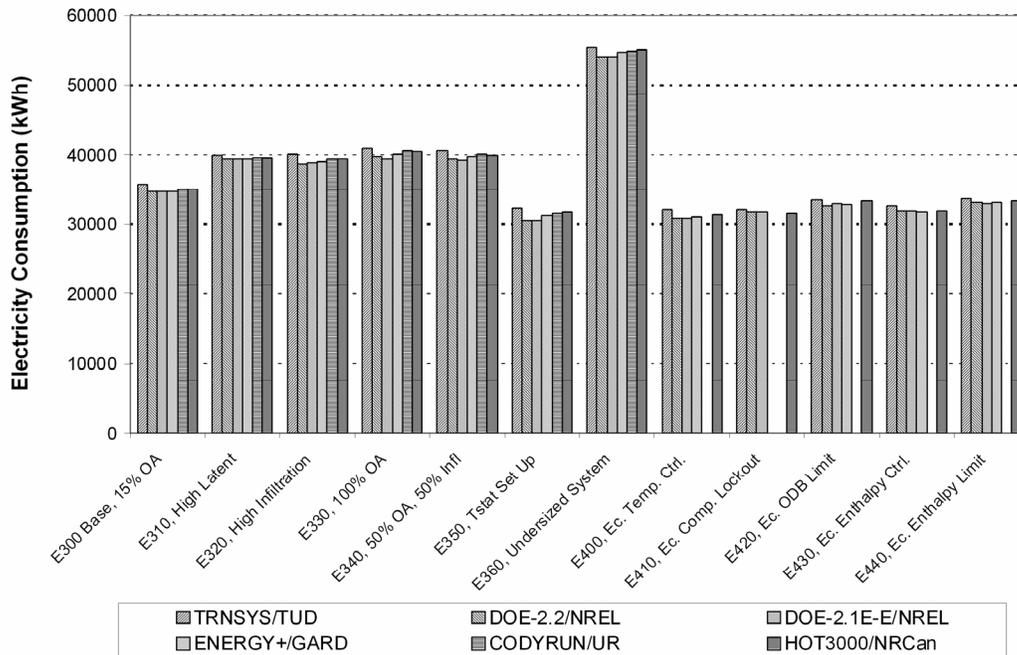
Figure 5. HVAC BESTEST Volume2: total electricity consumption, final simulation results (abbreviations along the x-axis are shorthand for the case descriptions; see Neymark and Judkoff [2004] for full case descriptions)

SUMMARY OF PREVIOUS AND CURRENT TESTING AND VALIDATION WORK

A summary of work in the areas of analytical verification, empirical validation, and comparative testing is available that cites and briefly summarizes approximately 100 articles and research reports covering work published from 1980 through mid-2001 (ASHRAE 2005; Neymark and Judkoff 2002). The following test suites have been completed and are in various stages of adaptation for Standard 140:

a) "Building Energy Simulation Test and Diagnostic Method for Heating, Ventilating, and Air-Conditioning Equipment Models (HVAC BESTEST), Fuel-Fired Furnace Test Suite" (Purdy and Beausoleil-Morrison 2003)

b) "Home Energy Rating System Building Energy Simulation Test (Judkoff and Neymark 1995b)

c) "International Energy Agency Building Energy Simulation Test and Diagnostic Method for Heating, Ventilating, and Air-Conditioning Equipment Models (HVAC BESTEST), Volume 2: Cases E300–E545" (Neymark and Judkoff 2004)

There are also a number of simulation test suites in various stages of completion that could eventually be included in Standard 140. These include, among others:

d) ASHRAE RP-1052, "Development of an Analytical Verification Test Suite for Whole Building Energy Simulation Programs – Building Fabric" (Spitler, Rees, and Xiao, 2001)

e) ASHRAE RP-865, "Development of Accuracy Tests for Mechanical System Simulation" (Yuill and Haberl 2002)
f) "RADTEST Radiant Heating and Cooling Test Cases" (Achermann and Zweifel 2003)
g) "Proposed IEA BESTEST Ground-Coupled Cases" (Deru, Judkoff and Neymark 2002)
h) ETNA BESTEST Empirical Validation Test Specification (Neymark et al. 2004)
i) "Daylighting – HVAC Interaction Tests for the Empirical Validation of Building Energy Analysis Tools" (Maxwell, Loutzenhiser, and Klaassen 2003)
j) "Economizer Control Tests for the Empirical Validation of Building Energy Analysis Tools" (Maxwell, Loutzenhiser, and Klaassen 2004)
k) A number of test suites that are being developed by National Renewable Energy Laboratory and researchers in International Energy Agency (IEA) member nations under auspices of IEA Solar Heating and Cooling Task 34 and IEA Energy Conservation in Buildings and Community Systems Annex 43 (IEA SHC 34/ECBCS 43; Judkoff and Neymark 2004).

CONCLUSION AND RECOMMENDATIONS

New material has been added to *ASHRAE Handbook—Fundamentals* (ASHRAE 2005) that describes a method to evaluate the accuracy of whole-building energy simulation programs, and to identify and diagnose differences in simulation predictions that may be caused by algorithmic differences, modeling limitations, coding errors, or input errors. This methodology, which was developed by the National Renewable Energy Laboratory starting over 20 years ago (Judkoff et. al., 1983), has been adopted by ANSI/ASHRAE Standard 140 (ANSI/ASHRAE 2001, 2004), *Method of Test for the Evaluation of Building Energy Analysis Computer Programs*. The current set of test cases included in Standard 140 occupy only a small area of the parameter space represented by the cells in the methodological test matrix described in this paper. Additional suites of test cases have been developed that can be considered for integration into Standard 140. We recommend more work to integrate such existing test suites into Standard 140.

Some high-priority areas for additional test cases are:

- Comparative Tests—Mechanical Equipment (Additional tests beyond those in HVAC BESTEST unitary cooling and heating equipment cases)
- Analytical Verification—Mechanical Equipment (Additional tests beyond those in RP-865 and HVAC BESTEST unitary cooling and heating equipment test cases)
- Empirical Validation—Mechanical Equipment (Additional tests beyond those in IEA SHC Task 22 and IEA SHC Task 34/ECBCS Annex 43).

More work to develop such test suites is recommended.

Standard 140 and/or the reports that comprise the test suites contained therein (commonly known as the NREL BESTEST reports), are being referenced and used by a growing number of code promulagation authorities throughout the world. ASHRAE Standard 90.1 (ANSI/ASHRAE/IESNA 2004, 2005) requires that software used for demonstrating performance compliance with Standard 90.1 be tested using ASHRAE Standard 140. Standard 90.1 is ASHRAE's consensus energy code for commercial buildings and for non-low-rise residential buildings. IEA BESTEST is also being used for simulation certification tests in The Netherlands (ISSO 2003), Australia (SEDA 2003, Pears 1998), New Zealand (Donn 2004), and Portugal (Maldonado 2005). As part of their building energy performance assessments under the European Community's Energy Performance Directive (European Union 2002), Austria, Denmark, Greece and The Netherlands are using a new software tool that includes algorithms that have been checked with BESTEST (Balaras, Poel, and van Crutchen, 2005). Also, CEN has utilized BESTEST to check their reference cooling load calculation general criteria of prEN ISO 13791 (CEN 2004a) and simplified methods of prEN ISO 13792 (CEN2004b). (Millet 2003) In the United States, the National Association of State Energy Officials (NASEO) Residential Energy Services Network (RESNET) has adopted Home Energy Rating System (HERS) BESTEST (Judkoff and Neymark 1995b) as the basis for certifying software to be used for home energy rating systems under the NASEO/RESNET national accreditation standard. (NASEO/RESNET 2002) This is indicative of the importance of validation methods for improving the state of the art in building energy software, and for helping to certify such software for use with home energy rating standards, building energy codes, building energy tax credits, and other building energy incentive programs. The importance of validation and test methods is further supported by a recent report comparing 20 whole-building energy

simulation tools. (Crawley *et al* 2005) The report indicates that 18 of the 20 tools reviewed have been tested with at least one of the two procedures currently included in ASHRAE Standard 140-2004, 7 of the tools have been tested with both procedures, and 9 of the tools have been tested with three additional procedures that are currently in various stages of adaptation for Standard 140.

Extensively reducing the energy intensity of buildings through better design is possible with the use of simulation tools. However, widespread use of building energy simulation software will not occur unless the design and engineering communities have confidence in these programs. The work described here represents a good start in the effort to develop rigorously validated building energy simulation tools.

ACKNOWLEDGEMENTS

REFERENCES

Achermann, M. and G. Zweifel. 2003. *RADTEST—Radiant Heating and Cooling Test Cases.* Horw-Lucerne, Switzerland: University of Applied Sciences Central Switzerland, Lucerne School of Architecture. http://www.iea-shc.org/task22/reports/RADTEST_final.pdf

ANSI/ASHRAE Standard 140-2001, *Standard Method of Test for the Evaluation of Building Energy Analysis Computer Programs*. (2001). Atlanta, GA: American Society of Heating, Refrigerating, and Air-Conditioning Engineers.

ANSI/ASHRAE Standard 140-2004. *Standard Method of Test for the Evaluation of Building Energy Analysis Computer Programs*. (2004). Atlanta, GA: American Society of Heating, Refrigerating, and Air-Conditioning Engineers.

ANSI/ASHRAE/IESNA Standard 90.1-2001, Addendum p, *Energy Standard for Buildings Except Low-Rise Residential Buildings.* (2004). Atlanta, GA: American Society of Heating, Refrigerating, and Air-Conditioning Engineers. References ANSI/ASHRAE Standard 140-2001.

ANSI/ASHRAE/IESNA Standard 90.1-2004, Addendum L, *Energy Standard for Buildings Except Low-Rise Residential Buildings.* (2005). Atlanta, GA: American Society of Heating, Refrigerating, and Air-Conditioning Engineers. References ANSI/ASHRAE Standard 140-2004. Addendum L has passed through the ASHRAE review process, and is expected to be published during mid-2006.

ASHRAE Handbook—Fundamentals. (2005). Atlanta, GA: American Society of Heating, Refrigerating, and Air-Conditioning Engineers. See Chapter 32, "Energy Estimating and Modeling Methods", section titled "Model Validation and Testing".

Balaras, C.A., B. Poel, and G. van Crutchen. (2005). *Software for Energy Performance Assessment of Existing Dwellings.* IBPSA News, Volume 15, Number 1. International Building Performance Simulation Association, April 2005.

Bland, B. (1992). "Conduction in Dynamic Thermal Models: Analytical Tests for Validation." *BSER&T* 13(4):197–208.

Bloomfield, D. (1988). *An Investigation into Analytical and Empirical Validation Techniques for Dynamic Thermal Models of Buildings*. Vol. 1, Executive Summary, 25 pp. SERC/BRE final report, Garston, Watford, UK: BRE.

Bloomfield, D. (1999). "An Overview of Validation Methods for Energy and Environmental Software." *ASHRAE Transactions* 105(2). Atlanta, GA: ASHRAE.

CEN. (2004a). prEN ISO 13791. *Thermal performance of buildings – Calculation of internal termperatures of a room insummer without mechanical cooling – General criteria and validation procedures.* Final Draft. Comité Européen de la Normalisation, Brussels.

CEN. (2004b). prEN ISO 13792. *Thermal performance of buildings – Calculation of internal termperatures of a room insummer without mechanical cooling – Simplified Methods.* Final Draft. Comité Européen de la Normalisation, Brussels.

Crawley, D., Hand, J., Kummert, M., Griffith, B. (2005). *Contrasting the Capabilities of Building Energy Performance Simulation Programs.* Washington, DC: U S Department of Energy ; Glasgow, Scotland, UK : University of Strathclyde ; Madison, WI : University of Wisconsin. http://gundog.lbl.gov/dirpubs/2005/05_compare.pdf

Deru, M., R. Judkoff, and J. Neymark. 2002. *Proposed IEA BESTEST Ground-Coupled Cases.* Golden, CO: National Renewable Energy Laboratory. Draft. July 2002.

Donn, M. (2004). May 2004 email communications with R. Judkoff and J. Neymark. Victoria, New Zealand: Victoria University.

European Union. (2002). *On the Energy Performance of Buildings.* Directive 2002/91/EC of the European Parliament and of the Council, Official Journal of the European Communites, December.

Guyon, G.; Palomo, E. (1999). "Validation of Two French Building Energy Analysis Programs Part 1: Analytical Verification." *ASHRAE Transactions* 105(2). Atlanta, GA: ASHRAE.

Hunn, B. et al. (1982). *Validation of Passive Solar Analysis/Design Tools Using Class A Performance Evaluation Data.* LA-UR-82-1732, Los Alamos, NM: LANL.

Irving, A. (January 1988). *Validation of Dynamic Thermal Models.* Energy, and Buildings, ISSN 0378-7788, Lausanne, Switzerland: Elsevier Sequoia, S.A.

ISSO. (2003). *Energie Diagnose Referentie Versie 3.0.* ISSO Publicatie 54. Rotterdam, The Netherlands: Institut voor Studie en Stimulering van Onderzoekop Het Gebied van Gebouwinstallaties (in Dutch).

Jensen, S., ed. (1989). The PASSYS Project Phase 1–Subgroup Model Validation and Development, Final Report –1986-1989. Commission of the European Communities, Directorate General XII.

Jensen, S.; van de Perre, R. (1991). "Tools for Whole Model Validation of Building Simulation Programs, Experience from the CEC Concerted Action PASSYS." *Proc. Building Simulation '91.* August 20–22, Nice, France. International Building Performance Simulation Association.

Judkoff, R. (1988). *Validation of Building Energy Analysis Simulation Programs at the Solar Energy Research Institute.* Energy and Buildings, Vol. 10, No. 3, p. 235. Lausanne, Switzerland: Elsevier Sequoia.

Judkoff, R., and J. Neymark. (1995a). *International Energy Agency Building Energy Simulation Test (BESTEST) and Diagnostic Method.* NREL/TP-472-6231. Golden, CO: National Renewable Energy Laboratory. http://www.nrel.gov/docs/legosti/old/6231.pdf (PDF 13.8 MB)

Judkoff, R., and J. Neymark. (1995b). *Home Energy Rating System Building Energy Simulation Test (HERS BESTEST).* NREL/TP-472-7332. Golden, CO: National Renewable Energy Laboratory. http://www.nrel.gov/docs/legosti/fy96/7332a.pdf (PDF 5.6 MB) http://www.nrel.gov/docs/legosti/fy96/7332b.pdf (PDF 1.9 MB)

Judkoff, R. (Operating Agent); Neymark, J. (2004). *IEA SHC Task 34 / ECBCS Annex 43, Testing and Validation of Building Energy Simulation Tools.* Annex Document. Paris, France: International Energy Agency: Solar Heating and Cooling Programme, and Energy Conservation in Buildings and Community Systems.

Judkoff, R.; Wortman, D.; Burch, J. (1983). *Measured versus Predicted Performance of the SERI Test House: A Validation Study.* SERI/TP-254-1953, Golden, CO: Solar Energy Research Institute (now National Renewable Energy Laboratory).

Judkoff, R.; Wortman, D.; Christensen, C.; O'Doherty, B.; Simms, D.; Hannifan, M. (1980). *A Comparative Study of Four Passive Building Energy Simulations: DOE-2.1, BLAST, SUNCAT-2.4, DEROB-III.* SERI/TP-721-837. UC-59c. Golden, CO: Solar Energy Research Institute, now National Renewable Energy Laboratory.

Judkoff, R., D. Wortman, B. O'Doherty, and J. Burch. (1983). *A Methodology for Validating Building Energy Analysis Simulations.* SERI/TR-254-1508. Golden, CO: Solar Energy Research Institute (now National Renewable Energy Laboratory).

Lomas, K. (1991). "Dynamic Thermal Simulation Models of Buildings: New Method of Empirical Validation." *BSER&T* 12(1):25–37.

Lomas, K.; Bowman, N. (1987). "Developing and Testing Tools for Empirical Validation," Ch. 14, Vol. IV of SERC/BRE final report, *An Investigation in Analytical and Empirical Validation Techniques for Dynamic Thermal Models of Buildings*. Garston, Watford, UK: BRE. 79 pp.

Lomas, K.; Eppel, H. (1992). "Sensitivity Analysis Techniques for Building Thermal Simulation Programs." *Energy and Building* (19)1:21–44.

Lomas, K.; Eppel, H.; Martin, C.; Bloomfield, D. (1994). *Empirical Validation of Thermal Building Simulation Programs Using Test Room Data*. Vol.1, Final Report. International Energy Agency Report #IEA21RN399/94. Vol. 2, Empirical Validation Package (1993), IEA21RR5/93. Vol. 3, Working Reports (1993), IEA21RN375/93. Leicester, UK: De Montfort University.

Maldonado, E. Porto, Portugal: University of Porto. (2005). *Energy Certification of Buildings in Portugal.* Presentation at Joint IEA ECBCS/SHC Executive Committee Meeting, Technical Day, June 15, 2005, Espinho, Portugal.

Maxwell, G.; Loutzenhiser, P.; Klaassen, C. (2003). *Daylighting – HVAC Interaction Tests for the Empirical Validation of Building Energy Analysis Tools.* Ames, Iowa: Iowa State University, Department of Mechanical Engineering.

Maxwell, G.; Loutzenhiser, P.; Klaassen, C. (2004). *Economizer Control Tests for the Emprical Validation of Building Energy Analysis Tools.* Ames, Iowa: Iowa State University, Department of Mechanical Engineering. Draft, June 2004.

Millet, J.-R. Paris, France: Centre Scientifique et Technique du Batiment. (2003). Personal Communications at IEA SHC Task Definition Workshop, April 24-25, 2003, Delft, Netherlands.

Morck, O. (June 1986). *Simulation Model Validation Using Test Cell Data*. IEA SHC Task VIII, Report #176, Thermal Insulation Laboratory, Lyngby, Denmark: Technical University of Denmark.

NASEO/RESNET. (2002). *Mortage Industry National Home Energy Rating Systems Accreditation Standards.* Oceanside, California, US: Residential Energy Services Network. June 15, 2002. Available from www.natresnet.com.

Neymark, J.; Girault, P.; Guyon, G.; Judkoff, R.; LeBerre, R.; Ojalvo, J.; Reimer, P. (2004). *ETNA BESTEST Empirical Validation Test Specification*. Golden, Colorado: J. Neymark & Associates; Moret sur Loing, France: Electricité de France. In collaboration with National Renewable Energy Laboratory, Golden, Colorado.

Neymark, J., and R. Judkoff. (2002a). *International Energy Agency Building Energy Simulation Test and Diagnostic Method for Heating, Ventilating, and Air-Conditioning Equipment Models (HVAC BESTEST) Volume 1: Cases E100-E200.* NREL/TP-550-30152. Golden, CO: National Renewable Energy Laboratory. http://www.nrel.gov/docs/fy02osti/30152.pdf (PDF 4.5 MB)

Neymark, J.; Judkoff, R. (2004). *International Energy Agency Building Energy Simulation Test and Diagnostic Method for Heating, Ventilating, and Air-Conditioning Equipment Models (HVAC BESTEST) Volume 2: Cases E300–E545.* NREL/TP-550-36754. Golden, CO: National Renewable Energy Laboratory. http://www.nrel.gov/docs/fy05osti/36754.pdf (PDF 15.4 MB)

Pears, A. (1998). *Rating Energy Efficiency of Non-Residential Buildings: A Path Forward for New South Wales.* Report for the Sustainable Energy Development Authority. Brighton, Victoria, Australia: Sustainable Solutions Pty Ltd. Available from www.abgr.com.au.

Purdy, J. and I. Beausoleil-Morrison. 2003. *Building Energy Simulation Test and Diagnostic Method for Heating, Ventilating, and Air-Conditioning Equipment Models (HVAC BESTEST): Fuel-Fired Furnace Test Cases.* Ottawa, Canada: Natural Resources Canada, CANMET Energy Technology Centre. http://www.iea-shc.org/task22/deliverables.htm

SEDA. (2003). *Guidelines for the Use of Simulation in Commitment Agreements.* Grosvenor Place, New South Wales, Australia: Sustainable Energy Development Authority.

Spitler, S. ; Rees, S.; Xiao, D. (2001). *Development of An Analytical Verification Test Suite for Whole Building Energy Simulation Programs – Building Fabric.* Final Report for ASHRAE 1052-RP. Stillwater, OK: Oklahoma State University School of Mechanical and Aerospace Engineering.

Yuill, G. and J. Haberl. (2002). *Development of Accuracy Tests for Mechanical System Simulation*. Final Report for ASHRAE 865-RP. Omaha, NE: University of Nebraska.

REPORT DOCUMENTATION PAGE

1. REPORT DATE *(DD-MM-YYYY)* July 2006	2. REPORT TYPE Conference Paper	3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE Model Validation and Testing: The Methodological Foundation of ASHRAE Standard 140; Preprint	5a. CONTRACT NUMBER DE-AC36-99-GO10337
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) R. Judkoff and J. Neymark	5d. PROJECT NUMBER NREL/CP-550-40360
	5e. TASK NUMBER BEC66051
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) National Renewable Energy Laboratory 1617 Cole Blvd. Golden, CO 80401-3393	8. PERFORMING ORGANIZATION REPORT NUMBER NREL/CP-550-40360

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) NREL
	11. SPONSORING/MONITORING AGENCY REPORT NUMBER

12. DISTRIBUTION AVAILABILITY STATEMENT
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161

13. SUPPLEMENTARY NOTES

14. ABSTRACT *(Maximum 200 Words)*
Ideally, whole-building energy simulation programs model all aspects of a building that influence energy use and thermal and visual comfort for the occupants. An essential component of the development of such computer simulation models is a rigorous program of validation and testing. This paper describes a methodology to evaluate the accuracy of whole-building energy simulation programs. The methodology is also used to identify and diagnose differences in simulation predictions that may be caused by algorithmic differences, modeling limitations, coding errors, or input errors. The methodology has been adopted by ANSI/ASHRAE Standard 140 (ANSI/ASHRAE 2001, 2004), Method of Test for the Evaluation of Building Energy Analysis Computer Programs. A summary of the method is included in the ASHRAE Handbook of Fundamentals (ASHRAE 2005). This paper describes the ANSI/ASHRAE Standard 140 method of test and its methodological basis. Also discussed are possible future enhancements to Standard 140 and related research recommendations.

15. SUBJECT TERMS
whole-building energy simulation; validation; testing; ashrae standard 140

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified	UL		19b. TELEPHONE NUMBER *(Include area code)*