Automated Comparison of Building Energy Simulation Engines

Ben Polly
Scott Horowitz
Chuck Booten
Neal Kruis
Craig Christensen

August, 2012

NREL/PR-5500-52651

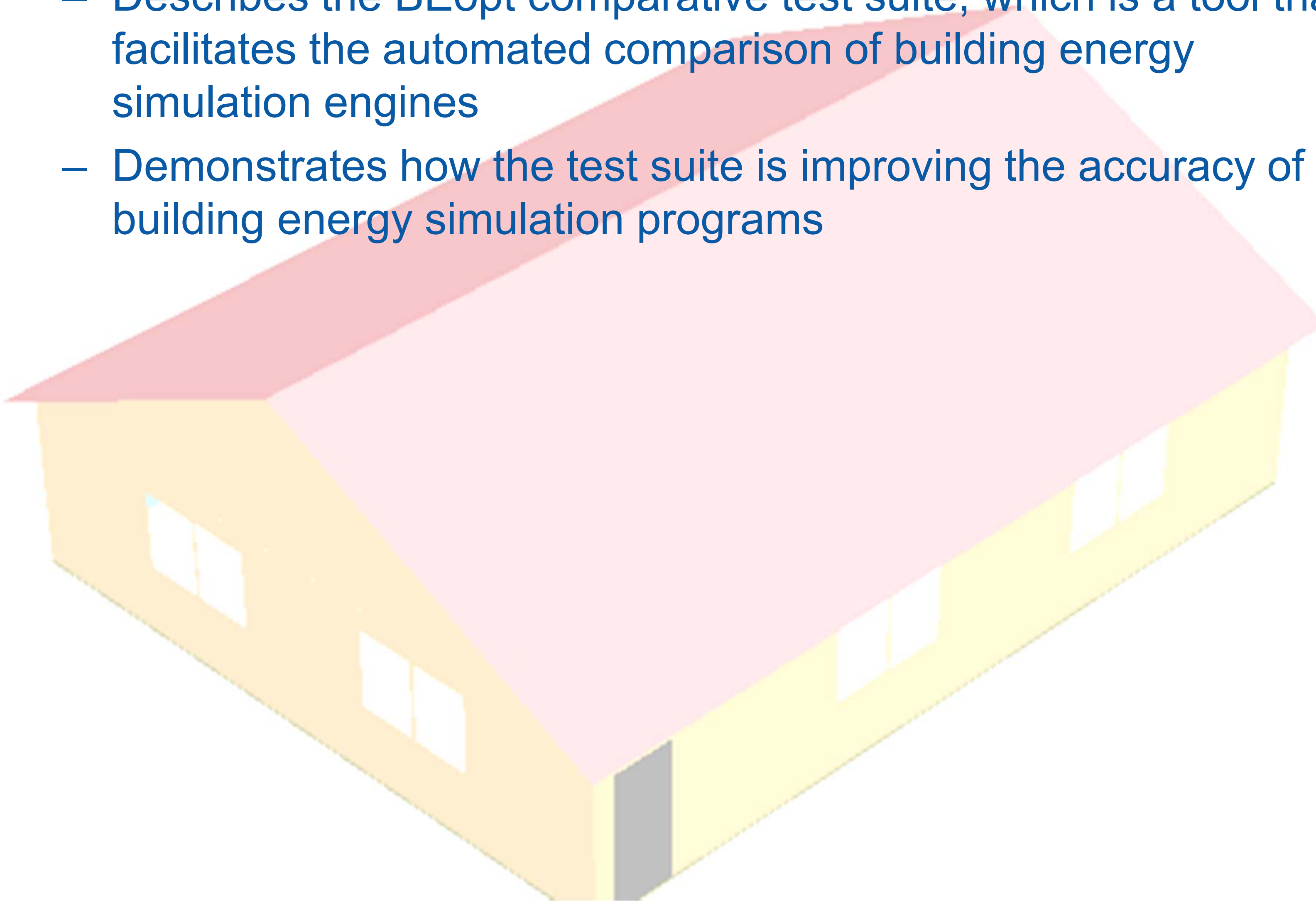
Table of Contents

1. Introduction

1. Introduction

This briefing package:

- Describes the BEopt comparative test suite, which is a tool that facilitates the automated comparison of building energy simulation engines
- Demonstrates how the test suite is improving the accuracy of building energy simulation programs



1. Introduction

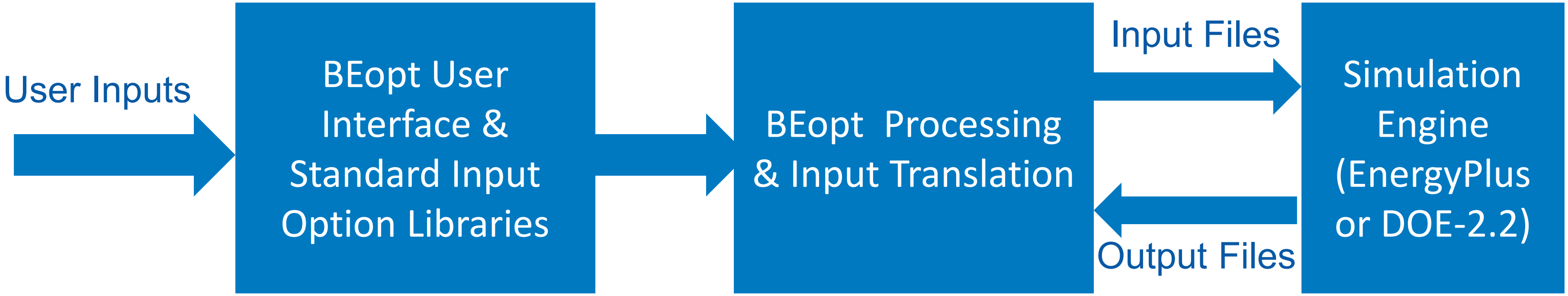
Motivation:

- Building energy simulation programs inform energy efficient design for new homes and energy efficient upgrades for existing homes
- Stakeholders rely on accurate predictions from simulation programs
- Previous research indicates that software tends to over-predict energy usage for poorly-insulated leaky homes
- NREL is identifying, investigating, and resolving software inaccuracy issues
- Comparative software testing is one method of many that NREL uses to identify potential software issues (more information on comparative testing and the need for the BEopt test suite can be found in Appendix A)

2. BEopt Comparative Test Suite

2. BEopt Comparative Test Suite: BEopt

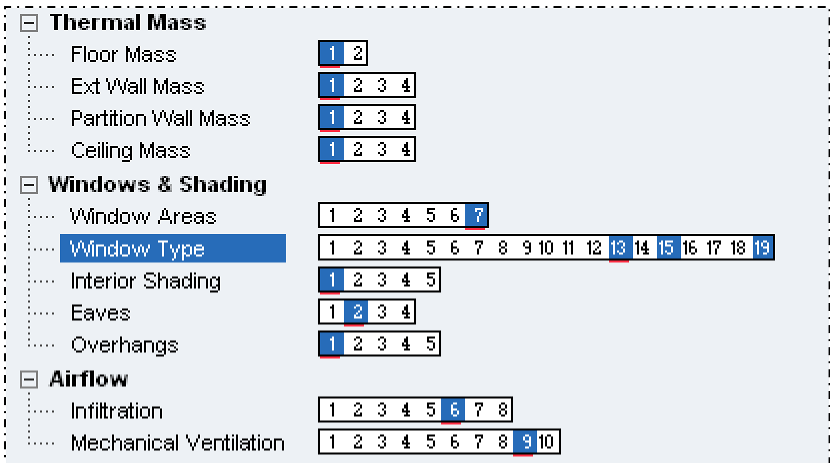
- BEopt is a building energy optimization software program developed by NREL
- BEopt is a “meta-program” that can interface with multiple simulation engines
- BEopt currently interfaces with DOE-2.2 and EnergyPlus



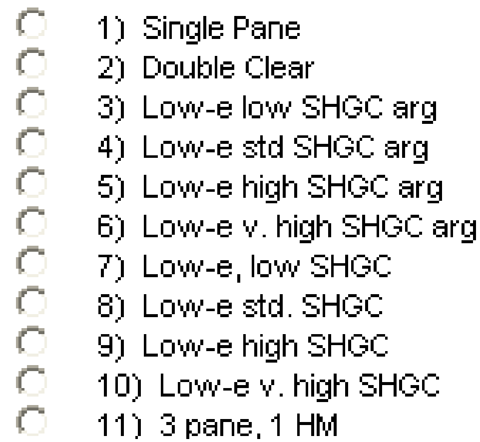
2. BEopt Comparative Test Suite: Approach

BEopt has many categories and options within each category:

Example of Categories



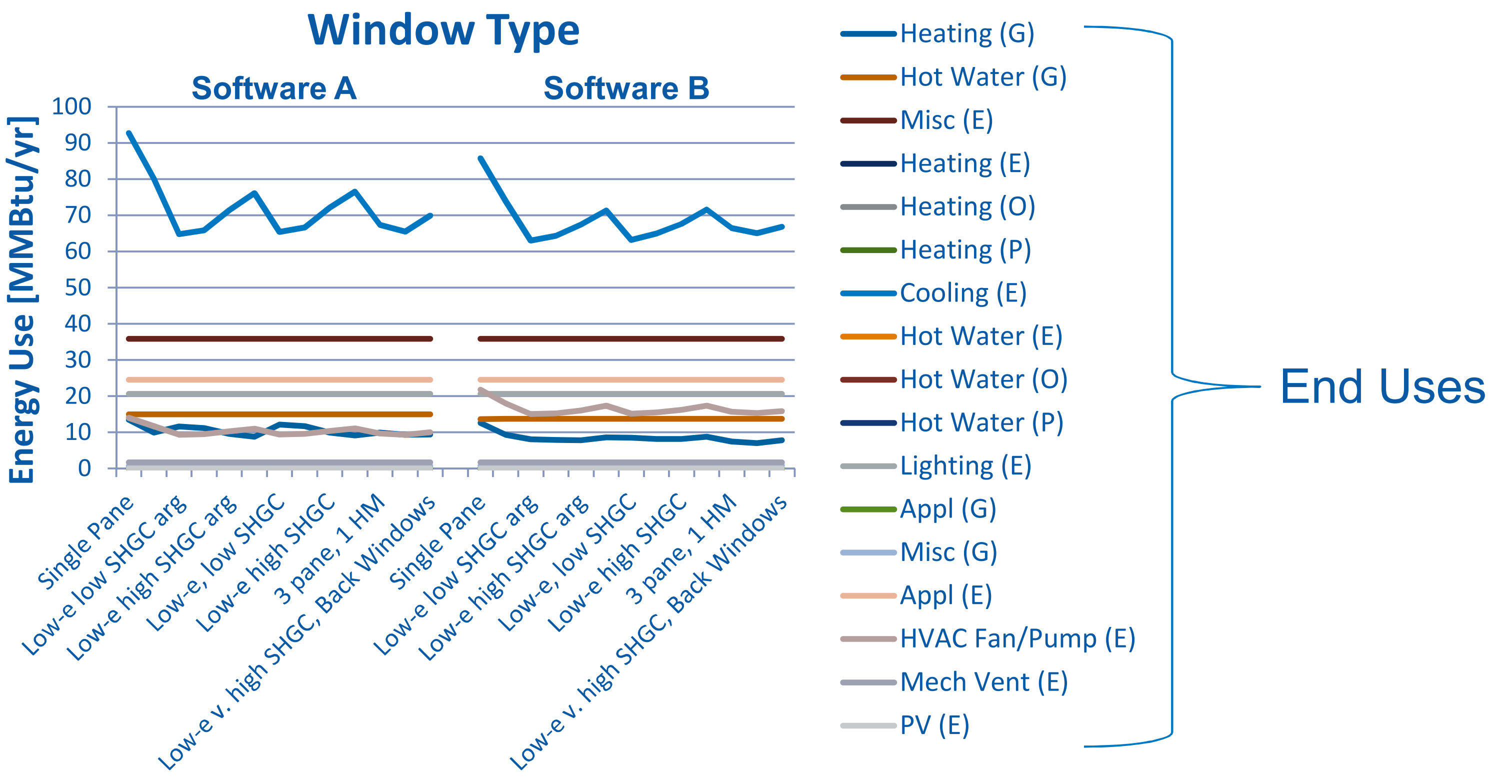
Example of Options for “Window Type” Category



BEopt Test Suite Approach: Automatically and systematically sweep through the different categories and the options within categories. The result is a sensitivity study over a large parameter space.

2. BEopt Comparative Test Suite: Approach

- The parameter sweeps can be performed with different simulation engines
- The results are generated in standard output formats and can be compared visually and numerically to identify discrepancies



2. BEopt Comparative Test Suite: Scenarios

In the standard test suite run, parameter sweeps are performed for the following six different scenarios, which cover a range of climates and building geometries:

Scenario	Location	# of Stories	Square Footage	Foundation Type	Attic Type	Garage?
1	Atlanta	2	1,800	Vented Crawlspace	Unfinished Attic	No
2	Atlanta	2	2,700	Unvented Crawlspace	Finished Attic	No
3	Phoenix	2	1,800	Slab	Unfinished Attic	Yes
4	Chicago	1	800	Unfinished Basement	Unfinished Attic	Yes
5	Chicago	1	3,000	Finished Basement	Unfinished Attic	Yes
6	Houston	2	2,400	Slab	Unfinished Attic	Yes

2. BEopt Comparative Test Suite: Buildings

The test suite consists of three building types:

Diagnostic Building: The diagnostic building is used for pinpointing the source of a discrepancy between two building models (primarily between the same building in DOE-2 and EnergyPlus). The idea behind this building is to zero-out the effects from other categories while running a parametric through the options within the category of interest.

New Construction Building: The new construction test building has a complete set of options typically found in new homes. When a parametric is run for a category, typical options for new home construction are used for all other categories.

Existing Building: The existing test building has a complete set of options typically found in existing homes. When a parametric is run for a category, typical options for existing homes are used for all other categories.

2. BEopt Comparative Test Suite: Buildings

Diagnostic Building Details:

- Super-insulated envelope
- Zero infiltration/ventilation
- Ideal systems for testing the envelope
- Sinusoidal (positive and negative) internal gains for testing HVAC equipment

Goal: Isolate the tested component

Important Note: In some cases dependencies exist where testing one element (e.g., floor mass) is dependent on the presence of another component (e.g., windows). The test suite accommodates these situations.

2. BEopt Comparative Test Suite: Buildings

New Construction and Existing Building Details

Goal: Test the software using detailed, realistic buildings.

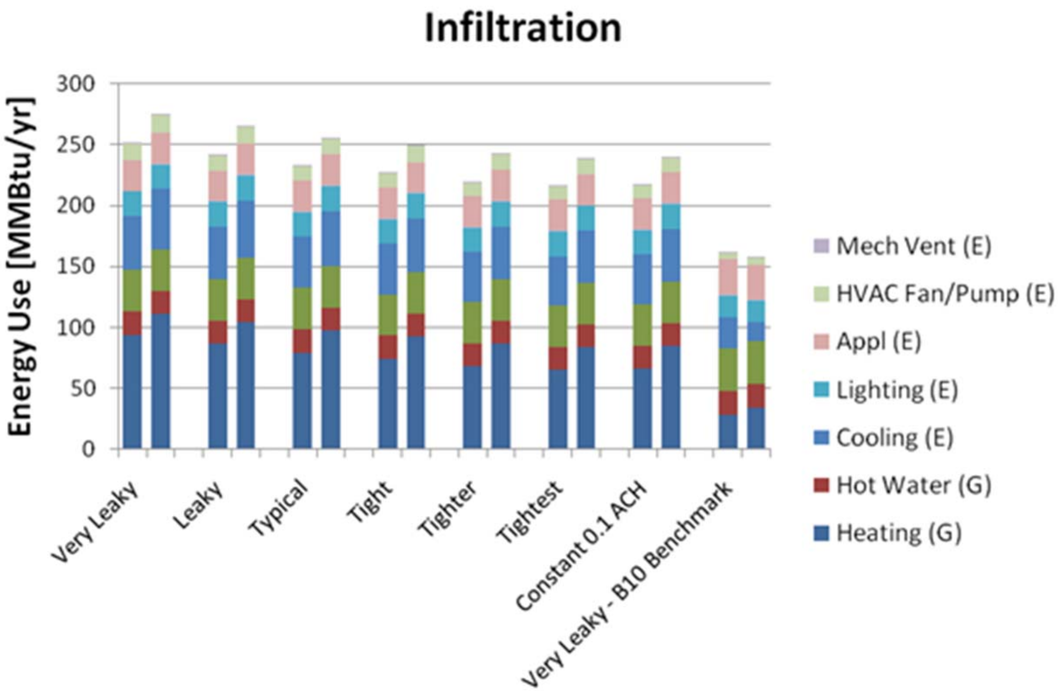
Important Note: The new construction and existing buildings are investigated separately because models and algorithms become more or less influential depending on the general efficiency of the building (e.g., window surface convection models are more influential for single-pane windows in existing homes than double-pane (Low-e) windows in new homes).

2. BEopt Comparative Test Suite: Buildings

Comparison of New Construction and Existing Building details:

	Existing	New Construction
Walls	Uninsulated	R-13
Attic	R-11	R-30
Basement Walls	Uninsulated	R-5
Crawlspace Ceiling	Uninsulated	R-13
Slab	Uninsulated	Uninsulated
Window Type	Single	Double Low-e
Window Area	15% of Wall Area	15% of Wall Area
Infiltration	SLA=0.00090	SLA=0.00050
Appliances/Water Heater	Old	Standard
Air Conditioner	SEER 10	SEER 13
Furnace	AFUE 78%	AFUE 78%
Ducts	Uninsulated, 30% leakage	R-6, 15% leakage

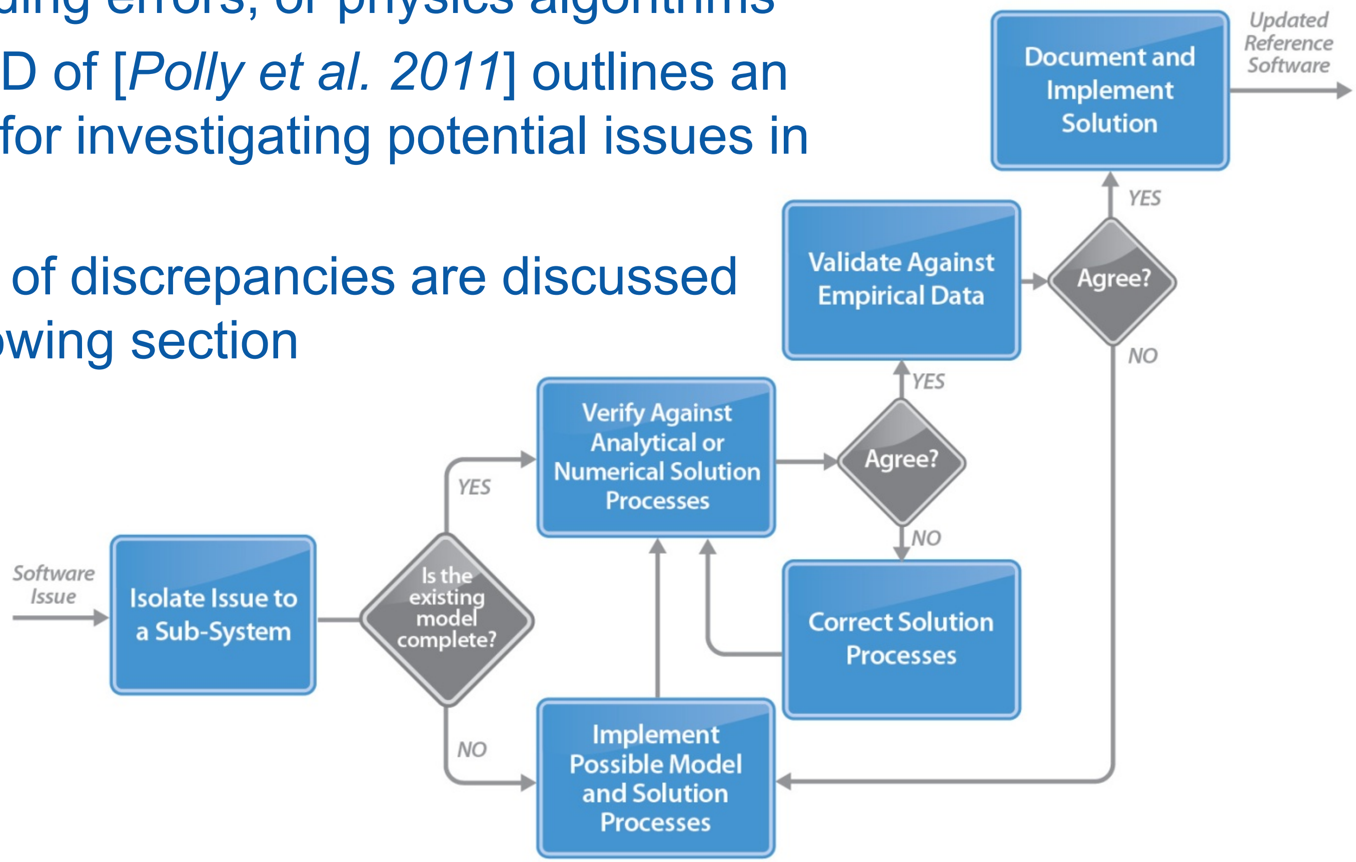
2. BEopt Comparative Test Suite: Output



- Energy consumption by end use can be compared for the many different test suite simulations
- It is important to examine magnitudes of differences (absolute and percentage) as well as trends within a building component (deltas between each option)
- Non-energy output can also be compared: equipment sizes, loads not met (hours/yr), etc.

2. BEopt Comparative Test Suite: Output

- When discrepancies are found, it takes time and effort to identify whether differences are due to non-equivalent simulation engine inputs, coding errors, or physics algorithms
- Appendix D of [*Polly et al. 2011*] outlines an approach for investigating potential issues in software[1]
- Examples of discrepancies are discussed in the following section



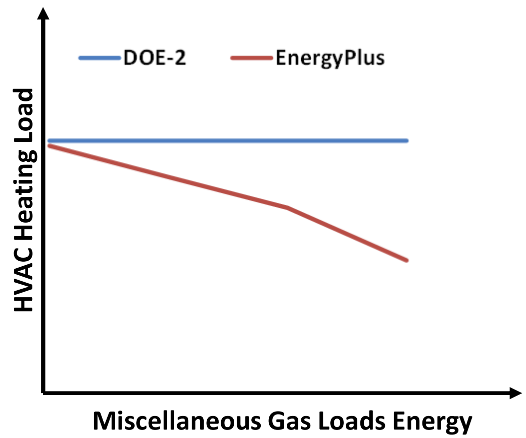
1. The approach is an application of the BESTEST methodology [*Judkoff et al. 2008*] to individual modeling issues in residential building energy analysis.

3. Example Test Suite Results

3. Example Test Suite Results

Example 1: Differences due to non-equivalent inputs

In this example, the test suite uncovered differences in HVAC heating load as a function of miscellaneous gas loads energy. The HVAC heating load should decrease with increasing misc. gas loads energy because gas loads (e.g., gas stoves) help meet total heating loads.



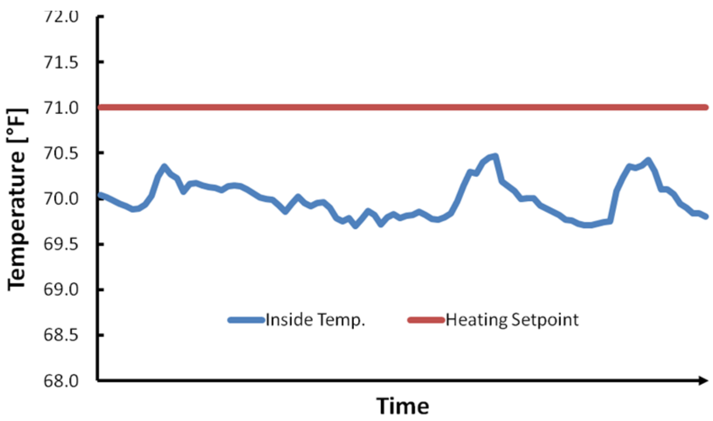
The issue was investigated and it was determined that a unit conversion problem in BEopt caused incorrect DOE-2 inputs for heat gain from miscellaneous gas loads.

3. Example Test Suite Results

Example 2: Differences due to coding error in simulation engine

The test suite compares the number of hours where the HVAC system is unable to meet the load (e.g., hours when the temperature in the house is below the heating setpoint).

An EnergyPlus test with a conditioned basement showed over 7,000 hours of unmet loads (80% of the year) versus typical unmet hours of less than 1%.

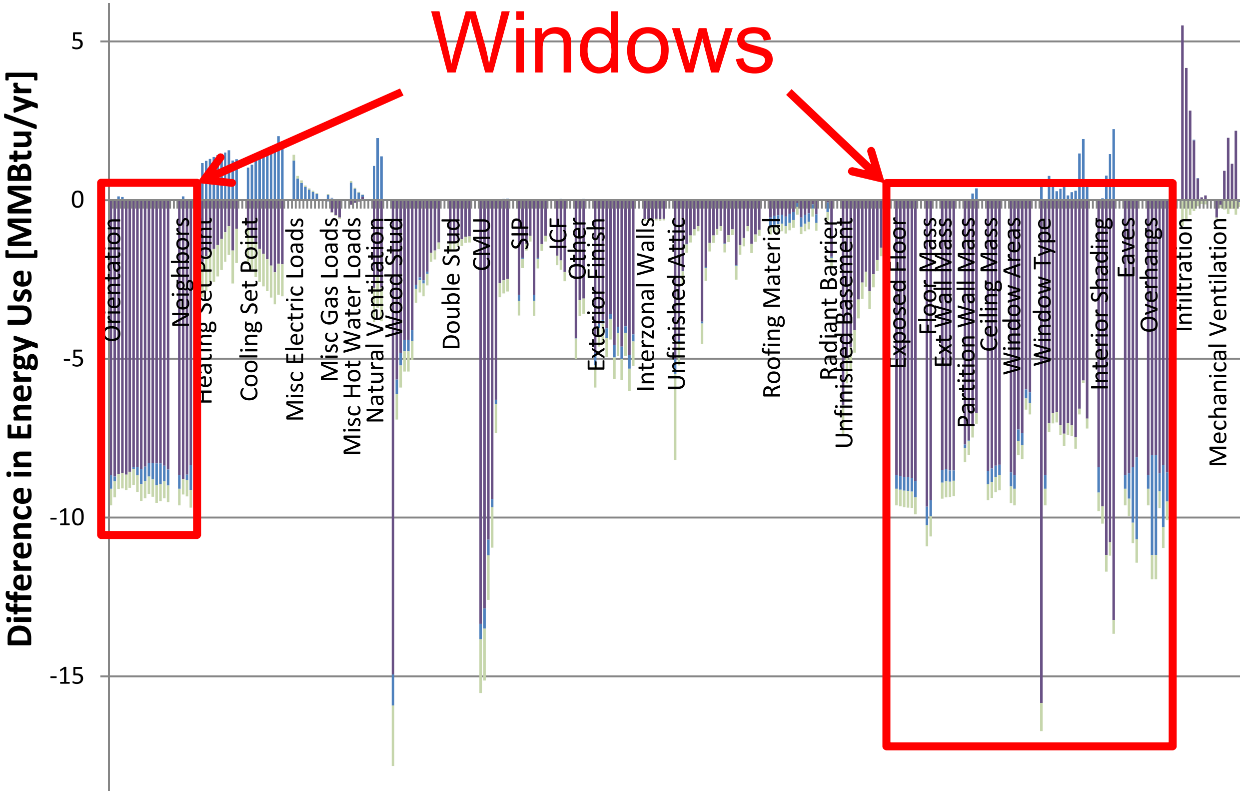


This led the EnergyPlus development team to fix a coding error (for version 5.0) related to the amounts of air delivered to individual zones served by a single system.

3. Example Test Suite Results

Example 3: Differences due to coding errors and physics algorithms in simulation engines

The test suite uncovered substantial differences between EnergyPlus (V6.0.0.023) and DOE-2.2 (V2.2-47h2) in the diagnostic building for windows-related categories:



3. Example Test Suite Results

Example 3: Differences due to coding errors and physics algorithms in simulation engines

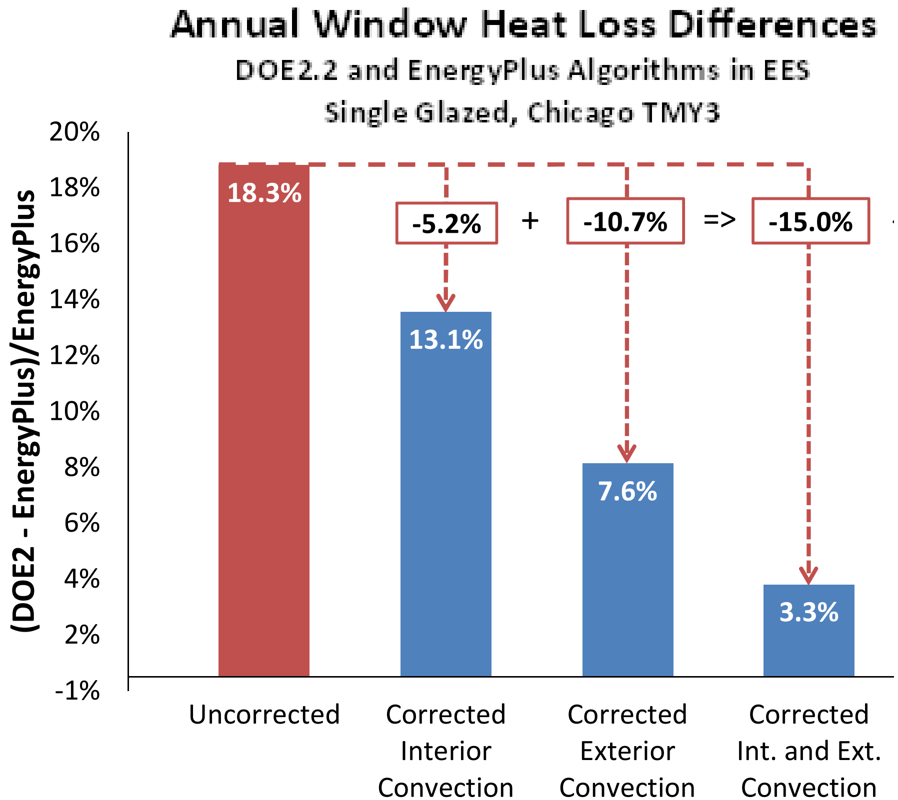
Rigorous analysis was performed to investigate and understand the differences for windows-related categories [*Kruis et al.* 2012]. Causes of differences include:

- Exterior forced convection:
  - EnergyPlus and DOE-2 exterior forced convection coefficients are calculated using regression coefficients that are inappropriate for use with near-surface wind speeds
  - DOE-2 incorrectly applies the weather station wind speed to calculate the heat transfer
- Interior convection:
  - EnergyPlus takes window height into account
  - DOE-2 interior convection model is not a function of window height
- Interior radiation:
  - EnergyPlus models interior radiative exchange between every surface
  - DOE-2 treats interior radiation as a thermal resistance between the window surface and the room air

3. Example Test Suite Results

Example 3: Differences due to coding errors and physics algorithms in simulation engines

Improvements to heat transfer coefficient algorithms were identified in both simulation engines that, if implemented, would reduce the difference in calculated window heat loss; the remaining difference is mostly explained by fundamental differences in how each engine models interior radiation.



Note: Results are for a small test building that was created to isolate modeling issues, though similar differences can be found for simulations of typical buildings.

4. Conclusions and Future Work

4. Conclusions

- NREL is identifying, investigating, and resolving inaccuracy issues in building energy simulation programs
- The BEopt comparative test suite has identified software accuracy issues
- The test suite systematically and automatically compares the DOE-2.2 and EnergyPlus simulation engines across a large range of simulation inputs
- Differences in output can be due to non-equivalent inputs, coding errors, and physics algorithms
- A detailed approach involving analytical verification and empirical validation may be needed to resolve modeling issues once they have been identified

4. Future Work

- Other key differences between EnergyPlus and DOE-2 will be documented and investigated
- If simulation coding errors are discovered, they will be documented and made available to the respective software development teams
- Other possible improvements to the BEopt Comparative Test Suite include:
  - Adding test cases to cover new technologies
  - Modifying existing test cases to better-isolate specific algorithms
  - Using synthetic weather data to excite specific physics algorithms, amplifying the effect of differences between simulation engines
  - Adding other simulation engines

5. References

Berry, L.; Gettings, M. (1998). "Realization Rates of the National Energy Audit." In *Proceedings of Thermal Performance of the Exterior Envelopes of Buildings VII. Clearwater, Florida: American Society of Heating, Refrigerating and Air-Conditioning Engineers.*

Kruis, N.; Booten, C.; Christensen, C. (2012). *Identifying and Resolving Issues in EnergyPlus and DOE-2 Window Heat Transfer Calculations. Golden, CO: National Renewable Energy Laboratory. NREL/TP- 5500-55787.*

Hendron, R.; Engebrecht, C. (2010). *Building America House Simulation Protocols. Golden, CO: National Renewable Energy Laboratory. NREL/TP-550-49246.*

Judkoff, R.; Neymark, J. (2006). "Model Validation and Testing: The Methodological Foundation of ASHRAE Standard 140." *ASHRAE Transactions 112(2):367–376. Atlanta, GA: American Society of Heating, Refrigerating and Air-Conditioning Engineers.*

Judkoff, R.; Wortman, D.; O'Doherty, B.; Burch, J. (2008). *A Methodology for Validating Building Energy Analysis Simulations. Golden, CO: National Renewable Energy Laboratory. NREL/TP-550-42059. Based on unpublished report of 1983 with same authors and title, previously referenced as SERI/TR-254-1508.*

Polly, B.; Kruis, N.; Roberts, D. (2011). *Assessing and Improving the Accuracy of Energy Analysis Methods for Residential Buildings*. *Golden, CO: National Renewable Energy Laboratory. DOE/GO-102011-3243.*

Appendix A. Comparative Software Testing

Appendix A. Comparative Software Testing

Differences between the predictions of two or more software programs can be caused by[1]:

- Input Differences
  - Building Inputs: Geometry, material physical properties, and characteristics of mechanical equipment
  - Occupant Inputs: Occupant behavior and occupant-controlled equipment settings
  - Site Inputs: Local weather, soil thermal properties, and adjacent structures/vegetation
- Software Differences
  - Physics Algorithms: Mathematical modeling of the physical behavior of the building and its equipment
  - Coding Errors: Typographical and logic errors inadvertently introduced into the software code

1. Based on Judkoff and Neymark (2006) and Berry and Gettings (1998)

Appendix A. Comparative Software Testing

Differences between the predictions of two or more software programs can be caused by[1]:

- Input Differences
  - Building Inputs: Geometry, material physical properties, and characteristics of mechanical equipment.
  - Occupant Inputs: Occupant behavior and occupant-controlled equipment settings.
  - Site Inputs: Local weather, soil thermal properties, and adjacent structures/vegetation.

Key Idea: Eliminate input differences to identify software differences!

- Software Differences
  - Physics Algorithms: Mathematical modeling of the physical behavior of the building and its equipment.
  - Coding Errors: Typographical and logic errors inadvertently introduced into the software code.

1. Based on Judkoff and Neymark (2006) and Berry and Gettings (1998)

Appendix A. Comparative Software Testing

Ideal Approach for Comparative Testing:

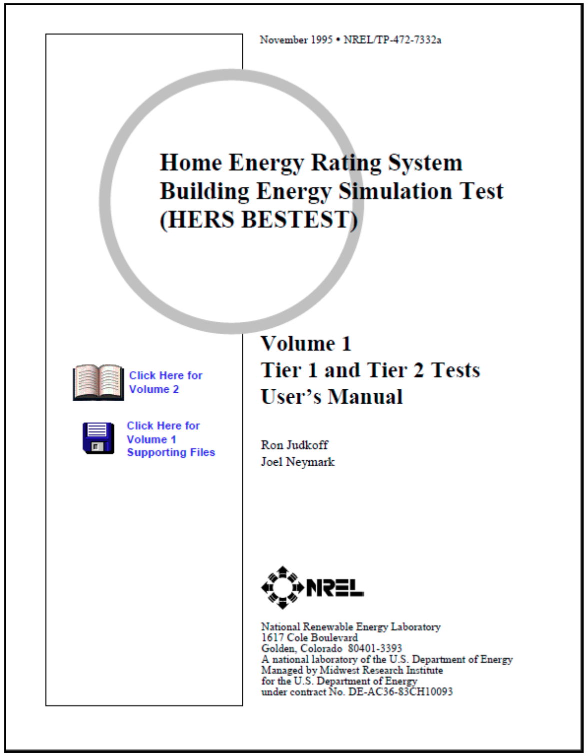
- Provide software programs with equal or equivalent inputs covering a comprehensive range of inputs (building characteristics, occupant behavior, and site conditions)
- Study the differences between simulation output, which are due to coding errors or fundamental differences in physics algorithms

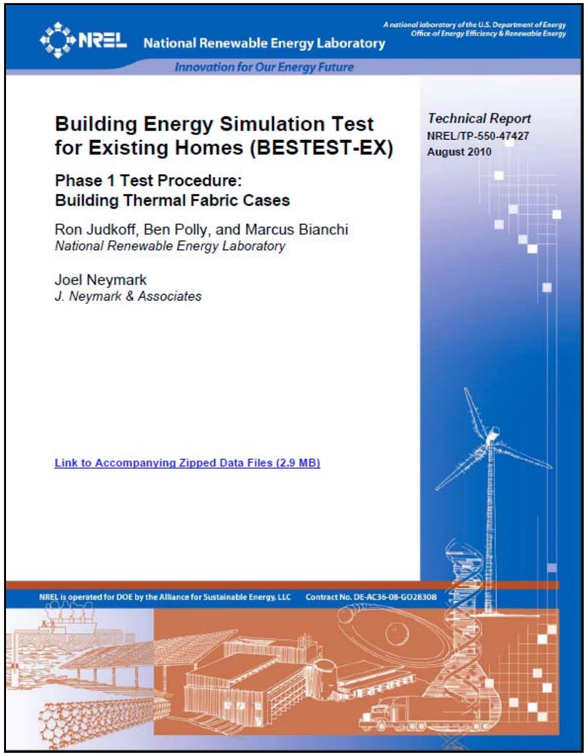
Important Ideas to Understand:

- In some cases it is difficult to provide equal or equivalent inputs
- Just because two or more software programs agree, does not necessarily mean they are “accurate,” but…
- If large discrepancies are found between programs, it is likely one or more of the programs is not accurately representing physical behavior of the building system—further investigation is needed!

Appendix A. Comparative Software Testing

- There are numerous existing comparative test suites
- Many tests were originally developed by NREL according to the Building Energy Simulation Test (BESTEST) methodology [*Judkoff and Neymark 2006*]
- BESTEST test suites are available to software developers





Appendix A. Comparative Software Testing

BESTEST Comparative Suites:

- Provide a variety of equivalent inputs so that many software programs can be tested
- Allow test-takers to compare their software to numerous other programs
- Have been automated by some users to track changes in software from one version to the next
- Are publically available through publications and technical standards

Key Idea: Develop a research tool that complements existing BESTEST suites by focusing on the automated comparison of commonly-used building energy simulation engines.

Appendix A. Comparative Software Testing

For research purposes, an automated test suite is needed:

Desired Strengths

- Allows for rapid comparison of widely-used building energy simulation engines to identify potential issues
- Covers a very large parameter space and can easily integrate new technologies
- Can apply detailed building and occupant descriptions (i.e. Building America House Simulation Protocols[1])
- Automatically creates simulation engine input files, reducing the probability of user input errors

Probable Limitations

- Not intended for generic use as a test suite for all building energy simulation programs
- Certain types of bugs in the process of automatically generating simulation engine input files will not be identified. These bugs can propagate to all simulation engines and will show up as agreement, despite underlying problems.

1. Hendron and Engebrecht (2010)