CONTENTS
ANSI/ASHRAE Standard 140-2023
Method of Test for Evaluating Building Performance Simulation Software

CONTENTS

CONTENTS

SECTION **PAGE**

CONTENTS

NOTE

Approved addenda, errata, or interpretations for this standard can be downloaded free of charge from the ASHRAE website at www.ashrae.org/technology.

(This foreword is not part of this standard. It is merely informative and does not contain requirements necessary for conformance to the standard. It has not been processed according to the ANSI requirements for a standard and may contain material that has not been subject to public review or a consensus process. Unresolved objectors on informative material are not offered the right to appeal at ASHRAE or ANSI.)

FOREWORD

This method of test (MOT) was developed to identify and diagnose differences in predictions from building performance simulation software that may be caused by algorithmic differences, modeling limitations, faulty coding, inadequate documentation, or input errors. These tests are part of an overall validation methodology described in Informative Annex B23. The procedures test software over a broad range of parametric interactions and for a number of different output types, thus minimizing the concealment of algorithmic differences by compensating errors. Different building performance simulation software programs, representing different degrees of modeling complexity, can be tested. However, some of the tests may be incompatible with some building performance simulation software.

The tests are a subset of all the possible tests that could occur. A large amount of effort has gone into establishing a sequence of tests that examines many of the physical and mathematical models relevant to simulating the energy performance of a building and its mechanical equipment. However, because building performance simulation software operates in an immense parameter space, it is not practical to test every combination of parameters over every possible range of function.

The tests consist of a series of carefully described test case building plan, material, and mechanical equipment specifications. Output values for the cases are compared and used in conjunction with diagnostic logic to determine the sources of predictive differences. The tests are designed to highlight the impact of specific building features or physical phenomena. As such, the tests will sometimes describe idealized conditions that do not represent typical modeling practices.

The test cases are divided into separate test classes to satisfy different levels of software modeling detail. Such classification allows more convenient citation of specific sections of Standard 140 by other codes and standards and by certifying and accrediting agencies. The Class I test cases are detailed diagnostic tests intended for simulation software capable of hourly or subhourly simulation time steps. The Class II test cases may be used for all types of building load calculation methods, regardless of time-step granularity, and are designed to represent single-family residential buildings. The Class I test cases are designed for more detailed diagnosis of simulation models than the Class II test cases. An overview of the test suites and methodologies that make up Standard 140 follows. Additionally, Section 4 of the standard provides a detailed road map of the test specifications and output requirements.

For Standard 140-2023, the existing material was reorganized and renumbered. The reorganization is editorial only and does not change the substance of the standard. The changes make it easier to add new test suites and for organizations to cite specific portions of the standard. An informative table (Standard140-2023_SectionSourceMap.pdf) illustrating how section numbers were remapped is included in the root folder of the accompanying electronic files.

CLASS I TEST PROCEDURES

The set of Class I tests included herein consist of

a. *software-to-software comparative tests, in which a program's results may be compared to itself or to the results of other programs in a consistent and repeatable manner, and*
b. *analytical verification tests, in which a program's results may be compared to the results from analytical, quasi-analytical, or verified numerical model solutions (this terminology is elaborated in a subsection below).*

In addition to comparative and analytical verification tests, the overall methodology for model validation and testing described in Informative Annex B23, in the 2021 ASHRAE Handbook—Fundamentals *(see Chapter 19, Section 8), and elsewhere (as cited in Annex B23, Section B23.1.1) includes empirical validation testing, where tested software models are validated to within the uncertainty of experimental input and output data. Such tests can be accommodated within the current title purpose and scope of Standard 140 but are not yet included. Additional research on this topic is recommended, as discussed in Informative Annex B23.*

The Class I test cases are based on a number of technical reports; see Annex B23, Section B23.1 for reference citations.

Weather Drivers Tests (Section 6)

These test cases are designed to test the ability to read and process standard format weather data files. Different weather files test the ability to handle different climates, hemispheres, latitudes, time zones, and ground reflectance. The ability to process both meteorological and solar data from weather files is evaluated. While a test suite that concerns reading and interpreting weather files may seem basic, there are potential algorithmic differences among programs related to psychrometrics, sky models, trigonometry for tilted surfaces, and differences in how instantaneous weather file data are mapped into the corresponding hour.

Building Thermal Envelope and Fabric Load Tests (Section 7)

Basic Comparative Tests of Sections 7.2.1 and 7.2.2

For these cases, the tested modeling physics includes conduction, convection, and radiative (solar and infrared) heat transfer associated with the following:

- *Various building surfaces and their thermal mass*
- *Windows and solar gains through windows*
- *External shading devices*
- *Internally generated heat*
- *Outside-air infiltration/ventilation*
- *Sunspaces*
- *Variations in thermostat control (deadband and setback)*

These test cases, along with the test cases of Section 7.2.3, were updated for Standard 140-2020 to address advancements in the state of the art in building performance modeling since the original publication for Standard 140-2001. Additionally, comments on the test specifications since its inception identified ambiguities, further necessitating an update to the original specifications. The update work built on the initial version of the test suite published in ASHRAE Standard 140-2001 through 140-2017 and includes additional test cases, diagnostic outputs, and updated informative example results. For supporting documentation of the update details, Standard 140-2020 provides informative marked revisions and an update summary in its accompanying files.

In-Depth Comparative Tests of Section 7.2.3

These test cases facilitate diagnosis by allowing excitation of specific heat transfer mechanisms, which are isolated to a greater degree than in the basic tests of Sections 7.2.1 and 7.2.2.

Ground-Coupled Slab Analytical Verification Tests of Section 8

These tests use the results of an analytical solution as a primary mathematical truth standard, and the results from a set of detailed verified numerical models for three-dimensional ground-coupled heat transfer as a secondary mathematical truth standard. These results are then used for comparing to the results of models typically embedded within building performance simulation software. (See section "The Importance of Analytical Solutions, Quasi-Analytical Solutions, Verified Numerical Model Results" below.)

Parametric variations from a steady-state base case include the following:

- *Harmonically varying ground surface temperature*
- *Floor slab aspect ratio*
- *Slab area*
- *Water table depth (depth of constant ground temperature)*
- *Slab interior and ground exterior-surface heat transfer coefficients*
- *Slab and ground thermal conductivity*

Space-Cooling Equipment Cases (Section 9)

The space-cooling equipment cases of Section 9 test the ability of programs to model the performance of unitary space-cooling equipment using manufacturer design data presented as empirically derived performance maps. Many building performance simulation software programs are designed to work with this type of data, and there is very little manufacturer's data that would support the alternative of first principles (direct physical system component) modeling.

Steady-State Analytical Verification Cases of Sections 9.2.1 and 9.2.2

These test cases have analytical and quasi-analytical solutions. The importance of such solutions is discussed in the subsection named to that effect below.

The test cases utilize a typical range of equipment performance data, where the following parameters are varied:

- *Sensible internal gains*
- *Latent internal gains*
- *Zone thermostat set point (entering dry-bulb temperature),*
- *Outdoor dry-bulb temperature*

Parametric variations isolate the effects of the parameters singly and in various combinations and isolate the influence of the following:

- *Part-loading of equipment*
- *Varying sensible heat ratio*
- *Dry-coil (no latent load) versus wet-coil (with dehumidification) operation*
- *Operation at typical Air-Conditioning, Heating, and Refrigeration Institute (AHRI) rating conditions*

Comparative Test Cases of Sections 9.2.3 and 9.2.4

These test cases utilize the following:

- *An expanded range of performance data*
- *An outdoor air mixing system*
- *Hourly varying weather data and internal gains*

These cases cannot be solved analytically but are more realistic. In these cases, the following parameters are varied:

- *Sensible internal gains*
- *Latent internal gains*
- *Infiltration rate*
- *Outdoor air fraction*
- *Thermostat set points*
- *Economizer control settings*

Through analysis of results, the influence of the following can also be isolated:

- *Part-loading of equipment*
- *Outdoor dry-bulb (ODB) temperature sensitivity*
- *Dry-coil (no latent load) versus wet-coil (with dehumidification) operation*

These cases help to scale the significance of simulation results disagreements more so than the steady-state cases.

Space-Heating Equipment Cases (Section 10)

These cases test the ability of programs to model the performance of residential fuel-fired furnaces. They are divided into two tiers.

- *Tier 1 analytical verification test cases (Sections 10.2.1 and 10.2.2) employ simplified boundary conditions and test the basic functionality of furnace models. These cases have analytical and quasi-analytical solutions. The importance of such solutions is discussed in the subsection named to that effect below.*
- *Tier 2 comparative tests (Section 10.2.3) employ more realistic boundary conditions, where additional specific aspects of furnace models are examined.*

The full set of space-heating test cases is designed to test the implementation of specific algorithms for modeling the following aspects affecting furnace performance:

- *Furnace steady-state efficiency*
- *Furnace part-load ratio*
- *Furnace fuel consumption*
- *Circulating fan and draft fan operation*
- *Thermostat setback*
- *Undersized capacity*

Air-Side Heating, Ventilating, and Air-Conditioning (HVAC) Equipment Cases (Section 11)

These cases test the ability of programs to model fundamental air distribution system mass flow and heat balance. These are steady-state analytical verification tests at a variety of constant zone and ambient conditions. The cases have analytical and quasi-analytical solutions. The importance of such solutions is discussed in the subsection named to that effect below.

The test systems include the following in order of increasing complexity:

- *Four-pipe fan coil (FC): This is a single-zone system with heating and cooling coils, zone air exhaust, and limited outdoor air (no economizer control), and it does not include a return air fan.*

- *Single-zone air conditioner (SZ): Same as FC system, except it adds an economizer and a return air fan.*
- *Constant-volume terminal reheat (CV): Same as SZ system, except it further applies multiple (two) zones, system supply-air temperature control, and terminal reheat coils.*
- *Variable-air-volume terminal reheat (VAV): Same as CV system, except it further applies a variable air-flow supply fan and terminal zone supply-air dampers.*

The test cases are conducted at five different sets of steady-state outdoor and zone conditions in heating, dry-coil cooling, and wet-coil cooling modes and with temperature and enthalpy economizer outdoor air control strategies applied to selected conditions.

CLASS II TEST PROCEDURES (SECTION 12)

The Class II (Section 12) test cases were adapted from HERS BESTEST, as cited in Annex B23, Section B23.1. These cases were developed in a more realistic residential context and have a more complex base building construction than the Class 1 test cases, which have more idealized and simplified construction to enhance diagnostic capability. The Class II test cases are based on a building plan for a single-story house with 1539 ft^2 of floor area, with one conditioned zone (the main floor), an unconditioned attic, a raised floor exposed to air (highly vented crawlspace), and typical glazing and insulation. This set of test cases formally codify Tier 1 and Tier 2 tests for certification of residential building performance simulation software.

Tier 1 (Section 12.2.1 and 12.2.2)

These cases test the ability of software to model building envelope loads in the base-case configuration with the following variations:

- *Infiltration*
- *Wall and ceiling R-values*
- *Glazing physical properties*
- *Area and orientation*
- *Shading by a south overhang*
- *Internal loads*
- *Exterior surface color*
- *Energy inefficient building*
- *Raised floor exposed to air*
- *Uninsulated and insulated slabs-on-grade and basements; the specification for the Class II floor slabs and below-grade basement walls is based on simplified methods from 1995, whereas the Class I ground-coupled slab cases of Section 8 allow testing of more detailed models.*

Tier 2 (Section 12.2.3)

These cases test the following additional elements related to passive solar design:

- *Variation in mass*
- *Glazing orientation*
- *East and west shading*
- *Glazing area*
- *South overhang*

To help avoid user input errors for the Section 12 test cases, the input for the test cases is simple, while remaining as close as possible to typical residential constructions and thermal and physical properties. Typical building descriptions and physical properties published by sources such as the National Association of Home Builders, the U.S. Department of Energy, ASHRAE, and the National Fenestration Rating Council are used for the Section 12 test cases.

Comparing Tested Results

The tests have a variety of uses, including the following:

a. *Comparing the predictions from other building performance simulation software programs to the example results provided in Informative Annexes B8 and B16 for Class I tests, Informative Annex B20 for Class II tests, and/or to other results that were generated using this MOT*
b. *Checking a program against a previous version of itself after internal code modifications to ensure that only the intended changes actually resulted*
c. *Checking a program against itself after a single algorithmic change to understand the sensitivity between algorithms*

d. Diagnosing the algorithmic sources and other sources of prediction differences (Diagnostic logic flow diagrams are included in Informative Annex B9.)
e. Comparing building performance simulation programs with each other to determine the degree of disagreement among them
f. Providing a set of procedures to evaluate the acceptability of software; Standard 140 is widely cited by codes, standards, and regulatory bodies as described in Informative Annex B23, Section B23.1.

Normative Annex A3 provides acceptance criteria for selected outputs for Class I tests. To comply, the software must demonstrate that a sufficient number of software results are within the ranges specified for each set of test cases.

Informative (nonmandatory) Annex B22 provides an example statistically based procedure for establishing acceptance range criteria to assess annual or seasonal heating and cooling load results for software undergoing the Class II tests contained in Section 12. Inclusion of this example is intended to be illustrative only for tests related to Section 12. However, certifying or accrediting agencies using Section 12 may wish to adopt procedures for developing acceptance range criteria for tested software.

The building performance simulation software programs used to generate example simulation results (see Informative Annexes, B8, B16, and B20) and to develop acceptance criteria (see Normative Annex A3 and Informative Annex B12) have been subjected to a number of analytical verification, empirical validation, and comparative testing studies. However, there is no such thing as completely validated building performance simulation software. All building models are simplifications of reality. The philosophy here is to generate a range of results from several programs that are generally accepted as representing the state of the art in building performance simulation software. To the extent possible, input errors or differences have been eliminated from the presented results. Thus, for a given case, the range of differences between comparative test results presented in Informative Annexes B8, B16, and B20 and in the acceptance criteria of Normative Annex A3 represents legitimate algorithmic differences (as defined in Informative Annex B11, Section B11.1.4) among these computer programs. For any given case, a tested program may fall outside this range without necessarily being incorrect. However, it is worthwhile to investigate the sources of substantial differences, as the collective experience of the authors of this standard is that such differences often indicate problems with the software or its usage, including, but not limited to the following:

a. User input error, where the user misinterpreted or incorrectly entered one or more program inputs
b. Inadequate or faulty documentation
c. A problem with a particular algorithm in the program
d. One or more program algorithms used outside their intended range

Also, for any given case, a program that yields values in the middle of the range established by the comparative test example results should not be perceived as better or worse than a program that yields values at the borders of the range.

Importance of Analytical Solutions, Quasi-Analytical Solutions, and Verified Numerical Model Results

In general, it is difficult to develop analytical verification test cases, but such cases are extremely useful. Under the classification of "analytical verification," we define three types of test case solutions: "analytical solutions," "quasi-analytical solutions," and "verified numerical models." Analytical solutions represent a "mathematical truth standard," while quasi-analytical solutions and verified numerical models represent "secondary mathematical truth standards" (as described in Informative Annex B23, Section B23.1.1.2; also see 2021 ASHRAE Handbook—Fundamentals*, Chapter 19, Section 8). For selected Class I test cases, Informative Annexes B8 and B16 provide analytical verification test results based on the above solution types, along with simulation results.*

For analytical solutions, given the underlying simplified physical assumptions in the case definitions, there is a mathematically correct solution for each case. For quasi-analytical solutions, the assumptions can be somewhat more realistic; however, there is also the possibility for human interpretation to yield solutions that are slightly different but with a much smaller range of disagreement than results from different simulation programs. Verified numerical models allow even more realistic assumptions and cases but must be subjected to a rigorous screening procedure to minimize the possibility of errors. Verified numerical models are first compared to all available analytical solutions and then compared to each other for cases that do not have exact analytical solutions. Once verified, these numerical solutions can be used to test other models as implemented within building performance simulation software. The ground-coupled slab-on-grade heat transfer test cases of Section 8 utilize an analytical solution and verified numerical models to extend the analytical verification method beyond the constraints inherent in analytical solutions. All three types of analytical verifi-

cation solutions provide a basis for greater diagnostic capability than the purely software-to-software comparative test method, and the verified numerical models allow more realistic boundary conditions to be used in the test cases than are possible with analytical solutions. See Informative Annex B23 for a more complete description of the analytical verification test methodology. Discussion is also included in 2021 ASHRAE Handbook—Fundamentals*, Chapter 19, Section 8.*

It is important to understand the difference between a "mathematical truth standard" and an "absolute truth standard." When applying mathematical truth standards, we only test the solution process for a model, not the appropriateness of the model itself; that is, we accept the given underlying physical assumptions while recognizing that these assumptions represent a simplification of physical reality. For example, a one-dimensional conduction model may be properly solved mathematically, but inappropriate where two-dimensional conduction dominates. By contrast, an "approximate truth standard" from an experiment tests both the solution process and the appropriateness of the model within experimental uncertainty. The ultimate or "absolute" validation truth standard would be comparison of simulation results with data from a perfectly performed empirical validation experiment, with all simulation inputs perfectly defined.

We include simulation results for the cases where analytical verification results (analytical, quasi-analytical, or verified numerical model solutions) exist. This allows simulationists to compare their relative agreement (or disagreement) versus the analytical verification results to that for other simulation results. Perfect agreement among simulations and analytical verification results is not necessarily expected because sometimes simulations cannot perfectly match the specified simplifying assumptions or boundary conditions required for developing the analytical verification solutions. The provided results give an indication of the degree of agreement that is possible between simulation results and the analytical verification results. Therefore, a tested program may disagree with analytical verification solutions without necessarily being incorrect. However, it is worthwhile to investigate the sources of such differences, as noted previously.

Supporting Files

The supporting electronic files to be used with Standard 140-2023 are laid out as described in README-140-2023.docx, provided in the root folder of the accompanying files package. Accompanying files are organized by a separate file folder for each set of tests. Each of the test-set-specific file folders is further subdivided by separate subfolders for normative files and informative files. Normative files include weather data, output report templates, and equipment performance data (for those test cases that apply such data). Informative files include example results, example entries for documentation reports, and other supporting information. The names of the files have been updated to reflect the new section numbers in Standard 140-2023. Electronic files can be downloaded online at ***www.ashrae.org/XXXXX****. Additional informative supporting files are located at* ***data.ashrae.org/XXXXX****. These include reference files that provide general informa-tion about Standard 140 and the analysis files for the weather drivers test suite results.*

1. PURPOSE

This standard specifies test procedures for evaluating the technical capabilities and ranges of applicability of software that simulates the performance of buildings and their systems.

2. SCOPE

These standard test procedures apply to software that simulates the performance of a building and its systems. While these standard test procedures cannot test all algorithms within building performance simulation software, they can be used to indicate major flaws or limitations in capabilities.

3. DEFINITIONS, ABBREVIATIONS, AND ACRONYMS

3.1 Terms Defined for This Standard

adiabatic: without loss or gain of heat. (***Informative Note:*** For example, an adiabatic boundary does not allow heat to flow through it.)

adjusted net sensible capacity: the gross sensible capacity less the actual fan power. (See *gross sensible capacity*.)

adjusted net total capacity: the gross total capacity less the actual fan power. (See *gross total capacity*.)

altitude: vertical elevation above sea level.

analytical solution: a mathematical solution of a model of reality that has an exact result for a given set of parameters and simplifying assumptions.

analytical verification: where outputs from a program, subroutine, algorithm, or software object are compared to results from a known analytical solution or to results from a set of closely agreeing quasi-analytical solutions or verified numerical models. (See *analytical solution*, *quasi-analytical solution*, and *verified numerical model*.)

annual heating load: the heating load for the entire one-year simulation period. (***Informative Note:*** For example, for hourly simulation programs, this is the sum of the hourly heating loads for the one-year simulation period.)

annual hourly integrated maximum zone air temperature: the hourly zone temperature that represents the maximum for the one-year simulation period.

annual hourly integrated minimum zone air temperature: the hourly zone temperature that represents the minimum for the one-year simulation period.

annual hourly integrated peak floor conduction: the hourly floor conduction that represents the maximum for the final year of the simulation period; used for tests of Section 8.

annual hourly integrated peak heating load: the hourly heating load that represents the maximum for the one-year simulation period.

annual hourly integrated peak sensible cooling load: the hourly sensible cooling load that represents the maximum for the one-year simulation period.

annual hourly integrated peak zone load: the hourly zone load that represents the maximum for the final year of the simulation period; used for tests of Section 8.

annual hourly 1°C zone air temperature bin frequency: the number of hours that the zone air temperature has values within a given bin (1°C bin width) for the one-year simulation period.

annual incident unshaded total solar radiation (diffuse and direct): the sum of direct solar radiation and diffuse solar radiation that strikes a given surface for the entire one-year simulation period when no shading is present. (***Informative Note:*** For example, for hourly simulation programs, this is the sum of the hourly total incident solar radiation for the one-year simulation period.)

annual mean zone air temperature: the average zone air temperature for the one-year simulation period. (***Informative Note:*** For example, for hourly simulation programs, this is the average of the hourly zone air temperatures for the one-year simulation period.)

annual sensible cooling load: the sensible cooling load for the entire one-year simulation period. (***Informative Note:*** For example, for hourly simulation programs, this is the sum of the hourly sensible cooling loads for the one-year simulation period.)

annual transmitted solar radiation (diffuse and direct): the sum of direct solar radiation and diffuse solar radiation that passes through a given window for the entire one-year simulation period. This quantity does not include radiation that is absorbed in the glass and conducted inward as heat. (***Informative Note:*** This