BSR/ASHRAE Addendum a to
ANSI/ASHRAE Standard 140-2023

First Public Review Draft

Proposed Addendum a to Standard 140-2023, Method of Test for Evaluating Building Performance Simulation Software

First Public Review Draft (April 2025)
(Draft shows Proposed Changes to Current Standard)

This draft has been recommended for public review by the responsible project committee. To submit a comment on this proposed standard, go to the ASHRAE website at www.ashrae.org/standards-research--technology/public-review-drafts and access the online comment database. The draft is subject to modification until it is approved for publication by the Board of Directors and ANSI. Until this time, the current edition of the standard (as modified by any published addenda on the ASHRAE website) remains in effect. The current edition of any standard may be purchased from the ASHRAE Online Store at www.ashrae.org/bookstore or by calling 404-636-8400 or 1-800-727-4723 (for orders in the U.S. or Canada).

The appearance of any technical data or editorial material in this public review document does not constitute endorsement, warranty, or guaranty by ASHRAE of any product, service, process, procedure, or design, and ASHRAE expressly disclaims such.

ASHRAE, 180 Technology Parkway, Peachtree Corners GA 30092

Proposed BSR/ASHRAE Addendum *a* to ANSI/ASHRAE Standard 140-2023, *Method of Test for Evaluating Building Performance Simulation Software*

(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 to Addendum *a*

This addendum adds time series results to the Standard 140 acceptance criteria for the thermal fabric test groups. The time series results included in the acceptance criteria are the daily hourly zone loads for February 1 and July 14 included in the standard output reports for the thermal fabric low-mass and high-mass test suite (Section 7).

Normative Annex A3 contains the ranges for specific test cases that will be applied to software results to determine if they fall in the acceptable ranges. The minimum number of range cases within the test group to pass has been updated to reflect the inclusion of more cases to the acceptance criteria.

Informative Annex B12 has been updated to include the procedure for developing the acceptance criteria ranges for time series results.

In addition, the range tables of Normative Annex A3 have been consolidated into fewer tables to make finding the acceptance ranges easier.

[Note to Reviewers: This addendum makes proposed changes to the current standard. These changes are indicated in the text by underlining (for additions) and strikethrough (for deletions) except where the reviewer instructions specifically describe some other means of showing the changes. Only these changes to the current standard are open for review and comment at this time. Additional material is provided for context only and is not open for comment except as it relates to the proposed changes.]

[Note.to.Reviewers¿Revisions.proposed.for.this.addendum.begin.with.Normative.Annex.A9¡]

NORMATIVE ANNEX A3
SOFTWARE ACCEPTANCE CRITERIA

This annex provides the acceptance criteria for the software. In order to comply, the software must demonstrate that a sufficient number of software results (see Section A3.2) are within the ranges specified in Section A3.1 for each set of test cases.

Software includes the calculation engine and the corresponding user interface. The testing of software only meets the requirements of this annex for that software and cannot be used as a proxy for documenting compliance of another software that uses the same calculation engine but a different user interface.

A3.1 Class I Software Ranges. Tables A3.1-1 through A3.1-9 ~~13~~ give the acceptance criteria ranges for the test groups. Ranges are given for ~~both~~ selected individual test-case results (e.g., Case 600); ~~and~~ differences between the results of two test cases, i.e., result sensitivities (e.g., Case 610 – Case 600); and daily hourly time series results (e.g., February 1 Case 600).

For single-value results (see Tables A3.1-1, A3.1-4, A3.1-7 through A3.1-9), a ~~A~~ software result is within a given range when the software result is greater than or equal to ~~or greater than~~ the lower limit and less than or equal to the upper limit shown for the given range in the tables. (Informative Note: The setting of these ranges is based on the calculations shown in Informative Annex B12, Section B12.1.1.1, "Acceptance Criteria Bounds Calculations", "Equations for Single-Value Results".)

For the time series ranges, comparison metrics (Pearson correlation coefficient (for curve shape) and Root Mean Square Error (RMSE) (for value scaling)) for the software results are calculated versus the representative curve results

(see Tables A3.1-2 and A3.1-5), and then these metrics are compared to the ranges (see Tables A3.1-3 and A3.1-6). A software result is within a given range if both the calculated Pearson correlation coefficient is above the limit in the table and the calculated RMSE value is below the limit in the table. (***Informative Note:*** The setting of these ranges is based on the calculations shown in Informative Annex B12, Section B12.1.1.2, "Acceptance Criteria Bounds Calculations", "Equations for Time Series Results".)

The criteria are organized by test suite as follows. Tables A3.1-1 through A3.1-6 ~~4~~ are related to the building thermal envelope and fabric load tests of Section 7. Table A3.1-7 ~~5~~ is related to the space-cooling equipment performance tests of Sections 9.2.3 and 9.2.4. Table~~s~~ A3.1-8 ~~6 and A3-7 are~~ is related to the space-heating equipment performance tests of Section 10. Table~~s~~ A3.1-9 ~~8 through A3-13 are~~ is related to the air-side HVAC equipment performance tests of Section 11.

Table A3.1-1 Low Mass Building~~—Annual Heating Load (MWh/yr)*~~ Single-Value Ranges

	Annual Heating Load (MWh/yr)*		Annual Sensible Cooling Load (MWh/yr)*	
Range Case	**Lower Limit**	**Upper Limit**	**Lower Limit**	**Upper Limit**
Case 600	3.75	4.98	5.00	6.83
Case 610 – Case 600	–0.14	0.29	–2.26	–0.80
Case 620 – Case 600	–0.08	0.40	–2.24	–1.64
Case 630 – Case 620	0.02	0.74	–1.68	–0.77
Case 640 – Case 600	–2.17	–1.22	–0.56	0.03
Case 650 – Case 600	n/a	n/a	–1.54	–0.95
Case 660 – Case 600	–1.07	–0.16	–3.09	–2.50
Case 670 – Case 600	0.25	2.98	0.05	0.84
Case 680 – Case 600	–2.54	–1.90	0.13	0.87
Case 685 – Case 600	0.33	0.77	2.70	3.31
Case 695 – Case 685	–2.38	–1.94	–0.21	0.44

* For output specified in Section 7.3.1.1 and 7.3.1.2 for heating and cooling loads, respectively.

~~**Table A3-2 Low Mass Building—Annual Sensible Cooling Load (MWh/yr)***~~

~~**Range Case**~~	~~**Lower Limit**~~	~~**Upper Limit**~~
~~Case 600~~	~~5.00~~	~~6.83~~
~~Case 610 – Case 600~~	~~–2.26~~	~~–0.80~~
~~Case 620 – Case 600~~	~~–2.24~~	~~–1.64~~
~~Case 630 – Case 620~~	~~–1.68~~	~~–0.77~~
~~Case 640 – Case 600~~	~~–0.56~~	~~0.03~~
~~Case 650 – Case 600~~	~~–1.54~~	~~–0.95~~
~~Case 660 – Case 600~~	~~–3.09~~	~~–2.50~~
~~Case 670 – Case 600~~	~~0.05~~	~~0.84~~
~~Case 680 – Case 600~~	~~0.13~~	~~0.87~~
~~Case 685 – Case 600~~	~~2.70~~	~~3.31~~

~~Case 695 – Case 685~~	~~-0.21~~	~~0.44~~

~~For output specified in Section 7.3.1.2~~

Table A3.1-2 Low Mass Building— Hourly Loads (kWh) with Heating > 0 and Cooling <0 Representative Curves Used to Calculate Comparison Metrics

	February 1							July 14					
Hour	Case 600	Case 640	Case 660	Case 670	Case 680	Case 685	Case 695	Case 600	Case 660	Case 670	Case 680	Case 685	Case 695
1	2.08	0.08	1.82	2.62	1.17	2.12	1.31	0.00	0.00	0.00	0.00	-0.05	-0.14
2	2.22	0.60	1.90	2.73	1.32	2.24	1.41	0.00	0.00	0.00	0.00	0.03	-0.07
3	2.28	0.95	1.94	2.82	1.42	2.29	1.48	0.00	0.00	0.00	0.00	0.13	-0.02
4	2.33	1.19	2.00	2.91	1.50	2.33	1.53	0.00	0.00	0.12	0.00	0.25	0.04
5	2.35	1.32	2.02	2.95	1.53	2.35	1.55	0.00	0.00	0.33	0.00	0.34	0.11
6	2.35	1.40	2.03	2.94	1.55	2.35	1.56	0.00	0.00	0.22	0.00	0.24	0.06
7	2.36	1.45	2.05	2.92	1.55	2.36	1.55	0.00	0.00	0.00	0.00	-0.10	-0.10
8	1.88	3.06	1.71	2.39	1.16	1.88	1.16	0.00	0.00	0.00	0.00	-0.46	-0.33
9	0.55	1.75	0.82	0.84	0.17	0.55	0.06	0.00	0.00	0.00	0.00	-0.91	-0.63
10	0.00	0.08	0.08	0.00	-0.10	-1.11	-1.67	-0.23	-0.08	-0.26	-0.19	-1.55	-1.15
11	-1.28	-0.67	0.00	-1.14	-1.98	-3.12	-3.67	-1.23	-0.78	-1.51	-1.07	-2.40	-1.87
12	-3.23	-2.69	-0.47	-3.27	-4.13	-4.52	-5.21	-2.22	-1.41	-2.61	-1.89	-3.14	-2.54
13	-4.42	-4.17	-1.63	-4.62	-5.31	-5.41	-6.11	-2.70	-1.85	-3.26	-2.41	-3.49	-3.00
14	-4.67	-4.59	-2.15	-4.89	-5.57	-5.43	-6.13	-2.75	-2.00	-3.31	-2.49	-3.51	-3.05
15	-4.05	-4.02	-1.87	-4.12	-5.05	-4.70	-5.51	-2.59	-1.90	-3.06	-2.26	-3.23	-2.78
16	-2.93	-2.92	-1.28	-2.78	-3.94	-3.59	-4.40	-2.21	-1.69	-2.53	-1.90	-2.85	-2.34
17	-1.33	-1.32	-0.38	-0.99	-2.35	-1.88	-2.73	-1.76	-1.41	-1.94	-1.51	-2.40	-1.95
18	-0.04	-0.04	0.00	0.00	-0.51	-0.06	-0.88	-1.37	-1.13	-1.46	-1.19	-1.99	-1.62
19	0.09	0.09	0.08	0.46	0.00	1.07	0.18	-0.98	-0.85	-1.02	-0.88	-1.59	-1.30
20	0.98	0.98	0.85	1.62	0.00	1.67	0.67	-0.57	-0.52	-0.55	-0.57	-1.15	-0.99
21	1.64	1.64	1.41	2.22	0.20	2.03	1.00	-0.23	-0.24	-0.14	-0.33	-0.76	-0.73
22	2.06	2.06	1.78	2.60	0.72	2.28	1.24	-0.02	-0.06	0.00	-0.17	-0.55	-0.55
23	2.32	2.32	2.01	2.88	1.08	2.44	1.42	0.00	0.00	0.00	-0.05	-0.38	-0.43
24	2.44	0.00	2.10	2.99	1.29	2.50	1.52	0.00	0.00	0.00	0.00	-0.25	-0.31

Table A3.1-3 Low Mass Building—Time Series Ranges

	Minimum Pearson Correlation Coefficient for February 1 and July 14 Hourly Loads*	Maximum Root Mean Square Error for February 1 Hourly Loads*	Maximum Root Mean Square Error for July 14 Hourly Loads*

Range Case	Lower Limit	Upper Limit	Upper Limit
Case 600	0.879	0.4194	0.1921
Case 640 (February 1 only)	0.879	0.6559	n/a
Case 660	0.879	0.3087	0.1697
Case 670	0.879	0.7881	0.3018
Case 680	0.879	0.3990	0.2097
Case 685	0.879	0.4643	0.3218
Case 695	0.879	0.2069	0.2001

* For loads output specified in Section 7.3.8. The same lower limit value applies to February 1 and July 14 loads, except there are no July 14 hourly load results for Case 640.

Table A3.1-4 ~~3~~ High Mass Building—~~Annual Heating Load (MWh/yr)*~~Single-Value Ranges

	Annual Heating Load (MWh/yr)*		Annual Sensible Cooling Load (MWh/yr)*	
Range Case	Lower Limit	Upper Limit	Lower Limit	Upper Limit
Case 900	1.04	2.28	2.35	2.60
Case 900 – Case 910	–0.52	–0.02	0.35	1.74
Case 920 – Case 900	1.51	1.92	0.08	0.48
Case 930 – Case 920	0.20	1.15	–1.19	–0.44
Case 940 – Case 900	–0.82	–0.37	–0.19	0.06
Case 950 – Case 900	n/a	n/a	–2.00	–1.56
Case 960 – Case 900	0.96	1.12	–1.81	–1.27
Case 980 – Case 900	–1.65	–1.00	1.09	1.41
Case 985 – Case 900	0.64	0.81	3.52	4.18
Case 995 – Case 985	–1.83	–1.07	0.63	1.15

* For output specified in Section 7.3.1.1 and 7.3.1.2 for heating and cooling loads, respectively.

~~**Table A3-4 High Mass Building—Annual Sensible Cooling Load (MWh/yr)***~~

~~Range Case~~	~~Lower Limit~~	~~Upper Limit~~
~~Case 900~~	~~2.35~~	~~2.60~~
~~Case 900 – Case 910~~	~~0.35~~	~~1.74~~
~~Case 920 – Case 900~~	~~0.08~~	~~0.48~~
~~Case 930 – Case 920~~	~~–1.19~~	~~–0.44~~
~~Case 940 – Case 900~~	~~–0.19~~	~~0.06~~
~~Case 950 – Case 900~~	~~–2.00~~	~~–1.56~~
~~Case 960 – Case 900~~	~~–1.81~~	~~–1.27~~

~~Case 980 – Case 900~~	~~1.09~~	~~1.41~~
~~Case 985 – Case 900~~	~~3.52~~	~~4.18~~
~~Case 995 – Case 985~~	~~0.63~~	~~1.15~~

~~For output specified in Section 7.3.1.2~~

Table A3.1-5 High Mass Building— Hourly Loads (kWh with Heating > 0 and Cooling <0 Representative Curves Used to Calculate Comparison Metrics

	February 1					July 14			
Hour	Case 900	Case 940	Case 980	Case 985	Case 995	Case 900	Case 980	Case 985	Case 995
1	0.98	0.00	0.00	1.13	0.42	-0.10	-0.22	-0.71	-0.64
2	1.08	0.00	0.00	1.23	0.54	-0.03	-0.16	-0.62	-0.58
3	1.18	0.00	0.00	1.33	0.63	0.00	-0.09	-0.53	-0.50
4	1.29	0.00	0.00	1.43	0.71	0.00	-0.02	-0.42	-0.43
5	1.37	0.00	0.00	1.50	0.77	0.00	0.00	-0.33	-0.35
6	1.44	0.00	0.10	1.57	0.81	0.00	0.00	-0.32	-0.37
7	1.51	0.00	0.22	1.62	0.84	0.00	-0.06	-0.40	-0.44
8	1.36	1.26	0.19	1.45	0.70	0.00	-0.16	-0.52	-0.54
9	0.79	1.58	0.00	0.87	0.25	-0.11	-0.27	-0.72	-0.67
10	0.27	0.84	0.00	0.35	-0.36	-0.36	-0.49	-1.04	-0.92
11	0.00	0.30	0.00	-0.25	-1.05	-0.66	-0.77	-1.43	-1.28
12	0.00	0.00	-0.11	-0.94	-1.80	-0.98	-1.05	-1.79	-1.58
13	0.00	0.00	-0.73	-1.59	-2.50	-1.22	-1.27	-2.04	-1.81
14	0.00	0.00	-1.17	-1.94	-2.79	-1.38	-1.40	-2.17	-1.89
15	-0.02	0.00	-1.35	-2.01	-2.83	-1.41	-1.39	-2.15	-1.89
16	-0.01	0.00	-1.28	-1.83	-2.72	-1.36	-1.28	-2.05	-1.78
17	0.00	0.00	-1.01	-1.36	-2.19	-1.26	-1.19	-1.94	-1.67
18	0.00	0.00	-0.47	-0.63	-1.40	-1.14	-1.08	-1.80	-1.55
19	0.00	0.00	-0.05	-0.19	-0.92	-1.01	-0.96	-1.63	-1.40
20	0.00	0.00	0.00	0.10	-0.65	-0.85	-0.81	-1.43	-1.22
21	0.00	0.09	0.00	0.34	-0.42	-0.70	-0.69	-1.27	-1.08
22	0.00	0.33	0.00	0.55	-0.23	-0.59	-0.59	-1.16	-0.98
23	0.08	0.54	0.00	0.71	-0.04	-0.49	-0.51	-1.05	-0.89
24	0.29	0.00	0.00	0.86	0.10	-0.37	-0.41	-0.94	-0.79

Table A3.1-6 High Mass Building—Time Series Ranges

	Minimum Pearson Correlation Coefficient for February 1 and July 14 Hourly Loads*	Maximum Root Mean Square Error for February 1 Hourly Loads*	Maximum Root Mean Square Error for July 14 Hourly Loads*
Range Case	**Lower Limit**	**Upper Limit**	**Upper Limit**
Case 900	0.879	0.3716	0.0967
Case 940 (February 1 only)	0.879	0.4298	n/a
Case 980	0.879	0.2969	0.0986
Case 985	0.879	0.4166	0.2053
Case 995	0.879	0.2944	0.2468

* For loads output specified in Section 7.3.8. The same lower limit value applies to February 1 and July 14 loads, except there are no July 14 hourly load results for Case 940.

Table A3.1-7~~5 Annual Total Cooling Energy Consumption (kWh/yr)*~~Cooling Equipment Single-Value Ranges

	Annual Total Cooling Energy Consumption (kWh/yr)*	
Range Case	**Lower Limit**	**Upper Limit**
CE300	33497	37023
CE310 – CE300	2671	6197
CE320 – CE300	2610	6136
CE330 – CE320	–766	2760
CE330 – CE300	3697	7223
CE340 – CE300	3172	6698
CE330 – CE340	–1393	2133
CE350 – CE300	–5354	–1828
CE360 – CE300	18084	21610
CE400 – CE300	–5783	–2179
CE410 – CE300	–5929	–371
CE420 – CE300	–3945	–419
CE430 – CE300	–4762	–1236
CE440 – CE300	–3694	–168

* For output specified in Section 9.3.2.1.1.1(a)

Table A3.1-8~~6 Furnace Input Energy (GJ)*~~Heating Equipment Single-Value Ranges

	Furnace Input Energy (GJ)*		Cumulative Fan Energy (kWh)*	
Range Case	**Lower Limit**	**Upper Limit**	**Lower Limit**	**Upper Limit**
HE210	49.86	55.11	279.55	308.98

HE220 – HE210	–5.48	–0.23	–30.21	–0.78
HE230 – HE220	–9.10	–3.86	186.29	216.78

* For output specified in Section 10.3.1.2 and 10.3.2.1 for furnace input energy and cumulative fan energy, respectively.

~~**Table A3-7 Heating Equipment Cumulative Fan Energy (kWh)***~~

~~**Range Case**~~	~~**Lower Limit**~~	~~**Upper Limit**~~
~~HE210~~	~~279.55~~	~~308.98~~
~~HE220 – HE210~~	~~–30.21~~	~~–0.78~~
~~HE230 – HE220~~	~~186.29~~	~~216.78~~

* ~~For output specified in Section 10.3.2.1~~

Table A3.1-9 Air-Side Equipment Single-Value Ranges

Range Case	Preheat-Coil Load (kWh/h)*		Heating-Coil Load (kWh/h)*		Cooling-Coil Total Load (kWh/h)*		Cooling-Coil Sensible Load (kWh/h)*		Zone 1 Reheat-Coil Load (kWh/h)*		Zone 2 Reheat-Coil Load (kWh/h)*	
	Lower Limit	Upper Limit	Lower Limit	Upper Limit	Lower Limit	Upper Limit	Lower Limit	Upper Limit	Lower Limit	Upper Limit	Lower Limit	Upper Limit
AE201	n/a	n/a	7.94	8.78	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
AE205	n/a	n/a	n/a	n/a	1.82	2.01	1.82	2.01	n/a	n/a	n/a	n/a
AE245 – AE205	n/a	n/a	n/a	n/a	0.20	0.39	0.20	0.39	n/a	n/a	n/a	n/a
AE206	n/a	n/a	n/a	n/a	2.54	2.85	1.62	1.79	n/a	n/a	n/a	n/a
AE226 – AE206	n/a	n/a	n/a	n/a	0.55	0.82	–0.23	–0.06	n/a	n/a	n/a	n/a
AE301	2.80	4.20	n/a	n/a	n/a	n/a	n/a	n/a	7.23	8.10	7.77	8.81
AE301 – AE401	–7.67	–6.16	n/a	n/a	n/a	n/a	n/a	n/a	2.70	3.60	2.87	3.93
AE305	n/a	n/a	n/a	n/a	8.77	9.85	8.77	9.85	1.92	2.46	2.02	2.71
AE305 – AE405	n/a	n/a	n/a	n/a	4.55	5.75	4.55	5.75	n/a	n/a	n/a	n/a
AE345 – AE305	n/a	n/a	n/a	n/a	–0.15	0.79	–0.15	0.79	–0.11	0.11	–0.12	0.12
AE345 – AE445	n/a	n/a	n/a	n/a	4.76	6.01	4.76	6.01	n/a	n/a	n/a	n/a
AE306	n/a	n/a	n/a	n/a	13.65	15.08	8.30	9.37	1.92	2.48	2.02	2.74
AE306 – AE406	n/a	n/a	n/a	n/a	5.42	6.86	4.64	5.73	n/a	n/a	n/a	n/a
AE326 – AE306	n/a	n/a	n/a	n/a	5.68	7.12	–0.91	–0.01	–0.11	0.11	–0.12	0.12
AE326 – AE426	n/a	n/a	n/a	n/a	11.21	12.65	4.27	5.31	n/a	n/a	n/a	n/a
AE401	9.75	10.77	n/a	n/a	n/a	n/a	n/a	n/a	4.27	4.72	4.63	5.12
AE405	n/a	n/a	n/a	n/a	3.90	4.31	3.90	4.31	0.00	0.00	0.00	0.00
AE445 – AE405	n/a	n/a	n/a	n/a	–0.15	0.26	–0.15	0.26	n/a	n/a	n/a	n/a
AE406	n/a	n/a	n/a	n/a	7.81	8.82	3.45	3.82	0.00	0.00	0.00	0.00
AE426 – AE406	n/a	n/a	n/a	n/a	0.20	1.02	–0.22	0.14	n/a	n/a	n/a	n/a

For output specified in Sections as follows. Preheat coil, 11.3.3.2(a),11.3.4.2; heating coil, 11.3.1.2(a), 11.3.2.2; cooling coil total, 11.3.1.2(d), 11.3.2.2, 11.3.3.2(d), 11.3.4.2; cooling coil sensible load 11.3.1.2(b), 11.3.2.2, 11.3.3.2(b), 11.3.4.2; zone 1 reheat coil, 11.3.3.2(f), 11.3.4.2; zone 2 reheat coil, 11.3.3.2(g), 11.3.4.2.

Table A3-8 Preheat Coil Load (kWh/h)*

Range Case	Lower Limit	Upper Limit
AE301	2.80	4.20
AE401	9.75	10.77
AE301 – AE401	–7.67	–6.16

* For output specified in Sections 11.3.3.2(a) and 11.3.4.2

Table A3-9 Heating Coil Load (kWh/h)*

Range Case	Lower Limit	Upper Limit
AE201	7.94	8.78

* For output specified in Sections 11.3.1.2(a) and 11.3.2.2

Table A3-10 Cooling Coil Total Load (kWh/h)*

Range Case	Lower Limit	Upper Limit
AE205	1.82	2.01
AE245 – AE205	0.20	0.39
AE206	2.54	2.85
AE226 – AE206	0.55	0.82
AE305	8.77	9.85
AE305 – AE405	4.55	5.75
AE345 – AE305	–0.15	0.79
AE345 – AE445	4.76	6.01
AE306	13.65	15.08
AE306 – AE406	5.42	6.86
AE326 – AE306	5.68	7.12
AE326 – AE426	11.21	12.65
AE405	3.90	4.31
AE445 – AE405	–0.15	0.26
AE406	7.81	8.82
AE426 – AE406	0.20	1.02

* For output specified in Sections 11.3.1.2(d), 11.3.2.2, 11.3.3.2(d), and 11.3.4.2

Table A3-11 Cooling Coil Sensible Load (kWh/h)*

Range Case	Lower Limit	Upper Limit
AE205	1.82	2.01
AE245 – AE205	0.20	0.39

AE206	1.62	1.79
AE226 – AE206	–0.23	–0.06
AE305	8.77	9.85
AE305 – AE405	4.55	5.75
AE345 – AE305	–0.15	0.79
AE345 – AE445	4.76	6.01
AE306	8.30	9.37
AE306 – AE406	4.64	5.73
AE326 – AE306	–0.91	–0.01
AE326 – AE426	4.27	5.31
AE405	3.90	4.31
AE445 – AE405	–0.15	0.26
AE406	3.45	3.82
AE426 – AE406	–0.22	0.14

* For output specified in Sections 11.3.1.2(b), 11.3.2.2, 11.3.3.2(b), and 11.3.4.2

Table A3-12 Zone 1 Reheat Coil Load (kWh/h)*

Range Case	**Lower Limit**	**Upper Limit**
AE301	7.23	8.10
AE401	4.27	4.72
AE301 – AE401	2.70	3.60
AE305	1.92	2.46
AE345 – AE305	–0.11	0.11
AE306	1.92	2.48
AE326 – AE306	–0.11	0.11
AE405	0.00	0.00
AE406	0.00	0.00

* For output specified in Sections 11.3.3.2(f) and 11.3.4.2

Table A3-13 Zone 2 Reheat Coil Load (kWh/h)*

Range Case	**Lower Limit**	**Upper Limit**
AE301	7.77	8.81
AE401	4.63	5.12
AE301 – AE401	2.87	3.93
AE305	2.02	2.71
AE345 – AE305	–0.12	0.12

~~AE306~~	~~2.02~~	~~2.74~~
~~AE326 – AE306~~	~~-0.12~~	~~0.12~~
~~AE405~~	~~0.00~~	~~0.00~~
~~AE406~~	~~0.00~~	~~0.00~~

* ~~For output specified in Sections 11.3.3.2(g) and 11.3.4.2~~

Table ~~A3-14~~ A3.2-1 Acceptance Range Pass Criteria

Test Group	Tables of Ranges	Number of Range Cases in Test Group	Minimum Number of Range Cases within the Test Group to Pass
Thermal Fabric Low Mass	A3.1-1, A3.1-3~~2~~	~~21~~34	~~18~~30
Thermal Fabric High Mass	A3.1-4~~3~~, A3.1-6~~4~~	~~19~~28	~~17~~25
Cooling Equipment	A3.1-7~~5~~	14	12
Heating Equipment	A3.1-8~~6, A3-7~~	6	5
Air-side Equipment	A3.1-9~~8, A3-9, A3-10, A3-11, A3-12, A3-13~~	54	48

A3.2 Number of Results within Acceptance Range to Pass a Test Group. Table ~~A3-14~~ A3.2-1 gives the number of range cases for which a software's results must be within the acceptable ranges to pass a test group. Software shall pass a test group when the count of results computed for each range case, from the corresponding tables of ranges, meets or exceeds the minimum count required in Table ~~A3-14~~A3.2-1.

Informative Note: See Informative Annex B12, Section B12.4 for a description of the procedures used to determine the numbers in the table.

A3.3 Submission of Results. Test results shall be provided in the normative output forms of Normative Annex A2. Submittals shall also include a complete set of reports, as described in Normative Annex A2, Attachment A2.8, with all report blocks completed.

INFORMATIVE ANNEX B12
DEVELOPMENT OF ACCEPTANCE CRITERIA

B12.1 Acceptance Criteria Bounds Calculations. The results from a reference set of software programs (see Section B12.1.2), also called "reference results" here, are used to calculate the upper and lower limits for the acceptance criteria.

B12.1.1 Equations for Setting Bound Limits. There are two different types of results included in the acceptance criteria: single-value results (e.g., annual totals) and time series results (e.g., hourly zone loads for a specific day). There are different calculation methodologies for each of these types of results.

B12.1.1.1 Equations for Single-Value Results. The upper and lower limits for each range case applied for acceptance criteria are set by selecting the lower minimum value and higher maximum value of the following:

a. Statistical Bounds (SB):

$$\text{Minimum: MinSB} = \text{Median}_{case} - \text{MAD}_{3sigma}$$
$$\text{Minimum: MaxSB} = \text{Median}_{case} + \text{MAD}_{3sigma}$$

where

Median_{case} ≡ median value of reference results for a given case

MAD_{3sigma} ≡ median absolute deviation (MAD) with 99.73% (3 sigma) confidence interval for the reference results of a given case

and

$$\text{MAD}(X) = \text{Median}[|X - \text{Median}(X)|]$$

for a normal distribution, we can relate MAD to the standard deviation as

$$MAD = \sigma\Phi^{-1}\left(\frac{3}{4}\right) = \sigma\sqrt{2}erf^{-1}\left(\frac{1}{2}\right) \approx 0.6745\sigma$$

where

σ ≡ standard deviation

Φ ≡ quantile function

erf ≡ error function

The bounds of the confidence interval can be set in terms of ± standard deviation from the average or ± MAD from the median. In this case, a 99.73% confidence interval results from three standard deviations (or 3 sigma), which leads to a MAD multiplier (MAD_{3sigma}) of 2.024.

b. Nonstatistical Bounds (NSB):

1. Where no quasi-analytical solution results exist for an acceptance criteria test group:

$$\text{Minimum: MinNSB} = \text{Median}_{case} - \text{Median}_{BC} \times 0.05$$
$$\text{Maximum: MaxNSB} = \text{Median}_{case} + \text{Median}_{BC} \times 0.05$$

where

Median_{BC} ≡ median value of reference results of base case (BC) for a given case

2. Where quasi-analytical solution results (QAS) exist for an acceptance criteria test group:

$$\text{Minimum: MinNSB} = \text{QAS}_{case} - \text{QAS}_{BC} \times 0.05$$
$$\text{Maximum: MaxNSB} = \text{QAS}_{case} + \text{QAS}_{BC} \times 0.05$$

where

QAS_{case} ≡ QAS value for a given case

QAS_{BC} ≡ QAS value of the base case (BC) for a given case

B12.1.1.2 Equations for Time Series Results

The time series results are 24 hourly results for a single day. A representative curve is needed to provide values for the comparison to software results. This curve is determined by taking the median value of the reference set of software results at each hour.

There are two different comparisons made with the time series results: the shape of the curve of the data and the scale of the values. For the shape of the curve, the Pearson correlation coefficient (pcc) is used, and for the scale of the values, the Root Mean Square Error (RMSE) is used.

The pcc measures the linear correlation between two sets of data. It is the ratio between the covariance of two variables and the product of their standard deviations; thus, it is essentially a normalized measurement of the covariance, such that the result always has a value between −1 and 1. A value of 1 denotes perfect correlation (the curves are exactly the same), -1 denotes a perfect inverse correlation (the curves are mirror images), and 0 denotes that there is no correlation between the curves.

$$pcc = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

Where
n is sample size (number of data points in each curve)
x_i, y_i are the individual sample points indexed with i (x being the representative curve and y being the software results being compared)
$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ (the sample) mean; and analogously for $\bar{y}$.

RMSE is the square root of the average of squared errors. The effect of each error on RMSE is proportional to the size of the squared error; thus, larger errors have a disproportionately large effect on RMSE. Consequently, RMSE is sensitive to outliers.

$$RMSE = \sqrt{\sum(x_i - y_i)^2 / n}$$

Since the pcc is an absolute measure of the agreement of the shape of the curves, the MAD procedure is not appropriate for determining agreement, and a significance test for the correlation coefficient is used instead. The P-Value of the pcc is the probability that the two curves are NOT in agreement (in this case, of the same shape), with greater values indicating greater disagreement. A P-value limit of 0.05 is used meaning that if the P-value of the computed correlation coefficient is less than 0.05, the two curves are considered to have the same shape. (A P-value of 0.05 approximates a 95% confidence interval based on a Student-T distribution.)

To find the P-value for a given test result compared to the reference, a test statistic, *t*, is computed as the ratio of the correlation coefficient to the standard error (σ_{pcc}) in the correlation coefficient.

$$t = \frac{pcc}{\sigma_{pcc}} = pcc\frac{\sqrt{(N-2)}}{\sqrt{1 - pcc^2}}$$

Where
pcc= Pearson correlation coefficient between the test result and the representative curve,
σ_{pcc} = standard error (sample standard deviation) applied to *pcc* over the number of data points (24) in each curve, and
N= number of reference test results sets used to create the representative curve.

The higher the correlation coefficient, the larger *t*. The probability (P-value) of obtaining this test statistic in error is determined from the student-T distribution with *N*-2 degrees of freedom. [1]

[1] https://en.wikipedia.org/wiki/Pearson_correlation_coefficient

With N = 5 samples used to create the representative curve, for P < 0.05 one needs t > 3.18 (https://en.wikipedia.org/wiki/Student%27s_t-distribution) which corresponds to pcc > 0.879. The student-T distribution was chosen in place of the normal distribution because of the limited number of samples used to create the representative curve. Thus, any program's results with a *pcc* > 0.879 are considered to be significantly correlated with the representative curve.

For the RMSE, the MAD procedure for single-value results (see Section B12.1.1.1) is used. However, the 3-sigma bounds used in the single-value comparisons were deemed too strict for the time series values. This was for two reasons: 1) that the time series comparisons were for the results for a single test case and not the difference between two test cases (i.e., differential sensitivity test results tend to subtract out overall noise that may be present in any case taken by itself) and 2) that there were 24 values being used rather than a single-value in the RMSE (i.e., larger number of values emphasizing dynamic variations has greater opportunity for disagreement at any single point in time). Subsequently, a 5-sigma bound is used for the RMSE comparison. This results in a MAD multiplier of 3.3725.

B12.1.2 Reference Software for Setting Bound Limits. The reference set of software for setting the bound limits is selected based on the following criteria:

- A single set of results per calculation engine is included (multiple user interfaces for the same calculation engine are not included).
- Results are submitted by the calculation engine copyright holder(s) or by a modeler designated by the calculation engine copyright holder(s), applying a user interface specified by the calculation engine copyright holder(s) if appropriate.
- The set of selected software should represent a diversity of modeling methods, such as control-based and load-based.
- The calculation engine should be actively, reasonably, and widely used and actively maintained.
- Software satisfies the following:
  - Able to run at a time step of an hour or less
  - Able to include hourly variations in occupancy, lighting power, miscellaneous equipment power, thermostat set points, and HVAC system operation
  - Include thermal mass effects
- If a software did not participate in the simulation trials for the initial development of a test suite, then prior to the submittal, its results are reviewed and undergo iterative rounds of review and revisions as determined during the submission review process. The reviewer comments and submitter responses are documented from each round, with a logical basis for corrections to programs or inputs included in the documentation. This review process only applies to the initial submittal of a software and not subsequent submittals of that software for version updates.

The results from a reference software

- include justification via Normative Annex A2, Attachment A2.8, Item G ("Report Block for Anomalous Results") if they fall outside of the established bounds and
- pass the minimum number of tests indicated for each test group in Table ~~A3-14~~ A3.2-1 for which a given reference software provided results.

The results from a reference software may be updated for each new public release of the calculation engine. For future continuous maintenance revisions of Standard 140, SSPC 140 may

- include any additional software in the acceptance criteria-bounds reference set that apply and meet the criteria for inclusion;
- remove any software from the reference set that no longer meet the criteria for inclusion; and
- update the acceptance criteria with the most recent results from the reference set of software.

The current reference set of software used to determine the acceptance criteria are (in alphabetical order): CSE, DeST, DOE-2.2, EnergyPlus, ESP-r, IDA ICE, IES-VE, TAS, and TRNSYS. Not all software are applied for all test groups.

B12.2 Acceptance Criteria Results Comparison. Given that the upper and lower acceptance bound limits for the test cases are determined from the reference set of software, the results from other software can be

compared against those limits. A software result is designated as an outlier for a given acceptance-criteria range case if that result is less than the lower limit or greater than the upper limit.

Tolerances will be used for comparisons to bounds.

a. A relative tolerance of 1% (i.e., numbers within ≤ 1% of each other are considered equivalent)
b. An absolute tolerance of 1E-4 (i.e., numbers within 1E-4 of each other are considered equivalent even if that is more than 1%)

B12.3 Acceptance Criteria Tests Included. Not all test suites and test cases included in Standard 140 are appropriate to be used in acceptance criteria. The following criteria are applied in determining which test results are used in the acceptance criteria:

- A test group has enough of the reference set of software that can adequately perform all of the tests within the test group to produce statistically meaningful acceptance ranges. (Beginning with Standard 140-2023, which first introduced the acceptance criteria, there are at least five reference-set software that can adequately perform the tests within each test group.)
- The tests are grouped to provide separate acceptance criteria for different sections of the standard.
- When available, acceptance-criteria tests include outputs such as annual total load, energy use, or other annual summations based on realistic, time-varying conditions.
- Differences between the results of two tests (i.e., result sensitivities) are preferred.
- A base case for each test suite or portion of the test suite should be included.

Based on these criteria, the following five test groups were selected for the acceptance criteria (selected outputs within each test group are also described, and Test Group 5 has separate selected outputs specific to its test cases):

1. Thermal Fabric Low Mass Test Group Range Cases
   Annual heating load and annual sensible cooling load
   Case 600
   Case 610 – Case 600
   Case 620 – Case 600
   Case 630 – Case 620
   Case 640 – Case 600
   Case 650 – Case 600 (cooling only)
   Case 660 – Case 600
   Case 670 – Case 600
   Case 680 – Case 600
   Case 685 – Case 600
   Case 695 – Case 685
   Daily hourly loads
   Case 600
   Case 640
   Case 660
   Case 670
   Case 680
   Case 685
   Case 695
2. Thermal Fabric High Mass Test Group Range Cases
   Annual heating load and annual sensible cooling load
   Case 900
   Case 900 – Case 910
   Case 920 – Case 900
   Case 930 – Case 920
   Case 940 – Case 900
   Case 950 – Case 900 (cooling only)
   Case 960 – Case 900

Case 980 – Case 900
Case 985 – Case 900
Case 995 – Case 985

Daily hourly loads

Case 900
Case 940
Case 980
Case 985
Case 995

3. Cooling Equipment Test Group Range Cases

Annual total cooling energy consumption

CE300
CE310 – CE300
CE320 – CE300
CE330 – CE320
CE330 – CE300
CE340 – CE300
CE330 – CE340
CE350 – CE300
CE360 – CE300
CE400 – CE300
CE410 – CE300
CE420 – CE300
CE430 – CE300
CE440 – CE300

4. Heating Equipment Test Group Range Cases

Furnace input energy and cumulative fan energy

HE210
HE220 – HE210
HE230 – HE220

5. Air-side Equipment Test Group Range Cases

Heating-coil load

AE201

Preheat-coil load

AE301
AE401
AE301 – AE401

Cooling-coil total load and cooling-coil sensible load

AE205
AE245 – AE205
AE206
AE226 – AE206
AE305
AE305 – AE405
AE345 – AE305
AE345 – AE445
AE306
AE306 – AE406
AE326 – AE306
AE326 – AE426
AE405
AE445 – AE405

AE406
AE426 – AE406

Reheat-coil loads—zone 1 and zone 2
AE301
AE401
AE301 – AE401
AE305
AE345 – AE305
AE306
AE326 – AE306
AE405
AE406

B12.4 Acceptance Criteria Overall Passing. Because there may be legitimate physical modeling reasons (see Informative Annex B11, Section B11.1.4) why a software's results would fall outside of the acceptance range for an individual test or difference between two tests, it is not reasonable to expect a software to pass all acceptance criteria for every test in a test group.

To determine the number of range cases for which a software's results need to be within the acceptance ranges to pass a test group, the number of range cases in the test group is multiplied by 0.9, and the resulting number is rounded down to the nearest integer. For example, for the thermal fabric low mass test group, there are ~~21~~ 34 range cases; ~~21~~ 34 × 0.9 = ~~18.9~~ 30.6; rounding down yields ~~18~~ 30 range cases.