BS 2021 1-3 SEPT BRUGES

INTERNATIONAL BUILDING PERFORMANCE SIMULATION ASSOCIATION

SimQuality
A novel test suite for dynamic building energy simulation tools

Andreas Nicolai[1], Stephan Hirth[1], Madjid Madjidi[2]
[1]TU Dresden, Dresden, Germany
[2]University of Applied Sciences, Munich, Germany

Abstract

Based on existing test suites like the ASHRAE 140 BESTEST and EN ISO 13791, a novel test suite for inter-model comparison of building energy simulation tools has been developed. The test cases investigate physical models and implementations individually and in combined tests, hereby identifying sources of deviations between tested tools. The methodology to build test cases incrementally and evaluate changes resulting from new features added to the tests, allows a much better isolation of error sources. The test suite is particularly focused on providing reference solutions for simulation tool developers. Another novel aspect of the test suite is a set of rules that ensures publication of models and simulation input data, such that results can be reproduced and verified easily by other users. Any software certified with the test suite can be considered to fulfill minimum requirements for the anticipated use cases. The article describes the test cases and variants formulated so far, and gives exemplary result for the tested tools NANDRAD, IDA-ICE, TRNSYS, TAS, Modelica/AixLib, THERAKLES and ETU-Simulation.

Key Innovations

- Novel inter-model comparison test suite for modern building energy simulation tools
- Test and evaluation criteria, including interactive quality tool compliance webpage

Practical Implications

The availability of independently verified and documented software tests may help increase confidence in numerical simulation tools. The standardization on physical problem description and model complexity increases comparability of results from different simulation tools.

Introduction to simulation tool validation

Since the oil crisis in the 1970ties numerous building energy performance simulation (BES) tools have been developed. These tools have since increased in complexity and number of incorporated physical submodels. As with any sophisticated software, with increasing complexity of the tools it becomes generally more difficult for engineering users to verify calculation results, or even do plausibility checks. Effectively, users have to trust that the model predictions are accurate, detailed enough and computed without errors. Tool validation can establish trust in software models and tools and help ensure high quality of planning tools. However, the term *validation* is not trivially defined. In fact, validation of simulation models is a fairly complicated and time-consuming task.

To make things even more difficult, a number of terms such as verification, benchmarking, regression testing, and others are used synonymously, and depending on the author, describe sometimes the same or different things. Hence, before we look into relevant literature and existing validation methods, we should attempt a definition of *validation*, or at least describe our description of its intended outcome.

A theoretical view to tool validation

When we (engineering users as well as scientists) use a simulation model, we generally do that to get answers to questions. For the authors a very simple and overall definition of validation might be:

"A validated simulation tool should correctly provide the requested answers within the desired accuracy."

The interesting terms in this statement are: *requested answers* and *desired accuracy* and *correctly*.

Now, with respect to multi-zone building energy performance simulations, there will be questions like:

- Do we get a summer-time overheating problem? This might be re-phrased into: How many overheating hours will be in critical zones, and where are those zones?
- What will be the net energy demand of the building? How much cooling/heating energy will be needed?
- What will be the primary energy demand?
- What is the peak load which has to be provided by the HVAC equipment?

https://doi.org/10.26868/25222708.2021.30766

When drilling down into the details of HVAC equipment sizing there will be many more. And in future, with the introduction of more and more renewable energy sources, there might be questions like:

- How much generated renewable energy is used by the building itself *directly*? How much is pushed into the grid and how much taken out again?
- Can we shift energy demand such that it better coincides with renewable energy supply?
- How much external carbon-based energy is utilized during nominal building operation?

Depending on the actual task at hand, the requirement on the simulation output detail can be quite different. For some purposes, monthly balances may be sufficient. Clearly, to answer the load-balancing questions related to renewable energy use, this will not be detailed enough. For many tasks, hourly balances/outputs are sufficient. And then there might be some questions requiring sub-hourly results, for example when dealing with HVAC control system design and parametrization. To summarize, the purposes to run a simulation model and the expectations towards the outcome are diverse, and so are the *requested answers*.

Similarly, the accuracy requirements might be different, and depend very much on the accuracy of available input data. For example, in early planning phases not much may be known about building façade materials or room usage/occupants patterns, so it may be in vain to expect accurate hourly temperature results. With progressing state of building planning and design, accuracy of input data increases, and so should the simulation results. Again, the level of achievable and required *accuracy* will be task-specific.

And finally, the term *correct* is very tricky to define. Models are always more or less detailed approximations to the reality. Clearly, different approximations (equation sets) will produce different results, even if the same input parametrization is used. If furthermore input parameters differ, for example when models require different sets of input data that need to be approximated/estimated/converted, the deviations of model outputs may further increase. To complicate things even further, the same set of mathematical equations may be solved using different algorithms, especially when numerical methods are involved. Hereby, additional numerical parameters are introduced that further increase differences in calculation results. These different (independent) sources of deviations may, if combined in different combinations, yield a substantial number of varying results with quite a large standard deviation. Assuming errors related to numerical techniques or number precisions are small, each result will be a *correct* solution for each combination of model equations, input data, calculation/numerical parameters and result interpretation.

Surprisingly, the interpretation of results is rather often a problem, especially when the tools documentation is not very clear about how results shall be interpreted. A very common problem when comparing simulation results is the handling of momentary values (i.e. values at a given time instant) vs. mean values over a given interval. Simulation models using hourly balances often produce mean values over each hour, whereas other simulation tools generate outputs for specific time points. Plotting the data will yield a characteristic shift (see Fig. 2 below)

Also, interpretation of input data can be troublesome. In the SimQuality test procedure we first used epw-files to define weather data. However, it turned out that due to ambiguous/unclear specification of the file standard there are different interpretations of time stamps, described by Crawley et al. (1999) and U.S. Department of Energy (2021).

Currently, the epw variants `2001,1,1,1,60,...` and `2001,1,1,1,0,...` are both being used, for example in the EnergyPlus weather directory (US Department of Energy, 2021). In simulation tools using hourly balances this is not a problem, since the 5th column is ignored and the provided hourly value is interpreted as mean value of the hour. For modern simulation tools, however, that use arbitrarily spaced input data, the usual procedure is to convert time stamps to specific points in time and use linear interpolation to get climate data at interim points. In the tools used in the SimQuality project we found 3 different interpretations for a time stamp in format `2001,1,1,1,60,...`:

```
(1)  01.01.2001 0:00 (begin of first hour)
(2)  01.01.2001 1:00 (end of first hour)
(3)  01.01.2001 2:00 (end of second hour)
```

Variant (3) was used in the Modelica library by using the calculation formula:

```
seconds_of_day = <hour column>*3600 +
                 <minute column>*60
```

In that light, care should be taken when using arbitrary epw files with Modelica building simulation models (AixLib/Buildings/... etc.). As a consequence of such problems, we ceased to use epw files in the *description* of the test cases. They are still being provided with the test cases for convenience, but with a very detailed description on how they should be interpreted, and with the obligation to the tool user to verify that the tool processes the data as intended (in a dedicated test case).

Given the large number of possible reasons for result deviations, it appears that the only way to define correctness properties for several tools would be to define a very narrow band of possible model equation variations and define the input parameters and solution procedures very well. This approach is followed to some extend by the VDI 2078/6007 (see literature review below). However,

we have to consider the substantial differences in numerical solution methods (from simple daily/hourly balances to time-accurate integration with adaptive time steps), present in current simulation tools. Also, model formulation and model complexity varies greatly in modern simulation tools. If the model formulation and required input parameters as well as solution procedure would be fixed in a test suite as it is done in VDI 2078/6007, this effectively disqualifies many equally suited simulation tools. For example, suppose you require hourly balances as solution procedure. Then, obtained results will be average temperatures for each hour. If you are to run the same test case with numerical methods of higher time resolution and higher integration accuracy, you may not meet the reference results defined for hourly balances. Thus, the much more detailed and accurate simulation tool could not be validated and test suites defined as in VDI 2078/6007 with strict specification of solution methods and model complexity are not meaningful for comparison of broad range of different simulation tools.

Indeed, it appears that none of the three properties *outputs/requested answers*, *accuracy/level of detail* or *correctness* can be easily defined. Instead, all appears to be problem-specific.

Hence, we need a consensus on the specific capabilities of individual model components with their level of detail being defined based on *typical application scenarios*. We could then classify tools as "suitable and validated for purpose XXX". To further reduce deviations due to input/output data interpretation, the description of the test cases should be *very detailed*. These considerations led to the SimQuality test suite described below.

Existing test suites and standards

The topic of tool validation is by far not a new one. Numerous papers/scientific works have been written on the subject, and most tool developments started with something similar in place. We will restrict the literature discussion only to standards/building codes or otherwise official documents dealing with validation of simulation tools.

EN ISO 13791:2012

The European standard EN ISO 13791 specifies the assumptions, boundary conditions, equations and validation tests for a calculation procedure, under transient hourly conditions, of the internal temperatures (air and operative) during the warm period, of a single room without any cooling/heating equipment in operation. In section 7 a validation procedure is given for transient BES tools consisting of 4 sections with 4 different test cases. In the first test case the focus lies on heat conduction through opaque elements, in the second on internal long-wave radiation heat exchanges, in the third the evaluation of short-wave radiation heat transfer (calculation of shaded area of a window due to external obstructions). Finally there is a test case concerning the validation procedure for the whole calculation method, where all physical effects are taken into account. This leads to numerous effects in the combined test case. It is generally not possible to determine specifically what causes a deviation of results if it is not detectable by the first three test cases. Furthermore the test suite only checks single zone models. Coupled zones and interzonal heat flows/exchange is not tested or validated at all.

EN ISO 52017-1:2018

In 2018, the EN ISO 13791 has been replaced by European standard ISO 52017-1. In that standard only a calculation procedure for the internal temperatures, heating and cooling in buildings is described. It contains a short section about validation, yet with no test cases, only refering the standard EN ISO 52016-1.

EN ISO 52016-1:2018

In this standard the most authoritative BESTEST cases have been included.

ANSI/ASHRAE 140 - BESTEST

The aim of the Building Energy Simulation Test (BESTEST) and ANSI/ASHRAE Standard 140 is to increase confidence in the use of building energy simulation programs. This is achieved by creating standardized and citable test procedures for validating, diagnosing, and improving the current generation of software. The BESTEST was originally published in the year 1995 and has been updated periodically (Judkoff (2013)). There are in total 39 test cases organized into a basic series and an in-depth series, with many different physical parameters and models to be tested incrementally. A series of different buildings are specified, from the thermally simple to the more realistic approximation. The specified test cases are defined such that the thermal properties, geometric proportions, and thermal responses are reasonable with respect to actual building envelope loads. Various BES tools already took part in the validation procedure and provided result data (ANSI/ASHRAE, 2017).

While the BESTEST is in our opinion a great test suite for quality checking and tool comparison. However, it suffers from some of the issues that apply to the other mentioned suites and standards. The in-depth (195 - 320) test cases do not test each physical model individually, which can result in deviations in detailed combined test cases (600 - 900). Also, only a rather small amount of different tools ran some of the in-depth test cases 195-220, which makes the cross-checking for them rather difficult (Fig. 1). Further, the use of only one climatic location (Denver), may not allow generalization of a tool's correct calculation when used. For example, in the southern hemisphere or in areas with very wide time zones, models may calculate the the suns position incorrectly and thus result in faulty radiation loads. Lastly, the reference for compliance testing (e.g. range of annual heating loads) is defined through aggregated results (mean, min, max) from a variety of tools. Result data is only for some test cases compared at different timesteps, which means that too large loads during some part of the simulation can be compensated by too small loads at other times.

Acceptance is defined through results from tools that have completed the test suite. While this is certainly a practical approach, this leads to the problem that if all programs have implemented a faulty/very simple model, a

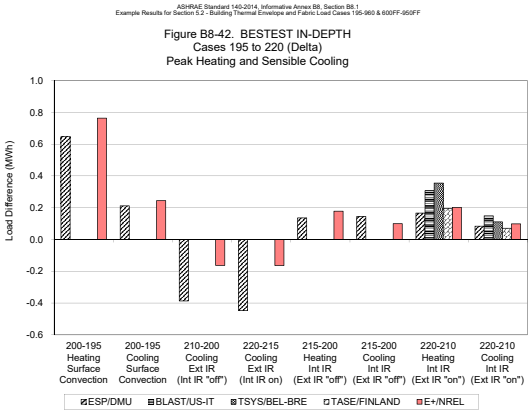

Figure 1: Examplary Delta for Test Cases 195 to 220: Is Energy Plus 9.4 (red) validated?

novel tool with a more detailed closer-to-reality model approach may yield results outside the previously established band of acceptance and thus, may not be validated.

To summarize, the BESTEST is currently the most sophisticated, cited and used suite for validating BES tools and definitely a good way to validate most simulation models.

VDI 6020/6007/2078

The VDI 6020 forms a validation procedure with various test cases for BES programs. The validation is carried out according to two different scenarios: In scenario 1 only thermal energy simulation without technical building equipment is validated. This part is completely described inside the VDI 6007. In scenario 2, BES Tools are validated with regard to their technical building equipment. These test cases are mainly described inside VDI 2078. It contains 16 different test cases with different geometries. There are several problems with the described validation procedure. Firstly, there are three different documents (main document, expansions and excel documents) with redundant and sometimes conflicting input data. When setting up a test, a lot of effort is needed to determine the correct input data for each test case (Schpfer et al. (2010)). In some cases, the documented input data is plain wrong or at least mismatching. For example, in Part 1 (VDI (2016b) p.49 Tab.A3) a window with an area of 10.5 m^2 is described, whereas a window area of 7 m^2 is also mentioned in Part 1 (VDI (2016b) p.47 Fig. 41), in the guideline itself (VDI (2016c) p.55, Tab.C1) and on the CD (VDI (2016a)) concerning the same test cases 01 - 07. In the same way, it is difficult to identify the correct reference results from the Excel sheets from CD1 (VDI (2016a)) needed to verify results.

Further, the validation suite actually tests a very specific model description and solution procedure. While this is normally a good approach to verify different implementations of the *exact same model and parameter set*, this also represents the strongest weakness of this test norm. The model to be tested is the specific VDI calculation model,based on the Beuken model. This model, however, simplifies the actual building physics strongly, hereby introducing in some situations rather large approximation errors. When using a very detailed physical model to calculate the prescribed VDI test cases, these approximation errors become apparent. Since the VDI validation results are rather tightly set around those obtainable with the VDI Beuken model, it is hence impossible to obtain a match with a more detailed model. There is also an option documented to validate a different calculation procedure, hereby using statistical methods. Yet the reference results are still being computed with the simplified VDI model, and hence more physically correct simulation results will still not always be accepted as correct. **Consequently, it would be meaningless to require a modern, physically detailed building energy simulation model to comply to the VDI validation norm.**

We conclude from our analysis of the VDI 6020 test norm, that this VDI guideline should indeed only be used to validate implementations of the VDI specific (Beuken) model (VDI (2015b, 2016c, 2015a)).

Empirical validation

There is, of course, also the option to validate tools against monitoring data. However, a large number of uncertainties and error sources may be introduced. Also, imprecise input data may give room for "fitting" results towards measurements, hereby masking potential modelling/software errors.

The SimQuality test suite

Test suite requirements/target audience

The SimQuality test suite targets two distinct user groups: simulation software developers and engineering users. Furthermore, the tests can be used as well-structured exercises within lectures in building simulation.

The first group should expect very detailed test descriptions (models and parameters) and results to check their own implementation against implementation errors. Since designing suitable test cases with a broad enough parameter variation can be time-consuming, this availability of a test series will effectively reduce development time and effort. For such a testing to be successful, the individual tests of the series should test individual models such, that occurring errors can directly be attributed to a specific model/parameter combination.

The second target group, the engineering users, will be primarily interested in the tool ratings and tool capability summary tables. With this information they can make informed decisions on which tool to use for which application case. Further, the test suite can speed up the learning curve for new simulation users, since test cases with individual modeling tasks can be downloaded, investigated (how are the model parameters set?), and used as means for self-checking. For this user group the webpage content will be most useful, as well as detailed descriptions of how tool results came together.

Scope

The scope of the SimQuality test suite (as far as the current research project is concerned) is the building envelope and general building physics. With increasing use of renewable energy sources we must ensure that the building's response to climate loads is represented well in a simulation program. This is particularly true for the rather mas-

sive building constructions used throughout Germany and Europe, where much of the indoor thermal comfort can (and should) be already maintained by suitable dampening of environmental conditions. The dynamic loads obtained from such a simulation are a critical input to any equipment model, so getting these right is our first priority.

Test methodology

The test suite is centered around typical application use cases. In such an application case, for example summer thermal comfort analysis, heating and cooling load analysis, a number of different model components play a critical role. Depending on the actual building/application scenario, the impact of the model components differs. Sometimes window models are more critical, sometimes it will be the shading, or the thermal construction models.

In order to generalise the validation status, we follow a similar approach as used in modern numerical time integration solvers. In order to control the global integration error, the local errors in each step are limited, thus limiting also the global error occurring over a number of steps. With respect to the SimQuality test suite we strive to limit the error occurring in each individual model component, thus also limiting the error occurring in any combination of said components. Hence, the test suite is composed of a series of tests for *individual model components*, alone.

The current test suite consists of the following tests:

TF00	weather data interpretation
TF01	solar position calculation
TF02	solar load calculation
TF03	heat transfer and storage - single-zone
TF04	heat transfer and storage - multi-zone
TF05	ventilation/infiltration
TF06	solar gains onto and heat transfer through opaque constructions
TF07	solar gains through windows and distribution in rooms
TF08	internal loads
TF09	outer shading (building geometry)
TF10	passive cooling (component activation)
TF11	complex case: summer overheating protection calculation
TF12	complex geometry case: real building geometry

Table 1: participating Editors in SimQuality

Tools	Editors
NANDRAD	Stephan Hirth, IBK TU Dresden
IDA ICE	Stefan Lehr, INNIUS D
TRNSYS	Julian Agudelo, HM Munich
THERAKLES	Andreas Nicolai, IBK TU Dresden
ETU Simulation	Rainer Rolffs, ETU Hottgenroth
Modelica/AixLib	Amin Nouri, RWTH Aachen
TAS	Rainer Strobel, PGMM

There is also a basic test case (TF0) that checks the correct import of provided climate data files, in case that the provided epw-files are being used/imported by a software. This was necessary due to the aforementioned specifica-

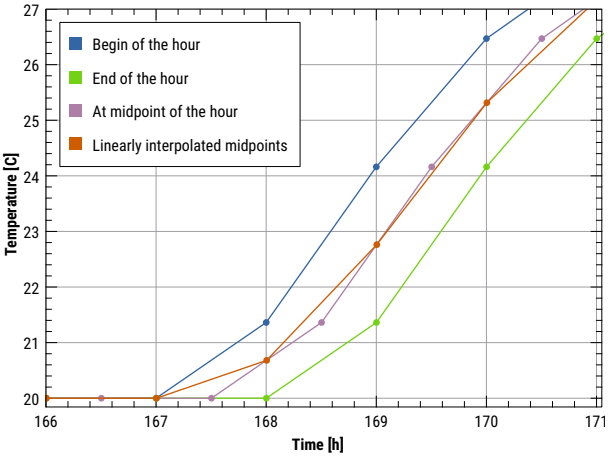

Figure 2: Illustration of the output interpretation/-time shift problem

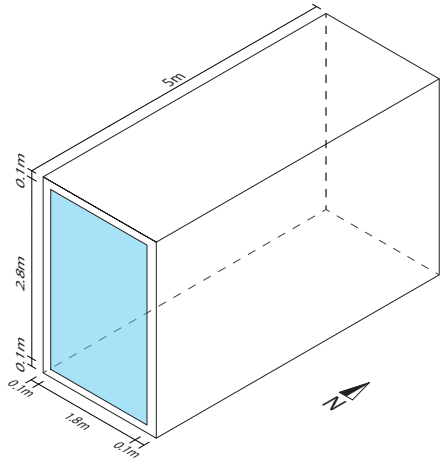

Figure 3: Geometry for the test case validating different window models

tion problems with the epw format. To avoid such errors we specifically describe our interpretation rule for the provided epw files in the test case.

Test procedure

Hereafter, the principle procedure is illustrated using test case 07. In this case, simple to detailed window models are incrementally tested and validated with increasingly more physical model components taken into account.

(07.1)	Opaque window with only heat conduction
(07.2)	Simple window model with constant SHGC-Value
(07.3)	Simple window model with angle dependent SHGC-Value
(07.4)	Detailed window model with two panes
(07.5)	Detailed window model with two panes and coating (lowE)
(07.6)	Detailed window model with two panes and coating (lowE) and longwave radiation
(07.7)	Detailed window model with two panes, longwave radiation and internal distribution

https://doi.org/10.26868/25222708.2021.30766

Result interpretation

When analysing results it is important to use a consistent output value interpretation. For time series data there are two basic output types: values at a given instant in time, and mean values over a certain time interval. The latter is frequently used in simulation models working with balances in a fixed grid and usually hourly steps. Modern time integration engines rather use variable adaptive step integration methods that adjust time steps based on error estimates. While it is easy to generate hourly mean values from these detailed integration engines, hourly values generated by balance methods needs some interpretation.

Figure 2 illustrates different options. If the balance model uses classic time integration formulation (e.g. explicit/implicit Euler), then the states are defined at begin and end of the hourly interval. Which of the curves "Begin of the hour" (blue) or "End of the hour" (green) is meant by the tool depends on the hour indexing method (which is hopefully documented).

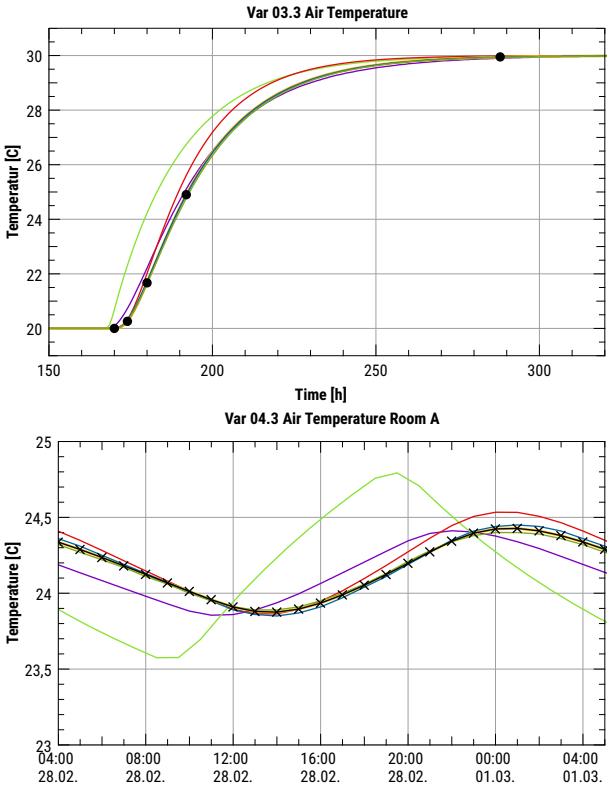

Figure 4: Examplary results from test cases 03 and 04, showing room air temperatures (see discussion in text)

Also possible is the interpretation as average value over the hour. However, instead of plotting these values at step function it is common practice to linearly interpolate between values placed at the mid-point of the hour (purple curve). When comparing such average results with reference values at begin/end of the hour, it is usually possible to compute values at begin/end of the hour via linear interpolation and use these values instead (red curve). This procedure is followed in the SimQuality evaluation procedure.

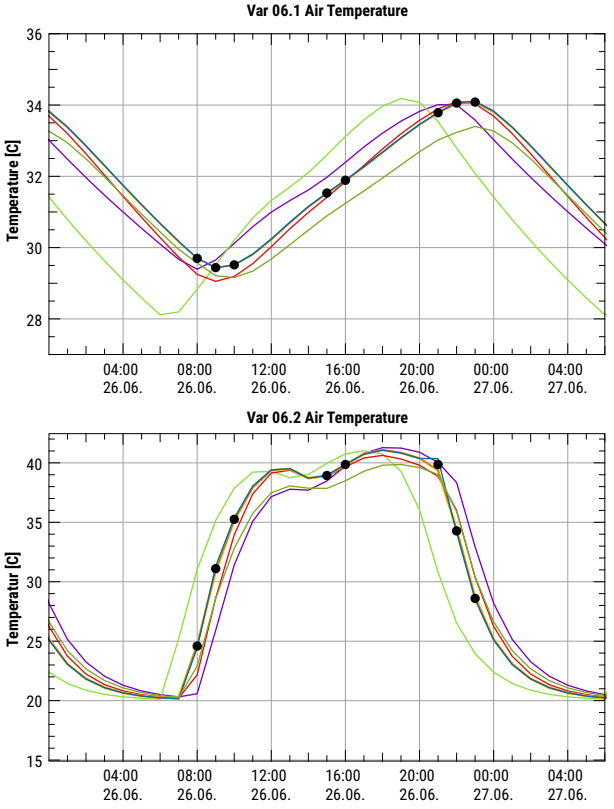

Figure 5: Room air temperature results for two variants of test case 06, with heavy construction (top), and light-weight construction (bottom).

Preliminary results

Since the test suite and research project is still underway, we will present a quick preview of already completed test cases. So far 7 tools have participated in the test suite.

Test cases 3 and 4 deal with correct representation of heat transfer and storage, particularly through massive multi-layered opaque constructions. Figure 4 shows two temperature curves from the test cases. Results from TF03 (top) show deviation of VDI tool (green) compared to majority of other tools. Results from TF04 (bottom) show results of first room in the two coupled room-scenario when quasi-steady state has been reached. Clearly, temperature oscillations are not captured very well by the VDI/Modelica tools (green and purple). The other tools compute rather similar temperatures.

Test case 06 evaluates how thermal loads from solar radiation and heat conduction on the outside of constructions reach the zone via heat conduction and result in a damped oscillation of the room air temperature. The variants of this test case differ by choice of construction materials and layers. Figure 5 shows computed room air temperatures in two cases, one (top) with heavy construction and the other (bottom) with light-weight construction. Most tools capture the characteristics in magnitude and time shift correctly. Again, the tools with simplified wall models show visible deviations. This effect becomes more pronounced for more massive constructions.

Test case 08 evaluates correct handling of thermal loads. As shown in Fig. 6, the impact of thermal transfer across

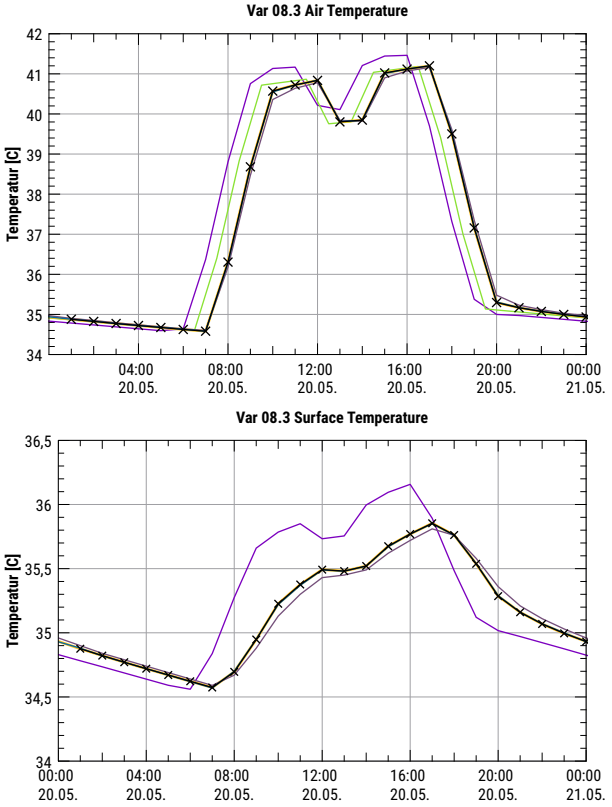

Figure 6: Results from test case 08 (internal loads), room air temperatures (top) and south wall inner surface temperatures.

the constructions is not so important. Rather, correct input of load schedules and inside surface transfer coefficients are critical to the case. Most tools capture the results rather well, as can be seen from the room air temperatures (Fig, top). The deviation of one tool (purple) appear in this diagram only a little off. However, the look to the surface temperature diagram (Fig. 6, bottom) indicates a likely problem with boundary condition parameters. Hence, for analysis of deviations it is often necessary to inspect several result data.

The complete result data set and all diagrams are available on the SimQuality webpage, see below.

Validation procedure

In order to check the correctness of the participating programs in SimQuality, the results are compared with analytical data and inter-program comparisons are performed. Reference values are given for each specific model approach. For example, different reference results apply for an isotropic radiation model then for anisotropic radiation models such as PEREZ. For test cases with analytically generated results, reference results are selected for certain time points and a certain tolerance (e.g. 0.5 K) is applied analogous to test case 1 from EN ISO 13791. For more advanced test cases where no analytical results are available, such as the validation of dynamic storage effects, the tools with the smallest deviations from previous test cases are selected and set as references. Again, references are set at specific time points and a certain tolerance is set. For test cases with strongly oscillating results with highly time dynamical effects, statistical parameters like RMSE are also compared over a longer period of time to better analyze the agreement of the curves. For each test case, a committee of the research partners participating in SimQuality currently decides which references apply. However, these are currently to be considered preliminary and are still being specified.

Dissemination and interactive data publication

A central part of the SimQuality validation and certification concept is an independent and verifiable documentation of tool validation. For this purpose a website and an interactive platform have been developed, *https://simquality.org* (English, available at the end of 2021) and *https://simquality.de* (German) with the following content:

- collection of standards, procedures, building codes, scientific works etc. on simulation tool validation for building components
- list of already defined validation suits for: thermal bridges, hygrothermal analysis, single and multi-zone building energy performance simulation, and HVAC component simulation models
- the SimQuality test suite description
- interactive platform to upload own validation data/-download published data

The ability to upload and store calculation results obtained with a range of tools helps to document and demonstrate capabilities of these tools and models. Since tools are continuously improved and developed, it is possible to upload data for different versions of a tool, hereby documenting improvements and enhancements/bug fixes in software programs.

Transparent tool validation through data publication

Currently, validation of tools is done either by software developers/companies themselves or by independent researchers or master/PhD students. The results of these tests are often just plain documents/published papers with more or less brief summaries of actual model parameters and excerpts of tool results. Part of this problem originates from inability to publish data alongside articles with most major journals. And given limited information it is often difficult to reproduce data, verify a validation claim or even find differences between own modelling attempts and those of (successful) others.

The publication of validation results on the SimQuality-website can improve transparency on tool validation in two ways:

- the actual simulation results (raw data) is being uploaded in a specified format, automatically processed and transformed into diagrams, and compared against the reference to generate the test certification
- users of the platform can download the published data and compare with own results
- the input data (project files, material data, climate date files, etc.) can be uploaded

Especially the latter feature is important. It should be possible to download the input data for a test case, install the software on own PCs, run the test cases and

get exactly the same results as published on the webpage. This also can be used to verify, that software code was not specifically tuned to run a given test case.

Quality Management

Since this is a public webpage open for arbitrary (registered) users, there are several rules governing the upload-/publication process:

- data is published as a *set of data* for a unique combination of *tool* (or tool set), *version*, and *editor/agent*
- data must comply to the test-specific format/completeness requirements (e.g. the variables and sampling times required etc.)
- data must run successful through the validation procedure, hereby generating a certain number of validation points/stars
- data can be complemented by tool input data sets (e.g. an archive with all mandatory data needed to re-run the simulation)
- data can be altered/withdrawn by the submitting user and the test case administrator (the latter only in cases of abuse/rule violation)

It might be possible that some user publishes results showing a failure of a tool (or just bogus results), while another user (possibly the author of the software with more in-depth knowledge) publishes results showing a validation. In this case, the results can be discussed in a community style manner.

The concept is that experts can register as public reviewers for certain programs and then rate (for example, on a scale of 0 to 5 stars) the uploaded data for their quality. The result data that exceeds a certain average rating (e.g. 3 stars) receives a certificate and is made publicly available on the platform. In addition, a program gets a rating by points depending on the level of deviation from the reference results, so that it finally obtains a gold, silver or bronze SimQuality award after fulfilling the whole validation procedure.

The review of results might also be possible within an umbrella organisation, such as the IBPSA.

Since validation data is persistently stored, accompanying articles/papers may use the website/data as reference for discussed data.

Summary

SimQuality is the name of a test suite for validation of modern building energy performance simulation tools and a platform for publication/documentation the validation status of software tools. The test suite's concept is based on defined application scenarios, which are broken down into smaller tests of individual model components. Each of the model components is tested for a variety of parameters. For example, solar load tests are done for a selection of locations all over the world; ventilation and internal load sub-models are tested with constant, scheduled and controlled/dynamically adapted loads, etc.

However, *SimQuality* is more than just a test suite. It is also a web platform that serves as central storage and documentation place for tool validation results. Herewith validation claims can be checked transparently and vendor-independently. Simulation users can inform themselves about what features are supported sufficiently well by a given tool, or search for tools capable of doing a desired analysis.

References

ANSI/ASHRAE (2017). *Standard Method of Test for the Evaluation of Building Energy Analysis Computer Programs (ANSI/ASHRAE Standard 140)*.

Crawley, D. B., J. W. Hand, and L. K. Lawrie (1999, August). Improving the weather information available to simulation programs. In *Proceedings from Building Simulation 1999*.

Judkoff, R. (2013). Twenty years on!: Updating the iea bestest building thermal fabric test cases for ashrae standard 140: Preprint. *Building Simulation 2013*.

Schpfer, T., F. Antretter, C. van Treeck, J. Frisch, and A. Holm (2010). Validation of building energy simulation models using vdi 6020. *BauSIM 2010*.

US Department of Energy (2021, January). Energyplus weather data.

U.S. Department of Energy (2021). Energyplus weather file (epw) data dictionary.

Verein Deutscher Ingenieure (2015a). *Calculation of thermal loads and room temperatures (design cooling load and annual simulation) (VDI 2078)*.

Verein Deutscher Ingenieure (2015b). *Calculation of transient thermal responseof rooms and buildingsModelling of rooms (VDI 6007)*.

Verein Deutscher Ingenieure (2016a). *CD1 - Requirements on methods of calculation tothermal and energy simulation of buildingsand plantsBuildings [Draft] (VDI 6020)*.

Verein Deutscher Ingenieure (2016b). *Part 1 - Requirements on methods of calculation tothermal and energy simulation of buildingsand plantsBuildings [Draft] (VDI 6020 Part 1)*.

Verein Deutscher Ingenieure (2016c). *Requirements on methods of calculation tothermal and energy simulation of buildingsand plantsBuildings [Draft] (VDI 6020)*.