ON INTER-MODEL COMPARISON EXERCISES OF WHOLE BUILDING HAM SIMULATION USING THE BESTEST BUILDING

Daniel Cóstola, Bert J.E. Blocken, Jan L.M. Hensen
Building Physics and Services, Eindhoven University of Technology, The Netherlands

ABSTRACT
This paper provides a discussion of the results of two inter-model comparison exercises from IEA-Annex 41, where combined heat and moisture transfer was calculated for three variations of the BESTEST building (based on cases 600 and 900 from IEA-ECBCS SHC-Task 12 / IEA-ECBCS Annex 21). Firstly, basic statistical techniques were used to post-process some results of IEA-ECBCS Annex 41, providing confidence intervals in accordance to a certain confidence level. Secondly, the role of moisture (as well as some aspects of the exercises, such as initialization and material properties) was analysed using recently developed and validated prototype software for the coupling of BES and BEHAM. The paper concludes that combined heat and moisture transfer has a minor influence in the temperature distribution, energy demand and peak power load in the three variations of the BESTEST building simulated. However, combined heat and moisture transfer has large influence on the indoor relative humidity of the BESTEST.

INTRODUCTION
Validation of whole building heat, air and moisture (HAM) simulation is complex task. Among other techniques, inter-model comparison is recognized as an important tool in the validation process, being extensively used in the validation of building element heat, air and moisture simulation (BEHAM) programs and building energy simulation (BES) programs.

The main sources of uncertainty in computational building performance simulation (CBPS) are: input uncertainties (material properties, patterns of use, etc) and modelling uncertainties (those resulting from the assumptions and simplifications adopted to describe physical processes). Validation of CBPS programs by inter-model comparison provides the opportunity to evaluate these programs without major concerns regarding input uncertainty, therefore highlighting the magnitude of modelling uncertainties.

Relevant examples of modelling uncertainty in whole-building HAM simulation can be found in the results of two comprehensive inter-model comparison exercises, carried out in the framework of the Energy Conservation in Buildings and Community Systems (ECBCS) Programme of the International Energy Agency (IEA). In both exercises, a very simple building (Figure 1) was used in the simulations. This geometry was simulated considering two main variations, one with lightweight construction (BESTEST case 600 building) and one with heavy-weight construction (BESTEST case 900 building) (Judkoff and Neymark 1995, Woloszyn and Rode 2008). In the next paragraphs, a few results obtained in these exercises are reproduced in order to illustrate the magnitude of modelling uncertainty in whole building HAM simulation.

The original BESTEST was developed as a cooperative project of IEA-ECBCS SHC-Task 12 and IEA-ECBCS Annex 21 to diagnose problems and validate BES programs (Judkoff and Neymark 1995). The main part of this test consists of a series of inter-model comparison exercises. Figure 2 shows results of annual heating energy demand predicted by different programs, for the BESTEST case 900 building. It is important to note that in these results, only heat transfer is taken into account, i.e. there is no air or moisture transfer through building elements apart from the fixed air change rate (Judkoff and Neymark 1995). Despite this specific focus and the precise definition of parameters, differences up to 74% can be observed between minimum (1.17 MWh) and maximum (2.04 MWh = 1.17 MWh x 1.74) results obtained from different programs. These results demonstrate the potential magnitude of modelling uncertainties in BES. In spite of these uncertainties, results obtained in the BESTEST demonstrated a certain level of agreement between...
the programs. Consequently these results were later used in the development of a standard for the validation of BES programs.

Several programs dealing with energy calculations for the whole building make the assumption that effects of air and moisture transfer through building elements can be neglected. This assumption was investigated by the IEA-ECBCS Annex 41 “Whole Building Heat, Air and Moisture Response” (from now on referred in this paper as Annex 41), in which the BESTEST case 900 building, among other buildings, was simulated by CBPS programs that take into account heat and moisture transfer (air transfer is not included in most simulations in Annex 41). Some modifications were introduced in the properties of materials used in some exercises of Annex 41 in relation to the original BESTEST, therefore results are not directly comparable. Nevertheless, the magnitude of modelling uncertainty obtained in both exercises can be compared.

Figure 3 shows results of Annex 41 for the BESTEST case 900, obtained by 13 different participants (the dashed lines indicate the range of values from Figure 2). These results demonstrate that the modelling uncertainty is even higher when air and moisture are taken into account. In the original BESTEST, results were published after a second round of simulations, so that programs with outlier results could be identified and coding mistakes could be corrected. The results in Figure 3 on the other hand were “blind”, i.e. no adjustment was performed in any of the programs. These results show variations of up to 370%, and this level of deviation can be observed as well in other results obtained in Annex 41, for variations of the BESTEST building and for other performance indicators (Woloszyn and Rode 2008). This spread puts to question the applicability of inter-model comparison for the validation of whole-building HAM simulation. Moreover, this spread is particularly large if one considers that liquid load, for example wind-driven rain, was not considered in this exercise of Annex 41, which could lead to even higher discrepancies.

Facing the results of Annex 41, it is clearly necessary to improve CBPS programs in order to reduce the uncertainty of whole-building HAM simulation. A step towards this reduction is the analysis of Annex 41 results, identifying elements that led to the large discrepancy between the proposed solutions.

Regarding the analysis of Annex 41 results, this paper addresses four main points:

- The extent to which modifications introduced in Annex 41 regarding properties of dry materials (particularly for the BESTEST case 600) compromise its comparison to the original BESTEST results.
- The applicability of Annex 41 results in the future validation of whole-building HAM simulation programs.
- The impact of moisture transfer through porous materials in the energy performance of BESTEST “simplified”, and BESTEST cases 600 and 900, as well as the physical explanations for such impact.
- Some of the possible reasons for the large spread in Annex 41 results.

**METHODOLOGY**

This paper addresses the two inter-model comparison exercises from IEA-ECBCS Annex 41 dealing with the BESTEST building: BESTEST “revised” (cases 600 and 900) and BESTEST “simplified” (Woloszyn and Rode 2008). These exercises are briefly described below.

The BESTEST “revised” exercise addressed the energy demand and peak power demand of two BESTEST cases: 600 and 900. The buildings used in this exercise are very similar to the original BESTEST cases 600 and 900 (Judkoff and Neymark 1995), except for a few modifications in the thermal conductivity and thermal capacity (dry material) for the case 600 (Woloszyn and Rode 2008).

The BESTEST simplified exercise addressed the calculation of temperature and humidity in the interior of the building and in the building envelope, as well as energy demand and peak power demand.
The BESTEST simplified building adopts the same geometry of the original case 600 building, but several simplifications are adopted, as summarized in Table 1. Simplifications are mainly focused on the climate (more moderate than the original BESTEST climate), constructions (single layer, porous material), fenestration and boundary condition of the floor (which is facing the external air, not the ground, i.e. the building is “floating in the air”). Results are reported only for the 5th of July (Woloszyn and Rode 2008).

**Table 1. BESTEST “simplified” summary**

<table>
<thead>
<tr>
<th>Location</th>
<th>Copenhagen, Denmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solar heat gain coefficient</td>
<td>1</td>
</tr>
<tr>
<td>Heating set-point</td>
<td>20 °C</td>
</tr>
<tr>
<td>Cooling set-point</td>
<td>27 °C</td>
</tr>
<tr>
<td>Walls/Floor/Roof</td>
<td>Aerated concrete (15 cm)</td>
</tr>
<tr>
<td>$U_{wall/floor/roof}$</td>
<td>1.0 W/m².K</td>
</tr>
<tr>
<td>$U_{window}$</td>
<td>3.0 W/m².K</td>
</tr>
<tr>
<td>Boundary condition - floor</td>
<td>External air</td>
</tr>
</tbody>
</table>

The methodology adopted in this study is described below. It is divided in three parts addressing the points listed in the introduction section.

**Modification of dry material properties (case 600)**

Simulations were conducted using the BES program ESP-r, using both the original BESTEST properties and the ones proposed by Annex 41. The aim is to understand if the modifications play a minor role and therefore results of Annex 41 could be compared to the original BESTEST results. The change in the properties deals with the thermal conductivity and thermal capacity of dry materials.

**Applicability of Annex 41 results**

Inter-model comparisons using the BESTEST building in Annex 41 led to wide ranges of results provided by different participants. In many cases (as in Figure 3 for example) the range of results provided by the participant is so wide that it is difficult to identify acceptable limits for these results. Therefore, the use of such wide ranges as validation benchmarks is largely compromised. However, a more representative and useful range of results could be obtained if outliers were removed from the data. In the context of this paper, outlier is “an observation that is numerically distant from the rest of the data”. Outliers are not necessarily wrong, but they are results that require special attention and further investigation in order to understand the source of differences between the outlier and the rest of the data. During the IEA-ECBCS Annex 41, 13 different solutions were provided using different whole-building HAM programs for the BESTEST case 600 and 900, and 12 solutions for the BESTEST “simplified”. From these solutions, a sub set was selected in this paper to construct “narrow ranges”, which represent the solution obtained by the majority of participants in the exercise. The narrow range was constructed by removing the two highest and the two lowest results from the set of results for each performance indicator (PI), leading to much clearer limits for acceptable results. The narrow range for cases 600 and 900 is derived out of 9 solutions, and for the BESTEST “simplified” it is derived out of 8 solutions. Narrow range of results can improve the validity of the inter-model comparison because it defines two groups of results that can be clearly contrasted: the models in and out the narrow range. The definition of these two groups facilitates the comparison between models, which is core of inter-model comparison. The definition of these two groups allows developers to identify potential bugs in codes with results out of the narrow range, but also allows evaluations on the impact of assumptions and simplifications in all models. For example, in a possible scenario a outlier result obtained using a model extensively validated using empirical data may lead to revisions in all other models. This sort of analysis, however, is beyond the scope of this paper.

It is also important to point out that the original BESTEST is a full diagnosis method, designed to reduce outliers due to mistakes and capable of identifying the actual source of disagreement between different models. However, the BESTEST approach was not used in the Annex 41, where only the cases 600 and 900 were used as inspiration for a much simpler inter-model comparison exercise.

**Impact of moisture transfer in BESTEST and reasons for deviations in Annex 41 results**

The impact of moisture in the BESTEST building cases 600 and 900 and in the BESTEST “simplified” was evaluated using a recently developed prototype software coupling two state of the art BES and BEHAM programs, ESP-r and HAMFEM (Cóstola 2011).

ESP-r is a leading scientific BES program (Clarke 2001). It is the result of more than three decades of continuous development; it has a simple graphic interface and extensive quality assurance procedures. Another positive point of the ESP-r structure is its modularity, so the code related to air flow network or HVAC systems is completely separate from the core module, responsible for the building heat balance.

HAMFEM is a 3D BEHAM program developed at the K.U.Leuven using the finite element method (Janssen et al. 2007). The code was validated using a combination of analytical solutions, inter-model comparison and experimental results. As many academic codes, the program is written in Fortran 90, has no graphical interface and no quality control by the program over the correctness of input data provided by users.

The external coupling between ESP-r and HAMFEM is implemented using TCP/IP sockets as inter-process...
communication. Information is exchanged between the programs every 6 minutes of simulation time. Two-way loose coupling is adopted and different time-steps are used in each program (multi-rate). The coupling was validated using a combination of analytical solutions, inter-model comparison (which are partially described in the present paper) and experimental results (Cóstola 2011). Figure 4 schematically represents the coupling between ESP-r and HAMFEM.

RESULTS

Modification of dry material properties (case 600)

Figure 5 shows a comparison of ESP-r stand-alone (i.e. no coupling with HAMFEM) results for the original BESTEST case 600 building and for the version of this building proposed by Annex 41. Results demonstrate that the modifications significantly affect the energy demand and peak power load, even when moisture transfer is not taken into account. Figure 5 demonstrates that results from the original BESTEST case 600 cannot be directly compared to results from Annex 41.

Figure 6 shows a comparison between dry material properties used in the original BESTEST and in Annex 41, where large differences in the thermal mass can be observed. Considering the increase in thermal mass in Annex 41, the reduction of all PIs observed in Figure 5 is indeed expected, and is consistent with the results from the original BESTEST (the higher thermal mass, the lower the energy demand and peak power load).

BESTEST cases 600 and 900 - Applicability of Annex 41 results

Figure 7 shows the ranges of results for the original BESTEST (cases 600 and 900), for the Annex 41 and the narrow range calculated in this paper using 9 out of 13 solutions. The narrow range shows that the variations in the results in IEA-ECBCS Annex 41 are due to outliers and that most programs agree with each other. For the case 900, the narrow range is very similar to the original BESTEST, demonstrating that moisture plays a minor role in the results. The case 600 shows a different behaviour regarding heating and cooling. For cooling, the narrow range is slightly larger than the original BESTEST and it is shifted towards lower values, which is consistent with results from Figure 5, as well as with the original BESTEST results. For heating, the narrow range is very similar to the original BESTEST, while a shift towards lower values would be expected according to results in Figure 5. Reasons for the discrepancy between results in Figure 5 and Figure 7 demands further investigations. This investigation cannot be carried out without deeper analysis on simulations carried out in Annex 41, which is beyond the scope of this paper.

BESTEST cases 600 and 900 - Impact of moisture transfer

Figure 8 shows results obtained for the BESTEST case 600 building (as described in Annex 41) using two-way loose coupling of BES and BEHAM. Firstly, it can be observed that most results are in agreement with the narrow range of results obtained using the other 9 CBPS programs. The exception is the heating energy demand, but Figure 7 demonstrated that solutions provided in Annex 41 might overestimate this PI. Secondly, the comparison between results obtained using conjugate heat and moisture transfer (BES and BEHAM coupling) and results obtained using only heat transfer (ESP-r stand-alone) indicate that moisture plays a minor role.
in this particular building and for these particular PIs. This fact can, in part, be attributed to the low impact of moisture in the thermal conductivity and thermal mass of this particular case, as shown in Figure 9. Moreover, the magnitude of latent heat flux in this building (under this particular weather) is negligible, as shown in Figure 10. In Figure 10 (A), the relation between moisture flux and latent heat is represented (assuming moisture at 20°C). The region highlighted in the figure indicates the range of values found in the interior and exterior surface of the BESTEST case 600 building using two-way loose coupling. For the interior surface, the values are negligible, while latent fluxes up to 12 W/m² are found at the exterior surface. The fluxes at the external surface are not high when compared to solar radiation, and Figure 10 (B) indicates that most of the time the latent heat flux at the external surface is in fact much lower than 12 W/m².

Although moisture accumulation and fluxes do not play a major role in energy demand and peak power demand, the initial moisture content in building elements (walls, roof, floor) significantly affects the results of BESTEST case 600. This effect can be observed in Figure 11, where the uncertainty due to moisture initialization is shown in comparison with the range of solutions described in Annex 41. The bold line shows results obtained in simulations using different initial moisture contents, from dry to saturate. Firstly, the impact of initialization is noticeable, particularly on energy demand. Secondly, it is possible to explain most of the variations in the different solutions solely by using this parameter. The importance of moisture initialization in the energy consumption was already demonstrated in previous research (e.g. Kong and Zheng, 2008). This importance can be explained by the variation in thermal conductivity/capacity and by the large moisture fluxes in saturated or in dry materials, which is particularly relevant in the first year.
Figure 8. Results for the BESTEST case 600 building in IEA-ECBCS Annex 41 using two-way loose coupling.

Figure 9. Variation in material properties due to moisture

Figure 10. Relation between latent heat and moisture flux (A) and the frequency of occurrence of different latent heat flux values (B).
BESTEST “simplified” - Impact of moisture transfer and applicability of Annex 41 results

Figure 12 shows results for the interior air temperature and exterior surface temperature of the roof, for the BESTEST “simplified” building on the 5th of July. Results of Annex 41 are divided in the full range and in the narrow range (8 out 12 solutions). These PIs are not strongly affected by moisture transfer through the building envelope, and consequently both ESP-r stand-alone and the two-way loose coupling of BES and BEHAM provide similar results. The results obtained in this paper agree with the narrow range of results from Annex 41. It is also noticeable that the narrow range of results from Annex 41 indicates good agreement between the 8 solutions for most of the PIs; hence the validity of this inter-model comparison increases significantly.

Figure 13 shows results of power demand for heating and cooling. As in Figure 12, results are not significantly influenced by moisture, i.e. ESP-r stand-alone results are very similar to the results of two-way loose coupling of BES and BEHAM. Both results show good agreement within the narrow range. It should be noticed that the agreement between solutions within the narrow range is much better for heating energy demand than for cooling.
CONCLUSIONS
Results provided in this paper allow the following conclusions to be made:

- The apparent large dispersion of results found in some inter-model comparisons carried out during the IEA-ECBCS Annex 41, such as the “simplified” BESTEST building and the BESTEST cases 600 and 900, can be mainly attributed to outlier results. Discarding these outliers demonstrated that most solutions are in reasonable agreement.
- Outlier results from IEA-ECBCS Annex 41 should be carefully analysed, in order to identify the sources of disagreement.
- The application of Annex 41 inter-model comparison exercises in the future validation of whole-building HAM simulation programs is possible using the narrow ranges provided in this paper.
- In the cases evaluated, moisture transfer does not play a major role in the following PIs: energy demand, peak power demand, interior air temperature, temperatures within building elements.
- In the cases evaluated, moisture transfer plays a major role in the calculation of interior air RH.
- Initialization of moisture content of walls plays a major role in the BESTEST case 600 inter-model comparison carried out during IEA-ECBCS Annex 41. Different initializations can, to a large extent, explain the variation in the solutions proposed for this exercise. This result is in agreement with previous studies.

- Results of Annex 41 and the original BESTEST cannot be compared due to relevant differences in the settings of these exercises.

ACKNOWLEDGEMENT
This research has been funded by the “Institute for the Promotion of Innovation by Science and Technology in Flanders” as part of the SBO-project IWT 050154 “Heat, Air and Moisture Performance Engineering: a whole building approach”.

REFERENCES