A state-of-the-art empirical round robin validation of heat, air and moisture (HAM) models

Dang, Xinyuan; Guimarães, Ana Sofia; Sarkany, Andreas; Laukkarinen, Anssi; Xu, Bingyu; Vanderschelden, Bruno; Rode, Carsten; Tenpierik, M.J.; Huijbregts, Z.; More Authors

DOI
10.1016/j.buildenv.2025.112867

Publication date
2025

Document Version
Final published version

Published in
Building and Environment

Citation (APA)
Dang, X., Guimarães, A. S., Sarkany, A., Laukkarinen, A., Xu, B., Vanderschelden, B., Rode, C., Tenpierik, M. J., Huijbregts, Z., & More Authors (2025). A state-of-the-art empirical round robin validation of heat, air and moisture (HAM) models. *Building and Environment*, *276*, Article 112867. https://doi.org/10.1016/j.buildenv.2025.112867

Building and Environment 276 (2025) 112867

Contents lists available at ScienceDirect

Building and Environment

journal homepage: www.elsevier.com/locate/buildenv

A state-of-the-art empirical round robin validation of heat, air and moisture (HAM) models

Xinyuan Dang [a,1,*], Ana Sofia Guimarães [ae,2], Andreas Sarkany [x,2], Anssi Laukkarinen [u,2], Bingyu Xu [y,ab,2], Bruno Vanderschelden [k,2], Carsten Rode [v,2], Changchang Xia [s,2], Chi Feng [f,2], Chris Whitman [d,2], Daan Deckers [a,2], Hanyu Yang [e,2], Heiko Fechner [w,2], Himanshu Sharma [g,2], Hua Ge [g,2], Huarong Xie [s,2], Huibo Zhang [q,2], Isabeau Vandemeulebroucke [k,2], Jakub Mazura [h,2], Jingchao Xie [b,2], Jiri Madera [i,2], John Grunewald [w,2], Juha Vinha [u,2], Kaat Janssens [k,2], Kazuma Fukui [m,2], Kostas Mourkos [n,2], Lin Wang [p,2], Marjorie Bart [aa,2], Martin Morelli [c,2], Martin Tenpierik [j,2], Masaru Abuku [l,2], Maurice Defo [p,2], Michele Bianchi Janetti [ac,2], Paul Klõšeiko [t,2], Pauli Sekki [u,2], Pavel Kopecký [h,2], Peter Matiasovsky [r,2], Riccardo Paolini [ad,2], Richard Slávik [o,r,2], Robert Cerny [i,2], Sina Akhavan Shams [g,2], Sophie Moissette [aa,2], Targo Kalamees [t,2], Teresa Stingl Freitas [ae,2], Thibaut Colinart [z,2], Thomas Bednar [x,2], Tommy Bunch-Nielsen [c,2], Vaclav Koci [i,2], Valentina Marincioni [y,ab,2], Villu Kukk [a,t,2], Weinan Gan [f,2], Xiaolin Chen [k,2], Xing Hu [q,2], Yohan Plantec [z,2], Yue Xie [b,2], Zara Huijbregts [j,2], Hans Janssen [a,1], Staf Roels [a,1]

[a] KU Leuven, Department of Civil Engineering, Leuven 3001, Belgium
[b] Beijing University of Technology, School of Architectural Engineering, Beijing 100124, China
[c] Bunch Bygningsfysik ApS, Vedbæk 2950, Denmark
[d] Cardiff University, Welsh School of Architecture, Cardiff CF103NB, UK
[e] China Academy of Building Research, State Key Laboratory of Building Safety and Built Environment, Beijing 100013, China
[f] Chongqing University, School of Architecture and Urban Planning, Chongqing 400045, China
[g] Concordia University, Department of Building, Civil and Environmental Engineering, Montreal H3M1G8, Canada
[h] Czech Technical University in Prague, Department of Architectural Engineering, Prague 16629, Czech Republic
[i] Czech Technical University in Prague, Department of Materials Engineering and Chemistry, Prague 16629, Czech Republic
[j] Delft University of Technology, Department of Architectural Engineering and Technology, Delft 2628BL, The Netherlands
[k] Ghent University, Department of Architecture and Urban Planning, Gent 9000, Belgium
[l] Kindai University, Faculty of Architecture, Osaka 577-8502, Japan
[m] Kobe University, Graduate School of Engineering, Kobe 657-8501, Japan
[n] Loughborough University, School of Architecture, Building and Civil Engineering, Loughborough LE113TU, UK
[o] Mendel University in Brno, Faculty of Forestry and Wood Technology, Brno 61300, Czech Republic
[p] National Research Council Canada, Construction Research Centre, Ottawa K1A0R6, Canada
[q] Shanghai Jiao Tong University, School of Design, Shanghai 200240, China
[r] Slovak Academy of Sciences, Institute of Construction and Architecture, Bratislava 84503, Slovakia
[s] Southeast University, School of Architecture, Nanjing 210096, China
[t] Tallinn University of Technology, Department of Civil Engineering & Architecture, Tallinn 19086, Estonia
[u] Tampere University, Faculty of Built Environment, Tampere 33720, Finland
[v] Technical University of Denmark, Department of Environmental and Resource Engineering, Kongens Lyngby 2800, Denmark
[w] TU Dresden, Institute of Building Climatology, Dresden 01069, Germany
[x] TU Wien, Faculty of Civil and Environmental Engineering, Vienna 1040, Austria
[y] UK Centre for Moisture in Buildings, London E152GW, UK
[z] Université Bretagne Sud, Institut de Recherche Dupuy de Lôme, Lorient 56100, France
[aa] Université de Rennes, IUT de Rennes - Département Génie Civil Construction Durable, Rennes 35704, France
[ab] University College London, Institute for Environmental Design and Engineering, London WC1H0NN, UK
[ac] University of Innsbruck, Department of Structural Engineering and Material Sciences, Innsbruck 6020, Austria

* Corresponding author at: BOKU 04 08, Kasteelpark Arenberg 40, 3001 Leuven, Belgium.
E-mail addresses: xinyuan.dang@kuleuven.be, dangxinyuan1994@gmail.com (X. Dang).

https://doi.org/10.1016/j.buildenv.2025.112867
Received 6 December 2024; Received in revised form 11 February 2025; Accepted 14 March 2025
Available online 15 March 2025

[ad] University of New South Wales, School of Built Environment, Sydney NSW2052, Australia
[ae] University of Porto, Faculty of Engineering, Porto 4200-465, Portugal

ARTICLE INFO

Keywords:
Validation
HAM model
Heat, air, and moisture
Hygrothermal response
Simulation
Round robin

ABSTRACT

Heat, air and moisture (HAM) models allow efficient simulation of the building components' hygrothermal behavior. However, specific model assumptions, simplifications and approximations, as well as users' preferences, biases and mistakes in the implementation of material properties, boundary conditions, etc., may yield divergences among results from different models. The lack of a standard framework for HAM model quality assessment results in inconsistent benchmark cases and assessment methods in previous studies. Thus, this state-of-the-art empirical round robin validation targets to test the robustness and the reliability of HAM models in predicting one-dimensional hygrothermal responses of building components under controlled boundary conditions. It ran from 2023 to 2024, was coordinated by KU Leuven, and achieved participation of 38 groups from 19 countries. A comprehensive experimental dataset serves as the "correct answer", and simulation results from other participants form "reference answers". Since the boundary conditions are simple and explicit, the material properties' implementation has the main impact on the simulated hygrothermal responses. Most models prove to be robust, particularly in the heat transfer prediction. The moisture transfer prediction, on the other hand, looks more challenging. Reliability is also achieved by most models, as the deviations between simulation and experimental results are reduced when actual measured material properties are implemented as inputs. However, inappropriate and/or incorrect implementations are also observed. More in-depth investigations are performed for a better understanding of HAM-simulation tools and achieving their better performance in predicting and interpreting the hygrothermal behavior of building components.

1. Introduction

Heat, air, and moisture (HAM) transfer is essential in the performance and degradation [1], the energy efficiency [2] and the climate resilience [3] of building materials and components. In addition to laboratory tests and field measurements, numerical HAM modeling role is becoming progressively more crucial when predicting and controlling the moisture accumulation in building envelopes [4]. While HAM models have strongly progressed [5–10] by integrating complex workflows for hygrothermal transport [11–14], questions remain regarding their robustness and reliability due to inherent uncertainties introduced by both model developers and model users [15,16]. Inaccuracies and incorrectness are sometimes observed that deserve more investigations. Nevertheless, a comprehensive quality assessment of HAM models is currently lacking. Or particularly, in previous studies, no standardized criteria have been developed to tackle the question "whether the model can correctly predict the reality". These studies vary in purposes, protocols and benchmarks, and some formulations (e.g., verification, validation, inter-model comparison, …) are even used interchangeably and incorrectly [17].

Assuming the models are correctly coded by the claimed algorithm equations as verified by the developers, there are two main approaches to assess the models' capability of reproducing the physical phenomena, namely, comparing simulations with experimental datasets or other models' predictions. Generally, experimental datasets are recognized as the "correct answer" in order to quantify the deviations between simulation and measurement results [18–35]. However, in addition to measurement errors and equipment constraints that may restrict their dependability and effectiveness, the inherent limitations of current available experimental datasets constitute a main issue. The primary limitations relate to the small scale of the employed configurations [24], which often target "sample" scale rather than "component" or "building" scales [36], and/or the monotony of the applied materials, for which single materials or materials with similar properties (particularly in moisture transport and storage) are used [37–39]. A subsequent limitation is that the benchmark experiments are often applied on simple hygrothermal transport processes or even single material characterization tests [40], instead of long-term comprehensive hygrothermal transfer processes [26]. A last (but not least) limitation is that most datasets of inputs and outputs are not always fully established (e.g., the assignment of default/reference values from standards or other sources to material properties and boundary conditions) and/or not completely documented, impeding the applicability of the datasets. All these deficiencies restrict the representativeness of these datasets and therefore hamper a further assessment of the capabilities and limitations of HAM models. Alternatively, inter-model comparison can be accepted as the "reference answer", and the consistency of simulation results from different models on the same scenario can infer reliability. For instance, the most recent extensive inter-model comparison collaboration, EU HAMSTAD [41], evaluates component-scale HAM simulation via five benchmark cases [42] without experimental evidence. This "peer review" approach has been generally employed, based on standard benchmarks, e.g., EU HAMSTAD [41] and EN 15026 [43], or based on specific cases [44–52]. However, similar trends in outcomes from different models may not always indicate correctness, as they might obscure common limitations and inaccuracies.

In conclusion therefore, most available experimental datasets may not accurately reflect real-world complexity, limiting the representativeness of the benchmarks, while pure inter-model comparison based on hypothetical scenarios similarly limits the effectiveness of assessment process. In response, this study uses a dedicated benchmark dataset based on thorough experimental results for a stepwise quality assessment of HAM models. The experiments were performed and documented at the Section of Building Physics and Sustainable Design of KU Leuven, Belgium [53,54]. The experimental benchmark dataset includes (I) measured hygrothermal responses of four different wall assemblies to controlled boundary conditions governed by a hot box-cold box (HB-CB) climate chamber, (II) full characterization results of hygrothermal material properties, and (III) measured surface transport coefficients. Along with the data completeness, this dataset is considered more representative than previous in-lab experimental datasets for component-scale hygrothermal validation, owing to the range of materials with different moisture transport patterns that are combined in multiple configurations. This dataset is employed in a round robin framework, by inviting HAM models to reproduce the benchmark experiment, yielding a state-of-the-art collaborative action entitled "Empirical validation of HAM-models based on a dedicated HB-CB experiment" [55] from 2023 to 2024. Researchers and practitioners worldwide simulated the

[1] Project coordinators.
[2] Project participants.

hygrothermal responses of the benchmark experiment via multiple HAM models (including personal codes). As a continuation of prior endeavors [41,42], this unprecedently extensive cooperation (over 70 contributors in 38 groups from 19 countries have participated) also adds significant value and credibility to bring about more general and stronger confidence in hygrothermal simulations.

This summative paper on the HAM model quality assessment consists of three prime parts. Firstly, the methodology for the round robin approach is brought. In the subsequent "overall assessment" part, the simulated hygrothermal responses from various groups are confronted with the measured datasets. In the concluding "impact analysis" part, the impact of implemented material properties (i.e., deviations between the actual and input values), as the key uncertainty in this case, on the simulated hygrothermal responses is evaluated. The approach, integrating experimental validation with inter-model comparison as well as combining result consistency and causal inference evaluation, seeks for more accurate simulation of complex hygrothermal processes in building envelopes.

2. Methodology

This initial section successively presents the experimental datasets that form the foundation of the round robin HAM quality assessment, the research scheme which has been adopted to stepwise execute the quality assessment, the participation overview on contributors and submissions to the round robin, and finally the performance indicators employed for the quality assessment.

2.1. Experimental datasets

The experimental datasets applied in this HAM model quality assessment stem from a dedicated benchmark experiment, wherein four different wall assemblies (WA) are subjected to a cold moist and a warm dry climate – provided by a cold box (CB) and hot box (HB) respectively – over an interval of two months [53]. During this experiment, temperatures and relative humidities as well as heat fluxes and moisture masses are regularly measured.

The configurations of the wall assemblies and the scheme of boundary conditions are shown in Figs. 1 and 2. The wall assemblies comprise three typical insulation materials with distinct moisture transfer properties [56], calcium silicate (CS, vapor-open & capillary-active), mineral wool (MW, vapor-open & non-capillary-active), wood fiber (WF, vapor-open), in different orders. An aluminum-foil-based vapor barrier (VB) is attached at the left surface in WA3 and at the WF/CS interface in WA4, to activate different moisture redistribution patterns. The separate material layers are each glued to an extruded polystyrene (XPS) insulation frame by polyurethane foam, ensuring that hygrothermal transport is predominantly one-dimensional. The boundary conditions are set at 2 °C & 80 % RH (in the CB, at the left side of the WA) and 28 °C & 54 % RH (in the HB, at the right side of the WA) respectively. The cold and hot boxes, and hence the boundary conditions for the test wall assemblies, are switched in the 748th hour, in the middle of the 60-day (1440 h) experiment, to induce different hygrothermal transfer processes. Temperature (T, °C), relative humidity (RH, %) and heat flux (HF, W/m^2) are continuously measured by sensors at interfaces or in the ambient air (whose positions are numbered in Fig. 1b), while distinct material layers are regularly detached and weighed to determine their moisture masses (MM, kg/m^2). The surface transport coefficients and hygrothermal material properties (for the same batch) are also measured. The experimental setup and datasets are thoroughly described and presented in [53,54].

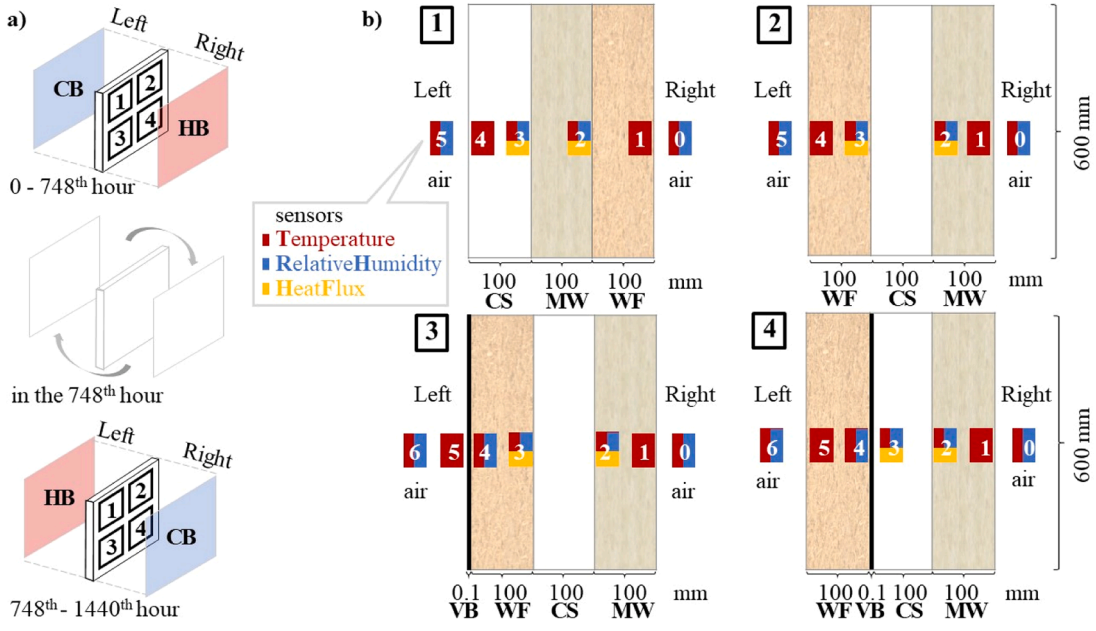

Fig. 1. Experiment design: a) scheme of the boundary conditions, and b) configurations of four wall assemblies with locations of sensors.

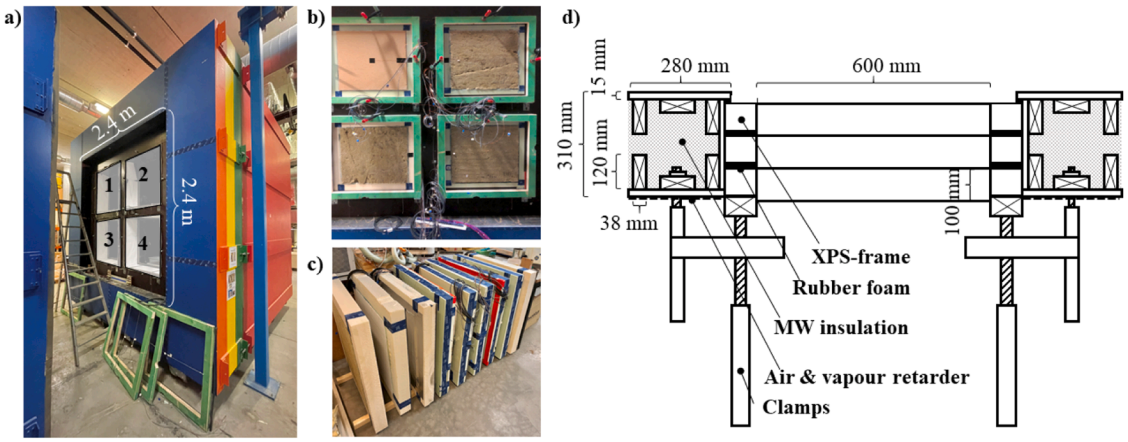

Fig. 2. Experiment setup: a) overview, b) front view, c) XPS-framed material samples of the wall assemblies, and d) schematic diagram (horizontal section).

Table 1
Data series provided (as inputs) and collected (as outputs) in the three stages.

Data series		Stage 1	Stage 2	Stage 3
Provided by the coordinator	Initial conditions	√	√	√
	Boundary conditions	√	√	√
	Surface transport coefficients	O	√	√
	Material characteristics	O	√	√
	Hygrothermal responses	×	×	√
Collected from the participants	Surface transport coefficients	O	O	√
	Material characteristics	√	√	√
	Hygrothermal responses	√	√	√

2.2. Research scheme

While HAM models often are largely equivalent with respect to the physics and mathematics [57,58], intrinsic limitations and user impacts collectively contribute to potential uncertainties. These intrinsic limitations stem from specific assumptions, simplifications, and approximations when implementing material properties [41,52], boundary conditions [4,59,60], and other dependent factors [61]. Assumptions may include material homogenization, ignorance of local defects, idealization of construction workmanship [62], etc. Simplifications may involve disregard of sorption hysteresis [63], fixed values for wind-driven rain and surface transport coefficient at the same height or for the same facade [60], simplified formula for estimating moisture diffusivity [64], etc. Approximations may comprise setting fixed values, linear interpolation, etc. The user impacts refer to the preferences, biases and/or mistakes during the selection of predefined parameters or data series and the customization of the implementation of the dependent factors [16]. Particularly for commercial software, users tend to choose materials in the models' databases with properties similar to the actual materials, but the target parameter in similarity may vary significantly (e.g., density, thermal conductivity, saturation moisture content, manufacturers/batches, etc.). They may also refer to different standards/references (or even arbitrary values) when assigning material properties, surface transport coefficients, solar radiation absorption/-reflective rates, etc. If not supported by sufficient expertise and double-checks, users are likely to create random errors unconsciously.

Given that it is not feasible to require that one benchmark scenario perfectly evaluates all possible uncertainties [15], this study primarily focuses on the implementation of material properties, and controls the other variables (e.g., uniform samples with negligible boundary effects, simple and quasi-steady-state boundary conditions, measured surface transport coefficients, etc.) as well as possible. Even then though, the allocation of noted deviations is still confounded because both intrinsic limitations and user impacts may affect the implementation of material properties in both commercial software and personal codes. For a more explicit analysis thus, a stepwise protocol is developed to test the robustness and reliability of the HAM models. A **robust simulation** ensures that the overall trends and patterns of the calculated hygrothermal responses remain consistent with reality, despite variations in implemented material properties (primary factor in this study), boundary conditions, etc. The reasons for variations include the selection of "similar" entries (e.g., with similar material properties to the claimed values by the producers) in the built-in databases of HAM models, the different processing methods for untested/indirectly-tested properties, the mismatch risks between observations and the reality due to human factors and technical constraints, etc. On the other hand, a **reliable simulation** refers to the dependability of delivering the expected result under given conditions [65]. When measured inputs (as many as possible) are "correctly" implemented, reliable outputs should exhibit a good (and expectedly better than the robustness phase) agreement and coherence with the measured datasets.

The first stage ("robustness test") is based on an "engineering perspective". Instead of providing material characterization results, only basic parameters from technical sheets are given (the claimed values of density ρ [kg/m^3], thermal conductivity λ [W/(m·K)], specific heat capacity c [J/(kg·K)], vapor diffusion resistance factor μ [-] and capillary absorption coefficient A_w [kg/(m^2·s$^{0.5}$)]). The "engineers" can either directly select material entries from the built-in databases of HAM models or generate their own material files. The estimated initial conditions (20 °C and 50 % RH: the lab room conditions at which the samples have been stored for months before the experiment) and measured boundary conditions are provided. The description for surface transport is limited to "*the HB, whose size is much larger than the cold box, simulates an indoor environment with just a small ventilator to mix the air, while the CB simulates the outdoor environment with four large ventilators creating a uniform air flow along the building component*", and therefore the assignment can be based on standards or experience. The second stage ("reliability test") takes a more academic view. In addition to the initial conditions and boundary conditions provided in the first stage, the rigorously measured material properties and surface transport

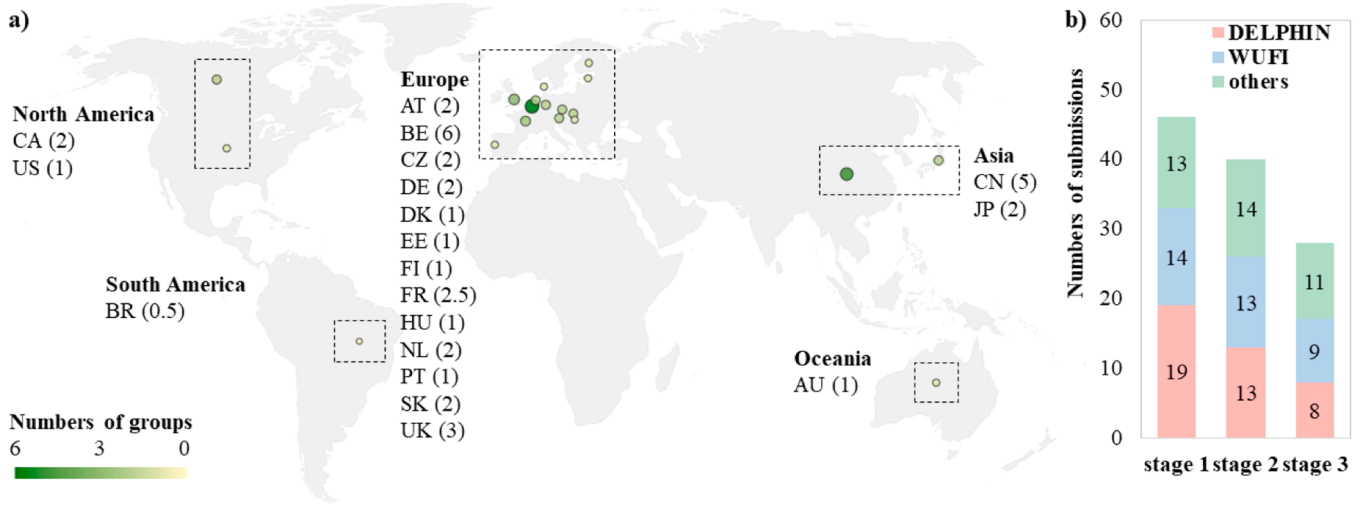

Fig. 3. Participation overview: a) numbers of the groups and b) numbers of submissions and employed models.

coefficients [53,54] are provided for translation into the forms required by different models. Although uncertainties remain in the translation, deviations are expected to be reduced due to the newly provided measured datasets. The third stage finally is an extra chance for participants who intend to fine-tune the implementations of material properties (mainly), algorithms (if applicable, particularly within personal codes) and other factors. The information and data provided in the previous stages are carried forward. The measured hygrothermal responses as well as simulation results from other groups are supplemented as benchmarks and references. It also allows the participants to diagnose and/or tackle difficulties in HAM modelling through trial and error, and in analyzing the sensitivity of different factors. Table 1 shows the data series provided (as inputs) to and required (as outputs) from the contributors over the three stages. The ticks (√), circles (O), and crossings (×) indicate "fully provided/required for submission", "limited information provided/optional for submission", and "not given", respectively.

Simulated hygrothermal responses, including temperature (T), relative humidity (RH) and heat flux (HF) at interfaces and moisture mass (MM) of each material layer (see Fig. 1b), are collected in time series for comparison.

2.3. Participation overview

This project received no direct funding, so all participants contributed voluntarily. The registration started in February 2023, initially attracting 55 groups from 22 countries. As Fig. 3a) shows though, the actual participation (in at least one stage of the tasks) involves 38 groups from 19 countries (for cross-national groups, the number only counts 1/N for each country, N = number of the countries): Austria (AT), Australia (AU), Belgium (BE), Brazil (BR), Canada (CA), China (CN), Czech Republic (CZ), Germany (DE), Denmark (DK), Estonia (EE), Finland (FI), France (FR), Hungary (HU), Japan (JP), The Netherlands (NL), Portugal (PT), Slovakia (SK), the United Kingdom (UK), and the United States (US). Groups may submit multiple results (from different models, or based on different trials, or by different persons). Each submission result is denoted by the group number (as assigned after the registration) and a letter (to distinguish different results from a single group). The entire process remains anonymous, meaning that one cannot link the numbers to the real participants. KU Leuven (BE), as the institutional coordinator, is responsible for task distribution, result collection and data analysis. Fig. 3b additionally illustrates the proportion of the employed models in the three stages of this project. Two commercial models, WUFI and DELPHIN, are the primary choices. Other models, including personal codes, have been employed in one third of the submissions.

2.4. Performance indicators

In addition to plotting the results in line charts, two statistical indices, the normalized root-mean-square-error (NRMSE) and Pearson correlation coefficient (r), are employed to quantify and visualize the deviations and correlations between the simulated and measured datasets as heat maps.

$$NRMSE = \frac{1}{m_{max} - m_{min}} \cdot \sqrt{\frac{\sum_{i=1}^{n} (m_i - s_i)^2}{n}} \tag{1}$$

$$r = \frac{\sum_{i=1}^{n} [(m_i - \overline{m}) \cdot (s_i - \overline{s})]}{\sqrt{\sum_{i=1}^{n} (m_i - \overline{m})^2} \cdot \sqrt{\sum_{i=1}^{n} (s_i - \overline{s})^2}} \tag{2}$$

where, i and n are the ordinal number and amount of the entries within one data series; m and s are the measured and simulated data; max and min are the maximum and minimum.

The root-mean-square-error (RMSE) describes the average absolute error of the differences between simulated and measured values, but this value is sensitive to the variation amplitudes. Therefore, its adapted form, normalized root-mean-square-error (see Eq. (1)) is employed to avoid misinterpretation and to ensure the same order of magnitude [66]

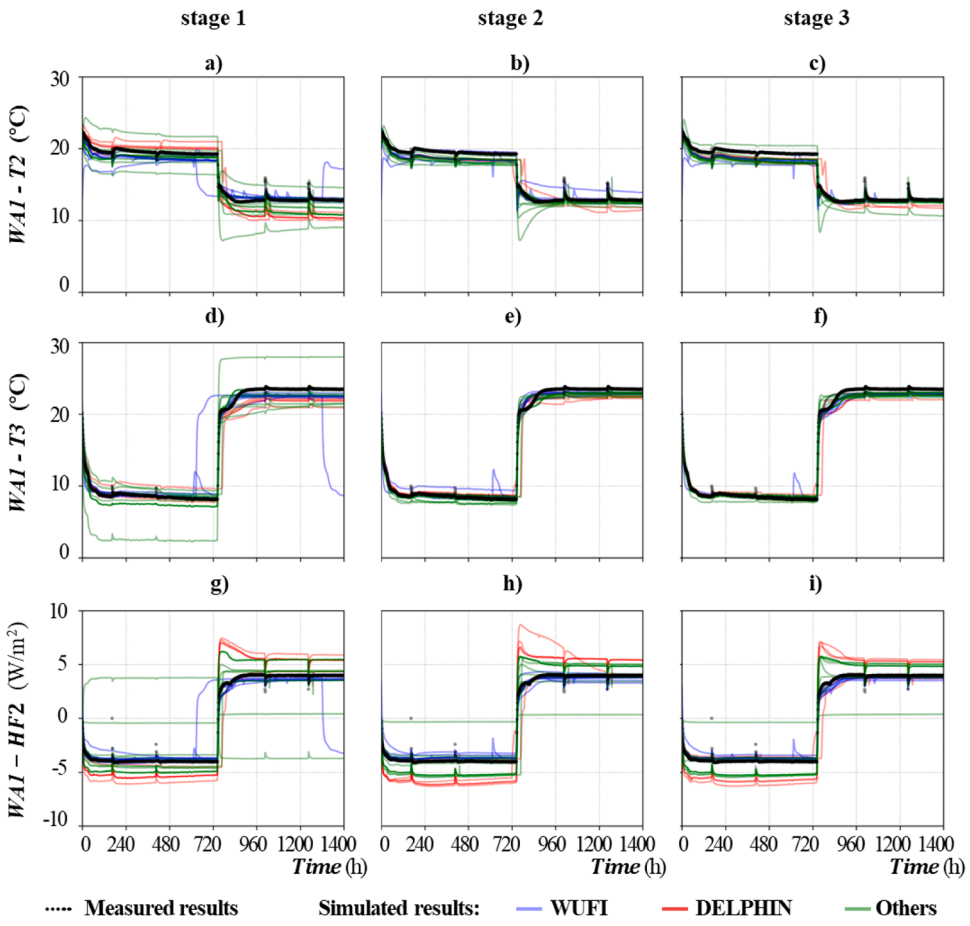

Fig. 4. Simulated heat-related responses of WA1.

across the variables. There is no standardized value for the NRMSE, but suggested thresholds of RMSE are 1 or 2 (for temperature) and 2 or 5 (for relative humidity) for "high" or "low" accuracy, respectively [67]. The Pearson correlation coefficient (see Eq. (2)) indicates the strength of a linear relationship between two data series, ranging from −1 (opposite trend) to +1 (similar trend). A near-zero value indicates no relation between both. There is no fixed threshold, but suggested values for building behavior are reported in literature, for instance, $r > 0.5$ [68] or its square $r^2 > 0.75$ [69]. As both indices require the same quantities in the paired data series, only non-null values are included for their calculation (since some measured data are missing, for instance, due to logger errors).

3. Results and discussion

The analysis is result-oriented and stepwise. Firstly, a global assessment is conducted based on the normalized root-mean-square-error (NRMSE) and the Pearson correlation coefficient (r) between the measured and simulated hygrothermal responses from all models in the three stages. The comparisons between the measured and simulated results of two wall assemblies are plotted in line charts to highlight deviations across variables, models, and stages. Secondly, the diversity in the implemented material properties in stages 1 and 2 is visualized, and its impact on the deviations in simulated hygrothermal responses is quantified. Finally, adjustments with valuable insights in stage 3 are summarized.

3.1. Overall assessment

Line charts in Figs. 4–7 confront the simulated hygrothermal responses of WA1 and WA4 with the measured datasets. These two wall assemblies are selected because most participants set the minimal deviations in WA1 as the optimization condition for fine-tuning in stage 3 while WA4 is representative for the other three wall assemblies, given its similar material order and the vapor barrier that makes the moisture transport more complex. To highlight the impact of the adjusted implementations, results from those groups who did not participate in all three stages are temporarily removed. Heat maps in Figs. 8 and 9 depict the normalized root mean square error (NRMSE) and Pearson correlation coefficient (r) between simulated and measured

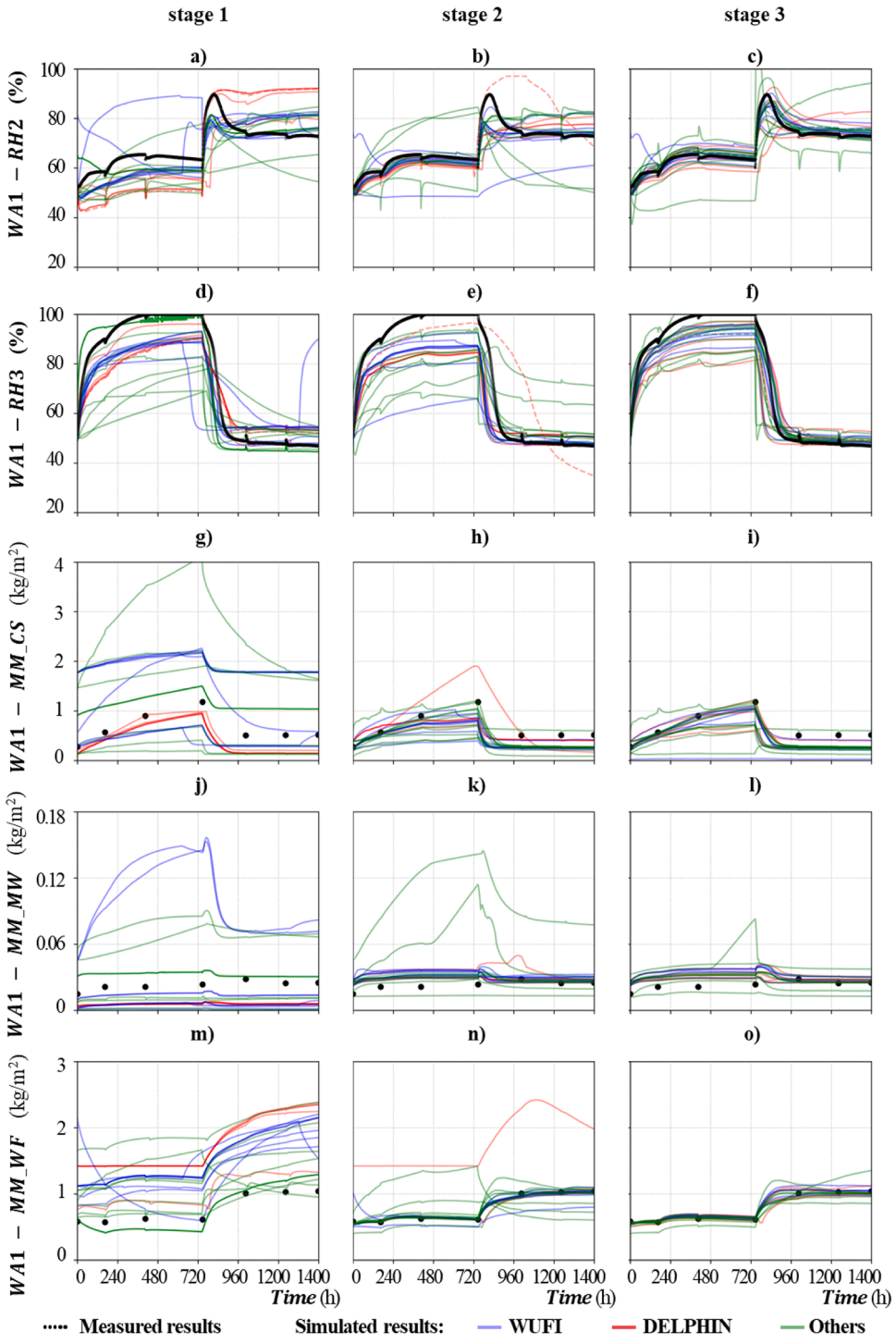

Fig. 5. Simulated moisture-related responses of WA1.

hygrothermal responses of the four wall assemblies in each of the three assessment stages. In each subgraph, each vertical grid corresponds to a different submission, which are grouped into three categories according to the employed models (DELPHIN, WUFI, and others). A star ("*") is assigned to empty boxes where the submitted data are missing, while an additional hyphen ("-") is marked in boxes with negative r values. As explained previously, a small NRMSE and a large r value ("bluer" boxes) indicates a good agreement between measurement and simulation. Note

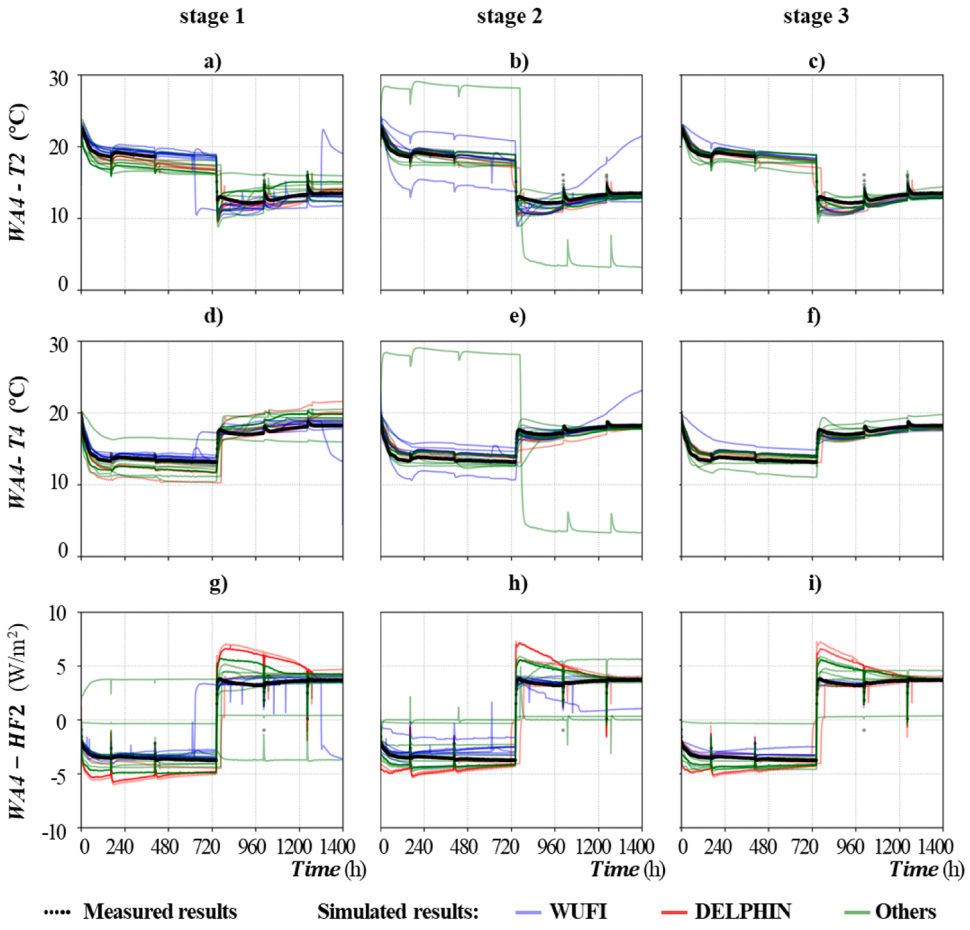

Fig. 6. Simulated heat-related responses of WA4.

that, three groups (using WUFI) only submitted the results of WA1 in stage 3, purely focusing on enhancing the predictive accuracy for this specific wall assembly through adjustments in their implementations.

There is a clear trend that the agreement between measurement and simulation increases from stage 1 over stage 2 to stage 3, observable from a horizontal view in Figs. 4–9. In stage 1, the flexibility to generate or select material properties and surface transport coefficients yields disparity in the NRMSE of the hygrothermal predictions (Fig. 8, left column) [16]. Even so, most outputs are in good correlation with the measured datasets, except for r of the moisture masses (MM) of CS and MW (see Fig. 9, left column). In stages 2 and 3, implementing the measured material properties as inputs effectively reduces the divergences. This optimization effect is much more obvious in the heat-related variables (T and HF) than in the moisture-related variables (RH and MM). In stage 3, nearly all heat-related variables satisfy the criteria NRMSE $<$ 0.1 and r $>$ 0.5. On the other hand, the moisture-related variables, particularly for MMs of MW, remain roughly uncorrelated, although RH predictions are improved. Complementarily, it seems that WUFI results are slightly better in line with the measured data than the outcomes from DELPHIN and other models.

While all in all good agreement and a considerable improvement over stages are observed for heat-related responses, the simulated heat fluxes from three model categories obviously deviate (Figs. 4g-i and 6g-i). Particularly around the 748th hour (when the boundary conditions are switched), some heat fluxes from DELPHIN (in red) experience a sharp increase before dropping, whereas most results from WUFI (in blue) remain closer to the measured results and have a smoother increase. Those from other models (in green) are mostly intermediate. To diagnose this disparity, Fig. 10 presents detailed outputs of the temperature profiles and heat fluxes in WA1, as simulated by the coordinator with WUFI and DELPHIN. There are roughly no differences between the temperature profiles in the two models when the boundary conditions are switched. This infers that the conduction heat flows simulated by WUFI and DELPHIN must be similar, and that these thus are not the origin of the noted disparity. In DELPHIN, the heat flux can be output as total flux and or as any of its components, including the conduction flux and the enthalpy fluxes associated with liquid convection, vapor diffusion and vapor convection [70]. In WUFI, on the other hand, only an unspecified “heat flux” can be output. The deviation between DELPHIN’s outputs and the measured results hence stems from

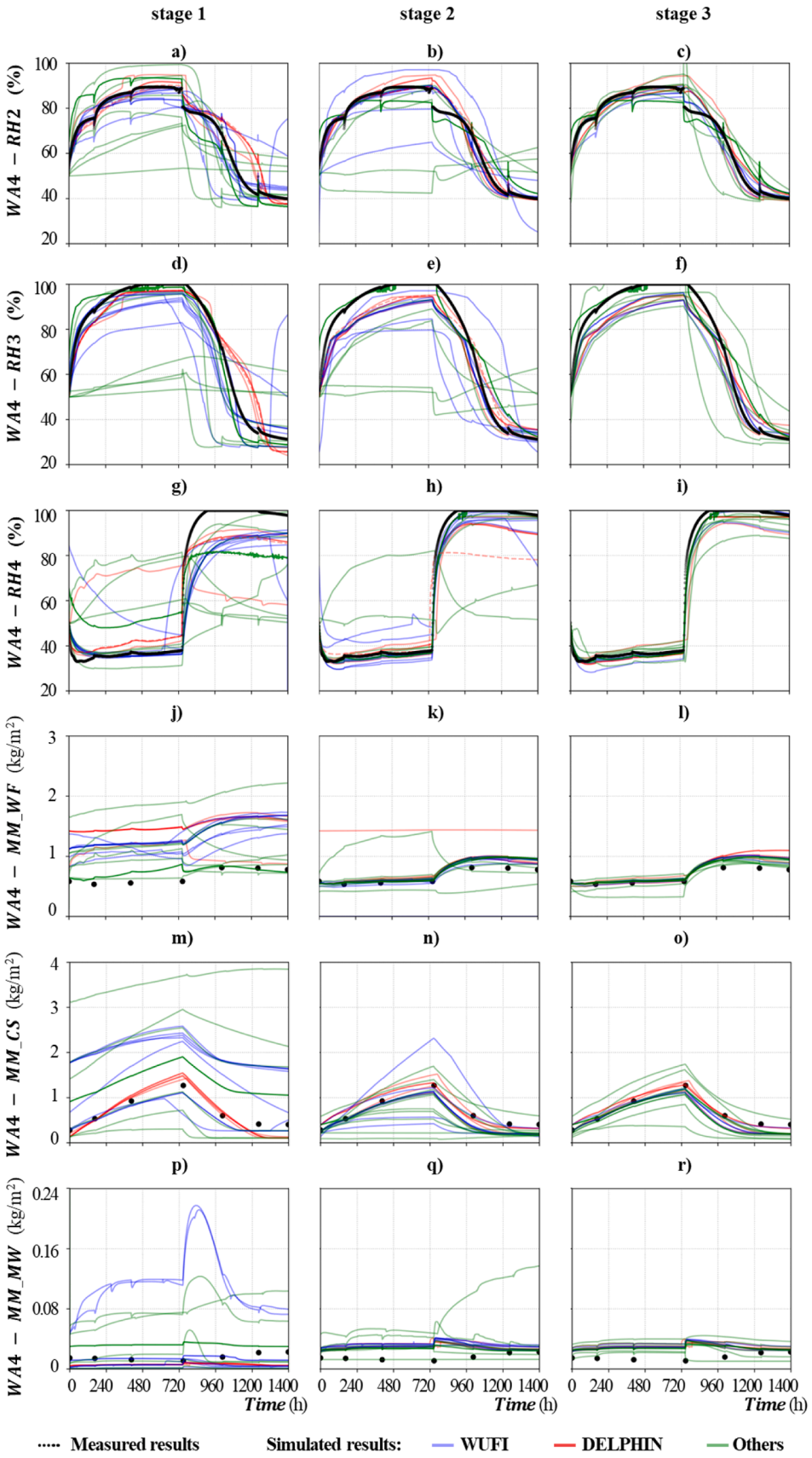

Fig. 7. Simulated moisture-related responses of WA4.

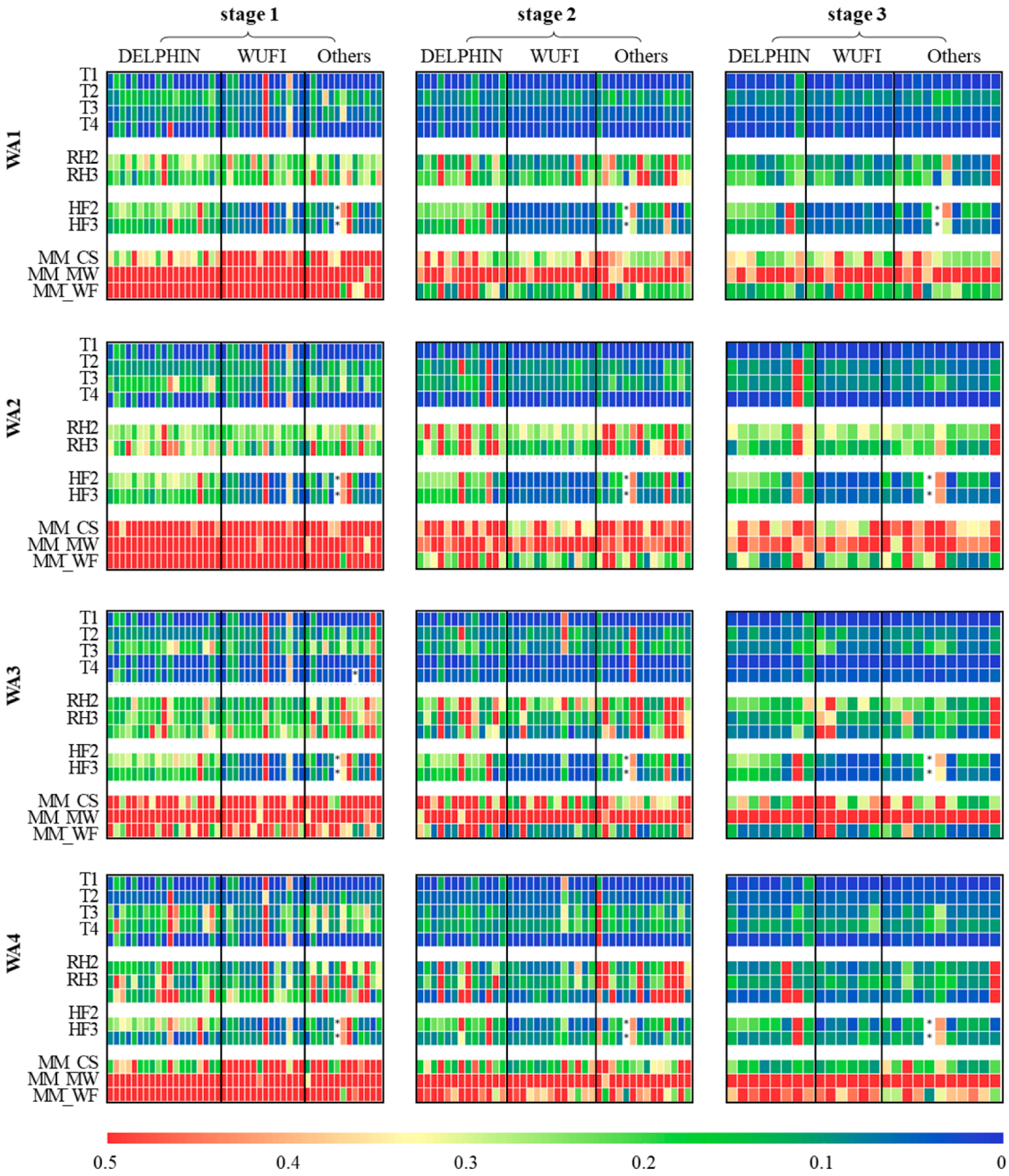

Fig. 8. Normalized root-mean-square-error (NRMSE) between the simulated and measured hygrothermal responses.

including enthalpy heat flux due to vapor diffusion in the output, because DELPHIN's conduction heat flux as well as WUFI's heat flux approximate to the measured heat flux quite well. A cautious conclusion that both models can predict the heat flow well can be drawn. Users should though always verify what is output by the models versus what is measured by the loggers, before assuming them to be equivalent.

The moisture-related responses express much larger deviations, but an overall trend agreement is achieved. The calculated moisture masses of CS are obviously piled into two model categories in stage 1: DELPHIN results (red curves) and WUFI results (blue curves) concentrate at different values, while the results from other models have a larger spectrum. Since this distinction is not obvious in the relative humidity courses, the limited material items in the databases and/or similarity of their properties are presumably the primary reasons (Figs. 5g and 7m). As measured material properties are provided in stage 2, the agreement in most moisture masses of CS and WF (Figs. 5h, i, n, o, and 9k, l, n, o) are improved, while there are still mismatches between the calculated and measured moisture masses of MW (Figs. 5k, l and 7q, r). Relative humidity or its equivalent parameters (vapor pressure, capillary pressure, water chemical potential [71], etc., which can be translated by physical and/or mathematical relations) is the actual potential in the moisture balance equation. As this quantity also heavily depends on the moisture flows from neighboring materials/air, a more accurate implementation of the material properties from stage 1 to stage 2 reduces the deviations between measured and simulated relative humidity results. It should also be noted that some relative humidity predictions roughly remain similar or become worse, presumably due to user impact. This is proven by the results calculated by fine-tuned models in stage 3, as

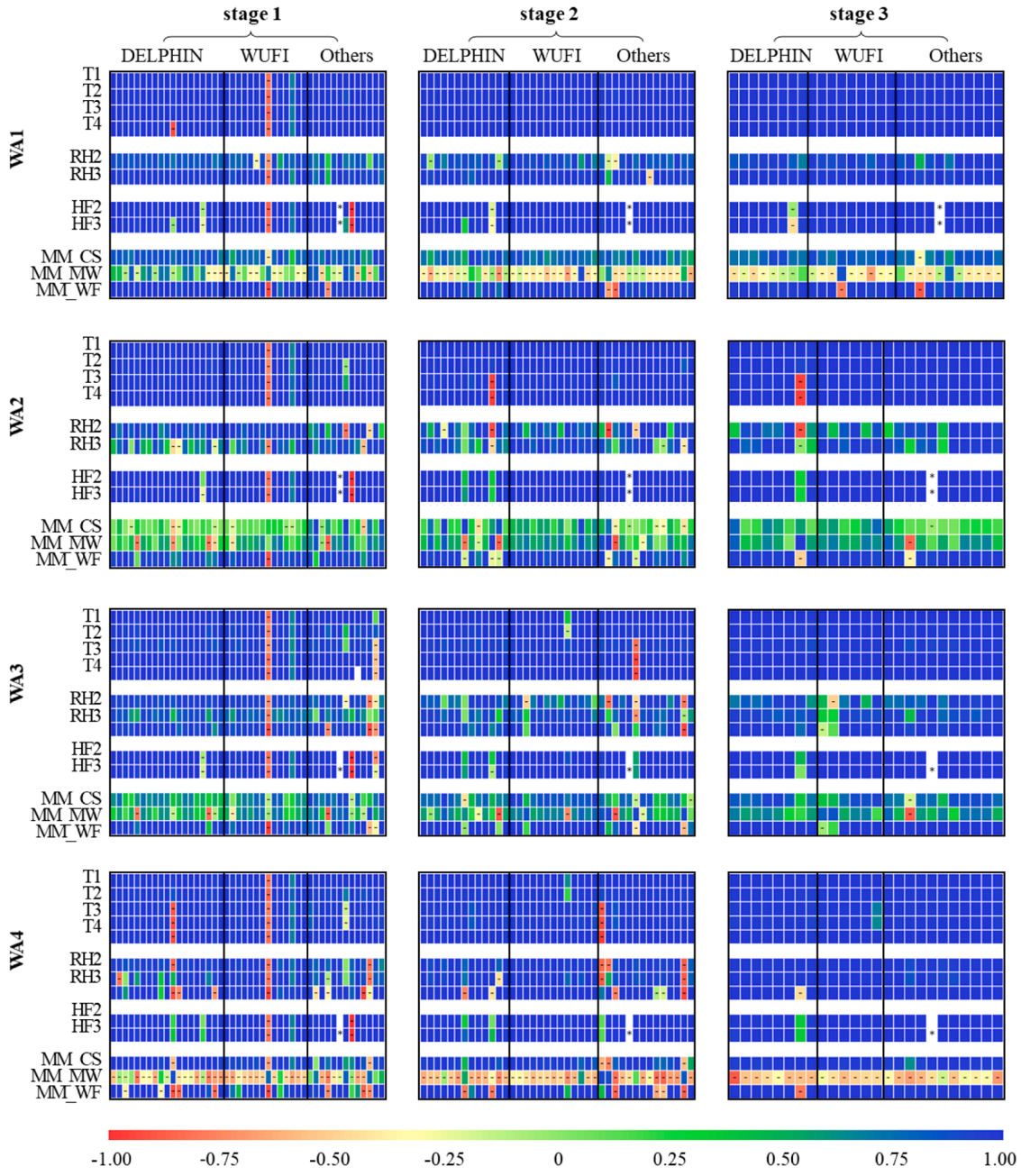

Fig. 9. Pearson correlation coefficient (r) between the simulated and measured hygrothermal responses.

mistakes might be diagnosed and corrected by the users themselves (Figs. 5b, c, e, f, and 7b, c, e, f, h, i).

It is obvious that the implemented material properties have a significant influence on the simulated moisture masses. Whether the materials are selected from the databases of HAM models or translated from the measured datasets, there are always uncertainties. That of the former is mainly the mismatch between the hygrothermal properties of the actual materials and the material entries in the databases. That of the latter corresponds to the translation from the measured material properties made by different tests (and some untested properties with suggested values) to the respective formulations required by the different models, during which specific assumptions, simplifications and approximations are made. In addition, human errors and preferences also significantly contribute [16].

It should be noted that some outputs likely include the data on wrong positions, for the opposite directions or wrong looping statements, as deduced by visual analysis by the coordinator. In addition, more obvious errors occur, such as relative humidity results between 0 and 1 (shown in dashed lines) instead of the required percentage and moisture contents in kg/m^3 instead of the required moisture masses in kg/m^2, judging from the trends and orders of magnitude. These kinds of deviations can be easily distinguished in this case but may be difficult to detect in more complicated scenarios and can lead to misinterpretation. Note that raw data are used for calculating NRMSE and r.

3.2. Impact analysis

3.2.1. Implementation of material properties in stage 1

Obvious deviations in simulated hygrothermal responses are observed in stage 1, in which detailed material properties are not provided. This reproduces an engineering scenario wherein the model user lacks measured hygrothermal properties, but one can turn to material

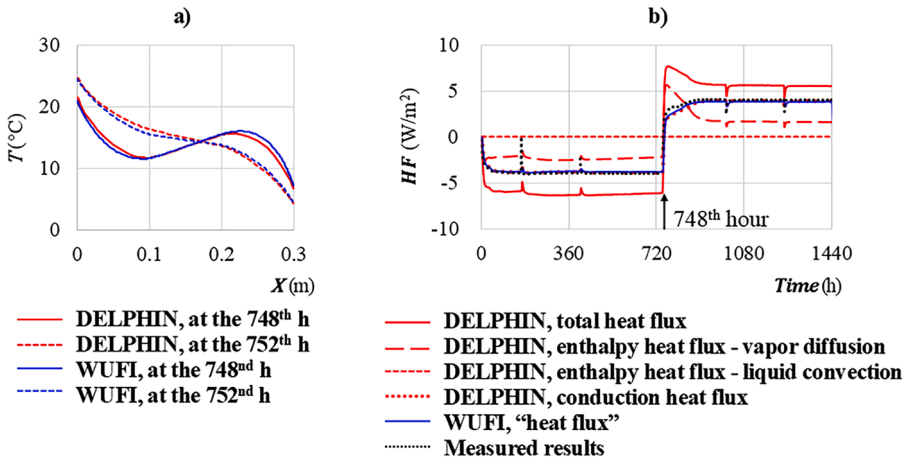

Fig. 10. Detailed outputs of heat-related responses in WA1: a) temperature profiles of WA1 at the 748th and 752nd hour, and b) detailed components of HF2.

Table 2
The claimed values of material properties in the technical sheets.

Material properties	CS	MW	WF	VB
Density, ρ [kg/m^3]	230 ~ 265	35	110	–
Specific heat capacity, c [J/(kg·K)]	1030	–	2100	–
Thermal conductivity, λ [W/(m·K)]	0.067 ~ 0.073	0.037	0.037	–
Vapor resistance factor, μ [-]	2	1	3	$> 5 \times 10^7$
Capillary absorption coefficient, A_w [kg/(m^2·s$^{0.5}$)]	0.85	–	≤ 1	–

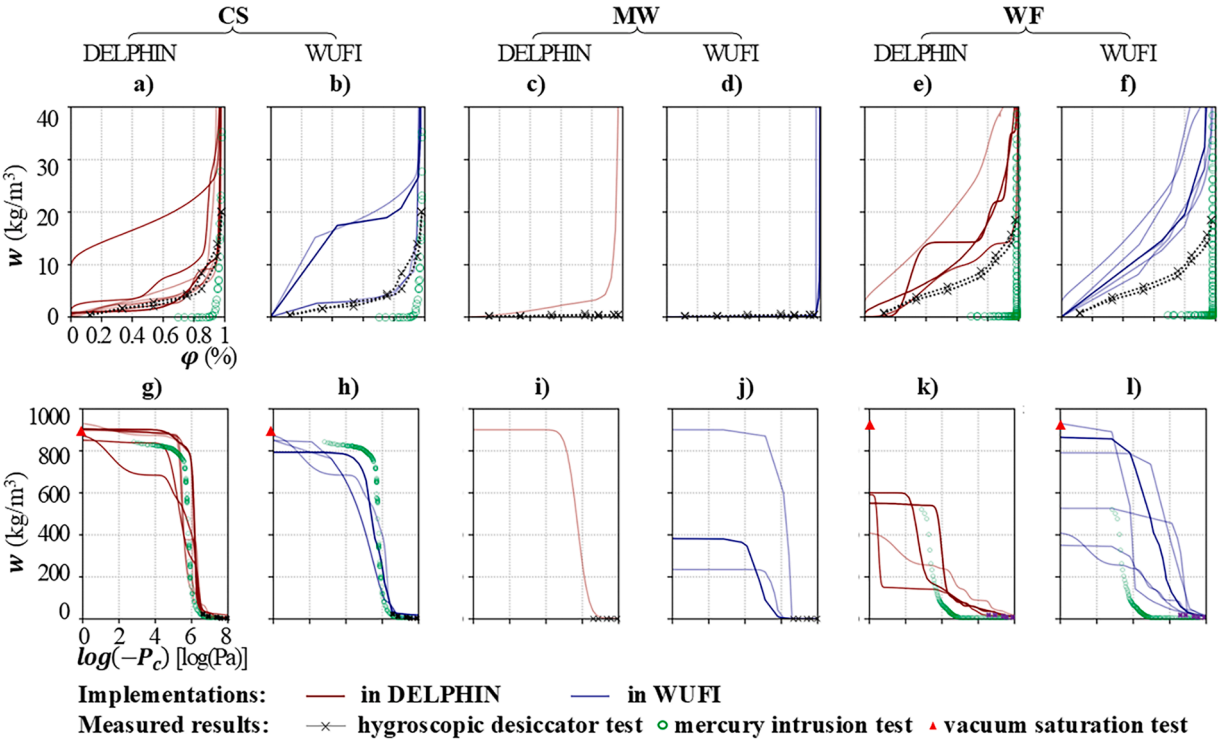

Fig. 11. The implemented: a) – f) sorption isotherms $w(\varphi)$, and g) – l) moisture retention curves $w(P_c)$ in DELPHIN (in red) and WUFI (in blue) in stage 1.

databases as provided by HAM models like DELPHIN and WUFI. Their built-in databases allow users to select materials with "similar" properties as the ones used in the experiment or to define new materials based on technical sheets from the manufacturers (provided for the participants as in Table 2). As some crucial hygrothermal properties (particularly moisture storage and transport functions) are not or only partially provided, the participants have flexibility in defining their criteria of "similarity". Note that though similarity in one property may coexist with deviations in other properties [16], and therefore this variation in selection may lead to deviations in predictions. Another uncertainty is that deviations are also observed between those declared values and those measured in the characterization tests [53,54].

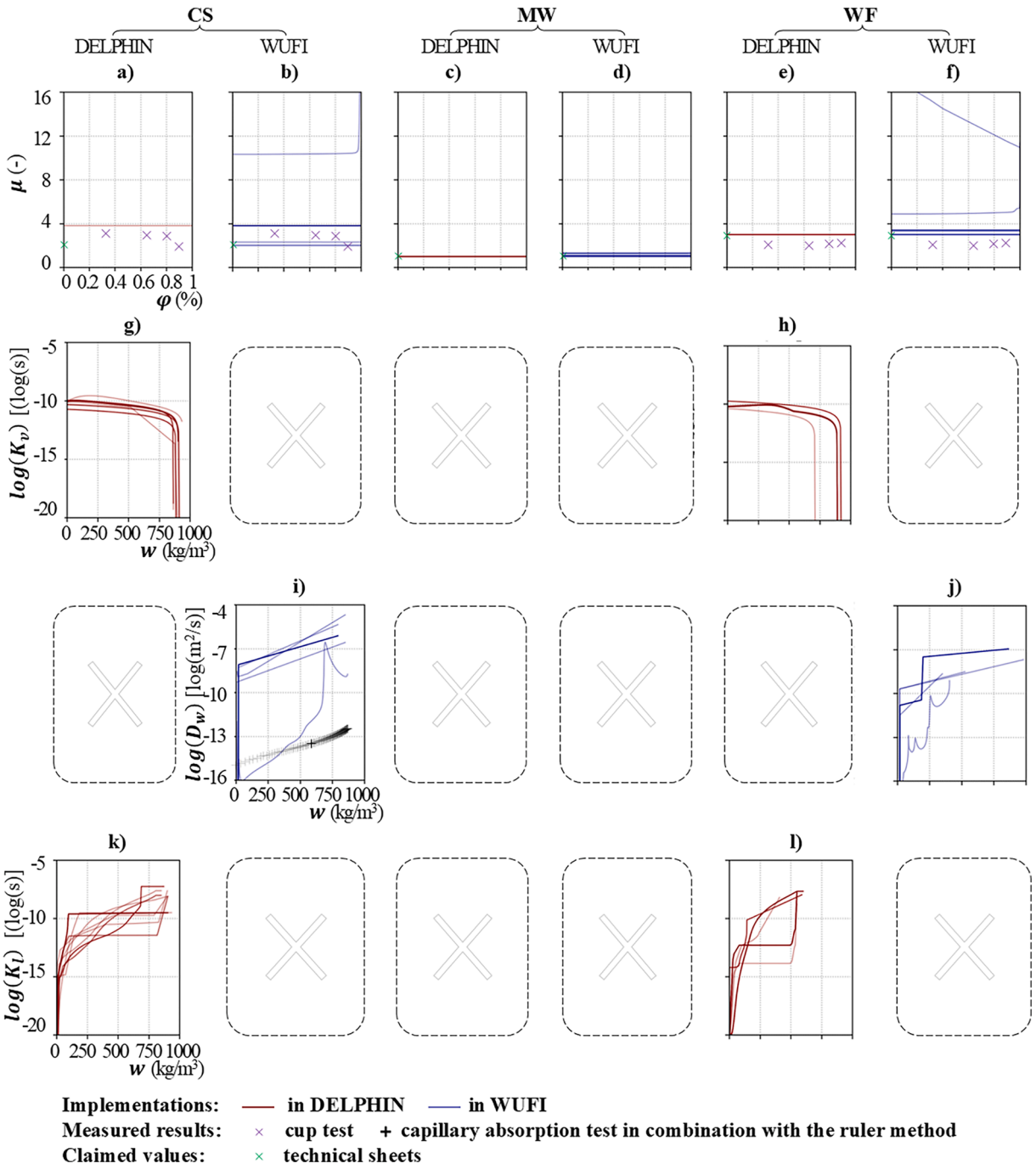

Fig. 12. The implemented: a) – f) vapor resistance factor $\mu(\varphi)$, g) - h) vapor permeability $K_v(w)$, i) - j) liquid diffusivity $D_w(w)$, and k) - l) liquid permeability $K_l(w)$ in DELPHIN (in red) and WUFI (in blue) in stage 1.

Figs. 11–12 illustrate the wide spectrum of the moisture storage and transport properties applied in DELHIN and WUFI during stage 1 and confront these with the measured properties. Note that those from the other models are not presented due to translation complexity under different driving potentials. For the moisture storage capacities of CS in DELPHIN and WUFI, the differences go up to 20 kg/m^3 in the hygroscopic range (see Figs. 11a&b) and 200 kg/m^3 in the over-hygroscopic range (see Figs. 11g&h), which may lead to significant deviations in relative humidity changes and moisture storage. Even larger deviations are found for MW (see Figs. 11c, d, i, j), a non-hygroscopic material. Due to the unique pore network, where capillary pressure is not evident, typical methods of describing moisture storage rarely apply to fiber-based materials. This leads to more assumptions, simplifications and approximations that generate more deviations as well. In some implementations, simple linear interpolation is applied to its moisture storage functions ending with arbitrary saturation values, while others assume MW to be purely "water-tight" and non-hygroscopic. Neither implementation is perfect in view of the simulated hygrothermal responses (Figs. 5j-l and 7p-r). Deviations can also be observed in moisture transport properties, even up to several orders of magnitude (Fig. 12).

3.2.2. *Implementation of material properties in stage 2*

In stage 2, for the simple single-valued parameters, most participants directly assign the measured values as inputs, but a few mistakes are made in some submissions (e.g., allocating density ρ of material A to material B). For the "indirect" parameters or "functions" (e.g., moisture

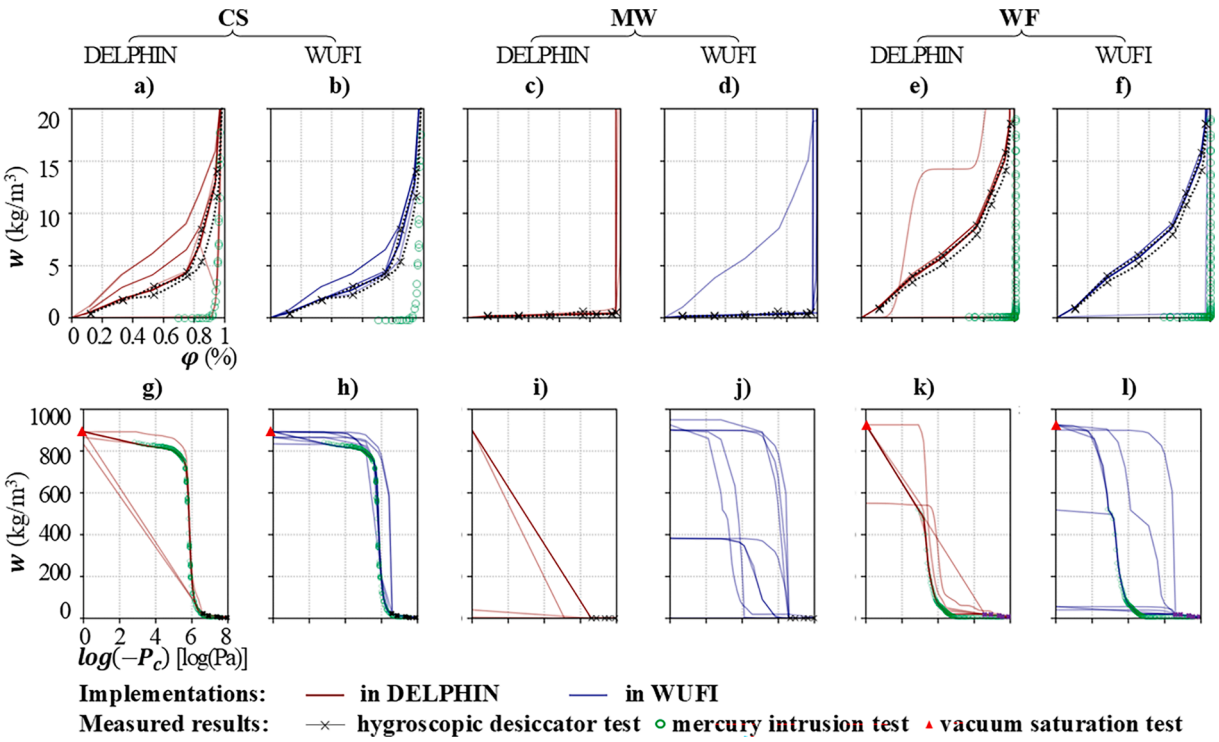

Fig. 13. The implemented: a) – f) sorption isotherms $w(\varphi)$, and g) – l) moisture retention curves $w(P_c)$ in DELPHIN (in red) and WUFI (in blue) in stage 2.

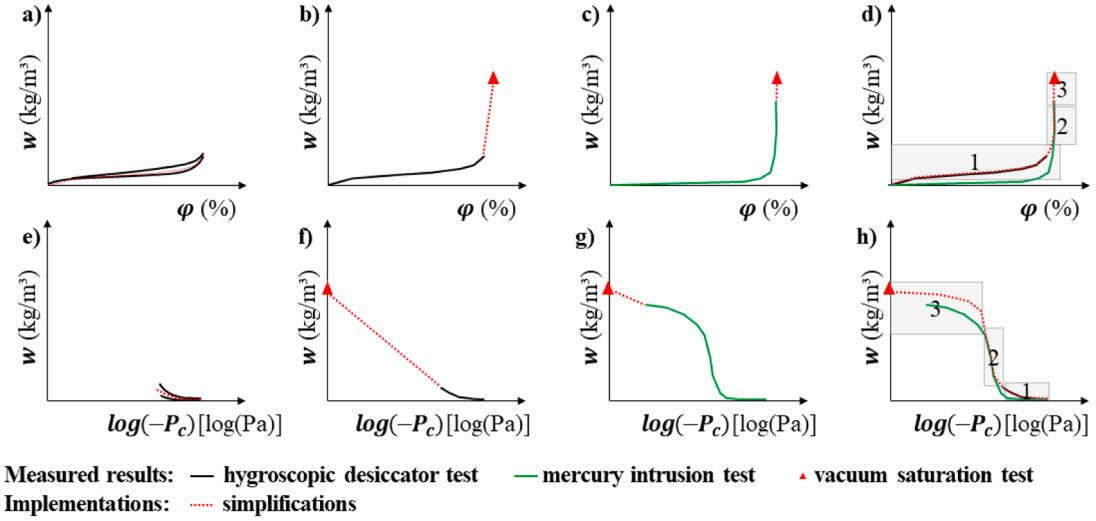

Fig. 14. Typical simplifications in implementing moisture storage properties: a) – d) sorption isotherms $w(\varphi)$, and e) – h) moisture retention curves $w(P_c)$.

storage and transport properties), there are much larger deviations as certain assumptions, simplifications and approximations as well as human errors might be made when translating the material properties into the forms required by different models. Considering the poor agreement in moisture-related outputs, the implemented moisture storage and transport coefficients in DELHIN and WUFI are confronted with the provided (measured or declared) properties in Figs. 13 and 14.

Fig. 13 illustrates the implemented sorption isotherms $w(\varphi)$ and moisture retention curves $w(P_c)$ in the two HAM models. The assignment for moisture storage function is compulsory for all materials including non-hygroscopic fiber-based materials (e.g., MW). In DELPHIN, either function can be applied, while WUFI only supports sorption isotherms. The coordinator has converted these into moisture retention curves, by applying Kelvin's law (with additional interpolation), to facilitate a comprehensive comparison. Due to the non-overlapping nature of the experimental results from the hygroscopic desiccator test, the mercury intrusion test and vacuum saturation test (due to inherent restraints of the experiments [72]), one needs to select among the provided results [53,54], for instance, by setting one dataset as the baseline (assumed to be the most accurate), while scaling or shifting partial data from other experiments. Fig. 14 illustrates typical simplification methods in implementing moisture storage properties in this case. Both WUFI and DELPHIN disregard the sorption hysteresis, forcing the users to apply single (ab-/de-sorption or their average) curve (Figs. 14a&e). This simplification seems to have little impact given that the hysteresis is fairly negligible for the materials involved here. The desiccator results

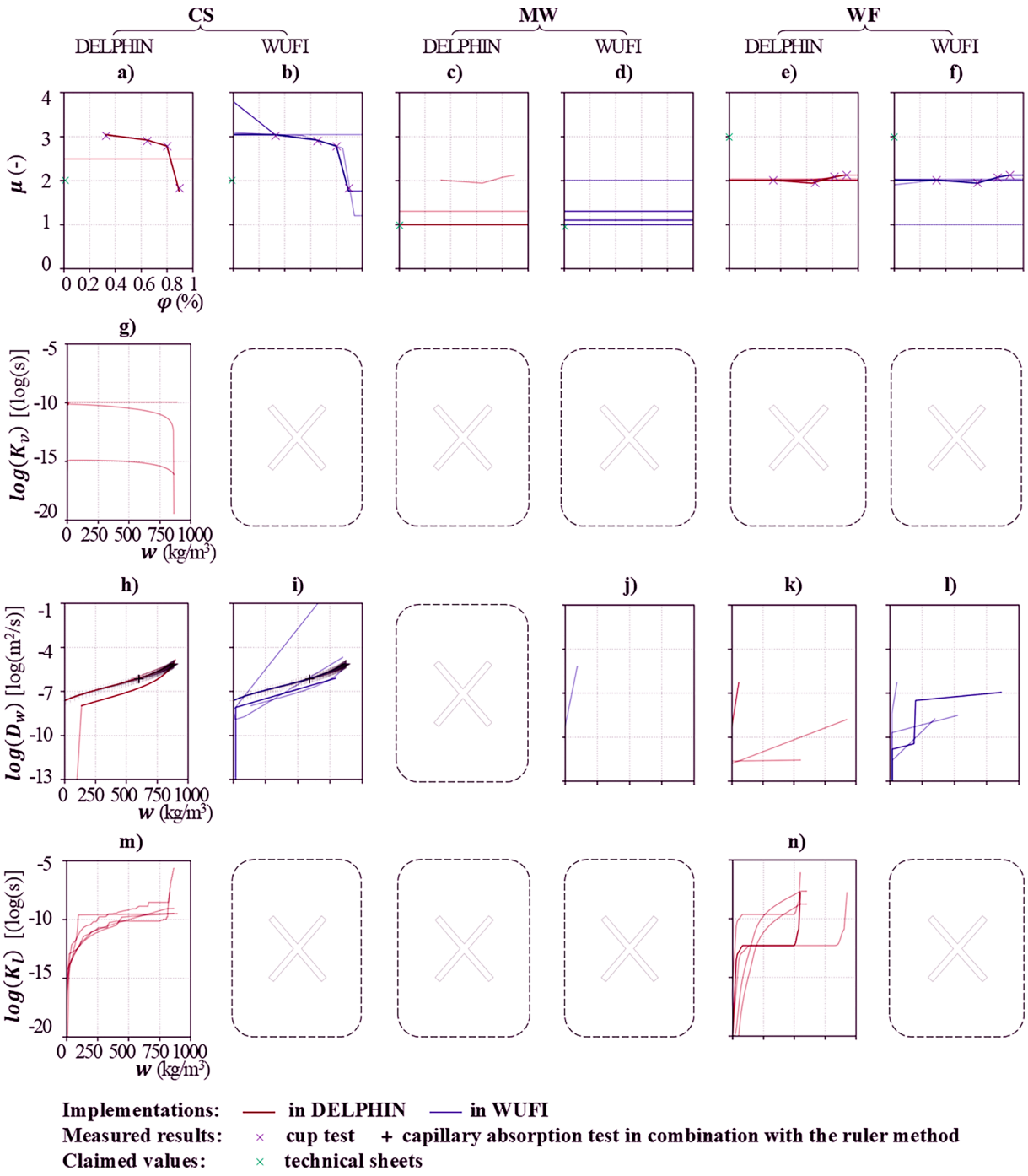

Fig. 15. The implemented: a) – f) vapor resistance factor $\mu(\varphi)$, g) vapor permeability $K_v(w)$, h) - l) liquid diffusivity $D_w(w)$, and m) - n) liquid permeability $K_l(w)$ in DELPHIN (in red) and WUFI (in blue) in stage 2.

are more straightforward and generally presumed to be more reliable for the hygroscopic range. Many participants simply connect the equilibrium moisture content at 97 % RH with the vacuum saturation moisture content when implementing sorption isotherms $w(\varphi)$ (Fig. 14b). Unless presented as a function of capillary pressure or other equivalent potentials (e.g., water chemical potential [71]) that can enlarge the scope of X-axis, the deviations in over-hygroscopic range are unlikely to be diagnosed. Similar linear connection may be applied when implementing moisture retention curves $w(P_c)$ (Figs. 13g, i, k, 14f). On the other hand, determining moisture storage properties in the over-hygroscopic range via the mercury intrusion test (or the pressure plate test, etc.) is more equipment-intensive and expertise-demanding [72], leading to uncertainties in data manipulation. Some participants predominantly rely on results from the mercury intrusion test due to its completeness, with connection to the saturation moisture content (Figs. 13a and 14c&g). Given the different sensitivity and reliability of different characterization tests, there is another implementation that processes measured datasets as a piecewise function: (1) in the hygroscopic range, applying measured results from the hygroscopic desiccator test, (2) in the "sharp changing" part, applying partial results from mercury intrusion test (with necessary scaling or shifting), and (3) connecting the saturation moisture content with the second part of the curve (Figs. 14d&h).

Water vapor transport properties can be written as vapor resistance

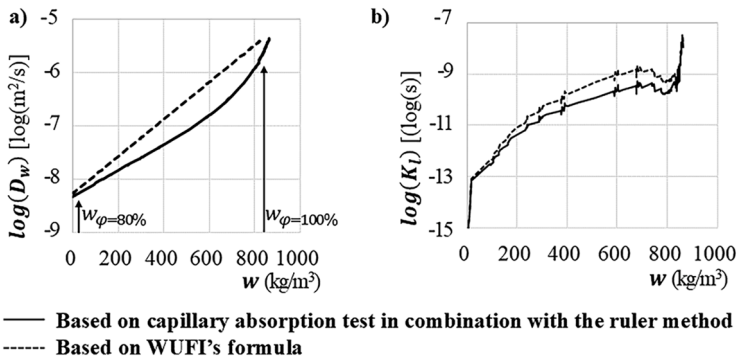

Fig. 16. Comparison between liquid transport coefficients of CS obtained from the ruler method and WUFI's simplification formula: a) liquid diffusivities and b) liquid permeabilities.

factor $\mu(\varphi)$ or translated into vapor permeability $K_v(w)$, while liquid transport properties can be expressed as liquid diffusivity $D_w(w)$ or liquid permeability (liquid conductivity) $K_l(w)$. All functions above are applicable in DELHIN, while only vapor resistance factor $\mu(\varphi)$ and liquid diffusivity $D_w(w)$ are applicable in WUFI. Fig. 15 illustrates these implementations in DELPHIN and WUFI. For vapor transport properties, the measured results from the cup test are directly assigned by most participants, while additional linear interpolation and averaging are applied in some models. A small number of DELPHIN users translate the measured vapor resistance factor at dry state μ_{dry} into $K_v(w)$ via Eq. (3) [42] (see Fig. 15g):

$$K_v = \frac{D_{va}}{\mu_{dry} R_v T} \cdot \frac{1 - \frac{w}{w_{sat}}}{0.503 \cdot \left(1 - \frac{w}{w_{sat}}\right)^2 + 0.497} = \frac{26.1 \times 10^{-6}}{\mu_{dry} \cdot 462 \cdot 293.15} \cdot \frac{1 - \frac{w}{w_{sat}}}{0.503 \cdot \left(1 - \frac{w}{w_{sat}}\right)^2 + 0.497} \quad (3)$$

where D_{va} [m^2/s] is vapor permeability in air; K_v [s] is vapor permeability; R_v [J/kg·K] is water vapor gas constant; T [°C] is temperature; w [kg/m^3] is moisture content; w_{sat} [kg/m^2] is saturation moisture content; μ_{dry} [-] is the vapor resistance factor at dry state.

The imlementation of liquid transport properties is subject to uncertainties stemming from both the users and the material properties' providers, as these properties are not directly measured but instead characterized with combined experiments to then mathematically "bridge the gaps". Due to deficiencies in the X-ray and NMR equipment, the provided moisture content profiles and liquid diffusivity of CS is obtained from capillary absorption experiment in combination with the ruler method [64], and may hence deviate from the actual properties. Presumably with this suspicion or other considerations, a proportion of participants assign other data series to CS (see Fig. 15i). In WUFI, two liquid diffusivities, D_{ws} for suction and D_{ww} for redistribution, are applied. The former is activated only when the structure is exposed to rain flux, otherwise (in this case) the latter is used [73]. As default, D_{ws} is generated by diffusivities at 0, 80 %, and 100 % relative humidity, while D_{ww} is based on D_{ws} at 0 and 80 % relative humidity and 0.1 times D_{ws} at 100 % relative humidity. A common simplified formula of D_{ws} [74] is recommended by WUFI developers as shown in (see Eq. (4)), and the corresponding liquid permeability K_l obtained by multiplication of liquid diffusivity with the slope of the moisture retention curve $\frac{\partial w}{\partial P_c}$ is shown in Eq. (5), which again can lead to considerable deviations.

$$D_{ws} = 3.8 \cdot \left(\frac{A_w}{w_f}\right)^2 \cdot 1000^{\frac{w}{w_f}-1} \quad (4)$$

$$K_l = D_w \cdot \frac{\partial w}{\partial P_c} \quad (5)$$

where, A_w is capillary absorption coefficient [kg/($m^2 \cdot s^{0.5}$)]; D_{ws} [m^2/s] is liquid diffusivity for suction; K_l [s] is liquid permeability; P_c [Pa] is capillary pressure; w [kg/m^3] is moisture content; w_f [kg/m^3] is the "free" saturation moisture content (normally equivalent to the capillary saturation moisture content w_{cap}).

Fig. 16 compares the liquid transport coefficients of CS obtained from capillary absorption experiment in combination with the ruler method and WUFI's simplified formula (Eq. (4)). WUFI asks for liquid diffusivity values at 80 % and 100 % RH to draw a linear curve of liquid diffusivity in logarithmic scale, which remains higher than that obtained

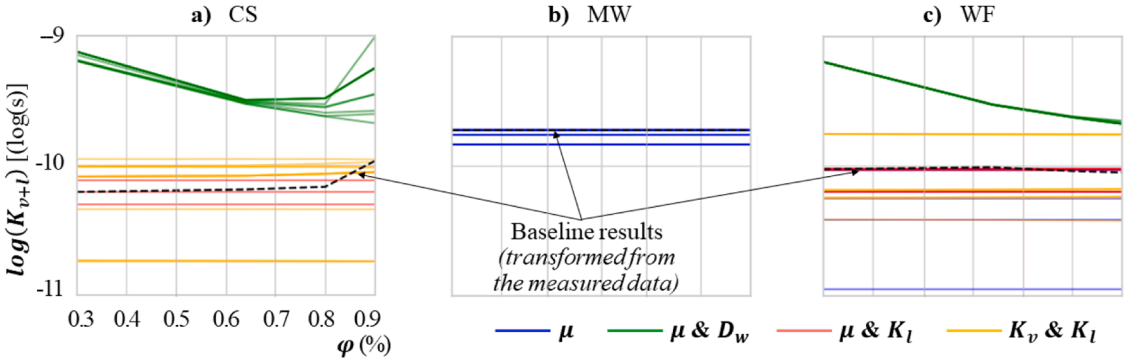

Fig. 17. Implemented equivalent permeability K_{v+l} in WUFI and DELPHIN in stages 1 and 2.

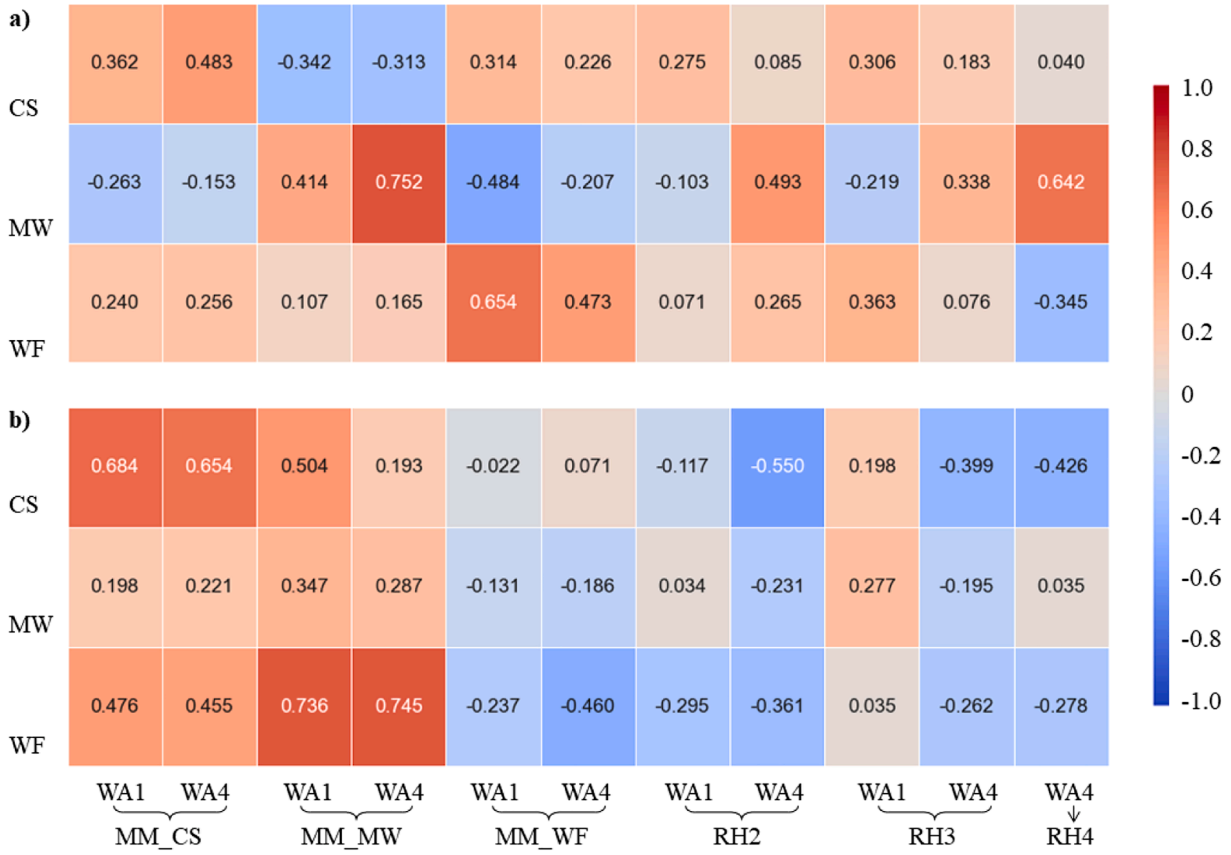

Fig. 18. Pearson correlation coefficient (r) between inputs and outputs in stages 1 and 2: a) r between (1) NRMSE between measured and implemented $w(\varphi)$ and (2) NRMSE between measured and simulated moisture-related responses, and b) r between (1) NRMSE between measured and implemented K_{v+l} and (2) NRMSE between measured and simulated moisture-related responses.

from the ruler method. This deviation leads to significant overestimation of the translated liquid permeability particularly in the over-hygroscopic range. Although MW is assumed to be non-capillary-active, some participants still assign values to their liquid transport coefficients (Figs. 15j). Due to the equipment deficiency, only the capillary absorption coefficient (A_w = 0.02 kg/(m^2·s$^{0.5}$)) of WF is measured and provided as the indicator for liquid transport, yielding different considerations: neglecting liquid transport properties or estimating a liquid diffusivity or a permeability (Figs. 15k, l, n).

Another issue is the double inclusion of liquid transport. Vapor transport is dominant in the dry cup test, while liquid transport often becomes non-negligible in the wet cup test. The latter should be split into vapor and liquid transport, to keep the values for pure vapor transport properties, or merged into a total moisture transport coefficient. As suggested by the WUFI developers, the implementation should be either based on a constant μ value (from the dry cup test) together with liquid transport properties, or alternatively it should use $\mu(\varphi)$ function without implementing additional liquid transport properties [75]. However, a considerable proportion of the participants apply both $\mu(\varphi)$ function and a liquid diffusivity. A similar concept should be complied within DELPHIN: the vapor permeability K_v (translated from dry cup test, considering pure vapor transport) should be accompanied with liquid permeability K_l. According to statistics, this double inclusion exists in 20 CS's material property files and 11 WF's material property files, out of the 26 WUFI/DELPHIN submissions in stage 2.

3.2.3. *Impact of input uncertainties on output accuracies in stages 1 and 2*

To quantify the impact of the deviations in the implemented material properties on the deviations in the simulated hygrothermal responses, two series of NRMSE results of WA1 and WA4 from stages 1 and 2 are combined. As the heat-related responses are often simulated relatively accurately, the assessment targets the moisture-related responses particularly, by relating the accuracy of the implemented moisture storage or transport properties to the accuracy of the simulated moisture-related responses. On the one hand, the NRMSE between the measured and implemented moisture storage properties (specifically, $w(\varphi)|_{\varphi<97\%}$, since most materials act mainly in the hygroscopic range) or moisture transport properties (specifically, a newly introduced parameter, equivalent vapor plus liquid permeability $K_{v+l}|_{\varphi<97\%}$, under the vapor pressure gradient) is used. On the other hand, the NRMSE between the measured and calculated moisture-related responses is applied. The Pearson correlation coefficient (r) of the two series of NRMSE then indicates which inputs have crucial impacts on which outputs, and how influential the impacts are. Some extreme outliers are removed according to the 1.5-IQR rule.

The equivalent permeability K_{v+l} under the gradient of vapor pressure P_v is applied to harmonize moisture transport properties in different forms and/or under different gradients (Eq. (6)). The four different implementation approaches of moisture transport properties in WUFI and DELPHIN are translated according to Eqs. (7–10), wherein δ_a is assumed 1.9×10^{-10} s and $P_{v,sat}$ is assumed 1706.3 Pa, at 15 °C. The translation results are confronted with the baseline results obtained by transforming the measured data from the cup test, as shown in Fig. 17.

$$g_{v+l} = -\left(K_v + K_l \cdot \left(\frac{\partial P_c}{\partial P_v}\right)\right)\nabla P_v = -K_{v+l}\nabla P_v \tag{6}$$

$$K_{v+l} \overset{\mu}{\Rightarrow} \frac{\delta_a}{\mu} \tag{7}$$

$$K_{v+l} \overset{\mu \ \& \ D_w}{\Rightarrow} \frac{\delta_a}{\mu} + D_w \cdot \left(\frac{\partial w}{\partial P_v}\right) = \frac{\delta_a}{\mu} + D_w \cdot \left(\frac{\partial w}{\partial \varphi}\right) \cdot \left(\frac{\partial \varphi}{\partial P_v}\right) = \frac{\delta_a}{\mu} + D_w \cdot \left(\frac{\partial w}{\partial \varphi}\right) \cdot \frac{1}{P_{v,sat}} \tag{8}$$

$$K_{v+l} \overset{\mu \ \& \ K_l}{\Rightarrow} \frac{\delta_a}{\mu} + K_l \cdot \left(\frac{\partial P_c}{\partial P_v}\right) \tag{9}$$

$$K_{v+l} \overset{K_v \ \& \ K_l}{\Rightarrow} K_v + K_l \cdot \left(\frac{\partial P_c}{\partial P_v}\right) \tag{10}$$

where, D_w [m^2/s] is liquid diffusivity; g_{v+l} [kg/(m^2·s)] is overall moisture flux; K_l [s] is liquid permeability; K_v [s] is vapor permeability; K_{v+l} [s] is the equivalent permeability for overall moisture transfers; P_c [Pa] is capillary pressure; P_v [Pa] is vapor pressure; $P_{v,sat}$ [Pa] is the saturation vapor pressure; w [kg/m^3] is moisture content; δ_a [s] is air permeability; φ [%] is relative humidity; μ [-] is vapor resistance factor.

Fig. 18 illustrates the correlations between (1) the NRMSE between measured and implemented moisture storage properties (Fig. 18a) or moisture transport properties (Fig. 18b) and (2) the NRMSE between measured and simulated moisture-related responses: a higher r value indicates a larger impact of the implemented hygric material property on the corresponding simulated moisture-related response. A first strong expression is that both moisture storage and transport properties have a significant impact on the moisture mass and relative humidity of primarily the material itself and sometimes of the neighboring materials. Specifically, implemented moisture storage properties influence the amount of moisture flows between materials and thereto relative humidity at the interfaces, while the moisture transport properties influence the speed of moisture flows. The impact of implemented WF's moisture transport property on its moisture mass is an exception. A possible explanation is that the moisture transport in WF also occur in over-hygroscopic condition, which are excluded in the correlation calculation.

This analysis underscores that the accuracy of moisture-related responses is highly dependent on the precise implementation of hygric material properties. Further sensitivity analysis can provide deeper insights into these associations. In terms of model validation, prior studies typically rely on temperature and relative humidity as key indicators. However, incorporating measured moisture mass as a benchmark is essential for validating simulated results influenced by hygric material properties. Without this, the credibility and reliability of the validation remain uncertain.

3.2.4. Insights from model fine-tuning in stages 3

The provided measured hygrothermal responses in stage 3 allow investigating whether the simulation results are close to the experimental observations. While there are also uncertainties in the measurement (e.g., logger errors, inherent limitations in experiment protocols, etc.), the measured hygrothermal responses can be used as a reference to modify the implementations for a "more satisfied" result. Since multi-factors are involved in most submissions, it is difficult to evaluate the impact on each specific adjustment with other variables controlled, but these adjustments with positive effects reported by the participants are summarized below.

Out of the 28 submissions, five remained unchanged as in stage 2. These contributors were satisfied with the agreement between simulation results and the measured datasets or avoided further adjustment to prevent artificial fitting without physical accuracy. Among the rest 23 complete submissions, a range of strategies are implemented to refine the simulation models. Most adjustments are made in moisture storage functions (10, the number of submissions, same below), liquid transport functions (10), surface transport coefficients (9), and thermal conductivities (5). The frequency order of the adjusted materials is CS (16), MW (10), WF (7) and VB (1). Some participants mentioned that an improved accuracy is observed after refining the grid discretization (6) and correcting sensor positions (2) to better capture the physical setup. Some participants took the chance to correct previous errors (5) in implementing surface transport coefficients, initial conditions, and boundary conditions as well as in outputting positions or units.

4. Conclusion

This study evaluates the robustness and reliability of HAM models in predicting one-dimensional hygrothermal responses of building components under controlled conditions. A comprehensive round robin action, involving 38 groups from 19 countries, highlights the influence of implemented material properties on simulation outcomes.

The first impression is that while good agreement is achieved in heat transport predictions across models, moisture transport predictions are highly sensitive to implemented hygric material properties. Variations in moisture storage and transport properties, stemming from user-defined inputs or database selection, lead to notable deviations in simulated hygrothermal responses, emphasizing the need for precise material characterization and implementation.

This three-stage approach, progressing from engineering-based robustness assessment to more rigorous reliability evaluation, provides insights into common challenges and opportunities for improving HAM modelling. In stage 1, discrepancies in moisture-related predictions arise from user impact on estimating material properties and database reliance. In stage 2, providing measured material properties reduces deviations but reveals challenges in physically and/or mathematically translating them accurately into model-specific forms. In stage 3, fine-tuning in assignment somehow improves accuracy, albeit limited by user expertise and persistent issues with material properties like moisture storage and transport functions. User-induced errors (including incorrect units or values, inconsistent parameter definitions, and wrong or incomplete boundary conditions) are identified but may remain undetected in complex scenarios and potentially yield misinterpretations.

It is crucial to combine experimental validation with inter-model comparison for a comprehensive model assessment. Moisture mass should be prioritized as a benchmark indicator due to its critical impact on moisture-related responses. A reliable implementation of moisture storage and transport properties should receive more attention, particularly when the properties are not directly measured or the measured results from different tests mismatch each other. Sensitivity analysis approach, which may mitigate the impact of intrinsic limitations of material characterization techniques and the uncertainties introduced by data manipulation, can be integrated with conventional validation and inter-model comparison methods.

For model users, the primary suggestion is to cautiously implement material properties. If choosing material properties from databases of HAM models or other sources, they should be aware that different batches of the same material available (in the market or in the databases) exhibit large deviations in moisture storage and transport properties, to which the hygrothermal responses of building components are sensitive. If performing material characterization tests on their own, they should note the intrinsic limitations in measurements as well as uncertainties in data processing. In addition, the gaps between actual and implemented material properties, and the usage differences in units, symbols, directions, orders of magnitudes, conceptions and definitions should be considered.

Admittedly, the virtual models can never perfectly reproduce the real world, and the varied approaches of interpretation among models make the model assessment complicated. While extensive models have been evaluated by this dedicated experimental dataset, it is still impossible to tackle all uncertainties (e.g., implementation of boundary conditions, material homogenization and dimensional reduction, etc.) in a single study. There is a clear need to extend this work with additional benchmark datasets and inter-model validation studies to achieve a more

reliable application of HAM models.

Author contribution

Xinyuan Dang: conceptualization, methodology, investigation, experiments, data processing and visualization, writing - original draft, funding acquisition; **Hans Janssen**: supervision, conceptualization, methodology, investigation, writing - review & editing; **Staf Roels**: supervision, conceptualization, methodology, investigation, experiments, writing - review & editing. **Others**: investigation.

Data availability

The full datasets of measured material properties and hygrothermal responses are described in https://doi.org/10.1007/s12273-024-1176-8 and available at https://doi.org/10.5281/zenodo.10998834.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The research is supported by China Scholarship Council (No. 202006090005), Estonian Ministry of Education and Research (No. TK230), and Estonian Research Council (No. PRG2732). This study can only achieve fruitful outcomes with the global interests in hygrothermal modelling as well as valuable suggestions and dissemination efforts of numerous researchers and practitioners. The institutional coordinator, Building Physics and Sustainable Design Section of KU Leuven, Belgium, would express sincere appreciation to all the participants.

References

[1] S. Panico, M. Larcher, V. Marincioni, A. Troi, C. Baglivo, P.M. Congedo, Identifying key parameters through a sensitivity analysis for realistic hygrothermal simulations at wall level supported by monitored data, Build Environ 229 (2023) 109969.
[2] E. Vereecken, L. Van Gelder, H. Janssen, S. Roels, Interior insulation for wall retrofitting–A probabilistic analysis of energy savings and hygrothermal risks, Energy Build 89 (2015) 231–244.
[3] I. Vandemeulebroucke, B. Vanderschelden, K. Janssens, S. Caluwaerts, N. Van Den Bossche, The impact of climate change on degradation in historical building envelopes: progress in research using hygrothermal models, J Cult Herit 70 (2024) 345–363.
[4] P. Gholamalipour, H. Ge, T. Stathopoulos, Wind-driven rain (WDR) loading on building facades: a state-of-the-art review, Build Environ 221 (2022) 109314.
[5] B.H. Vos, Condensation in flat roofs under non-steady-state conditions, Building Science 6 (1971) 7–15.
[6] H. Hens, PhD Thesis, KU Leuven, Leuven, Belgium, 1976.
[7] G.H. Galbraith, PhD Thesis, University of Strathclyde, Glasgow, the UK, 1992.
[8] M. Krus, Moisture transport and storage coefficients of porous mineral building materials. Theoretical Principles and New Test Methods, IRB-Verlag, Stuttgart, 1996.
[9] M. Funk, K.G. Wakili, Driving potentials of heat and mass transport in porous building materials: a comparison between general linear, thermodynamic and micromechanical derivation schemes, Transp Porous Media 72 (2008) 273–294.
[10] G.A. Scheffler, Validation of Hygrothermal Material Modelling Under Consideration of the Hysteresis of Moisture Storage, Dresden University of Technology, Dresden, Germany, 2008.
[11] Canadian Mortgage and Housing Corporation, Review of Hygrothermal Models For Building Envelope Retrofit analysis. Research Highlights. Technical Series 03-128, Canadian Mortgage and Housing Corporation, Ottawa, 2003.
[12] M. Woloszyn, C. Rode, Tools for performance simulation of heat, air and moisture conditions of whole buildings, Building Simulation 1 (1) (2008) 5–24.
[13] J.M.P.Q. Delgado, N.M. Ramos, E. Barreira, V.P. De Freitas, A critical review of hygrothermal models used in porous building materials, J Porous Media 13 (3) (2010).
[14] A. Brambilla, A. Sangiorgio, Moisture and Buildings: Durability Issues, Health Implications and Strategies to Mitigate the Risks, Woodhead Publishing, 2021, pp. 129–152.
[15] H.L. Hens, Combined heat, air, moisture modelling: a look back, how, of help? Build Environ 91 (2015) 138–151.
[16] X. Dang, H. Janssen, S. Roels, User impact as an uncertainty in hygrothermal simulations: insights from a round robin test, in: U. Berardi (Ed.), Multiphysics and Multiscale Building Physics. IABP 2024. Lecture Notes in Civil Engineering, Multiphysics and Multiscale Building Physics. IABP 2024. Lecture Notes in Civil Engineering, 552, Springer, Singapore, 2025, pp. 68–75.
[17] S. Cornick, W. Maref, F. Tariku, Verification and validation: Establishing Confidence in Hygrothermal Tools, National Research Council of Canada, 2009.
[18] H.M. Künzel, K. Kiessl, Drying of brick walls after impregnation, Bauinstandsetzen 2 (1996) 87–100, in German.
[19] I. Budaiwi, R. El-Diasty, A. Abdou, Modelling of moisture and thermal transient behaviour of multi-layer non-cavity walls, Build Environ 34 (5) (1999) 537–551.
[20] S. Roels, W. Depraetere, J. Carmeliet, H. Hens, Simulating non-isothermal water vapour transfer: an experimental validation on multi-layered building components, Journal of Thermal Envelope and Building Science 23 (1) (1999) 17–40.
[21] T. Kalamees, J. Vinha, Hygrothermal calculations and laboratory tests on timber-framed wall structures, Build Environ 38 (5) (2003) 689–697.
[22] P. Talukdar, S.O. Olutmayin, O.F. Osanyintola, C.J. Simonson, An experimental data set for benchmarking 1-D, transient heat and moisture transfer models of porous building materials – Part II: experimental, numerical and analytical data, Int J Heat Mass Transf 50 (23–24) (2007) 4915–4926.
[23] M. Steeman, M. Van Belleghem, M. De Paepe, A. Janssens, Experimental validation and sensitivity analysis of a coupled BES–HAM model, Build Environ 45 (10) (2010) 2202–2217.
[24] C. James, C.J. Simonson, P. Talukdar, S. Roels, Numerical and experimental data set for benchmarking hygroscopic buffering models, Int J Heat Mass Transf 53 (19–20) (2010) 3638–3654.
[25] M. Van Belleghem, M. Steeman, A. Willockx, A. Janssens, M. De Paepe, Benchmark experiments for moisture transfer modelling in air and porous materials, Build Environ 46 (4) (2011) 884–898.
[26] T.Z. Desta, J. Langmans, S. Roels, Experimental data set for validation of heat, air and moisture transport models of building envelopes, Build Environ 46 (5) (2011) 1038–1046.
[27] X.G. Guo, Y.M. Chen, Y.Q. Deng, Development and experimental validation of a one-dimensional dynamic hygrothermal modeling based on air humidity ratio, Journal of Central South University 19 (3) (2012) 703–708.
[28] R. McClung, H. Ge, J. Straube, J. Wang, Hygrothermal performance of cross-laminated timber wall assemblies with built-in moisture: field measurements and simulations, Build Environ 71 (2014) 95–110.
[29] K. Carbonez, N. Van Den Bossche, H. Ge, A. Janssens, Comparison between uniform rain loads and point sources to simulate rainwater leakage with commercial HAM-models, in: Proceedings of the International Symposium on Building Pathology (ISBP 2015), Porto, Portugal, 2015, pp. 1–8.
[30] T. Colinart, P. Glouannec, Accounting for the temperature dependence in the sorption isotherms in the prediction of hygrothermal behavior of hemp concrete, in: Proceedings of Central European Symposium on Building Physics, Dresden, Germany, 2016, pp. 1–8.
[31] T. Colinart, D. Lelièvre, P. Glouannec, Experimental and numerical analysis of the transient hygrothermal behavior of multilayered hemp concrete wall, Energy Build 112 (2016) 1–11.
[32] W. Dong, Y. Chen, Y. Bao, A. Fang, A validation of dynamic hygrothermal model with coupled heat and moisture transfer in porous building materials and envelopes, Journal of Building Engineering 32 (2020) 101484.
[33] M. Asli, E. Sassine, F. Brachelet, E. Antczak, Hygrothermal behavior of wood fiber insulation, numerical and experimental approach, Heat and Mass Transfer 57 (2021) 1069–1085.
[34] S. Schroderus, V.M. Lähteenmäki, A. Haapala, F. Fedorik, Validation of Hygrothermal Numerical Simulation with Experiment for Future Climate Control, Scandinavian Simulation Society, 2022, pp. 156–162.
[35] M. Amorim, V.P. de Freitas, I. Torres, Influence of moisture on the energy performance of retrofitted walls-experimental assessment and validation of an hygrothermal model, International Journal of Architectural Heritage 18 (3) (2024) 477–491.
[36] T. Colinart, D. Lelièvre, P. Glouannec, Experimental and numerical analysis of the transient hygrothermal behavior of multilayered hemp concrete wall, Energy Build 112 (2016) 1–11.
[37] P. Talukdar, S.O. Olutmayin, O.F. Osanyintola, C.J. Simonson, An experimental data set for benchmarking 1-D, transient heat and moisture transfer models of hygroscopic building materials. Part I: experimental facility and material property data, Int J Heat Mass Transf 50 (23–24) (2007) 4527–4539.
[38] H. Rafidiarison, R. Remond, E. Mougel, Dataset for validating 1-D heat and mass transfer models within building walls with hygroscopic materials, Build Environ 89 (2015) 356–368.
[39] M. Salonvaara, P. Boudreaux, A. Desjarlais, F. Antretter, E. Werling, Validation of hygrothermal simulations with wall performance experiments in an environmental chamber, in: E3S Web of Conferences 172, EDP Sciences, 2020, p. 04010.
[40] C. James, C.J. Simonson, P. Talukdar, S. Roels, Numerical and experimental data set for benchmarking hygroscopic buffering models, Int J Heat Mass Transf 53 (19–20) (2010) 3638–3654.
[41] O. Adan, H. Brocken, J. Carmeliet, H. Hens, S. Roels, C.E. Hagentoft, Determination of liquid water transfer properties of porous building, materials and development of numerical assessment methods: introduction to the EC HAMSTAD project, Journal of Building Physics 27 (4) (2004) 253–260.
[42] C.E. Hagentoft, A.S. Kalagasidis, B. Adl-Zarrabi, S. Roels, J. Carmeliet, H. Hens, J. Grunewald, M. Funk, R. Becker, D. Shamir, O. Adan, H. Brocken, K. Kumaran, R. Djebbar, Assessment method of numerical prediction models for combined heat,

air and moisture transfer in building components: benchmarks for one-dimensional cases, Journal of thermal envelope and building science 27 (4) (2004) 327–352.
[43] EN 15026. (2004). Hygrothermal performance of building components and building elements - assessment of moisture transfer by numerical simulation.
[44] M. Nofal, M. Straver, K. Kumaran, Comparison of four hygrothermal models in terms of long-term performance assessment of wood-frame constructions, in: Proceedings of the 8th conference of Canadian building science solutions to moisture problems in building enclosures, Toronto, Canada, 2001, pp. 118–138.
[45] M. Qin, R. Belarbi, A. Aït-Mokhtar, L.O. Nilsson, Coupled heat and moisture transfer in multi-layer building materials, Construction and building materials 23 (2) (2009) 967–975.
[46] M. Qin, A. Aït-Mokhtar, R. Belarbi, Two-dimensional hygrothermal transfer in porous building materials, Appl Therm Eng 30 (16) (2010) 2555–2562.
[47] L. Corcoran, A. Duffy, S. Rouholamin, A hygrothermal analysis of international timber frame wall assemblies tested under temperate maritime climatic conditions, in: Proceedings of Central European Symposium for Building Physics 2013, Vienna, Austria, 2013, pp. 1–8.
[48] S. Sassi, H. Ge, M. Horvat, Validating of wind-driven rain modules for HAM Tools, in: Proceedings of Thermal Performance of the Exterior Envelopes of Buildings XII, Florida, ASHRAE, 2013, pp. 1–9.
[49] B. Hejazi, N.R. Sakiyama, J. Frick, H. Garrecht, Hygrothermal simulations comparative study: assessment of different materials using WUFI and DELPHIN software, in: Proceedings of the 16th IBPSA International Conference, Rome, Italy, 2019, pp. 2–4.
[50] D. Chung, J. Wen, L. James Lo, Development and verification of the open source platform, HAM-tools, for hygrothermal performance simulation of buildings using a stochastic approach, Build Simulation 13 (3) (2020) 497–514.
[51] M. Defo, M. Lacasse, A. Laouadi, A comparison of hygrothermal simulation results derived from four simulation tools, Journal of Building Physics 45 (4) (2022) 432–456.
[52] X. Dang, H. Janssen, S. Roels, Hygrothermal modelling of one-dimensional wall assemblies: inter-model validation between WUFI and DELPHIN, in: Journal of Physics: Conference Series 2654, IOP Publishing, 2023, 012040.
[53] X. Dang, H. Janssen, S. Roels, A comprehensive benchmark dataset for the validation of building component heat, air, and moisture (HAM) models, Building Simulation 17 (11) (2024) 1893–1907.
[54] X. Dang, H. Janssen, S. Roels, Empirical validation of HAM-models based on a dedicated HB-CB experiment_datasets [Data set]. Zenodo. doi:10.5281/zenodo.10998834.
[55] Empirical validation of HAM-models based on a dedicated HB-CB experiment. (2023). Call for participation. doi:10.13140/RG.2.2.10053.83682.
[56] X. Dang, H. Janssen, S. Roels, Comparative simulations on hygrothermal performance of calcium silicate and wood fiber as capillary active internal insulation materials, in: Proceedings of the 2nd International Conference on Moisture in Buildings, London, UK, 2023, pp. 1–2.
[57] H. Hens, (1991). IEA Annex 24: heat, air and moisture transfer through new and retrofitted insulated envelope parts (Hamtie).
[58] S. Roels, IEA Annex 41 Whole Building Heat, Air, Moisture Response. Experimental Analysis of Moisture Buffering, KU Leuven, Belgium, 2008.
[59] H. Janssen, B. Blocken, J. Carmeliet, Conservative modelling of the moisture and heat transfer in building components under atmospheric excitation, Int J Heat Mass Transf 50 (5–6) (2007) 1128–1140.
[60] X. Dang, E. Vereecken, H. Janssen, S. Roels, Impact of semi-empirical methods implemented in heat, air, and moisture (HAM) models on predicted wind-driven rain (WDR) loads and hygrothermal responses, Build Environ (2024) 111770.
[61] T. Busser, J. Berger, A. Piot, M. Pailha, M. Woloszyn, Comparison of model numerical predictions of heat and moisture transfer in porous media with experimental observations at material and wall scales: an analysis of recent trends, Drying Technology (2018).
[62] E. Vereecken, S. Roels, Hygric performance of a massive masonry wall: how do the mortar joints influence the moisture flux? Construction and Building Materials 41 (2013) 697–707.
[63] M. Maaroufi, F. Bennai, R. Belarbi, K. Abahri, Experimental and numerical highlighting of water vapor sorption hysteresis in the coupled heat and moisture transfers, Journal of Building Engineering 40 (2021) 102321.
[64] P. Ren, C. Feng, H. Janssen, Hygric properties of porous building materials (V): comparison of different methods to determine moisture diffusivity, Build Environ 164 (2019) 106344.
[65] ASHRAE. (2014). Guideline 14 measurement of energy, demand, and water savings, 4, 1–150.
[66] D. Chakraborty, H. Elzarka, Performance testing of energy models: are we using the right statistical metrics? Journal of Building Performance Simulation 11 (4) (2018) 433–448.
[67] H.E. Huerto-Cardenas, F. Leonforte, N. Aste, C. Del Pero, G. Evola, V. Costanzo, E. Lucchi, Validation of dynamic hygrothermal simulation models for historical buildings: state of the art, research challenges and recommendations, Build Environ 180 (2020) 107081.
[68] Paolo, B., Roberta, P., & Alessandro, P. (2013). Simulazione energetica degli edifici esistenti: guida alla definizione di modelli calibrati. EPC srl. https://www.epc.it/Prodotto/Editoria/Libri/libro-Simulazione-energetica-degli-edifici-esistenti/2003. (In Italian).
[69] ANSI/ASHRAE. Guideline 14-2002 measurement of energy and demand savings, 2002.
[70] DELPHIN. Outputs in DELPHIN 6.1. https://www.bauklimatik-dresden.de/delphin/2nd/doc/DELPHIN6_Outputs/Tutorial_4_D6_Outputs_en.html.
[71] M. Matsumoto, PhD thesis, Kyoto National University, Japan, 1978.
[72] C. Feng, H. Janssen, Hygric properties of porous building materials (VII): full-range benchmark characterizations of three materials, Build Environ 195 (2021) 107727.
[73] Wufi-Wiki. Details: liquid transport coefficients. https://www.wufi-wiki.com/mediawiki/index.php/Details:LiquidTransportCoefficients.
[74] H.M. Künzel, Simultaneous Heat and Moisture Transport in Building components. One-and two-Dimensional Calculation Using Simple Parameters, IRB-Verlag, Stuttgart, 1995, p. 25.
[75] WUFI forum. Liquid transport coefficient. https://www.wufi-forum.com/viewtopic.php?f=32&t=1923&p=5654&hilit=resistance+factor#p5654.