

# Detailed cross comparison of building energy simulation tools results using a reference office building as a case study

**Citation for published version (APA):**

Magni, M., Ochs, F., de Vries, S. B., Maccarini, A., & Sigg, F. (2021). Detailed cross comparison of building energy simulation tools results using a reference office building as a case study. *Energy and Buildings*, 250, Article 111260. <https://doi.org/10.1016/j.enbuild.2021.111260>

**DOI:**

[10.1016/j.enbuild.2021.111260](https://doi.org/10.1016/j.enbuild.2021.111260)

**Document status and date:**

Published: 01/11/2021

**Document Version:**

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

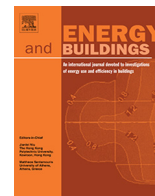
[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.



# Detailed cross comparison of building energy simulation tools results using a reference office building as a case study



Mara Magni<sup>a,\*</sup>, Fabian Ochs<sup>a</sup>, Samuel de Vries<sup>b</sup>, Alessandro Maccarini<sup>c</sup>, Ferdinand Sigg<sup>d</sup>

<sup>a</sup> University of Innsbruck, Unit for Energy Efficient Buildings, Innsbruck, Austria

<sup>b</sup> Eindhoven University of Technology, Eindhoven, the Netherlands

<sup>c</sup> Aalborg University Copenhagen, Denmark

<sup>d</sup> Rosenheim Technical University of Applied Sciences, Germany

## ARTICLE INFO

### Article history:

Received 18 February 2021

Revised 28 May 2021

Accepted 4 July 2021

Available online 7 July 2021

### Keywords:

Statistical indices

Computational cost

Modelling approaches

Cross comparison

Building Simulation

## ABSTRACT

Building Energy Simulation (BES) tools play a key role in the optimization of the building system during the different phases, from pre-design through commissioning to operation. BES tools are increasingly used in research as well as in companies. New BES tools and updated versions are continuously being released. Each tool follows an independent validation process but rarely all the tools are compared against each other using a common case study. In this work, the modelling approaches of widespread dynamic simulation tools (i.e. EnergyPlus, TRNSYS, Simulink libraries CarnotUIBK and ALMABuild, IDA ICE, Modelica/Dymola and DALEC), as well as PHPP (a well-known quasi-steady-state tool), are described and the results of all the tools modelling the same characteristic office cell, defined within the IEA SHC Task 56, are compared on a monthly and hourly basis for the climates of Stockholm, Stuttgart and Rome. Unfortunately, different tools require different levels of input detail, which are often not matching with available data, hence the parametrization process highly influences the quality of the simulation results. In the current study to evaluate the deviation between the tools, frequently used statistical indices and normalization methods are analysed and the problems related to their application, in a cross-comparison of different tools, are investigated. In this regard, the deviation thresholds indicated by ASHRAE Guideline 14-2014 are used as a basis to identify results that suggest an acceptable level of disagreement between the predictions of a particular model and the outcomes of all models. The process of reaching a good agreement between all tools required several iterations and great effort on behalf of the modellers. To aid the definition of building component descriptions and future references for inter-model comparison a short history of the executed steps is presented in this work. Together with the comparison of the results of the tools, their computational cost is evaluated and an overview of the modelling approaches supported by the different tools for this case study is provided aiming to support the users in choosing a fit-for-purpose simulation tool.

© 2021 Elsevier B.V. All rights reserved.

## 1. Introduction

A large number of Building Energy Simulation (BES) tools, with different focus and level of detail, have been developed in the last six decades [1,2]. A comprehensive list of BES is provided in [3] and [4]. BES tools are used in research and increasingly in building design, construction, commissioning and operation for accelerating and improving the design and planning process, optimizing building performance, developing building controls, testing new products and evaluating the market potential of novel concepts.

\* Corresponding author at: Technikerstraße 13, 5. Stock, A-6020, Innsbruck, Austria.

E-mail address: [Mara.Magni@uibk.ac.at](mailto:Mara.Magni@uibk.ac.at) (M. Magni).

Moreover, ever-stricter building codes and energy standards have stimulated the usage of BES [5].

### 1.1. Model complexity: approximation versus estimation

The two main aspects under which the different tools can be classified are first, the complexity of the mathematical models depending on the purpose and focus of the tool, and second, the possibility to access the BES source code (which is particularly relevant in the research field). When the source code is accessible and can be easily modified, the user can tailor the equations to

## Nomenclature

### Acronyms and abbreviations:

av	Average
av > 0	Average including only values higher than zero
BC	Boundary Conditions
BES	Building Energy Simulation
CD	Cooling Demand
DAL	DALEC
EP	EnergyPlus
h	Hourly
HD	Heating Demand
HTC	Heat Transfer Coefficient
Inf	Infiltration losses
iqr	Interquartile range
m	Monthly
MAE	Mean Absolute Error
MBE	Mean Bias Error
ME	Mean Error
MOD	Modelica
nm	Normalization means
NMAE	Normalized Mean Absolute Error
NMBE	Normalized Mean Bias Error
NRMSE	Normalized Root Mean Square Error
PHPP	Passive House Planning Package
RMSE	Root Mean Square Error
ROM	Rome
SHGC	Solar Heat Gain Coefficient
SIM_BO	Simulink library developed by the University of Bologna (ALMABuild)
SIM_IBK	Simulink library developed by the University of Innsbruck (CarnotUIBK)
SOL	Solar gains
std	Standard deviation
STO	Stockholm
STU	Stuttgart
TMY	Typical Meteorological Year
TR	Transmission losses
TRN	TRNSYS
TZ	Thermal Zone
Vent	Ventilation losses

### Variables and Parameters:

F	View Factor [-]
H	Radiosity [W/m <sup>2</sup> ]
I	Solar irradiation [W/m <sup>2</sup> ]
K	Constant heat exchange coefficient [W/(m <sup>2</sup> K)]
N	Number of considered data points [-]

$\dot{Q}$	Heat flux [W/m <sup>2</sup> ]
r	Reference results
R	Thermal resistance [W/(m <sup>2</sup> K)]
R <sup>2</sup>	Coefficient of Determination [%]
s	Simulated results
T	Temperature [K]
v	Wind Velocity [m/s]
$\alpha$	Absorption factor [%]
$\beta$	Constant [-]
$\varepsilon$	Emissivity [%]
$\vartheta$	Temperature [°C]
$\sigma$	Stefan–Boltzmann constant $5.670\,373\,(21) \times 10^{-8}$ [W/(m <sup>2</sup> K <sup>4</sup> )]
$\tau$	Solar transmittance [%]

### Subscripts:

amb	Ambient
av	Average
av > 0	Average considering only numbers higher than zero
C	Convective
CD	Cooling Demand
ce	Convective external
ci	Convective internal
ext	External
g	Global
gnd	Ground
h	Hourly
HD	Heating Demand
hor	Horizontal
i	i <sup>th</sup> time step
inf	Infiltration
int	Internal
m	Monthly
p1	Point one
p2	Point two
R	Radiative
re	Radiative external
ri	Radiative internal
se	Surface external
si	Surface internal
sky	Sky
sol	Solar
south	South
tr	Transmission
vent	Ventilation
win	Windows

perfectly suit the specific case study increasing the degree of detail<sup>1</sup>. Unfortunately, this will also lead to an increase in the total simulation time, modelling effort and skill that is required. Dynamic simulation tools such as TRNSYS, Modelica/Dymola, IDA ICE, EnergyPlus, and Simulink, focus on the transient behaviour of systems and allow detailed analysis. However, their application requires a high

level of user expertise and expert knowledge of many input parameters. Moreover, the modelling and simulation process can be time-consuming in terms of computational cost [5] and modeller effort. In addition, the suitability of a particular level of model complexity, or a particular modelling tool, can only be assessed in relation to the objectives and constraints of the simulation study. In the present study, in order to develop a meaningful conclusion to a specific problem, both the degree of detail of the model as well as the accuracy of the available input data are the defining aspects. As a result, the potential overall error in performance predictions is a function of both the degree of approximation of the physical phenomena involved as well as the degree of estimation of uncertain input parameters. Therefore, the process of minimization of the overall

<sup>1</sup> This is possible only in the case of white and grey box models and not in black-box models [81]. White box models require detailed knowledge of the physical process while black box models do not require full knowledge of the system and are developed using a data-driven approach. Grey box model preserve the physical description of the system, but their parameters are estimated using system identification methods [81].

error involves balancing the complexity of the model with the information, time and other resources available to the simulation study [6].

In general, the process of model parametrisation is error-prone since it is easy to make unrepresentative assumptions or just commit user errors in preparing building input files. Fortunately, the number of inputs to be parametrized can be reduced by using a tool with the degree of detail required by the modelling phase. Typing errors might be avoided by increasing the usage of automated procedures. In this regard, Building Information Modelling (BIM) might have the capability to lead to a more automated process and thus reduce user mistakes. Nevertheless, further development in terms of availability and agreement in information transfer is required to apply BIM to BES in practice (see [7;8]).

In order to assess the extent to which different simulation tools fit the need of a particular simulation study, modellers require an overview of the complexity or degree of abstraction of the component models that are applied in the different tools as well as the input parameters, skill effort and time that are required to use them. In this context, Crawley et al. [3] comprehensively and extensively reported the modelling features for a large number of available tools, which helps the BES user to select the right tool according to the purpose. However, this kind of BES tools comparison needs to be updated since new component models are constantly being developed. In addition, new versions and tools are available such as Matlab/Simulink toolboxes (e.g. CARNOT toolbox [9], CarnotUIBK [10] and ALMABuild [11]), Modelica libraries [12], DALEC [13], PHPP [14]. Unfortunately, using a tool with a high degree of detail might not only be time-consuming to create the simulation model, but also for the computation itself. Certainly, some particular aspects can be analysed only if the model describes them. For example, a detailed comfort analysis can only be done if the models calculate the temperature distribution in the room [15]. The aspect of the computational cost has been qualitatively mentioned in [16–19]. Quantitative comparisons in terms of the computational time of different tools have been proposed in [20] including DALEC, Radiance, Relux and EnergyPlus, in [21] comparing eQUEST against EnergyPlus, in [22] including EnergyPlus and DOE2 and in [23] where TRNSYS and Modelica have been compared. Yet, this is an important topic especially in the expanding field of simulation-based optimization [24] and studies presenting updated comparison of the computational cost including a wide range of BES tools are needed.

The results of such a comprehensive analysis that also show the impact on the computational time, would be a helpful instrument for supporting the user in the decision of which tool to be used for a certain study.

### 1.2. Validation and trust in BES

According to Feist [25], the five major reasons for the deviations between different simulation models are: (1) different algorithms, (2) numerical errors (errors in the calculation), (3) programming errors (errors in implementation), (4) non-identical inputs, and (5) different processing of the weather data that is used by the BES. Additionally, the results are influenced by different physical approaches with different levels of detail and different numerical schemes. As an example, the thermal zone can be modelled with a 2-star or star-node approach and the wall model with the finite difference method or transfer function. To gain trust in the building simulation, it is important to validate the models. Therefore, many validation studies have been published over the years and the level of attention on this subject has increased significantly in the recent years. According to Judkoff et al. [26] and [27], the results of BES tools can be validated against measurements (empirical validation), analytical solution (analytical validation), or against other

codes (cross-validation). In this context, an analytical solution is only available for simple configurations (e.g. wall model and verification of system simulations [7,8]). Empirical validations involving different tools are presented in [16,28], and [29]; this kind of validation has to deal with the measurement uncertainties and is usually limited to a specific amount of data for a limited period of time. Whereas cross-validation allows to minimize the input uncertainties and allows analysis of many different configurations with a high level of detail [5,30,31,32]. Nevertheless, the main drawback of the cross-validation approach is that it is difficult to precisely define the set of reference results.

The BESTEST methodology, established by the ANSI/ASHRAE Standard 140–2017 [33] and later updated by [34], EN 15255 [35] and EN 15265 [36] and German guidelines such as VDI 6020 [37], VDI 2078 [38] and VDI 6007 [39], describe different test cases against which the implemented algorithms or models have to be tested to correct code errors, modelling limitations or input errors eventually present in the tool [40]. Comparative values are given for the evaluation of the calculated results. However, only specific parameters are analysed within each BESTEST case. The spread between the minimum and maximum thresholds, suggested by BESTEST for the tool validation, can be as large as 47% [34] in some test cases.

Moreover, in both empirical and cross-validations, when only heating and cooling demands are considered and all the other contributions are discarded of the thermal balance (i.e., thermal losses, solar gains, internal gains, ventilation and infiltration losses and temperatures of the surfaces and air), it is difficult to interpret the results since all possible error sources are acting simultaneously [40]. Thus, conclusions about the model accuracy have to be derived, preventing as much as possible the offsetting errors by considering as many as possible outputs and different time scales. This is difficult to achieve with empirical validation since the available measurements refer to a limited time frame and do not cover the majority of the outputs of a simulation tool. For these reasons in the current work a detailed cross-comparison, focused on a solar-driven, reference office test cell that includes shading and ventilation control logic, is carried out that considers all the components of the energy balance on a monthly and hourly basis for one year of simulation.

### 1.3. Statistical indices for the quantification of the deviations among the results of the different tools

The definition of accuracy indices is required to quantify the deviations between time series and therefore to assess the goodness of the results of a model against other simulations or measured data. As highlighted in [27], there is the need to find appropriate system performance indices. In this field, Uniform Methods Project [41] and ASHRAE Guideline 14–2014 [42] are widely recognized guidelines aiming at establishing a method for measuring the accuracy of building models as well as the Federal Energy Management Program (FEMP) [43] and the International Performance Measurement and Verification Protocol (IPMVP) [44], which refer to ASHRAE Guideline 14–2014 [42]. These documents suggest thresholds, which are different for monthly or hourly calibration, for the Normalized Mean Bias Error (NMBE) and Normalized Root Mean Square Error (NRMSE). In addition, it has been suggested to analyze a time frame of at least one year and to compare the simulation results with the utility bills and/or spot measurements. However, as highlighted in [45] the indoor condition and temperatures are not addressed in the suggested calibration approaches. Numerous papers using calibration methods are available in the literature, as reported in [45] nevertheless, varying approaches have been applied by different researchers in the previous works due to lack of a standard procedure. These

approaches use different indicators, applied to different variables with varying resolution and time frames to quantify the model accuracy. As an example [16] and [29] calculate the statistical indices (i.e. Coefficient of determination ( $R^2$ ), Root Mean Square Error (RMSE) and NRMSE are used in [16]; Mean Error (ME), RMSE, standard deviation and maximum error are used in [29]) based on temperatures, while [46] bases its analysis on the energy consumption of the building calculating the Mean Bias Error (MBE) and the NRMSE for each month and the whole year using hourly data. In general, it is not uncommon to find literature referring to the same index with different names or definitions as highlighted in [47]. In the current work, the main statistical indices and normalization means that are used are critically analysed to help shed some light on the difficulties that are encountered so a new normalization method can be proposed.

As reported in [4] a long list of BES tools exist, which includes tools from different countries. However, in the current work, EnergyPlus [48], TRNSYS [49], Matlab/Simulink (ALMABuild [11] and CarnotUIBK [10]), the Modelica/Dymola buildings library [50], and IDA ICE [51] are analyzed along with DALEC [20] and PHPP [14], which can be classified as pre-design tools. Almost all these tools have already undergone a validation process following the BESTEST method (i.e. EnergyPlus in [52], ALMABuild [11], Modelica [53], IDA [54], TRNSYS [34]), and all the tools except the Simulink libraries ALMABuild and CarnotUIBK have been compared against measured data in [16,28,29,20] and [55]. However, they have not been compared against each other using a solar-driven building as a common test case. Finally, only a few recent studies have presented a quantitative and comprehensive comparison in terms of computational time including a wide variety of dynamic building simulation tools.

With each tool, the reference office cell was modelled starting from the same description (specially tailored on the TRNSYS input). Even though almost all the tools are successfully validated using the BESTEST methodology (i.e. EnergyPlus in [52], ALMABuild [11], Modelica [53], IDA [54], TRNSYS [34]), high deviations between the results of the analysed tools were experienced during the first iterations of the cross-comparison process. Indeed, to reach a good agreement between the tools, several iterations were necessary due mainly to a lack of equivalent inputs in the building description, user mistakes and different modelling approaches of the tools under analysis.

## 2. Methodology

In this section, the building model used for this case study is described, and the modelling features of the different tools are introduced and the parametrization process is explained. Moreover, the statistical indices are described and the boundary conditions, applied for the evaluation of the computational cost, are reported.

### 2.1. Boundary conditions

Three different locations Rome (ROM), Stuttgart (STU) and Stockholm (STO), characterized by different European climates (Mediterranean, oceanic climate and humid continental climate respectively, according to the Köppen climate classification [56]) were considered in this study. Table 2.1 shows the annual average ambient temperature ( $\bar{\vartheta}_{amb,av}$ ), annual global irradiation on a horizontal surface ( $I_{g,hor}$ ) and annual irradiation on a south-oriented vertical surface ( $I_{south}$ ) for each location.

All the tools use the same Typical Meteorological Year (TMY2) for each location. Therefore, the ambient temperature and global irradiation on the horizontal were the same (see [57] for more

**Table 2.1**

Main boundary conditions: yearly average ambient temperature ( $\bar{\vartheta}_{amb,av}$ ), yearly global irradiation on a horizontal surface ( $I_{g,hor}$ ) and yearly irradiation on a south-oriented vertical surface ( $I_{south}$ ).

Location	$\bar{\vartheta}_{amb,av}$ [°C]	$I_{g,hor}$ [kWh/m <sup>2</sup> ]	$I_{south}$ [kWh/m <sup>2</sup> ]
Rome	15.8	1632	1253
Stuttgart	9.9	1101	889
Stockholm	7.8	952	884

details). Nevertheless, it is important to highlight that DALEC calculates the  $I_{south}$  using a diffuse isotropic sky model<sup>2</sup> described in [58] whereas all the other tools implement the Perez model (i.e. EnergyPlus, Simulink, IDA ICE and Modelica used the Perez 1990 model [59] while TRNSYS used the Perez 1999 model (see [60] page 7–99). Therefore, Simulink, EnergyPlus, TRNSYS, Modelica and IDA ICE present negligible deviations in terms of  $I_{south}$ , while DALEC underestimates  $I_{south}$  in all the climates (see Figure 2.5). To perform a fair comparison, the solar gains of DALEC have been aligned with the solar gains of the other tools by calibrating the shading control thresholds (see Section 2.4).

### 2.2. Building description

The reference building was chosen to be representative of a typical European office cell located on the middle floor of a high-rise building. Figure 2.1 shows the office cell, which has a heated area of 27 m<sup>2</sup> and a volume of 81 m<sup>3</sup>. All the surfaces are considered adiabatic, except for the façade oriented towards the South (with a window-to-wall ratio of 60%) where ambient boundary conditions are applied and solar active technologies such as daylighting systems can be installed<sup>3</sup>. Shadings from adjacent obstacles were not considered, whereas external movable shading, able to block 70% of the incoming radiation, was activated when direct solar radiation impinging the south façade was higher than 120 W/m<sup>2</sup>.

Table 2.2 reports the heat transfer coefficient (HTC) of the opaque wall element and the characteristics of the windows such as HTC, Solar Heat Gain Coefficient (SHGC) and the solar transmittance ( $\tau_{sol}$ ) for the three climates. The HTC of the window can be calculated as the area-weighted HTC of the frame and the glass. The internal walls were typical plasterboard walls, while the exterior wall consists of a three-layer structure with different insulation thicknesses depending on the climate.

User behaviour (e.g., occupancy, appliances and lighting) was taken into account by employing hourly resolution profiles that show different user behaviour for weekday and weekends [61] (see Figure 2.2).

The natural infiltration rate was assumed to be constant and the air change per hour was set to 0.15 ACH. A fresh air supply of 40 m<sup>3</sup>/h per person was provided by a mechanical ventilation system with a sensible heat recovery of 70% efficiency. A bypass of the heat recovery was activated when the temperature of the zone was higher than 23 °C and the ambient temperature was lower than the indoor temperature (see Figure 2.3).

A simplified ideal all-air heating and cooling system was used by all the models. The set-point for the indoor convective temperature applied for the heating and cooling control was

<sup>2</sup> A 'uniform sky' defines a sky distribution which has a constant luminance over the hemisphere. This isotropic distribution is therefore independent on the height and azimuth of the sun and is only scaled with respect to the intensity. In this model the diffuse solar irradiation on a vertical surface is equal to the diffuse solar irradiation on the horizontal divided by 2.

<sup>3</sup> Not applied in the present comparison. A preliminary study including the comparison of SIM\_BO, EP and TRN modelling the office cell equipped with photovoltaic panels and air to air heat pump is presented in [82] Section 2.1.5.

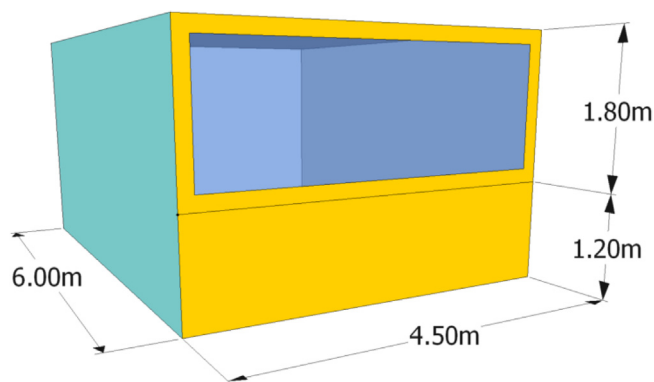


Figure 2.1. Representation of the reference office zone.

21 °C and 25 °C, respectively. When the convective temperature was between 21 °C and 25 °C neither the cooling system nor the heating system is activated.

A summary of the building data and boundary conditions is reported in the appendix in Table C1 and further information is reported in [57].

### 2.3. Modelling features of the tools

The tools analysed in this work had different modelling features as follows:

- EnergyPlus™ v.9.3 (EP) is a building energy simulation program that is used to model energy consumption, lighting, plug and process loads and water use in buildings [62]. The development of the program was funded by the United States Department of Energy’s (DOE) Building Technologies Office and managed by the National Renewable Energy Laboratory (NREL). The software is open-source and is collaboratively developed by NREL, various other DOE national laboratories, academic institutions and private firms;
- TRNSYS 18 (TRN) is a commercial transient system simulation program, developed at the University of Wisconsin, based on a component approach with a modular structure. The TRNSYS library includes a detailed multi-zone building model and components for HVAC systems, renewable energy systems, etc. [49];
- CarnotUIBK (SIM\_IBK) and ALMAbuild (SIM\_BO) are two open-source Matlab/Simulink libraries, compatible with CARNOT Toolbox [9], developed by the University of Innsbruck (AT) [10] and the University of Bologna (IT) [11] respectively;
- IDA ICE v.4.8 (IDA) is a commercial software, developed by EQUA Simulation AB. It is focused on detailed and dynamic multi-zone simulation applications for the study of thermal indoor climate as well as the energy consumption of the entire building involving envelope, HVAC systems, plant and control strategies [63]. The model inputs are described using the equation-based Neutral Model Format language [64];

Table 2.2  
Main properties of the south-oriented façade.

Properties	Rome (Italy)	Stuttgart (Germany)	Stockholm (Sweden)
$HTC_{ext,wall}$ [W/(m <sup>2</sup> K)]	0.80	0.40	0.30
$HTC_{win}$ [W/(m <sup>2</sup> K)]	1.26	1.35	0.90
SHGC [%]	0.33	0.59	0.63
$\tau_{sol}$ [%]	0.26	0.43	0.46

- DALEC (DAL) is a free web tool developed by Bartenbach (AT), University of Innsbruck (AT) and Zumtobel (AT). The main focus of DALEC is on combined thermal and lighting building simulations in early design phases [13];
- MODELICA (MOD) is a non-proprietary, object-oriented, equation-based language for modelling complex physical systems developed by Modelica Association. Currently, several Modelica libraries exist for building components and HVAC systems, and these are continuously being upgraded. In this work, the Buildings library v.5.0.1 was used together with the Modelica standard library [50]. The Dymola modeling and simulation environment (v. 2020x) was also used;
- PHPP v.9.1 Passive House Planning Package is a commercial quasi-steady-state calculation tool, developed as a spreadsheet by the Passive House Institute, for use by architects, engineers and planning experts [14].

The different tools implement different models with different levels of detail to approach the numerical solution of the building system using different equations. An extensive description of the equations implemented in TRN, IDA and EP is provided in [16], a description of the mathematical model used in DAL is provided in [20,58] and of MOD in [12]. SIM\_BO and SIM\_IBK do not provide user manuals, while CARNOT toolbox provides only an introduction to the library in [65].

Table 2.3 provides a summary of the approaches applied in this work (in black) and available (in grey) in each tool for different aspects of the building model. The model of the thermal zone is used to calculate the convective and radiative temperature. The most detailed approach would be a CFD simulation considering a distribution of the convective as well as the radiative temperature within the room. Such a model requires a high computational cost and therefore, it is not usually applied in annual energy simulation [66]; however MOD and IDA offer the possibility of using a simplified CFD approach<sup>4</sup> (see [67;68], respectively). EP, TRN, SIM\_BO, IDA and MOD, for convex and closed volume, can model the air in the room as one unique node while calculating the radiative exchange between the surfaces using view factors. EP, TRN, SIM\_BO and IDA can additionally calculate the view factors between the internal surfaces and a matrix of points in the room, whose location is defined by the user, allowing the calculation of the mean radiative temperature field in the room. The two-star node approaches [69] (implemented in SIM\_IBK, SIM\_BO) include a convective node and a radiative node (the long-wave radiative exchange between the surfaces is modelled using the star network). In the simplified calculation mode, TRN implements a star network (see [70] Figure 5.4.1–7) where an artificial temperature node is used to consider the parallel energy flow from the inside wall surface to the zone air by convection and the long-wave radiation exchange between the surfaces. DAL is based on the Standard ISO 13790:2008 [71] where the room thermal balance is solved considering three nodes and both the air and mean radiant temperatures are calculated [20]. In DAL the nodes are connected through a specific coupling conductance defined in the Standard ISO 13790:2008 [71]. The total thermal capacity of the walls and air volume is connected to the node representing the mean radiant temperature. PHPP is a quasi-steady-state tool based on Standard ISO 13790:2008 [71] that calculates losses and gains on monthly basis considering a fixed set point temperature (different for winter and summer energy balances).

The convective heat transfer coefficient (see Table 2.3 and Table 2.4) is considered as a constant (Table 2.4: C1/C2) in PHPP and DAL. Contrariwise EP, TRN, IDA and MOD offer both constant

<sup>4</sup> Not included in the current study.

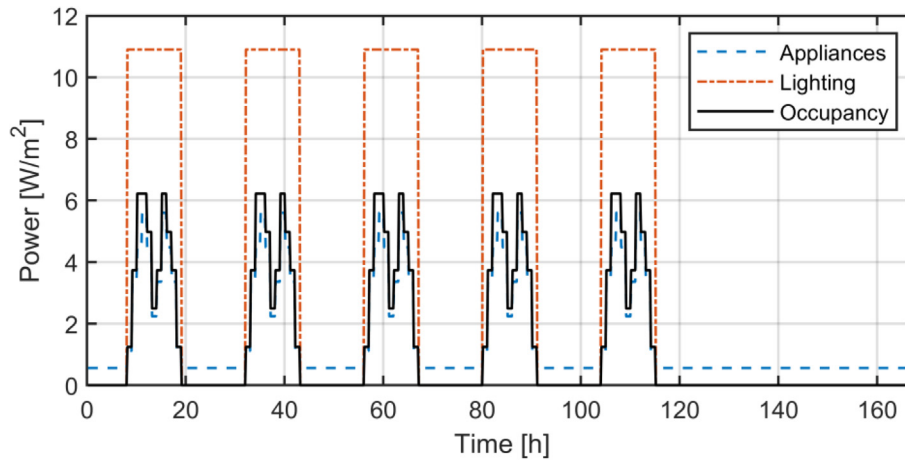


Figure 2.2. Internal gains due to appliances, lighting and occupancy.

(Table 2.4: C1/C2) and equation-based (Table 2.4: C3) calculation. The radiative exchange coefficient can be:

- Constant when an overall HTC is considered (Table 2.4: R1);
- Linearized (Table 2.4: R2);
- Proportional to the temperature difference at the fourth power (Table 2.4: R3); and
- Calculated considering the view factors between the surfaces (Table 2.4: R4).

Different window models are implemented in the analysed tools, in particular, EP, TRN, SIM\_BO, IDA and MOD perform a thermal balance over each pane of the window (see Figure 2.4a), while DAL, SIM\_IBK and PHPP are based on a simplified window thermal model where the transmission losses are calculated by using a constant heat transfer coefficient (see Figure 2.4b).

EP, TRN, SIM\_BO, IDA, MOD consider the solar transmittance and the distribution of absorbed solar radiation across the multi-layer glazing system as a function of the angle of incidence. EP, IDA, MOD explicitly describe this behaviour by recursively solving the transmittance, reflectance and absorptance of solar radiation through each layer, while TRN and SIM\_BO by using pre-processed data of multi-layer window system from the LBNL-Window software [72] (see [73] for the WINDOW technical documentation). A description of the implementation of the detailed window modelling approach in EP can be found in [74] and [75]. EP, MOD and IDA ease the investigation of particular fenestration systems (e.g. including glass coatings) and require more detailed input. SIM\_IBK uses a standard correction function (different for single, double and triple panes windows) of the SHGC at normal incidence, to different incidence angles. DAL discretizes the sky into 145 patches and, for every patch applies a pre-calculated correction factor to the SHGC at normal incidence given by the user (see [20]).

EP and TRN model the opaque structures with the transfer function method, whereas both Simulink libraries, IDA and MOD imple-

ment the finite difference method. For more details about these different approaches see: [69] where the theory behind these methods is explained and [76], which provides a detailed comparison between transfer function and finite difference methods. DAL and PHPP implement a simplified model of the walls, based on the HTC of the structures. PHPP is a quasi-steady-state calculation tool where the capacity of the building is used only to calculate the utilization factor of the internal and solar gains as described in the Standard ISO 13790:2008 [71]. In DAL the mass of the building is lumped in one capacity as described in the Standard ISO 13790:2008 [71].

The model used for the adiabatic structures influences the active capacity of the building. Both Simulink libraries and IDA apply the same boundary conditions (BC) to both sides of the structure, TRN additionally applies the same thermal resistance between the surface and the air and radiative nodes on both sides. EP implements a null thermal flux on the external side of the adiabatic structure and MOD in the middle of the structure. PHPP and DAL consider the thermal mass in one unique node belonging to the Thermal Zone (TZ).

The isotropic sky model, used in DAL [58], assumes that the diffuse radiation from the skydome is uniform across the sky and predicts lower solar radiation availability on the south façade (see Figure 2.5), compared to the other tools implementing the Perez sky model (see Section 2.1).

The internal gain can be defined as an hourly profile in almost all tools except for PHPP where the internal gain is considered as a constant and in DAL where the internal gain can be defined with a constant value throughout the day, as shown in Figure 2.6.

The control status of the shading and ventilation bypass is defined in every time step for all tools except for PHPP where constant shading and ventilation rate are applied for the summer and winter balance calculation (see Section 2.4, Table 2.7 and Table 2.8). To clarify this aspect, the control status of the shading simulated with SIM\_BO for the climate of Stockholm together with the sunset and sunrise time for each day of the year is reported in

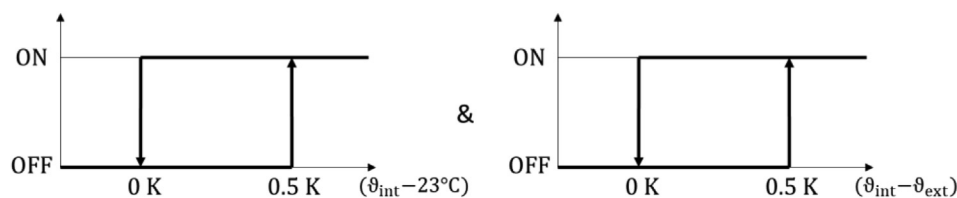


Figure 2.3. Control logic of the ventilation bypass.

**Table 2.3**  
Summary of the features of the mathematical models employed in each tool within this case study (black points) or are available (grey points).

		EP	TRN	SIM_IBK	SIM_BO	IDA	MOD	DAL	PHPP
Thermal zone model	Standard ISO 13790 (monthly)								●
	Standard ISO 13790 (hourly)							●	
	Star node model			●					
	Star network		●						
	Two-star node model			●	●	●			
	One convective node and surf. to surf. radiative exchange	●	●		●	●	●		
Convective heat transfer	More convective nodes and surf. to surf. radiative exchange				●	●	●		
	Constant (see Table 2.4 line C1)	●	●		●		●	●	●
	Internal const.; External prop. to wind speed (see Table 2.4 line C2)		●		●		●		
	Internal f( $\Delta T$ ) (Eq. (2-5)); External Constant (Eq. (2-2))		●				●		
Radiative exchange	Internal f( $\Delta T$ ); External prop. to wind speed (see Table 2.4 line C3)	●	●	●		●	●		
	Radiative node - Constant (see Table 2.4 line R1)	●	●					●	●
	Radiative node - Linearized (see Table 2.4 line R2)				●				
	Radiative node - Proportional to $\Delta T^4$ (see Table 2.4 line R3)		●	●		●			
Window thermal model	Surf. to Surf. - Based on view Factor (see Table 2.4 line R4)	●	●		●	●	●		
	Constant HTC			●				●	●
	Constant thermal resistance (see Figure 2.4b)			●		●			
Window optical model (short wave radiation)	Thermal balance of each pane (see Figure 2.4a)	●	●		●	●	●		
	Isotropic SHGC (see Figure 2.4b)			●		●			●
	Anisotropic SHGC (see Figure 2.4b)			●		●		●	
Shading thermal model	Anisotropic optics + solar absorption at each pane (see Figure 2.4a) <sup>1</sup>	●	●		●	●	●		
	No interaction with the window		●	●	●	●	●	●	●
	Constant additional thermal resistance					●			
Shading optical prop (short wave radiation)	Thermal balance of the air gap (shading - glass)	●	●			●	●		
	Constant reduction factor		●	●	●	●	●	●	●
	Optical properties of the shading layer	●	●			●	●	●	●
Distribution of the solar gains in the TZ	Added directly to the capacity of the TZ		●					●	●
	Proportional to the surface area		●	●					
	Complex calculation based on the reflection of the surfaces	●	●		●	●	●		
Wall model	Constant HTC (no capacity)	●		●				●	●
	Transfer function		●			●			
	Finite difference	●		●	●	●	●		
	Hygrothermal model	●		●		●			
Adiabatic structure model	Capacity of the adiabatic structure added to the building node							●	●
	Null thermal flux outside	●					●		
	Null thermal flux in the middle						●		
	Same BC inside and outside and same heat exchange coeff.		●						
	Same BC inside and outside			●	●	●	●		
Sky model	Isotropic			●	●	●	●	●	●
	Perez	●	●	●	●	●	●		●
Internal gain	Constant	●	●	●	●	●	●		●
	Daily profile	●	●	●	●	●	●	●	
	Hourly profile	●	●	●	●	●	●		
Control of shading and ventilation bypass	Constant	●	●	●	●	●	●	●	●
	Dynamic	●	●	●	●	●	●	●	
Calculation time step	Constant	●	●	●	●	●	●	●	●
	Variable			●	●	●	●		

<sup>1</sup> EP, IDA and MOD require as inputs the solar reflectance for each side of each pane and the solar transmittance of each pane, while TRN and SIM\_BO require only the solar absorption on each pane and the overall solar transmittance calculated using an external subsystem simulation (LBNL-Window [72]).

Figure 2.7. PHPP performs only monthly balances and the shading coefficient can be given only as a constant that might be different for the summer and winter balance (see Table 2.8). Therefore, the graph reported in Figure 2.7 would be completely black for the PHPP and the value represented by the black colour would be the constant shading coefficient given as input to the PHPP. The same applies to the control of the bypass of the ventilation heat recovery.

Concerning the solver settings, the calculation time step can be defined as constant or can be variable in Simulink, MOD and IDA while it is fixed with the time length defined by the user in TRN

and EP. Whereas DAL and PHPP are based on a fixed-time interval of one hour and one month, respectively.

Subportions of the models of the two Simulink libraries could be quite easily exchanged even when this is not foreseen from the structure of the tool itself. The SIM\_BO library contains a detailed window model, which is missing in the original SIM\_IBK library that has only a simplified model with constant thermal resistance. Therefore, in the current work the complex window model developed in SIM\_BO (see Table 2.3, lines: Window thermal model and Window optical model, column SIM\_BO) has been used also in the SIM\_IBK model and the convective and radiative exchange coefficients of SIM\_IBK

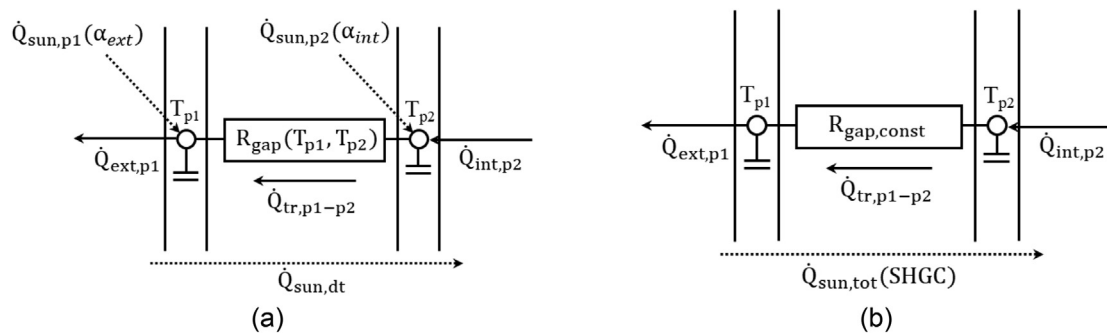


**Table 2.4**

Internal and external convective (subscripts ci and ce) and radiative (subscripts ri and re) exchange equations where: K is a constant heat exchange coefficient [W/(m<sup>2</sup>K)], v is the wind speed [m/s], T and  $\vartheta$  are temperatures in K and °C respectively,  $F_{ij}$  is a view factor between the surfaces i and j, H is the radiosity [W/m<sup>2</sup>],  $\epsilon$  is the emissivity and  $\sigma$  the Stefan-Boltzmann constant 5.670 373 (21) × 10<sup>-8</sup> [W/(m<sup>2</sup> K<sup>4</sup>)]. The subscripts C and R mean convective and radiative, si and se are the internal and external surfaces, amb is ambient and gnd is ground.

	Internal side	Eq.	External side	Eq.
C1	$q_{ci} = K_{ci}(\vartheta_{si} - \vartheta_c)$	2-1	$q_{ce} = K_{ce}(\vartheta_{se} - \vartheta_c)$	2-2
C2	$q_{ci} = K_{ci}(\vartheta_{si} - \vartheta_c)$	2-3	$q_{ce} = (4v + 4)(\vartheta_{se} - \vartheta_c)$	2-4
C3	$q_{ci} = K_{ci}  s_i - c  (s_i - c)$	2-5		
R1	$q_{ri} = K_{ri}(\vartheta_{si} - \vartheta_R)$	2-6	$q_{re} = K_{re}(\vartheta_{se} - \vartheta_{amb})$	2-7
R2	$q_{ri} = 4\epsilon\sigma T_R^3(\vartheta_{si} - \vartheta_R)$	2-8	$q_{re} = [T_{amb}^3 4\epsilon\sigma(\vartheta_{amb} - \vartheta_{se})] + [(T_{sky}^4 - T_{se}^4)F_{sky}\epsilon\sigma]$	2-9
R3	$q_{ri} = \epsilon\sigma(T_{si}^4 - T_R^4)$	2-10	$q_{re} = \epsilon\sigma[F_{gnd}(T_{gnd}^4 - T_{se}^4) + F_{sky}(T_{sky}^4 - T_{se}^4) + F_{amb}(T_{amb}^4 - T_{se}^4)]$	2-11
R4	$q_{ri} = \sum_j F_{ij}(H_i - H_j)^{(1)}$	2-12		

<sup>1</sup>Different methods are applied for the detailed radiative model within the different tools. TRN implements the Gebhart matrix, EP the ScriptF method similarly, in IDA a longwave absorption matrix is used to calculate the net absorbed longwave radiation (these approaches are reported in [16]) and MOD [83] and SIM\_BO [15] are using the radiosity approach.



**Figure 2.4.** Sketch of the window models addressing the solar distribution and the thermal resistance between the panes. In (a) a thermal balance over each pane is performed and the solar gains are distributed over each pane and in (b) a constant thermal resistance is used and the total solar gains are calculated using the SHGC.

**Table 2.5**

Parameter set that needed to be adapted for each tool.

	EP	SIM_IBK	SIM_BO	IDA	MOD	DAL	PHPP
Window optical properties	●				●		
Window HTC				●		●	●
Building Capacity				●		●	●
Ventilation Rate	● <sup>1</sup>			●		●	●
Internal Gain						●	●
Shading control						●	●

<sup>1</sup> In this case study the ideal loads air system model is used in EP. In this model heating and cooling energy are convectively supplied in sufficient quantity to meet zone loads. This model offers default controls for heat recovery and outdoor air delivery that differ from the controls described for this case study (see Section 2.1). EMS scripts were therefore used in EP to control heat recovery bypass and heat gains from ventilation frost protection. More detailed HVAC models could have been selected in EP. However, these would have required further estimation of unknown input parameters.

(see Table 2.3, lines: Convective heat transfer and Radiative exchange, column SIM\_IBK) have been used in SIM\_BO.

2.4. Parametrization process

Unfortunately, the building description, written using the input required by TRN [57], is either too detailed (this is the case for DAL

and PHPP) or too simplified (for EP, IDA, MOD) and therefore equivalent inputs had to be found. The list of inputs that had to be adapted is reported in Table 2.5. Since SIM\_BO uses the same input as TRN both tools do not need any parametrization.

As reported in [77] the window properties and heat transfer model have a substantial influence on the results, which is particularly evident in this case study. The model of the

**Table 2.6**

HTC glass and façade reported in the report, calculated with SIM\_BO and used in DAL, PHPP and IDA.

	REPORT		SIM_BO		IDA	DAL	PHPP
	HTC-glass [W/(m <sup>2</sup> K)]	HTC-façade [W/(m <sup>2</sup> K)]	HTC-glass [W/(m <sup>2</sup> K)]	HTC-façade [W/(m <sup>2</sup> K)]	HTC-glass [W/(m <sup>2</sup> K)]	HTC-façade [W/(m <sup>2</sup> K)]	HTC-façade [W/(m <sup>2</sup> K)]
Rome	1.29	1.08	1.30	1.08	1.21	0.97	1.08
Stuttgart	1.40	0.97	1.59	1.05	1.45	0.98	1.03
Stockholm	0.81	0.66	1.29	0.88	1.22	0.87	0.85

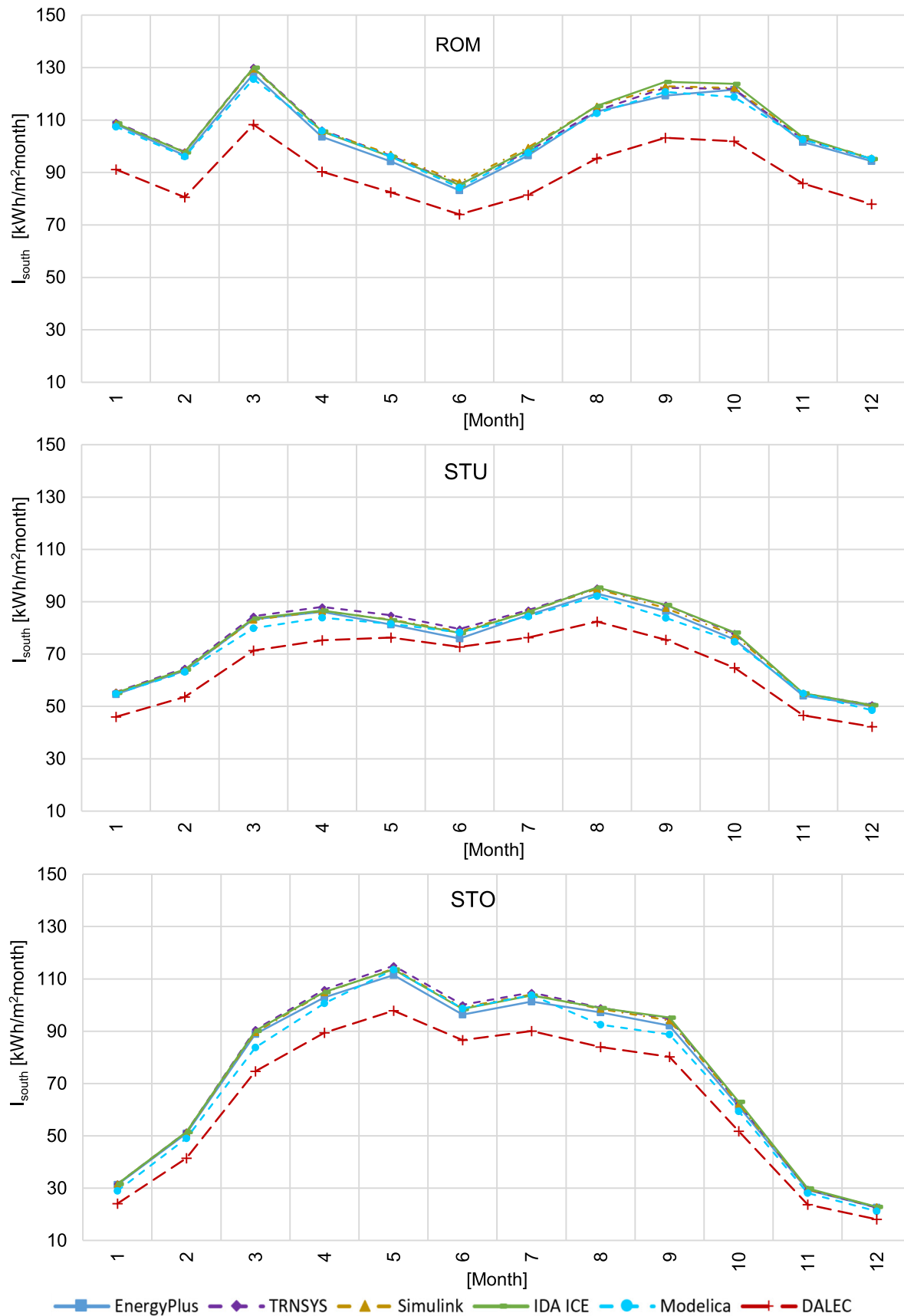
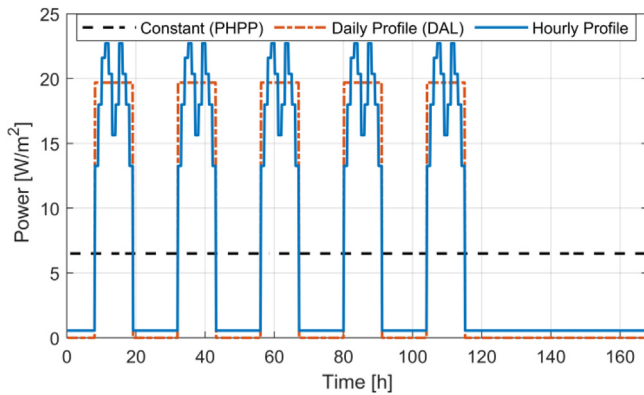


Figure 2.5. Solar irradiation on the south-oriented facade for the climate of Rome (ROM), Stuttgart (STU) and Stockholm (STO).



**Figure 2.6.** Total internal gain profiles, including occupancy, lighting and appliances, used in the PHPP (as a constant average value), in DAL (as a daily profile) and all the other dynamic simulation tools (as hourly profiles).

fenestration system in EP and MOD requires optical and thermal properties to be assigned for each glazing pane. Yet, the description of the input, defined for TRN [57], only describes the optical behaviour of the overall glazing system as a function of the angle of incidence. To obtain the detailed glazing properties for EP and MOD, the solar transmittance and reflectance of the glazing panes were varied and the results were compared with SIM\_BO, which uses the same input as TRN. In EP this was done using an exhaustive search approach and in MOD a trial and error process was followed. The detailed glazing properties were then selected that gave the smallest difference ( $\leq 0.1 \text{ kWh/m}^2$ ) in terms of the predicted annual sums of the solar energy that is transmitted by the glazing system, and the energy that is absorbed at each of the two panes. The resulting inputs that were found for the window models in MOD and EP are the same for ROM while small differences for the window properties in STO and STU are present, as reported in the appendix (see Figure A1).

Table 2.6 reports the HTC of the glass and façade, defined in a report from the current study [57], derived from SIM\_BO and the one used in IDA, DAL and PHPP. SIM\_BO dynamically calculates the HTC of the window since it depends on the boundary conditions, therefore, the constant HTC reported in Table 2.6, for SIM\_BO, is calculated by finding a constant HTC of the glass, which delivers the same annual thermal losses calculated with the dynamic HTC. The discrepancy between the HTC calculated by Simulink and the one indicated in the report [57], which is especially high in Stockholm, is caused by the fact that the report gives this input considering fixed boundary conditions (i.e.,  $\vartheta_{\text{amb}} = 26.7 \text{ }^\circ\text{C}$ , no solar irradiation, Wind speed = 6.71 m/s, Internal temperature = 21 °C).

In IDA, both complex and simplified window models could be used. However, the complex window model has to be used with a complex shading model involved in the thermal balance of the window while the simplified window model can be used with a simplified shading model that does not affect the thermal balance

of the window. The shading device is described just as a reduction factor of the incoming solar radiation impinging the window in this case study [57]. Therefore, all the tools except for EP (where this was not easily modified), do not consider the thermal effect of the shading layer on the thermal balance of the window. Hence, in IDA the simplified approach was used too due to this strong simplification of the fixed shading factor. IDA uses a simplified window model with two panes that use a constant thermal resistance (the fenestration system is always simplified as two-pane glazing). The value of this constant has been derived from the annual dynamic simulation of SIM\_BO.

In IDA it is not possible to increase the capacity of the air node (as suggested in [57]) therefore the thermal mass was artificially increased by adding a wall (i.e., area: 20 m<sup>2</sup>; thickness: 0.05 m; density: 900 kg/m<sup>3</sup>; specific heat capacity: 987 J/kgK) with high thermal conductivity (i.e. 10000 W/mK) and heat transfer coefficient (i.e., 20000 W/m<sup>2</sup>K). Moreover, IDA controls the mass flow rate of the supply and exhaust air while in [50] the volume flow rate is given and specified as constant. Therefore, the efficiency of the heat recovery system was adjusted (+4% compared to [50]) to compensate for the slightly unbalanced volume flows of the supply and exhaust air.

DAL is an open-access online tool where the code cannot be modified by the user. To adapt the DAL calculation to the building description, an unpublished Matlab version of the code was used. The results of the original web DALEC and the calibration process are reported in Section 3.4. Concerning the online version the following parts have been modified:

- The lumped capacity of the model was calibrated by comparing the hourly heating and cooling demand with SIM\_BO;
- Ventilation rate: in the original calculation the ventilation system is always active, while in the calibrated version it is active only during the presence of people (as described in [57]);
- Shading model: in the original version, the shading is modelled considering optical properties dependent on the solar incident angle while in the calibrated version a constant reduction of the incoming solar gain was applied. Given the fact that DAL applies an isotropic sky model, resulting in lower solar irradiation toward the south façade, the shading threshold is modified to make the solar gain comparable; and
- The overall HTC of the window and wall is selected to minimize the difference in transmission losses between DAL and SIM\_BO.

PHPP and DAL base their calculation on a constant HTC for the wall and window therefore, especially for Stockholm, it was necessary to parametrize the (Table 2.6).

Table 2.7 documents the effective volume flow (considering the 70% efficiency of the heat recovery) elaborated by the ventilation unit during the occupied time (from 8 until 18 see Figure 2.2) and the increased volume flow when the bypass is activated. The values used in DAL are found by minimizing the deviation in terms of cooling demand and ventilation losses with SIM\_BO.

PHPP calculation is based on monthly balances therefore only one constant ventilation value for the summer and one for the

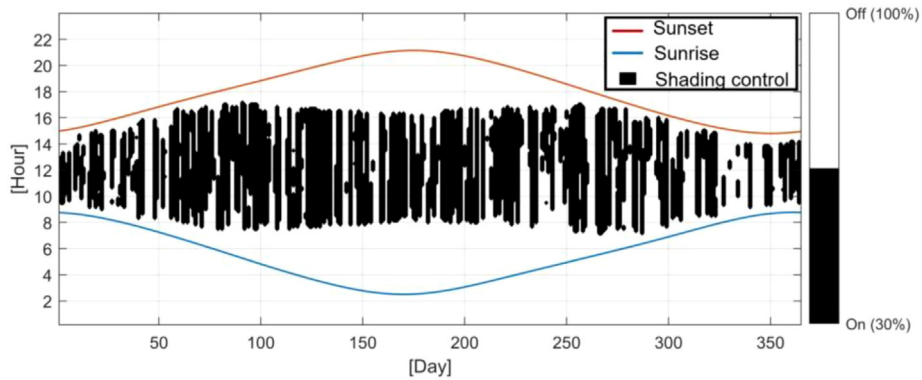
**Table 2.7**

Effective volume flow of the mechanical ventilation system and volume flow when the bypass is activated described in the report and used in DAL and the average volume flow used in PHPP for the winter and summer balance.

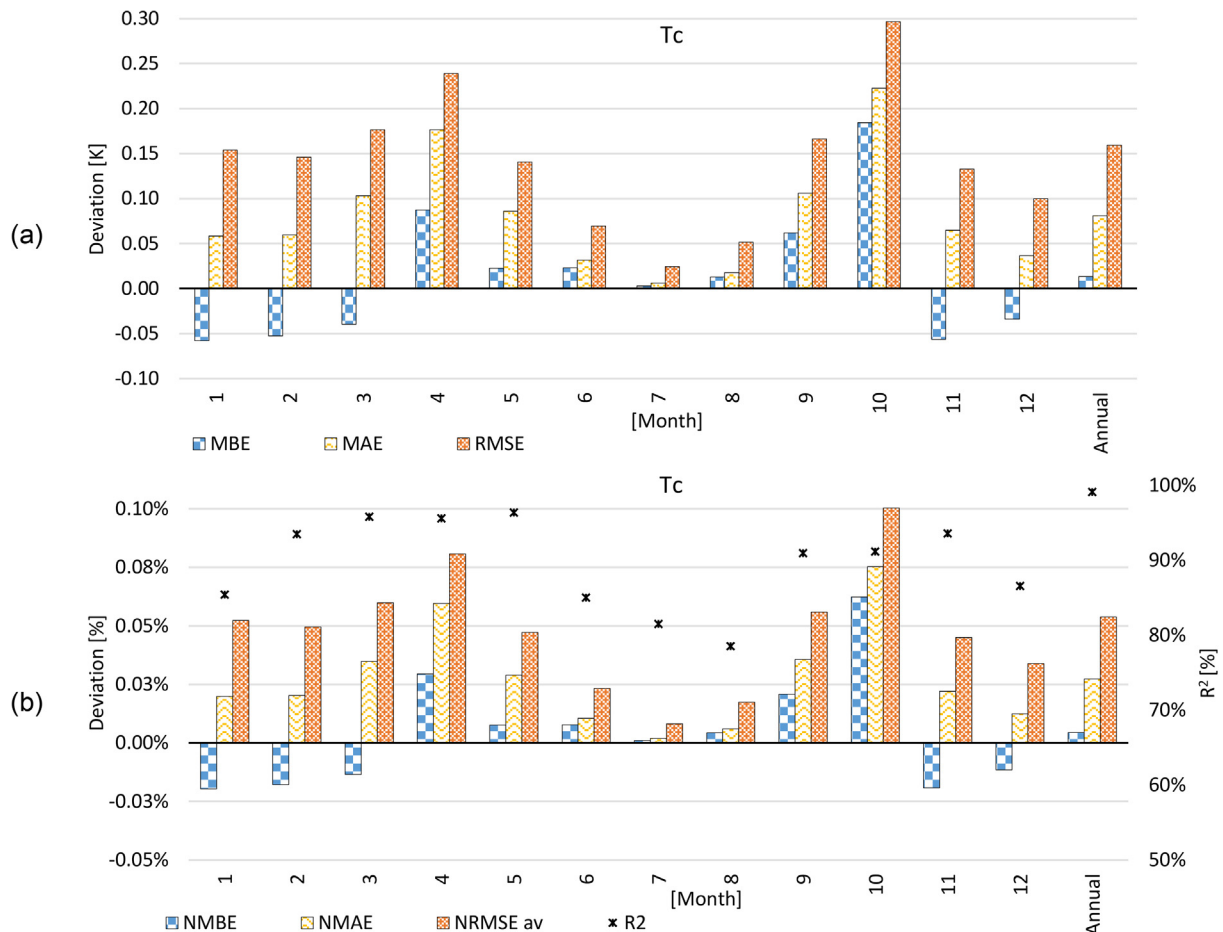
	Report / SIM_BO		DAL		PHPP	
	Effective volume flow [m <sup>3</sup> /h <sub>presence</sub> ]	Volume flow bypass [m <sup>3</sup> /h]	Effective volume flow [m <sup>3</sup> /h <sub>presence</sub> ]	Volume flow bypass [m <sup>3</sup> /h]	Winter [m <sup>3</sup> /h]	Summer [m <sup>3</sup> /h]
Rome	0.44	1.48	0.50	1.46	0.20	0.40
Stuttgart			0.42	1.57	0.15	0.41
Stockholm			0.40	1.48	0.15	0.41

**Table 2.8**  
Winter and summer ratio between solar gains and incident solar radiation derived from Simulink and applied in PHPP.

	SIM_BO		PHPP	
	Winter	Summer	Winter	Summer
	[kWh <sub>solargain</sub> /kWh <sub>incident</sub> ]		[kWh <sub>solargain</sub> /kWh <sub>incident</sub> ]	
Rome	0.112	0.120	0.107	0.112
Stuttgart	0.254	0.315	0.238	0.295
Stockholm	0.277	0.291	0.265	0.280



**Figure 2.7.** Hourly shading control simulated with SIM\_BO for the climate of Stockholm. The blue and red lines represent sunrise and sunset time, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Figure 3.1.** Non-normalized (a) and normalized (b) statistical indices applied for the calculation of the deviation between the simulated convective temperature with SIM\_IBK and the reference convective temperature calculated as the median value of all tools for each time step, for the climate of Stockholm.

winter can be used, disregarding the occupancy schedule. In practice, describing the ventilation with detailed characteristics (e.g. mechanical ventilation with heat recovery of this case study) might be not possible because it is extremely complicated to guess the corresponding ventilation rates without the results of the dynamic simulation as a reference. On the contrary, if the mechanical ventilation system is not yet defined it might be easier to model it with PHPP instead of tools requiring detailed inputs.

Table 2.8 reports the ratio between the solar gain and the solar radiation impinging the south glass area as a result of the SIM\_BO simulation and the input used in PHPP. The SHGC of the window, as well as the shading and dirt factor, are included in this ratio. The shading coefficient in PHPP is during the winter period 0.60, 0.50, 0.48 in Rome, Stuttgart and Stockholm, respectively and during the summer period 0.58, 0.38, 0.45 in Rome, Stuttgart and Stockholm, respectively.

### 2.5. Statistical indices

In this work, statistical indices were used to quantify the deviations between the results of the different tools. This evaluation was conducted using monthly and hourly datasets for all the components of the energy balance (heating, cooling, ventilation plus infiltration, transmission losses and solar gain) as well as for the convective temperature. Since no absolute reference such as measurements is considered within this work, it is important to define a benchmark against which the results can be compared and the median value of all tools will be used for each parameter.

ASHRAE Guideline 14-2014 [42], FEMP [43] and IPMVP [44], describe a method for the validation of the building model against measurement and suggest limits for the NMBE shown in Eqs. 2-14, NRMSE (using the average as a normalization means) in Eqs. 2-18 and the R<sup>2</sup> in Eqs. 2-19 to verify the accuracy of the models. The calibration criteria given by these standards are: ±5% for the monthly NMBE, 15% for the monthly NRMSE, ±10% for the hourly NMBE, 30% for the hourly NRMSE and > 0.75 for the R<sup>2</sup> ([47], Table 1).

In Eqs. (2-13) to (2-19) r<sub>i</sub> represents the reference value for the i<sup>th</sup> time step, calculated as the median of the results of all the tools in each considered time step (i.e. hourly or monthly), while s<sub>i</sub> is the simulated value for a particular tool at the i<sup>th</sup> time step, N is the number of considered data,  $\bar{r}$  is the average of the reference values (i.e., median of all the tools for each time step) and nm is a normalization means. All the indices reported in Table 2.9 are calculated for each tool using hourly or monthly data. The MBE and its normalized value NMBE are good indicators of the overall bias but suffer from the cancellation effect, therefore at least one additional index is needed. ASHRAE Guideline 14-2014, FEMP and IPMVP suggest using the NRMSE based on the average value as a normalization means. The Mean Absolute Error (MAE) also provides analogous information as the RMSE, but is easier to interpret since

**Table 2.9** non-normalized (Mean Bias Error, Mean Absolute Error, Root Mean Square Error) and normalized statistical indices (Normalized Mean Bias Error, Normalized Mean Absolute Error, Normalized Root Mean Square Error, Coefficient of Determination).

Non-normalized indices		Normalized Indices	
$MBE = \frac{\sum_{i=1}^N (s_i - r_i)}{N}$	(2-13)	$NMBE = \frac{\sum_{i=1}^N (s_i - r_i)}{\sum_{i=1}^N r_i} [\%]$	(2-14)
$MAE = \frac{\sum_{i=1}^N  s_i - r_i }{N}$	(2-15)	$NMAE = \frac{\sum_{i=1}^N  s_i - r_i }{\sum_{i=1}^N r_i} [\%]$	(2-16)
$RMSE = \sqrt{\frac{\sum_{i=1}^N (s_i - r_i)^2}{N}}$	(2-17)	$NRMSE = \frac{1}{nm} \sqrt{\frac{\sum_{i=1}^N (s_i - r_i)^2}{N}}$	(2-18)
		$R^2 = 1 - \frac{\sum_{i=1}^N (s_i - r_i)^2}{\sum_{i=1}^N (r_i - \bar{r})^2}$	(2-19)

each deviation influences linearly the MAE while the RMSE is more sensible to the outliers.

The RMSE is scale-dependent [78] therefore can be used as a measure of accuracy for a specific dataset but not between different datasets. The normalization of the RMSE eases the comparison between datasets with different scales but it should be computed only for data based on a scale with an absolute zero (e.g., Kelvin, not Celsius or Rankine, not Fahrenheit). Although the average has been commonly used and is suggested by [42,43] and [44] as a normalization means, this is not an appropriate method when the dataset contains a large number of zeros since the average is then close to zero and the NRMSE is no longer meaningful as highlighted in [78]. Therefore, it is noteworthy to mention that the normalization process may be based on different normalization means, such as average, range, interquartile range (not affected by outliers) or standard deviation (affected by outliers). Inconsistencies and problematics related to the application of statistical indices are explained in detail in Section 3.1.

### 2.6. Evaluation of the computational cost

For the evaluation of the computational cost, each tool runs the corresponding model on the same local workstation with the following specifications:

- Processor: Intel® Core™ i5-8350U CPU @ 1.7 GHz;
- 4 physical cores;
- 8 logical processors;
- RAM: 16.0 GB;
- OS: Windows 10, 64 bit.

During the simulation, no other applications were running in the background except the operating system. Each computation was repeated 10 times and the median value of the CPU times was used for the comparison. Since the outputs are treated differently in each tool, the computational cost was evaluated including and excluding the processing of the output to analyse the computational cost of the model itself and to be able to address the additional time needed for the preparation of the output. 28 days of pre-simulation time is considered for all the tools. Table 2.10 reports the simulation set-up used in each tool. Figure 3.7 shows the comparison of the runtimes.

## 3. Results and discussion

In this section, inconsistencies and problems related to the application of statistical indices are explained and the monthly and hourly results of all the tools are reported together with their deviations. In addition, the results of the computational cost are presented. Furthermore, the influence of the parametrization process on the results of IDA and PHPP is reported.

### 3.1. Challenges related to the usage of statistical indices

The statistical indices introduced in Section 2.5 were analysed to explain their limits and their application within this work. The literature review highlighted that different approaches are applied for the determination of deviations in the building calibration and validation field and they differ mainly by the statistical indices that are used, and the time frame and reference variables on which they are calculated. Another aspect that is not often mentioned is the normalization of the statistical indices, which can be problematic especially when the normalization approaches zero. To show these aspects, the statistical indices were calculated using the deviations of the convective temperatures and heating power between the

**Table 2.10**  
Simulation settings for the different tools.

	EP (TZ) <sup>1</sup>	TRN	SIM_IBK	SIM_BO	IDA	MOD	DAL
Time step	Fixed	Fixed	Variable	Variable	Variable	Variable	Fixed
Maximum time step [min]	10	10	10	10	10	10	60
Solver	Third-order finite difference approximation	Modified-Euler method	implicit Runge-Kutta (ode23tb)	implicit Runge-Kutta (ode23tb)	See [79]	Cvode	Crank-Nicholson

<sup>1</sup> EP implements different algorithm for the numerical solution of the thermal zone and of the HVAC system. The time step is fixed for zone heat balance calculation and variable for HVAC system simulation.

hourly results of SIM\_IBK and the median of the hourly results of all the tools. The MBE in Eq. (2-13), and MAE in Eq. (2-15), and the RMSE Eq. (2-17) were calculated based on the deviations for the climate for Stockholm; the hourly results of SIM\_IBK with respect to the reference results concerning the convective temperature (see Figure 3.1a); and the heating power (see Figure 3.2a). While the normalized statistical indices such as the NMBE in Eq. 2(-14), the NMAE in Eq. (2-16) and the NRMSE in Eq. (2-18) and the  $R^2$  in Eq. (2-19) were reported in Figure 3.1b applied to the convective temperature, and Figure 3.2b as applied to the heating power.

The normalization of the RMSE was calculated using the average ( $av$ ), average of the absolute value higher than zero ( $av > 0$ ), the range, the interquartile range ( $iqr$ ) and the standard deviation ( $std$ ). In addition, the statistical indices were calculated across different time periods using an hourly resolution. These time periods can vary from a few days (e.g., periods in which measurements are available) [16], to a month [46] to a year [5]. Here the results are presented considering monthly and annual time frames.

Figure B.4 shows the distribution of the absolute deviation of the internal convective temperature for all the tools considering hourly data for a one-year simulation. It can be noticed that the deviation range for SIM\_IBK is  $\pm 0.5$  K and only for 2% of the time which can be considered as a good agreement. Figure 3.5b shows a box plot of the reference convective temperature distribution for each month and considers the data set of the whole year. Figure 3.1a shows the MBE, MAE and RMSE for each month and also considers the data for the whole year. It can be noticed that the MBE is the only indicator giving information about the sign of the deviation and that the MAE and RMSE follow the same trend but the RMSE is always higher than MAE since it is more sensible to the outliers. Figure 3.1b reports the normalized indices and the  $R^2$ , in this case, NMBE, NME and NRMSE\_av are ranging between  $-0.02\%$  and  $+0.1\%$ . The same indices calculated using the temperature in Celsius would vary between  $-0.27\%$  and  $+1.3\%$ . Since the internal convective temperature is always higher than zero, NRMSE\_av is equal to  $NRMSE\_av > 0$ . In this case, it is not possible to normalize the RMSE for the  $iqr$  or  $std$  during the winter and summer months since the temperature is close to the set point most of the time (as highlighted in Figure 3.5) therefore  $iqr$  and  $std$  are set equal to zero. These normalizations could only be used when the hourly data of the whole year were considered. The NRMSE\_range can also be misleading since this index would be higher in months where the temperature is constant (i.e., range close to zero). The  $R^2$  is lower in the months where the temperature is close to the set point (January, June, July, August and December) since in these periods the reference temperature is always near to the average temperature (see Eq. (2-19)).

Figure 3.5 shows the distribution of the reference heating power used for the calculation of the SIM\_IBK deviation.

Figure 3.2 shows the indices considering the hourly data of heating power on monthly and yearly time frames. In addition, one additional column is added considering hourly data for the whole year only when the heating system is active. The monthly absolute deviation indices (i.e, MAE and RMSE) tend to be lower

in April and May where the heating system is working less. The normalization of the statistical indices is also non-trivial for the heating power, in fact in April and October (see Figure 3.2b) where the heating power trends toward zero almost all the normalization means trend to zero as well ( $std$ ,  $iqr$ , and  $av$ ). This leads to really high relative deviations. The only normalization means that allow a stable evaluation of the deviation with relative indices is the range and  $av > 0$ . Another reasonable option might be to exclude transition periods (i.e. April and October) but this would require a subjective definition of the threshold under which the heating demand is too low for the calculation of the deviations. This problem arises also when a building has either a high envelope quality or is placed in a warm climate, i.e. if it has a really low heating demand (e.g. in this study for the climate of Rome).

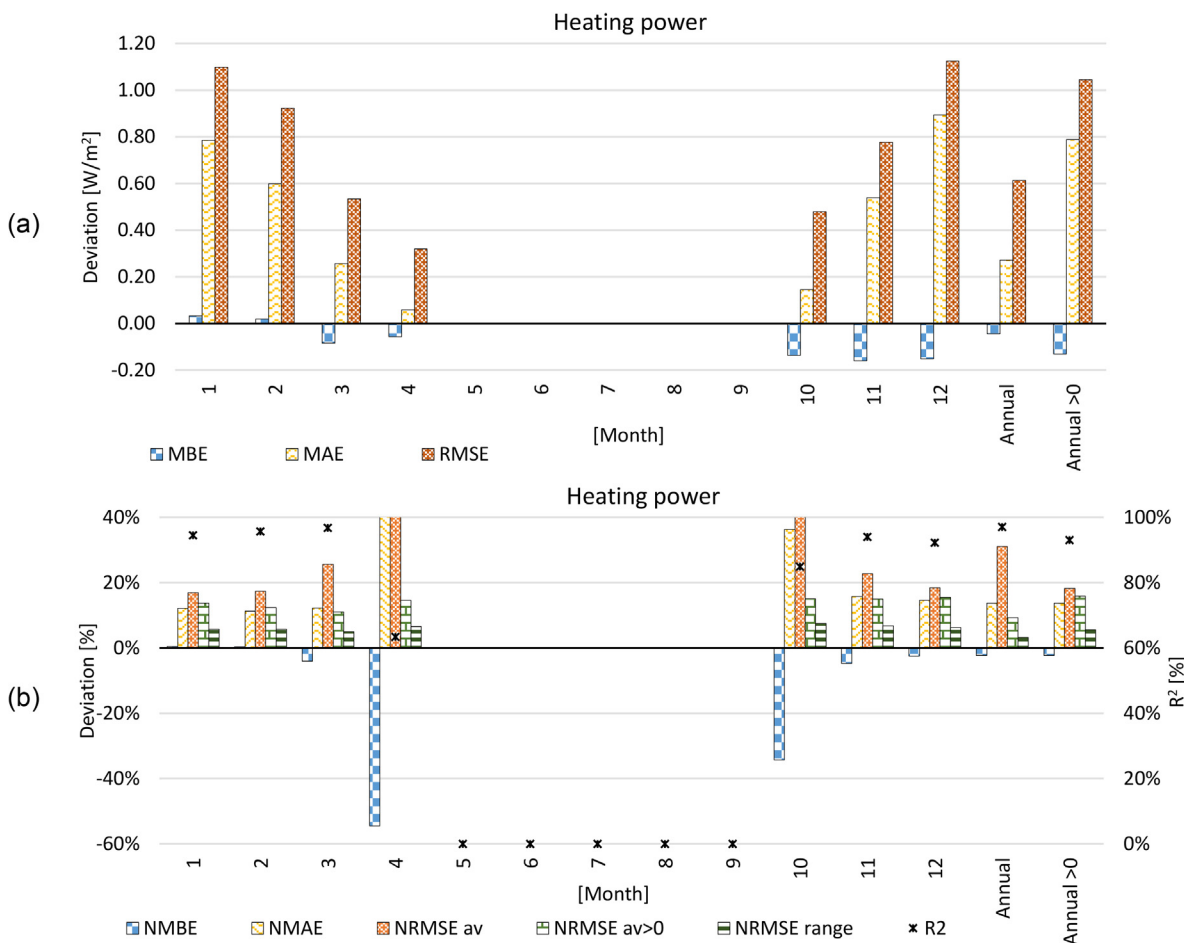
By considering the whole year as a time frame, the NRMSE\_av is 31.1% which would even be outside the thresholds suggested by ASHRAE Guideline 14-2014 [42] even though looking at Figure B.3a it can be noticed that SIM\_IBK records deviations in heating power only 1% of the time of the whole year simulations and the deviation range between  $\pm 5$  W/m<sup>2</sup>. The problem, in this case, is that the average heating power for the whole year trends towards zero since in many periods the building does not need any heating power (see Figure 3.5a). Calculating the deviation including only the heated periods ( $annual > 0$ ), leads to a much lower NRMSE\_av (18.2%), thanks to a higher mean value (see Figure 3.5a), even though the RMSE is higher since the number of samples is lower (see Figure 3.2a).

This approach would require selecting the periods over which the deviations are calculated leading to a subjective decision process. Another possibility is to consider all the available data and normalize the sum of the squared errors with the average of the reference values including only reference data higher than zero ( $NRMSE\_av > 0$ ).

The NMBE is not influenced by this problem since the sum of the reference values (i.e. in this case reference heating powers) at the denominator does not change considering or not values equal to zero.

After this analysis, it can be stated that:

- The MBE/NMBE is needed to show the sign of the deviation, but another index has to be used since the MBE suffers from the cancellation effect;
- The RMSE and MAE provide similar information but the RMSE is more sensitive to the outliers than MAE where each error increases linearly the MAE;
- The  $R^2$  is not a good index when the analysed variable is mainly constant;
- The normalization of the statistical indices is a complex step and has to be performed carefully. In the current work this work the  $NRMSE\_av > 0$  will be used for the analysis of temperatures, power and energy;
- Having a larger dataset (hourly data for the whole year) makes the evaluation of the deviations more robust. Therefore when hourly data are considered the indices should be calculated over the whole year;



**Figure 3.2.** Non-normalized (a) and normalized (b) statistical indices applied for the calculation of the deviation between the simulated heating power with SIM\_IBK and the reference heating power calculated as the median value of all tools for each time step, for the climate of Stockholm.

- The absolute thresholds suggested by ASHRAE Guideline 14–2014 can lead to unjust conclusions regarding the validity of cases with low energy demand (i.e. high-quality buildings or warm climates).

### 3.2. Results of the tool Cross-Comparisons

In this section, the monthly and hourly results of all the tools is presented that considers all the climates (i.e. Rome, Stuttgart and Stockholm) and the corresponding deviations.

#### 3.2.1. Monthly results

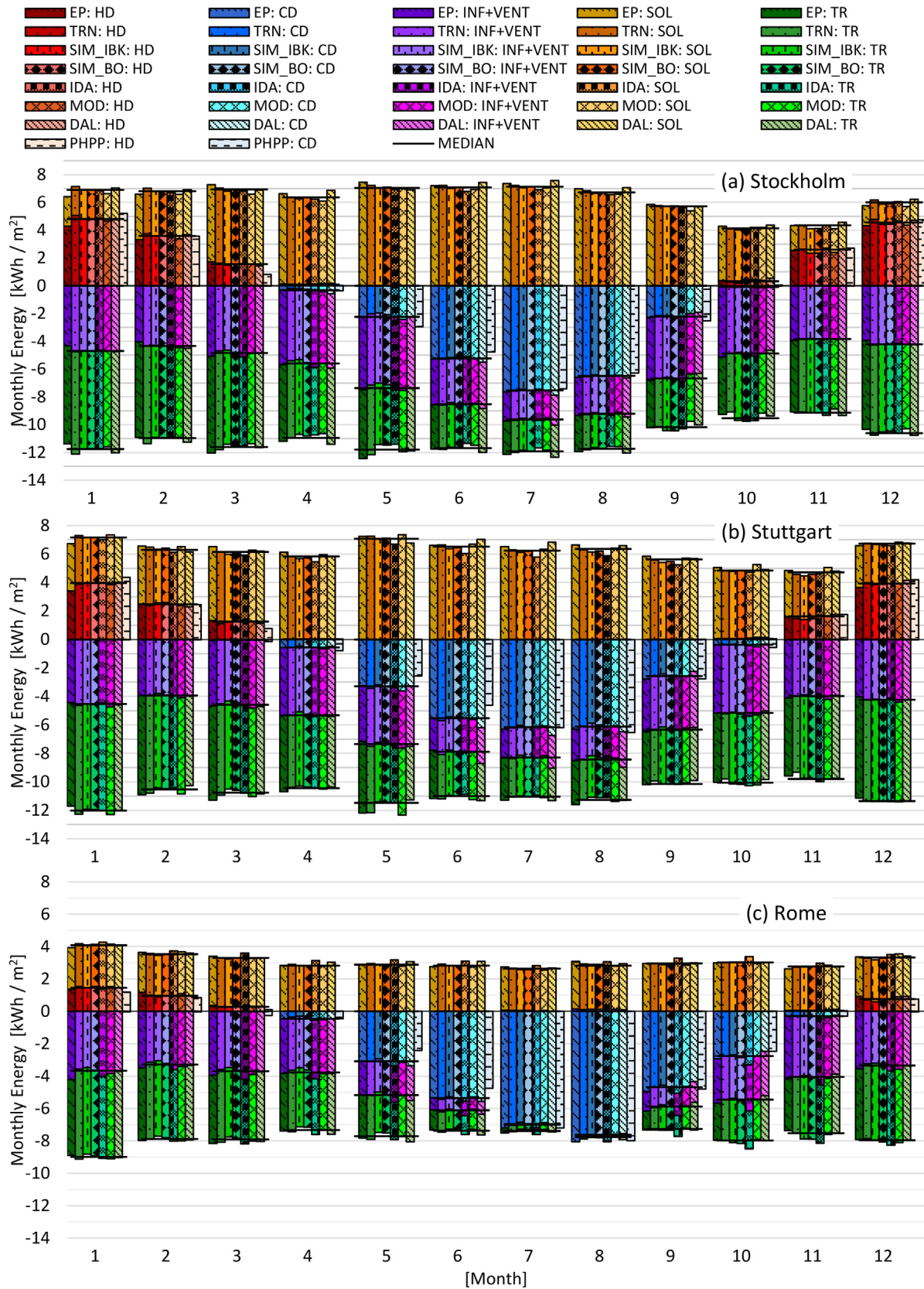
Figure 3.3 shows the monthly heating (HD) and cooling (CD) demands, solar gain (SOL), ventilation summed with the infiltration (INF + VENT) and transmission (TR) losses for all the tools except for the PHPP where only HD and CD are reported, for climates of Stockholm, Stuttgart and Rome. The median value of each component of the thermal balance is presented in Figure 3.3 as a solid black line. The monthly values of the internal gains are excluded from the figure because all the tools consider the same monthly energy. Figure 3.3 shows that an overall good agreement between all the tools is reached, after the parametrization phase (see Sections 2.4 and 3.4). The PHPP was aligned with the other tools in the coldest winter and hottest summer months, but it could not exactly predict the HD and CD during the interim season. In all the locations, SIM\_IBK had slightly lower infiltration and ventilation losses during the first part of the year. Considering the climate of Stockholm, DAL and EP had slightly higher solar gains and

MOD slightly lower than the median value. Regarding the ventilation and infiltration losses, EP recorded lower losses in winter and slightly higher in summer, while IDA had higher ventilation losses in summer with respect to the median value. The effect of the capacity of the adiabatic structures generated deviations in the transmission losses in spring and autumn especially for the climate of Stockholm. In Stuttgart, EP was slightly above the median value concerning the solar gains, while IDA lower. MOD had slightly higher solar gains during winter and DAL during the summer with respect to the median value. Regarding the climate of Rome, IDA had higher solar gains during the whole year compared to the other tools and DAL only during summer.

All the component of the thermal balance are acting together and the results are the heating and cooling demands. In this specific case study, the control of the ventilation system allows to reduce the overheating problem (or cooling demand) compensating higher solar gains with increased ventilation losses.

#### 3.2.2. Hourly results

Figure 3.4 shows the hourly average convective temperature ( $\vartheta_c$ ), ventilation and infiltration losses ( $\dot{Q}_{Inf+Vent}$ ), solar gain ( $\dot{Q}_{SOL}$ ), heating ( $\dot{Q}_{HD}$ ) and cooling power ( $\dot{Q}_{CD}$ ) for all the dynamic simulation tools for climate of Stockholm, for four different periods (48 h each) to represent the typical behaviour in all the seasons. From the convective temperature plots, it can be noticed that DAL is responding slower than the other tools in the free-floating periods since it considers only one lumped capacity. The



**Figure 3.3.** Monthly heating (HD) and cooling demand (CD), ventilation and infiltration losses (VV), transmission losses (TR), solar gains (SOL) and median values for the climates of Stockholm (a), Stuttgart (b) and Rome (c) for all the tools.

convective temperature in IDA is reacting faster than other tools and in MOD it is rising slightly earlier than other tools during the reported spring period (this is not happening systematically every

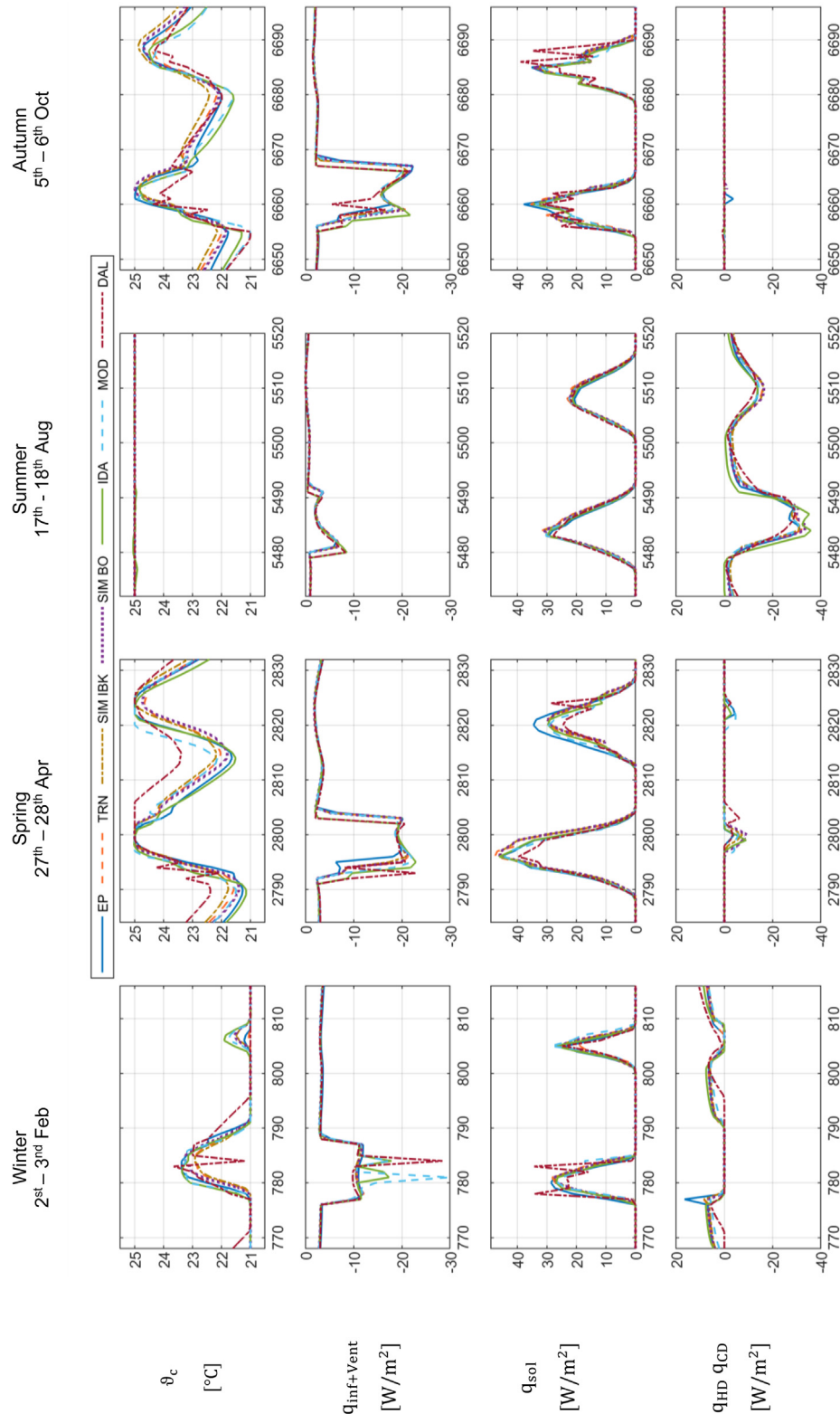
day). Regarding the ventilation losses it can be noticed that DAL and IDA are starting one hour earlier than other tools in spring, summer and autumn since they consider the daylight-saving time.



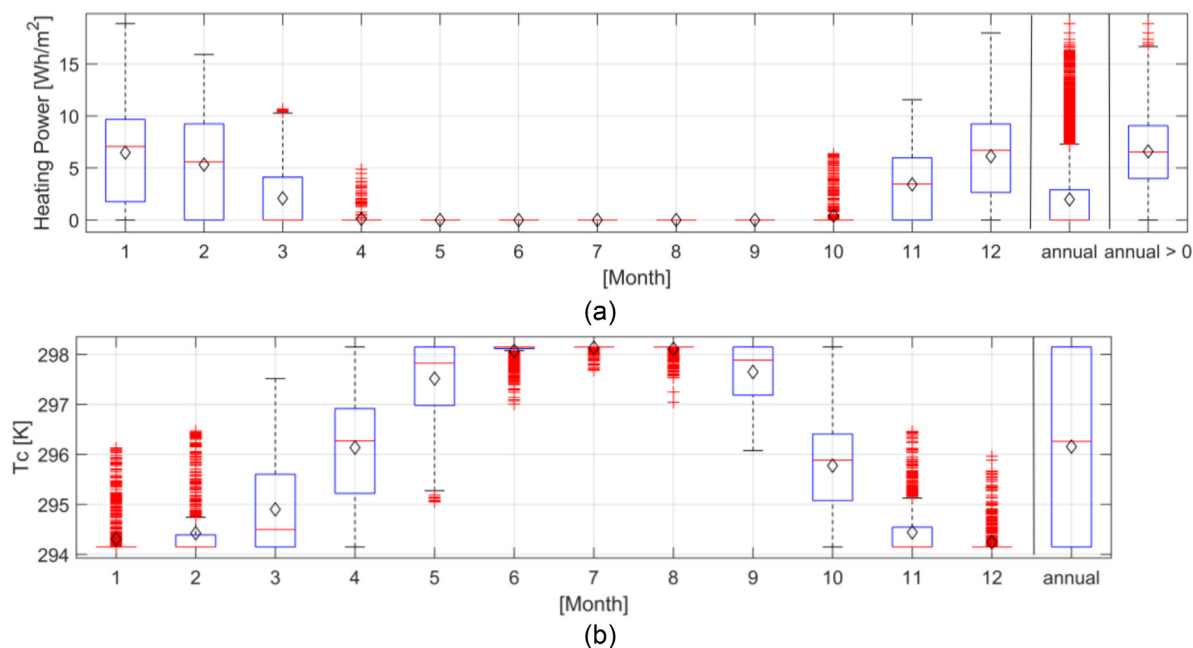
In addition, EP has a slightly different control of the ventilation heat recovery bypass, which is activated slightly later than other tools in autumn and spring. The results show that solar gains are in good agreement as well as the cooling demand, even though DAL has sometimes higher peaks of the solar gains compared to

the other tools due to the different sky model and shading control. Figures for the climates of Stuttgart (STU) and Rome (ROM) are reported in Appendix B (Figures B.1 and B.2).

Figure 3.5 shows boxplots<sup>5</sup> of the reference (i.e., median) heating power (a) and convective temperature (b) for the climate of Stock-



**Figure 3.4.** Hourly results (i.e., convective temperature, ventilation plus infiltration losses, solar gain, heating and cooling power) for all the dynamic simulation tools for the climate of Stockholm for winter, spring, summer and autumn periods (x-axis is hour of the year).



**Figure 3.5.** Box plot of the reference (i.e. median value of all tools for each time step) (a) heating power on monthly basis, for the whole year (i.e., “annual”) and for the whole year considering only periods where the heating system is switched on (i.e. “annual > 0”); (b) convective temperature on monthly basis and for the whole year (i.e., “annual”). The red lines represent the median and the black rhombuses the mean. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

holm considering monthly and annual data. In Figure 3.5a there is an additional box plot of the heating power considering the whole year, for those periods when only the heating system is activated (“annual > 0”). It can be noticed that the median heating power is equal to zero and the mean is close to zero if the whole year is considered (i.e., “annual”) while excluding the periods in which the heating system is switched off (i.e., “annual > 0”) the median is close to the mean value and they are both higher than zero. Problems might arise from this aspect concerning the normalization of the statistical indices as reported in Section 3.1 in more detail.

### 3.2.3. Deviations

Figure 3.6 presents the NMBE and NRMSE<sub>av</sub> > 0 calculated on an hourly (h) and monthly (m) basis using the energy gain and losses associated with heating, cooling, infiltration, ventilation for the climate of Stockholm (a), Stuttgart (b) and Rome (c). The thresholds<sup>6</sup> for the NMBE and NRMSE are reported in Section 2.5 and are shown in Figure 3.6 using dot-dash lines. It is noteworthy, that PHPP only calculates heating and cooling demand on a monthly basis as standard output.

For the climate of STO and STU (see Figure 3.6a and b) DAL exceeds the threshold of 30% for the hourly NRMSE calculated on the solar gains, though the other indices are clearly within the thresholds. This is due to the combined effect of the different sky model used in DAL and the calibration of the shading threshold which allow reaching the same solar gain from the energy point of view, but with a different hourly distribution with respect to the other tools (see Figure 3.4 and Figure B.1). IDA slightly exceeds the solar gains NMBE<sub>m</sub> in STU even though the other components of the thermal balance remain acceptable, while in ROM the higher solar gains also cause high ventilation losses and cooling demand, which leads to the thresholds (i.e., NMBE<sub>m</sub> for solar gains and cooling demand and NMBE<sub>m</sub>/h, NRMSE<sub>av</sub> > 0<sub>m</sub>/h for the ventilation losses) being exceeded. In ROM also SIM\_IBK slightly

exceeds the NMBE<sub>m</sub> for the ventilation losses, even though all the other components of the balance are within the thresholds.

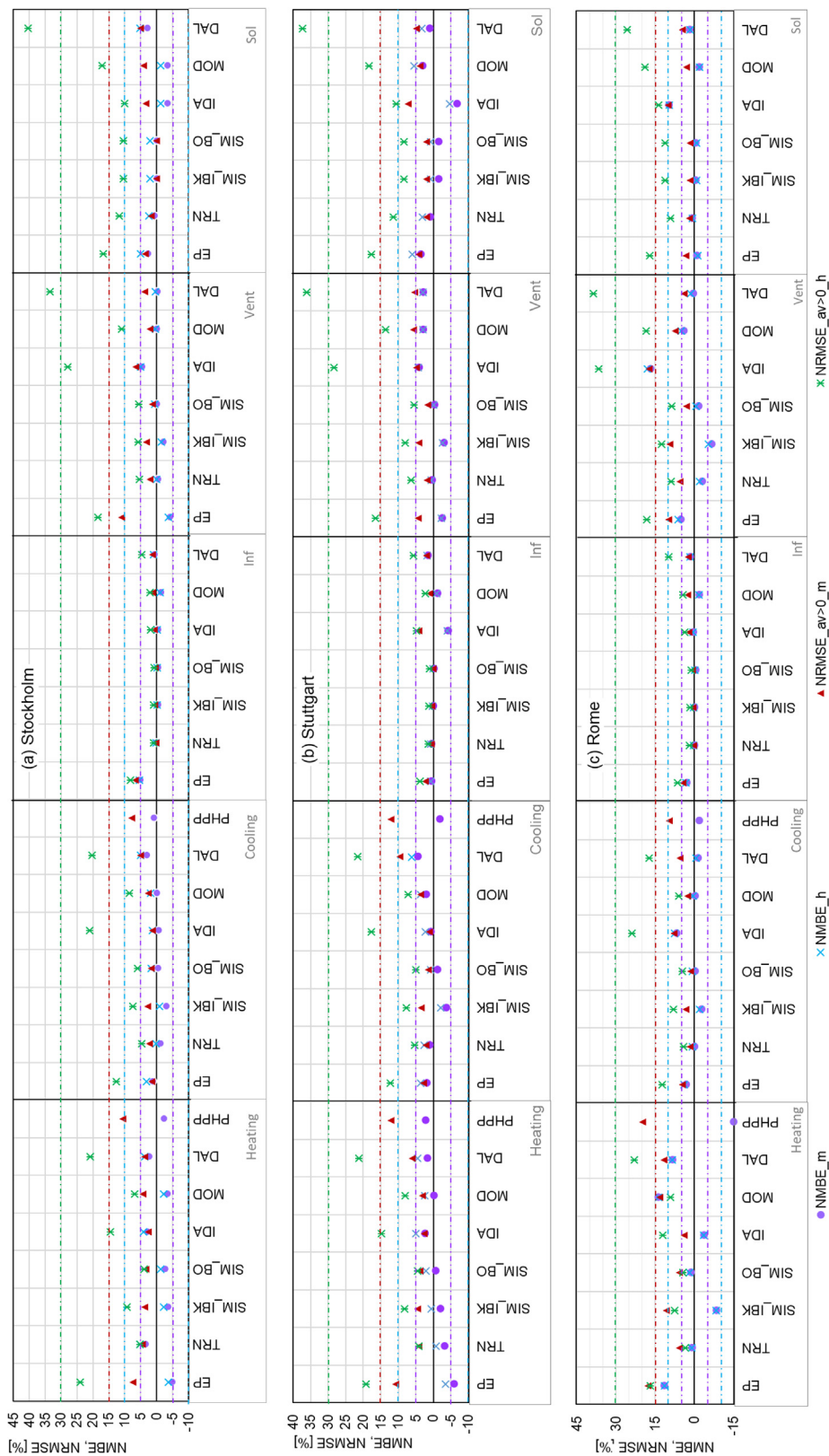
The results for monthly heating demand in ROM show that TRN, SIM\_BO and IDA are within the thresholds while EP, DAL, SIM\_IBK, MOD, DAL and PHPP are not. However, the HD in ROM is so low that it could be disregarded. In most cases, the hourly NRMSE<sub>av</sub> > 0 is higher than the monthly value, while the hourly and the monthly NMBE deliver the same information. The results show that it is important to take both the NMBE and the NRMSE into consideration since in some cases the overall annual energy gains and losses might be in agreement, but the deviations considering hourly data are not. For example, the calculation of these indices for the solar gain in DAL can be observed. The hourly NRMSE<sub>av</sub> > 0 calculated for the ventilation losses is particularly high for IDA and DAL because these two tools are the only two considering the shift between the daylight saving time and summer time, which leads to a shift in the schedule of the ventilation system (as can be seen in Figure 3.4, Figure B.1 and Figure B.2). For sake of better readability, R<sup>2</sup> was not reported in the graph but is always above 75%, (the minimum required by ASHRAE Guideline 14-2014), except for the heating demand in ROM for DAL where it is equal to 61% and for the ventilation losses of DAL in Stockholm, Stuttgart and Rome where it is equal to 67%, 64% and 72%, respectively. Concerning the convective temperature, the agreement between the tools is always within the thresholds in all the climates.

Overall, it can be stated that a good agreement between the tools is reached and that all the models are quite reliable after the parametrization process.

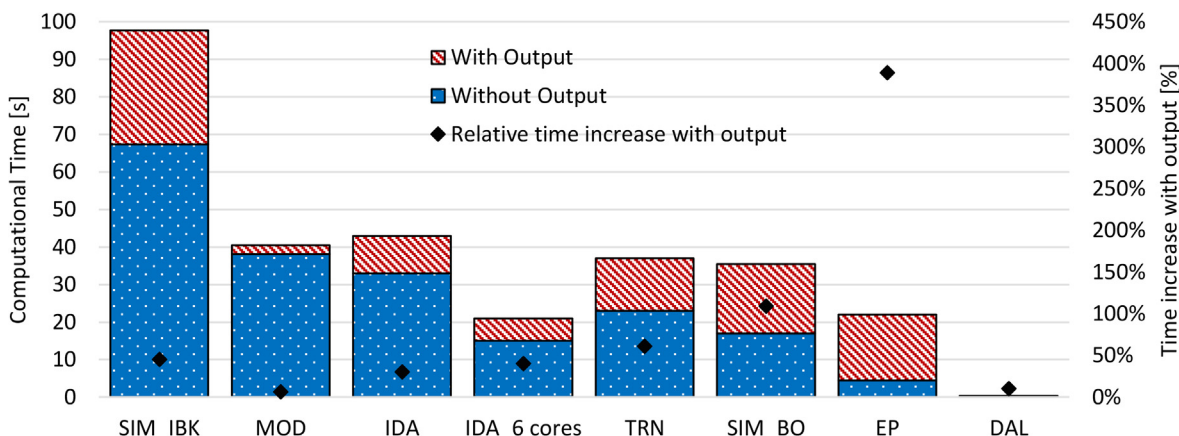
### 3.3. Computational cost

In Section 2.6 the boundary conditions in which the computational cost of each tool is calculated are defined. The PHPP is an instantaneous calculation tool therefore the fastest and is not included within this comparison. Figure 3.7 shows the computational cost for each tool without the post-processing of the output (in blue) and including it (in red). The filled diamond represents

<sup>6</sup> ±5% for the monthly NMBE, 15% for the monthly NRMSE, ±10% for the hourly NMBE, 30% for the hourly NRMSE and >0.75 for the R<sup>2</sup> [47].



**Figure 3.6.** Hourly (h) and Monthly (m) NMBE and NRMSE<sub>av > 0</sub> calculated for Heating and cooling demand, ventilation and infiltration losses and solar gains considering the climate of Stockholm (a), Stuttgart (b) and Rome (c). The horizontal dot-dash lines represent the calibration criteria suggested by [42]: ±5% for the monthly NMBE, 15% for the monthly NRMSE, ±10% for the hourly NMBE, 30% for the hourly NRMSE.



**Figure 3.7.** Computational time of each tool without the processing phase of the output (in blue) and the additional time required for the output preparation (in red). The dots represent the relative additional time when the outputs are prepared. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the relative additional time required when the output file needs to be prepared. The number and format of the outputs are different in the different tools. This is important to consider because the number of outputs considered is a user-defined variable that depends on the level of detail the user needs to analyse the results. For these reasons, the computational time is shown with and without output-processing to make a fair comparison of the calculation effort required by the model itself and, at the same time, to give an insight into the time required for the simulation to generate the output file.

The computational time of MOD is comparable with IDA, while EP and DAL can run an annual simulation in only a few seconds. IDA ICE can decrease to half the simulation time by splitting the annual simulation into various time slices, assigning them to different processing cores in the computer, while the other tools (TRNSYS, Simulink, Modelica and EnergyPlus) can profit from the multiple cores when more simulations run simultaneously (e.g., parametric sweep), but not on a single simulation for this model configuration. Among the investigated tools, DAL is the fastest but it must be highlighted that it is the only tool based on an hourly time step and can be used only for predefined simple geometries (i.e., a shoebox) for which the view factor- and daylight matrices are pre-calculated. Although EP is one of the fastest tools, it requires a similar amount of time, for the elaboration of the output, than TRN and SIM\_BO. The MOD and IDA simulations are faster than the other tools for the preparation of the output file.

Another aspect that is not depicted in Figure 3.7 is the time required for setting up the model. TRNSYS, IDA ICE, user interfaces for EnergyPlus, both Simulink libraries and PHPP can import the geometry information from a 3D drawing (i.e., gbXML, idf, dxg, IFC, etc. . .), easing the definition of the building model. For Modelica/Dymola such features are currently being developed within IBPSA project 1 [80]. Nevertheless, the requirements of the 3D drawing are different for the different tools therefore the interoperability is not straightforward. Though, this problem is supposed to be addressed by the improvements in the field of BIM to BEM in the near future.

### 3.4. Results pre and post-parametrization for DAL and PHPP

This section describes the effects of the parametrization process (see Section 2.4) on the results of PHPP and DAL, reporting the results of the non-parametrized simulation (i.e., DAL\_online and PHPP\_nonpar) and of the parametrized simulation (i.e. DAL\_par and PHPP\_par), which correspond to the final results of DAL and PHPP reported in Sections 3.2.1, 3.2.2 and 3.2.3.

With pre-design tools, it is not straightforward to model aspects such as hourly internal gain profiles, dynamic shading and ventilation bypass control as required by the description of this case study. The average value of the power due to the internal gain (see Figure 2.6) can be easily determined by averaging the hourly profile over the day for DAL (19.7 W/m<sup>2</sup>) and over the week for the PHPP (6.5 W/m<sup>2</sup>). The standard effective ventilation rate needs to also be recalculated for both PHPP and DAL online since both tools are based on a constant rate while, according to the description of the building the ventilation system, is switched on only during the working time. The additional ventilation rate due to the dynamic control of the bypass is considered in DAL online while in PHPP it is only possible to set this option for the summertime since the control is based on monthly ambient temperature instead of hourly as in DAL. The shading can be controlled as required by this case study in DAL online but the solar irradiation on the south facade used by DAL is different from the other tools since an isotropic sky model is applied instead of the Perez model, therefore the solar gains are also different. In PHPP only a summer and winter shading value can be given as input and in this case study, it would be difficult to guess this constant without knowing the results of the other dynamic simulation tools.

In PHPP\_nonpar the standard assumptions are used during winter (25% reduction) and the reduction factor described in the report (70% reduction) is used for the summertime as temporary sun protection. In DAL online and PHPP\_nonpar, the HTC of wall and window described in the report are used. DAL\_par compared to DAL\_online implements an improved ventilation model where only during the occupied time the ventilation system works, moreover the shading control threshold, the HTC of the facade and the building capacity are parametrised against the results of the other dynamic simulation tools.

PHPP\_par implements a parametrized shading coefficient for summer and winter, ventilation rate and HTC (see Table 2.8). Table 3.1 reports the annual heating demand (HD) and cooling demand (CD) of PHPP and DAL pre and post parametrization, showing the relative deviation against the reference HD and CD (median of the results of all the tools). Figure 3.8 shows the monthly heating and cooling demand, infiltration plus ventilation and transmission losses, solar gains and average convective temperature simulated with DAL and PHPP pre and post parametrization and the reference value derived as a median of the results of all the tools for the climate of Stockholm. PHPP delivers only monthly heating and cooling demand, all the other components of the energy balance can be extrapolated by the user, but especially in spring and autumn, this process is non-trivial. Therefore,

**Table 3.1**

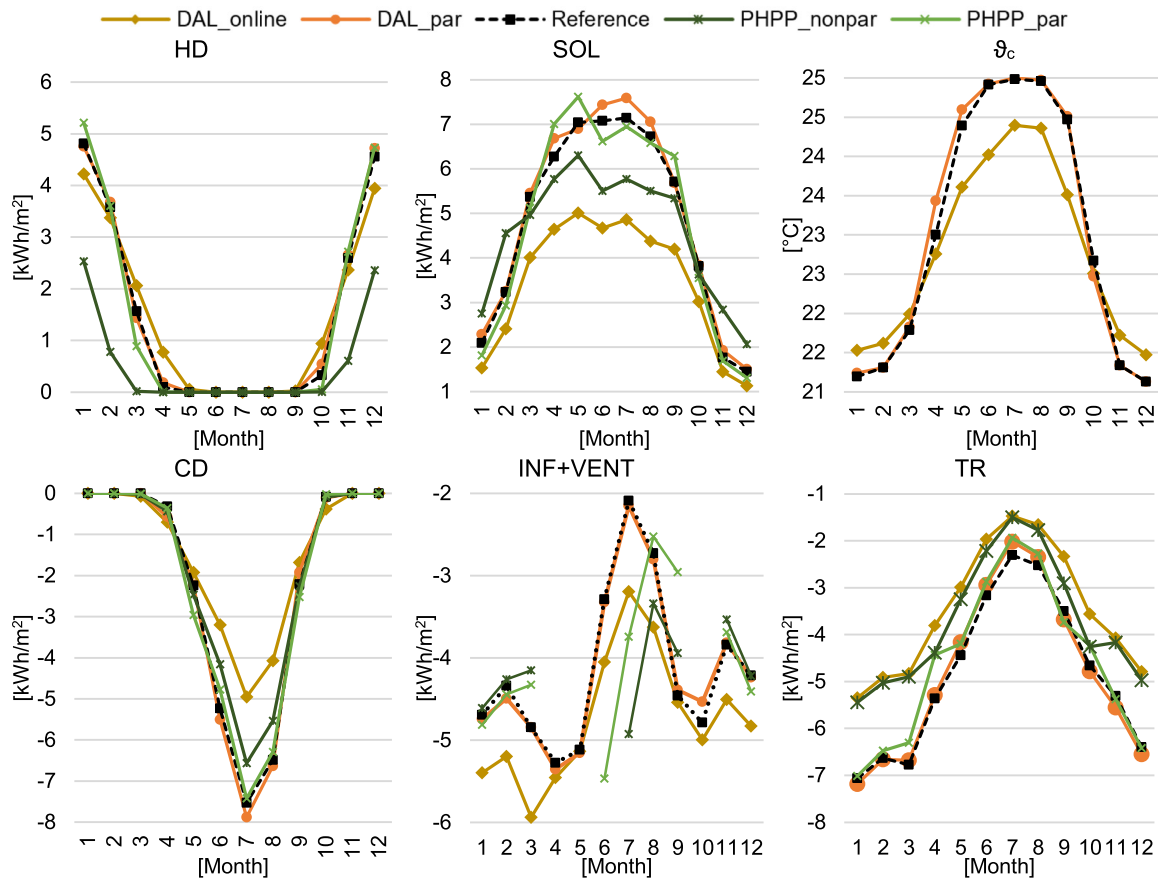
Annual Heating and Cooling demand (HD, CD) calculated with DALEC and PHPP pre and post parametrization compared against the reference heating and cooling demand (median of all the tools), for the climate of Stockholm.

	HD [kWh/(m <sup>2</sup> )]	$\Delta$ Ref [%]	CD [kWh/(m <sup>2</sup> )]	$\Delta$ Ref [%]
DAL_online	17.8	1%	-17.0	-30%
DAL_par	18.0	2%	-24.9	3%
PHPP_nonpar	6.3	-64%	-21.5	-11%
PHPP_par	17.2	-2%	-24.3	1%
<b>Reference</b>	<b>17.5</b>	<b>-</b>	<b>-24.1</b>	<b>-</b>

these months are not reported in Figure 3.8. It can also be noticed that the HTC of the window plays an important role in the transmission losses and that after the parametrization both tools are in agreement with the reference results. The standard assumptions of PHPP lead to an overestimation of the solar gain in winter and an underestimation in summer and to underestimate the ventilation losses, especially during summer. All these components lead to lower CD and HD of the PHPP\_nonpar. DAL\_online underestimates the solar gain throughout the year because of the different calculation of the solar radiation impinging the south façade (see Figure 2.5), this is improved in DAL\_par by modifying the threshold of the shading control. The ventilation losses in DAL\_par are fitting the results of SIM\_BO since a schedule for the activation of the mechanical ventilation is introduced. It is also noteworthy to mention that in DAL\_online the lower losses (transmission and ventilation) are compensated by lower solar gains resulting in an HD matching with the reference HD (see Table 3.1).

**4. Lessons learned**

Numerous iterations were necessary to reach a good agreement among the different tools even though the case study is well described in a comprehensive report and the reference building is geometrically simple. The challenges encountered in this process are only partly avoidable since user mistakes are difficult to foresee. Part of the challenge could be prevented by having a more detailed description of the building, avoiding misinterpretation and the entire parametrization process. An important issue that must also be highlighted is that since the tools implement different models, the outputs and the required inputs are defined differently. Defining equivalent input parameters for each model is a non-trivial task considering that different models describe physical phenomena using different levels of abstraction and at different levels of scale. A lesson that the authors derived from the parametrization process is that it is recommended to define input



**Figure 3.8.** Monthly heating (HD) and cooling (CD) demand, infiltration plus ventilation (Inf+Vent) and transmission (TR) losses, solar gain (SOL) and convective temperature (∂c) simulated with DAL online, DAL parametrized (DAL\_par), PHPP non parametrized (PHPP\_nonpar) and parametrized (PHPP\_par) compared against the reference results for the climate of Stockholm.

parameters starting at the most detailed scale and the lowest level of abstraction amongst the tools that are going to be compared. Using the window system as an example, the process started with a reference office description where the window was described using properties at the level of the glazing system (TRN, SIM\_BO) and detailed properties at the level of the individual window pane (EP, MOD) had to be derived through an exhaustive search process. The process of defining equivalent input parameters could have proceeded faster, however, if properties were originally defined at the window pane level and then used to derive equivalent properties at a higher level of abstraction based on subsystem simulation.

The window example also illustrates the problem of aligning simulation outputs. In tools with a simplified window, only the total solar gain can be given as output, while in other tools solar gains are reported using only directly transmitted solar radiation or the solar radiation absorbed at interior surfaces. Here, clearly defining the boundary conditions of different energy contributions was an essential part of aligning the simulation results of the different tools. The following paragraphs will give an overview of the user-related input errors and modelling decisions that were identified throughout the calibration process as sources of discrepancies between the different tools. Additionally, the magnitude of these discrepancies at different stages of calibration will be presented.

User mistakes experienced within this case study, were as follows:

- Flipped order of the layers of surface constructions;
- Wrong interpretation of the HTC of the window (HTC-glass instead of HTC-window). HTC-window is the results of the area-weighted HTC of the frame and the glass;
- Wrongly selected weather file where different data for the same location is used;
- Wrong starting day of the internal gain profile; and
- Wrong wall thermal capacity.

Wrong interpretations of the office description that could have been prevented by a more detailed description were made regarding the:

- Control of the mechanical ventilation;
- Control of the shading;
- Ventilation air volume flow;
- Heat recovery from fans of the ventilation system; and
- Control of the pre-heater used to avoid ice formation in the ventilation heat recovery system.

Moreover, information that was initially missing in the report were implemented differently in the different tools, i.e.:

- Air density (as a function of temperature or constant);
- Absorption and emission coefficient of opaque structures; and
- The convective exchange coefficients (as a constant or function of the temperature difference);

Comparing only the modelled heating and cooling demand, was not enough to find the reasons for the deviations between the tools. Often, higher gains might be balanced by higher losses hardly affecting the heating and cooling demand and therefore the zone heat balance was investigated. Additionally, isolating particular heat transfer phenomena assisted in identifying the cause of deviations. For instance, simulations were executed:

- Without solar radiation;
- Without windows;

- With a permanently activated, or no shading device;
- Using constant convective heat transfer coefficients, identical in all models, rather than the detailed algorithms available in some tools;
- Without internal gains;
- With different modelling approaches of the adiabatic structures;
- Without ventilation system.

All these cases were not executed with all the tools, but only where deviations needed to be identified with further analysis. Along with the different tests, an increasing number of outputs were analysed and compared (e.g., Internal surface temperatures, all the details of the balance of windows and walls, all the components of the weather data, distribution of the internal gains and solar gains, ventilation and infiltration losses, etc. ...).

The window model has a key role, in this case study, since it represents the main source of transmission losses. Therefore, the parametrization of the input for both simplified and detailed window models was a time-consuming process. Within this work, using the given constant HTC for the window in Stockholm deliver an HD 37% lower than using the detailed window model.

Table 4.1 reports the annual heating (HD) and cooling (CD) demand of each tool for the climates of Rome (ROM), Stuttgart (STU) and Stockholm (STO) in the initial iteration (V1), an intermediate step (V2) and the final results presented in this paper (V3). For each case, the maximum and minimum annual NMBE are reported and on the right-hand side, the dispersion of the annual HD and CD is represented by box plots for each iteration. The first iteration is represented by V1, here the dispersion of the results is important since many user mistakes and wrong interpretation of the office description are present. Within the V2 the situation is improved, but still not acceptable since the spread of the results is still high.

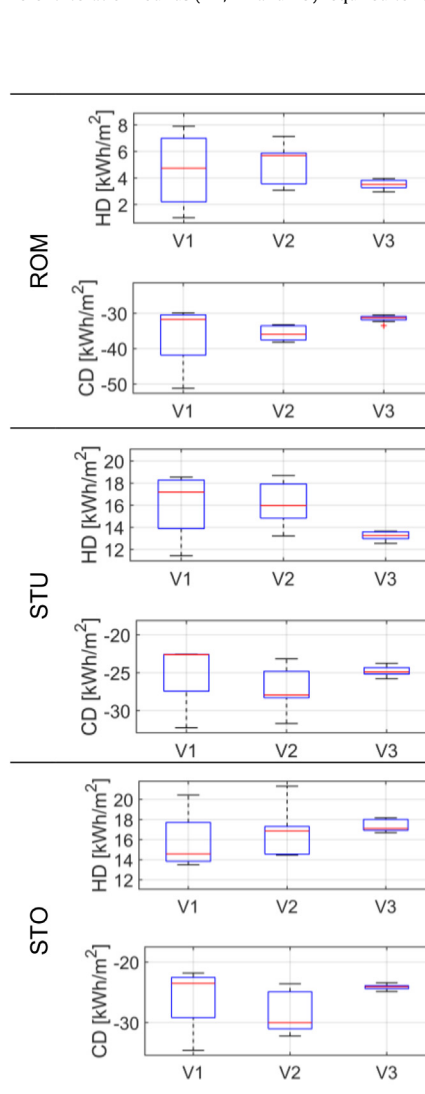
With V3 a good agreement is reached thanks to the recognition of the user mistakes and the parametrization of the window model. From V1 to V2 the window properties were parameterized in EP, the set point of the anti-freezing was corrected in TRN, the shading model, the volume of the TZ, the starting day of the occupancy profile, the adiabatic model of the opaque structure and the weather file were corrected in SIM\_IBK, and the HTC of the window in DAL. From V2 to V3 an improved parameterization of the window properties was carried out for MOD and EP, the convective coefficient calculation was modified to use the same equations in EP, TRN and MOD. In SIM\_IBK was introduced the window model from the library of SIM\_BO and the order of the construction layers of the adiabatic ceiling and floor was corrected in MOD and TRN.

## 5. Conclusions

In this work, the modelling approaches of well-known dynamic simulation tools, EnergyPlus, TRNSYS, IDA ICE, Modelica/Dymola the new Matlab/Simulink libraries, CarnotUIBK and ALMABuild as well as the predesign tools DALEC and PHPP were described and the tools compared against each other using a typical office cell located in Rome, Stuttgart and Stockholm as a reference building. To compare the results of the different tools, commonly used statistical indices and normalization means were analysed and finally the Normalized Mean Bias Error and Normalized Root Mean Square Error were used to assess the degree of agreement of the results with the median value used as a reference. The thresholds suggested by ASHRAE Guideline 14–2014 were used as a basis for this evaluation. The normalization process of the statistical indices was non-trivial and the encountered challenges were presented and thoroughly highlighted within this paper.

**Table 4.1**  
Annual Heating and Cooling demand for each tool in the climate of ROM, STU and STO through the different iteration rounds (V1, V2 and V3) required to reach a good agreement.

	V1 (04/2018)		V2 (04/2019)		V3 (04/2021)		
	HD	CD	HD	CD	HD	CD	
	[kWh/m <sup>2</sup> ]						
ROM	EP	1.0	-51.2	3.6	-36.4	3.9	-32.3
	TRN	3.4	-32.4	3.5	-33.3	3.5	-31.4
	SIM_IBK	7.9	-29.9	5.8	-38.2	3.2	-30.5
	SIM_BO	-	-	3.1	-33.4	3.5	-31.3
	IDA	-	-	-	-	3.4	-33.5
	MOD	-	-	7.1	-34.0	4.0	-31.3
	DAL	6.1	-31.0	5.9	-35.9	3.8	-31.0
	PHPP	-	-	5.7	-37.9	3.0	-30.8
	<b>MEDIAN</b>	<b>4.7</b>	<b>-31.7</b>	<b>5.7</b>	<b>-35.9</b>	<b>3.5</b>	<b>-31.3</b>
	Max NMBE	67%	61%	25%	6%	12%	7%
Min NMBE	-79%	-6%	-46%	-7%	-16%	-2%	
STU	EP	11.4	-32.3	15.8	-28.1	12.5	-25.2
	TRN	18.0	-22.6	18.7	-23.2	12.9	-24.9
	SIM_IBK	18.5	-22.6	16.0	-31.7	13.0	-23.8
	SIM_BO	-	-	13.2	-23.9	13.2	-24.4
	IDA	-	-	-	-	13.7	-24.9
	MOD	-	-	17.1	-27.9	13.3	-25.2
	DAL	16.4	-22.6	18.2	-28.4	13.6	-25.8
	PHPP	-	-	14.5	-27.6	13.6	-24.2
	<b>MEDIAN</b>	<b>17.2</b>	<b>-22.6</b>	<b>16.0</b>	<b>-27.9</b>	<b>13.3</b>	<b>-24.9</b>
	Max NMBE	8%	43%	17%	13%	3%	4%
Min NMBE	-34%	0%	-17%	-17%	-5%	-5%	
STO	EP	13.5	-34.6	17.0	-32.2	16.7	-24.4
	TRN	20.4	-23.3	21.3	-23.8	18.2	-23.8
	SIM_IBK	15.0	-23.8	14.5	-31.0	16.9	-23.4
	SIM_BO	-	-	17.4	-23.6	17.1	-24.0
	IDA	-	-	-	-	18.1	-24.0
	MOD	-	-	14.5	-30.0	16.9	-24.1
	DAL	14.2	-21.8	16.9	-28.2	18.0	-24.9
	PHPP	-	-	14.6	-31.0	17.1	-24.3
	<b>MEDIAN</b>	<b>14.6</b>	<b>-23.5</b>	<b>16.9</b>	<b>-30.0</b>	<b>17.1</b>	<b>-24.1</b>
	Max NMBE	40%	47%	26%	7%	6%	3%
Min NMBE	-7%	-7%	-14%	-21%	-2%	-3%	



After many iterations, it can be stated that overall a good agreement between all the dynamic tools was reached. The annual results from the PHPP program was also aligned with the annual results of all the other tools. At the beginning of this process, the relative deviation of the heating and cooling demands predicted by the different tools was from + 61% to -34%, which was reduced to + 7% to -5% after many simulation rounds (excluding the heating demand in Rome which is almost negligible). The deviations that were experienced were mainly due to difficulties in defining equivalent input parameters for the models based on different approaches; user mistakes; and misinterpretation of the building description. User mistakes were clearly difficult to avoid, while different interpretations of the building description were mainly caused by an insufficient description. At the same time a tedious parametrization phase was required for those tools which were either more simplified (e.g., PHPP and DALEC) or more complex (e.g. Modelica and EnergyPlus) compared to the building description that was written using TRNSYS as a reference. The parametrization of inputs such as the ventilation rate and constant shading allowed the comparison to reach a good agreement even for simplified tools like PHPP, where dynamic control logics (i.e., shading and ventilation control) required for this case study, cannot be modelled. The results of DALEC and PHPP pre and post-parametrization for the climate of Stockholm were reported and

the deviation of PHPP annual heating demand compared to the reference median value was reduced from -64% to -2% and from -11% to 1% for the cooling demand. The different tools involved in this study allow analysis with different degrees of detail and at the same time, they have very different computational costs. An overview of the modelling approaches available within the different software packages was given and the computational time required by each model used for this case study was reported, to support users in choosing a simulation tool that fits their purpose.

The computational cost was quantified by running the simulations using the different tools on the same computer. DALEC and PHPP were both almost instantaneous. However, these were also simplified compared to the other simulation software packages. EnergyPlus ran for about 5 sec and was faster compared to the other tools. TRNSYS and ALMABuild were in the same range around 20 sec, whereas Modelica/Dymola and IDA ICE were slightly slower with 38 and 33 sec respectively. CarnotUIBK was the slowest with 67 sec.

**CRedit authorship contribution statement**

**Mara Magni:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Visualization. **Fabian Ochs:** Supervision, Project

administration, Funding acquisition, Writing - review & editing. **Samuel de Vries:** Software, Validation, Writing - review & editing. **Alessandro Maccarini:** Software, Writing - review & editing. **Ferdinand Sigg:** Software, Writing - review & editing.

**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**

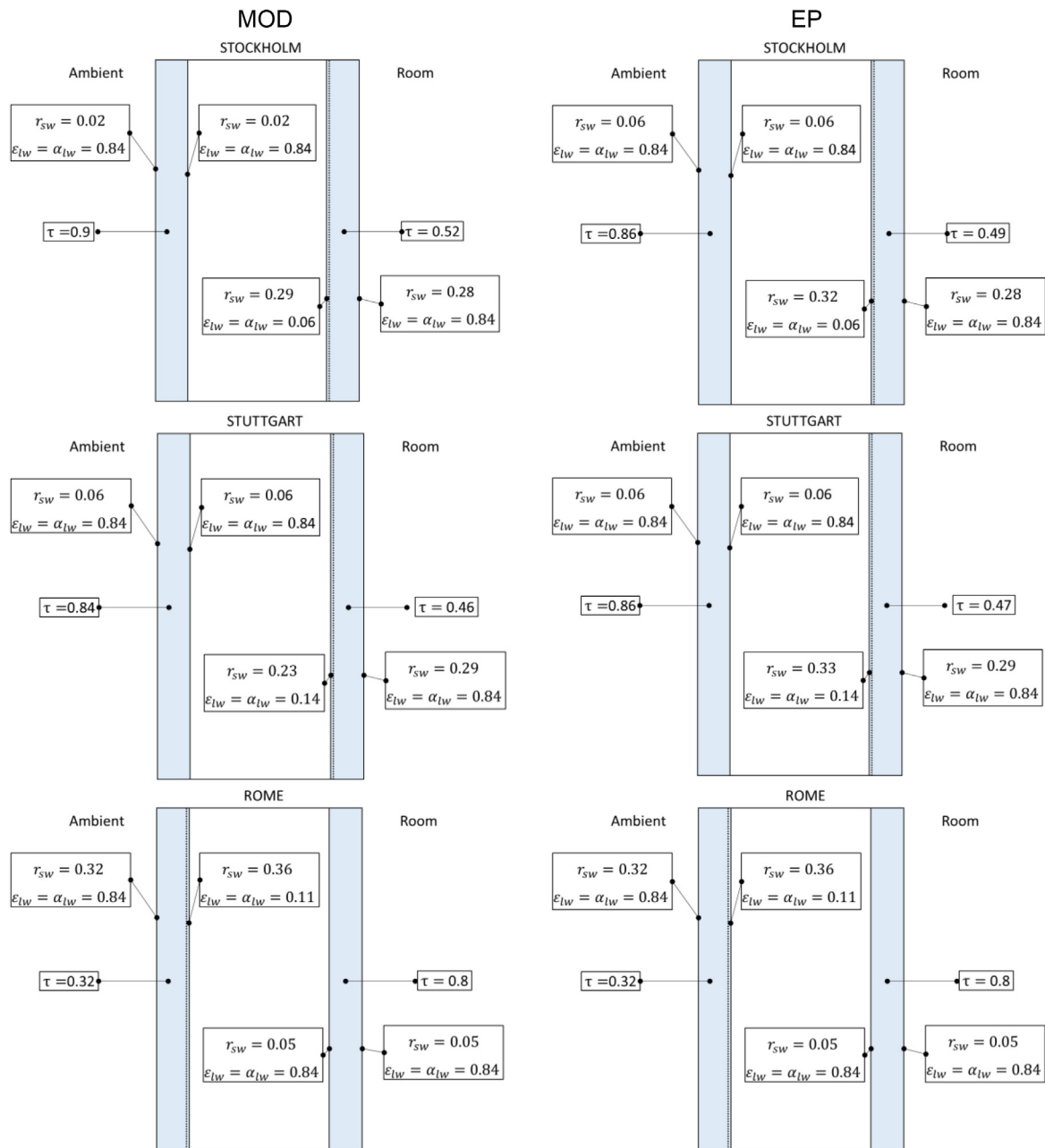
This work was performed within the framework of IEA SHC Task 56 international project.

We thank Abdulrahman Dahash and Shandilya Apeksha for an internal review of the paper, Hauer Martin and Plörer Daniel for supporting the simulations with DALEC, EURAC for the collaboration within the IEA SHC Task 56 project. Thanks to the University of Bologna for allowing us to use ALMABuild, Nicola Franzoi for the fruitful discussion about the statistical indices, Toni Calabrese and Ellika Taveres-Cachat for contributing in creating the TRNSYS and IDA ICE models.

**Appendix**

*A. Detailed window properties*

In **Figure A.1** the window properties used in MOD and EP in the climates of Rome, Stuttgart and Stockholm are reported.



**Figure A1.** Detailed optical window properties used in EP and MOD in each climate.



B. Hourly plot

Figure B.1 and Figure B.2 show the hourly average convective temperature ( $\vartheta_c$ ), ventilation and infiltration losses ( $\dot{Q}_{Inf+Vent}$ ), solar gain ( $\dot{Q}_{SOL}$ ), heating ( $\dot{Q}_{HD}$ ) and cooling power ( $\dot{Q}_{CD}$ ) in 4 representing

periods for winter, spring, summer and autumn for all the dynamic simulation tools considering the climate of Stuttgart and Rome, respectively. In Figure B.1 and Figure B.2 can be noticed that:

- DAL is responding slower than the other tools in the free-floating periods since it considers only one lumped capacity;

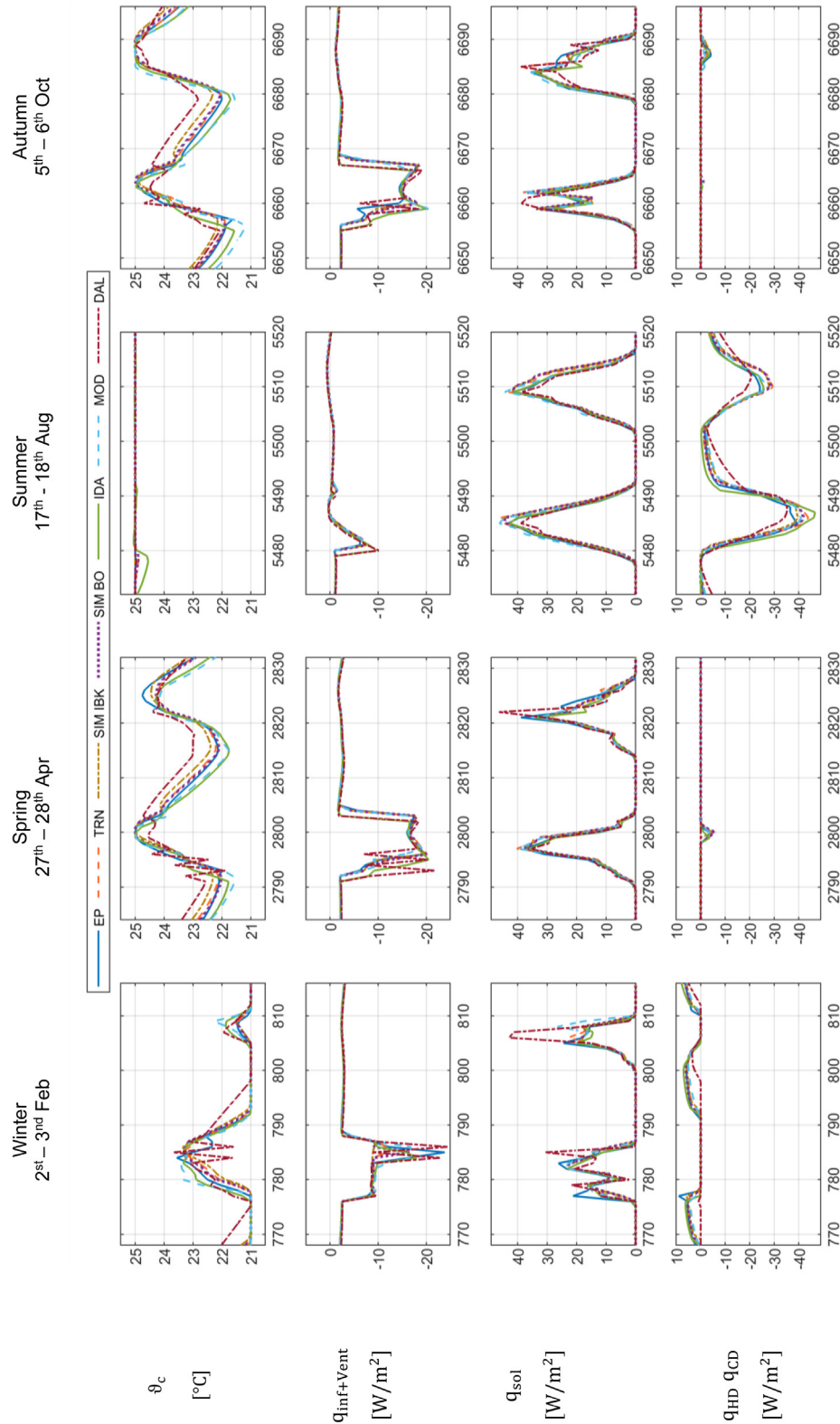
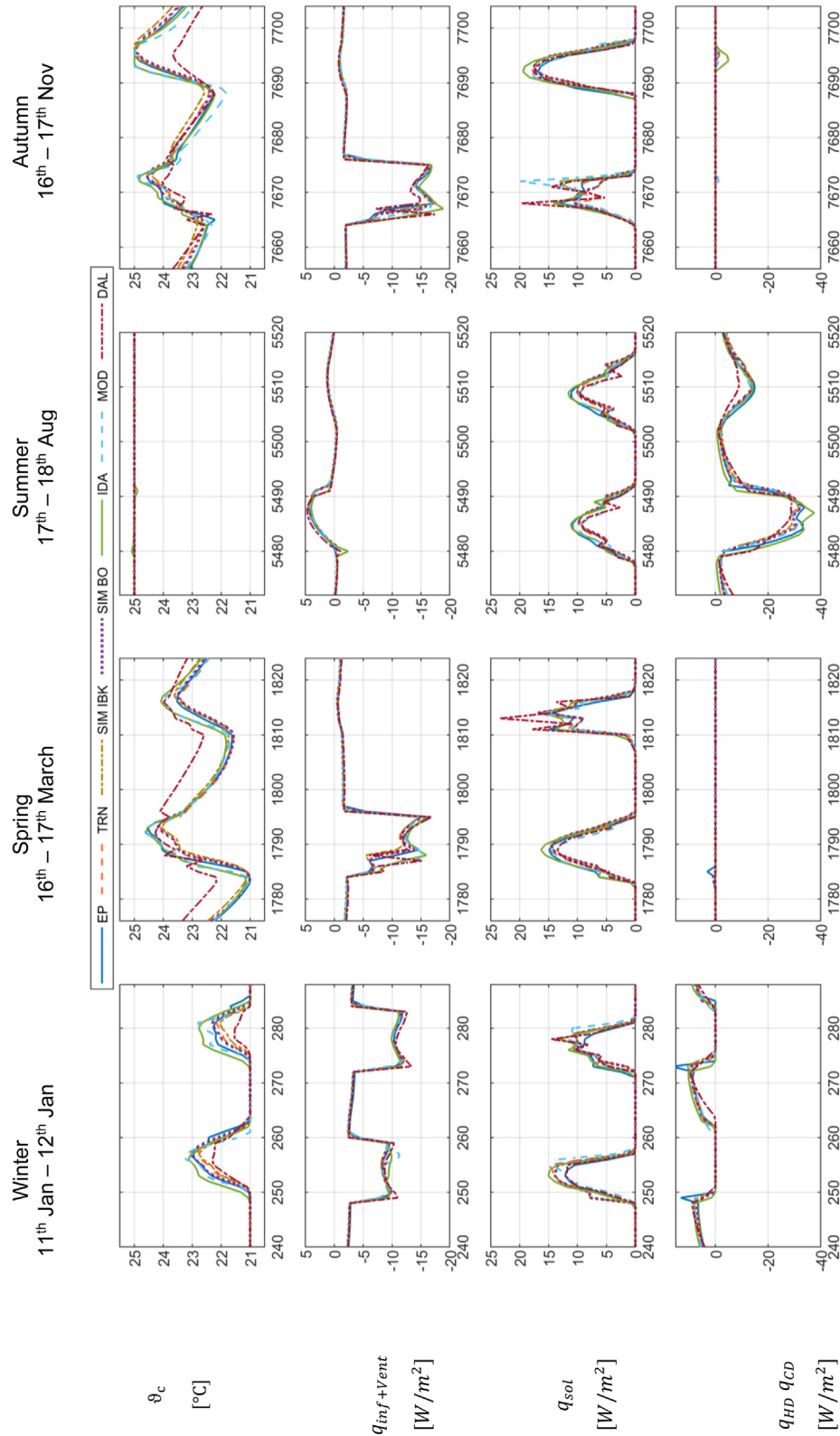
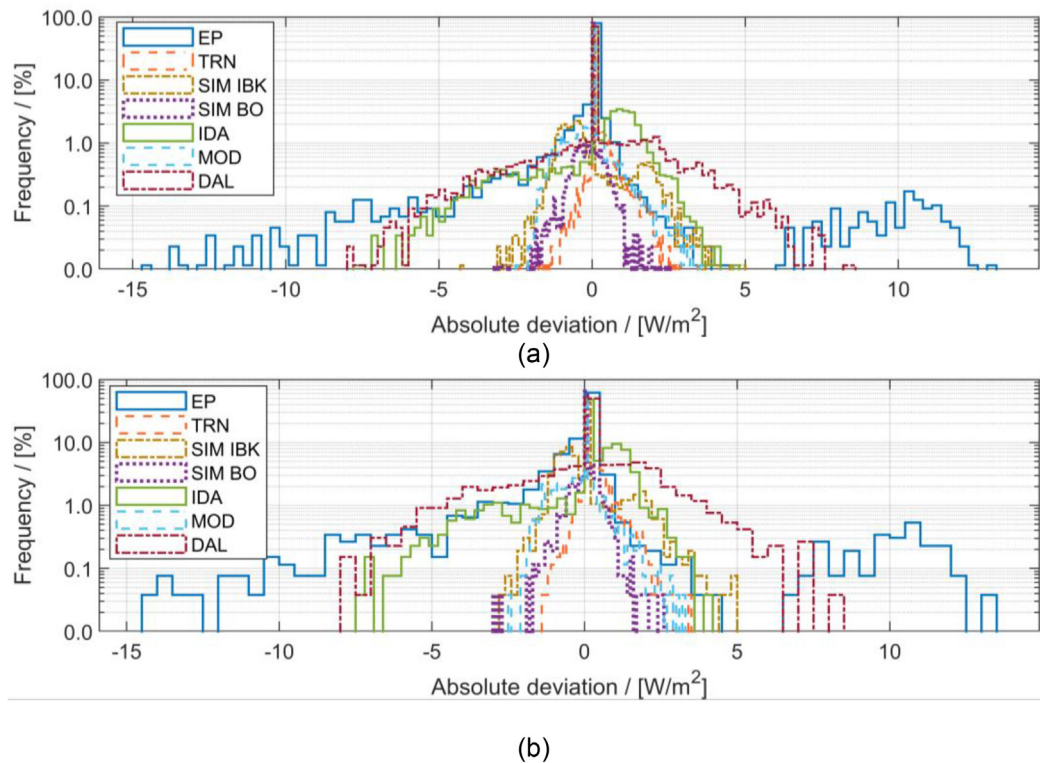


Figure B1. Hourly results (i.e. convective Temperature, ventilation plus infiltration losses, solar gain, heating and cooling power) for all the dynamic simulation tools considering the climate of Stuttgart and 4 representing periods for winter, spring, summer and autumn.

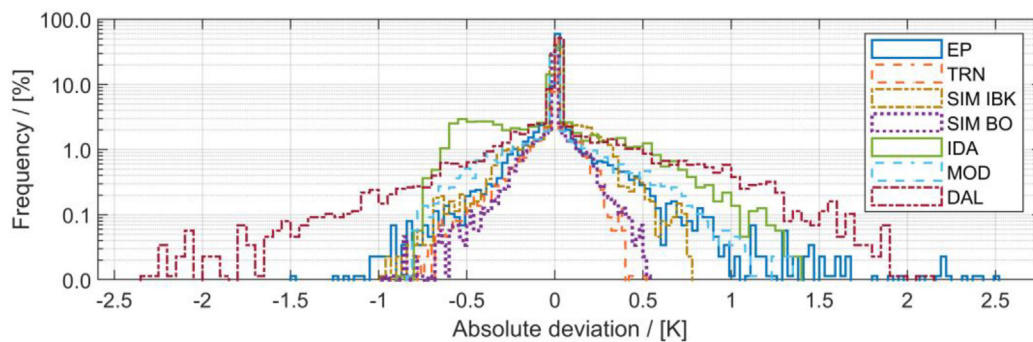


**Figure B2.** Hourly results (i.e. convective Temperature, ventilation plus infiltration losses, solar gain, heating and cooling power) for all the dynamic simulation tools considering the climate of Rome and 4 representing periods for winter, spring, summer and autumn.

- The convective temperature in IDA and MOD is reacting faster than other tools;
- The ventilation system is starting one hour earlier in DAL and IDA than other tools in spring, summer and autumn since DAL and IDA consider the daylight-saving time;
- DAL has sometimes higher peaks of the solar gains compared to the other tools due to the different sky model and shading control;
- In winter EP has heating power peaks due to non-completely ideal control.



**Figure B3.** Absolute deviation of the hourly heating power for all the tools for one-year simulation in the climate of Stockholm, (a) considering the whole year, (b) considering only periods in which the heating system is working. The deviation is calculated as the hourly heating power of the tool minus the hourly reference power (median of all the tools).



**Figure B4.** Absolute deviation of the hourly convective temperature for all the tools for one-year simulation in the climate of Stockholm. The deviation is calculated as the hourly convective temperature of the tool minus the hourly reference convective temperature (median of all the tools).

The histogram reporting the distribution of the deviations can give precious information regarding the sign of the deviation and how often it occurs.

Figure B.3 shows the distribution of the absolute deviations of the heating power for all the tools considering the climate of Stockholm including the whole year (a) and only periods in which the heating system is active (b). It can be noticed that the distribution of the deviations for the different tools is the same in graphs a and b but the frequency of the deviation is drastically higher when only periods in which the heating system is working are considered.

Figure B.4 shows the distribution of the absolute deviations of the convective temperature for all the tools considering the climate of Stockholm.

From Figures B.3 and B.4 it can be easily noticed that IDA and EP reported a non-symmetrical distribution of the deviations. In particular, IDA reported a lower convective temperature ( $-0.5$  K)

more often than other tools. Starting from this information and analysing the hourly plot for the whole year, it was possible to see that the convective temperature of IDA in summer during the night is decreasing slightly faster than the other tools. For EP it can be noticed that the distribution of the convective temperature deviation (see Figure B.4) is shifted towards the positive side and the distribution of the deviation of the heating power is shifted towards the negative side (see Figure B.3). Analysing the hourly plot for the whole year, it was possible to notice that the ventilation control of EP on some days during winter caused lower ventilation losses with respect to the other tools. The peaks of the heating power of EP, visible in Figure 3.4 (hour 775), caused the frequency peak correspondent to  $10$  W/m<sup>2</sup> in Figure B.3. The high dispersion in both convective temperature and heating power deviations for DAL is due to the simplified capacity model used within this tool (see Sections 2.3, 2.4, and 3.2.2).

### C. Summary of reference office cell description and boundary conditions

In Table C1 a short description of the office building inputs and of the applied boundary conditions is provided.

**Table C1**  
Summary of reference office cell description and boundary conditions.

Geometry	Dimensions Facade orientation Window to wall ratio: Frame to window ratio:	width: 4.5 m; depth: 6 m; height: 3 m (27 m <sup>2</sup> ) South 60% 25%
Fenestration	Stockholm: Stuttgart: Rome:	HTC <sub>gl</sub> : 0.81 W/(m <sup>2</sup> K), HTC <sub>frame</sub> : 1.18 W/(m <sup>2</sup> K), τ <sub>sol</sub> : 0.46, SHGC: 0.63, CEN HTC <sub>gl</sub> : 1.40 W/(m <sup>2</sup> K), HTC <sub>frame</sub> : 1.18 W/(m <sup>2</sup> K), τ <sub>sol</sub> : 0.43, SHGC: 0.59, CEN HTC <sub>gl</sub> : 1.29 W/(m <sup>2</sup> K), HTC <sub>frame</sub> : 1.18 W/(m <sup>2</sup> K), τ <sub>sol</sub> : 0.26, SHGC: 0.33, CEN
Shading Facade	Stockholm: Stuttgart: Rome:	Active when direct solar radiation on the south façade is higher than 120 W/m <sup>2</sup> HTC <sub>wall</sub> = 0.3 W/m <sup>2</sup> K HTC <sub>wall</sub> = 0.4 W/m <sup>2</sup> K HTC <sub>wall</sub> = 0.8 W/m <sup>2</sup> K
Ceiling, walls, floor Internal gains	People: Occupancy: Lighting: Equipment:	Mixed: heavy weight floor/ceiling, lightweight walls 3 (variable occupancy), 120 W/pers. Weekdays: 8:00–19:00, 3 people (variable occupancy), 120 W per person 10.9 W/m <sup>2</sup> (hourly profile) 7.0 W/m <sup>2</sup> (hourly profile)
HVAC and settings	Infiltration: Ventilation: Set points:	ACH: 0.15 Constant during occupied hours, 40 m <sup>3</sup> /(h*pers.) Sensible heat recovery with controlled bypass, efficiency: 70% Lower set point: 21 °C, Upper set point: 25 °C (constant)
Weather		TM2, Stockholm, Stuttgart, Rome

### References

- [1] H. Wang, Z. Zhai, Advances in building simulation and computational techniques: a review between 1987 and 2014, *Energy Build.* 128 (2016) 319–335, <https://doi.org/10.1016/j.enbuild.2016.06.080>.
- [2] V.S.K.V. Harish, A. Kumar, A review on modeling and simulation of building energy systems, *Renew. Sustain. Energy Rev.* 56 (2016) 1272–1292, <https://doi.org/10.1016/j.rser.2015.12.040>.
- [3] D.B. Crawley, J.W. Hand, M. Kummert, B.T. Griffith, Contrasting the capabilities of building energy performance simulation programs, *Build. Environ.* 43 (4) (2008) 661–673, <https://doi.org/10.1016/j.buildenv.2006.10.027>.
- [4] Building Energy Software Tools, Best Directory | Building Energy Software Tools, (2020). [https://www.buildingenergysoftwaretools.com/?\\_\\_cf\\_chl\\_jschl\\_tk\\_\\_=333a9062c0deb41f0aedbc09d4de96cab9d74a2d-1595925926-0-Ac80m\\_md0jCDx9IWuUyq88WYrQj-imLS12aG2ZaEQAA0YJNQ0m3L4gPuZsqdDSQZjPvfeimai8ls\\_wyDsp2znHqB0uwrDsw9iKK61KqCSarbSfBzG2zeBoLKS2vw74sxlA8bi7](https://www.buildingenergysoftwaretools.com/?__cf_chl_jschl_tk__=333a9062c0deb41f0aedbc09d4de96cab9d74a2d-1595925926-0-Ac80m_md0jCDx9IWuUyq88WYrQj-imLS12aG2ZaEQAA0YJNQ0m3L4gPuZsqdDSQZjPvfeimai8ls_wyDsp2znHqB0uwrDsw9iKK61KqCSarbSfBzG2zeBoLKS2vw74sxlA8bi7) (accessed July 28, 2020).
- [5] T. Zakula, M. Bagaric, N. Ferdelji, B. Milovanovic, S. Mudrinic, K. Ritosa, Comparison of dynamic simulations and the ISO 52016 standard for the assessment of building energy performance, *Appl. Energy* 254 (2019) 113553, <https://doi.org/10.1016/j.apenergy.2019.113553>.
- [6] M. Trčka, J.L.M. Hensen, Overview of HVAC system simulation, *Autom. Constr.* 19 (2) (2010) 93–99, <https://doi.org/10.1016/j.autcon.2009.11.019>.
- [7] S. Kota, F.J.F. Stipo, W. Jeong, J.B. Kim, J.L.B. Alcocer, M.J. Clayton, W. Yan, J.S. Haberl, Development of a reference building information model for thermal model compliance testing-part I: Guidelines for Generating Thermal Model Input Files, *ASHRAE Trans.* 122 (2016) 256–266.
- [8] F. Farias, S. Kota, W. Jeong, J.B. Kim, J.L.B. Alcocer, J.S. Haberl, M.J. Clayton, W. Yan, Development of a reference building information model (BIM) for thermal model compliance testing-part ii: test cases and analysis, *ASHRAE Trans.* 125 (2019) 750–764.
- [9] CARNOT Toolbox - File Exchange - MATLAB Central, (2020). <https://de.mathworks.com/matlabcentral/fileexchange/68890-carnot-toolbox> (accessed August 19, 2020).
- [10] D. Siegle, E. Leonardi, F. Ochs, A new MATLAB Simulink Toolbox for Dynamic Building Simulation with B.I.M. and Hardware in the Loop compatibility, *Proc. Build. Simul. 2019 16th Conf. IBPSA.* (2019) 2651–2658, <https://doi.org/10.26868/25222708.2019.210641>.
- [11] J.P. Campana, G.L. Morini, BESTEST and EN ISO 52016 benchmarking of ALMABuild, a new open-source simulink tool for dynamic energy modelling of buildings, *Energies* 12 (2019) 2938, <https://doi.org/10.3390/en12152938>.
- [12] Berkeley National Laboratory, Open source library for building and district energy and control systems, (2020). <https://simulationresearch.lbl.gov/modelica/> (accessed November 30, 2020).
- [13] Bartenbach, University, of Innsbruck, DALEC, Zubotel Lighting, 2020, <http://dalec.uibk.ac.at/#room> (accessed November 30, 2020).
- [14] Passivhaus Institut, (2020). [https://passiv.de/en/04\\_phpp/04\\_phpp.htm](https://passiv.de/en/04_phpp/04_phpp.htm) (accessed August 25, 2020).
- [15] Mara Magni, Jean Pierre Campana, Fabian Ochs, Gian Luca Morini, Numerical investigation of the influence of heat emitters on the local thermal comfort in a room, *Build. Simul.* 12 (3) (2019) 395–410, <https://doi.org/10.1007/s12273-019-0506-8>.
- [16] D. Mazzeo, N. Matera, C. Cornaro, G. Oliveti, P. Romagnoni, L. De Santoli, EnergyPlus, IDA ICE and TRNSYS predictive simulation accuracy for building thermal behaviour evaluation by using an experimental campaign in solar test boxes with and without a PCM module, *Energy Build.* 212 (2020) 109812, <https://doi.org/10.1016/j.enbuild.2020.109812>.
- [17] L. Giraud, R. Bavière, C. Paulus, Modeling of solar district heating: a comparison between TRNSYS and MODELICA, in: EuroSun, Aix-les-Bains, 2014: p. 11. <https://doi.org/10.18086/eurosun.2014.19.06>.
- [18] W. Tian, A review of sensitivity analysis methods in building energy analysis, *Renew. Sustain. Energy Rev.* 20 (2013) 411–419, <https://doi.org/10.1016/j.rser.2012.12.014>.
- [19] A.S. Solmaz, A critical review on building performance simulation tools, *Int. J. Sustain.* 12 (2019) 7–21.
- [20] Matthias Werner, David Geisler-Moroder, Bert Junghans, Oliver Ebert, Wolfgang Feist, DALEC—a novel web tool for integrated day- and artificial light and energy calculation, *J. Build. Perform. Simul.* 10 (3) (2017) 344–363, <https://doi.org/10.1080/19401493.2016.1259352>.
- [21] Hema Sree Rallapalli, A Comparison of EnergyPlus and eQUEST Whole Building Energy Simulation Results for a Medium Sized Office Building, Arizona State University, 2010.
- [22] Tianzhen Hong, Fred Buhl, Philip Haves, Stephen Selkowitz, Michael Wetter, Comparing computer run time of building simulation programs, *Build. Simul.* 1 (3) (2008) 210–213, <https://doi.org/10.1007/s12273-008-8123-y>.
- [23] D. Kim, W. Zuo, J.E. Braun, M. Wetter, Comparisons of building system modeling approaches for control system design, in: *Build. Simul.* (2013) 3267–3274.
- [24] A.T. Nguyen, S. Reiter, P. Rigo, A review on simulation-based optimization methods applied to building performance analysis, *Appl. Energy.* 113 (2014) 1043–1058, <https://doi.org/10.1016/j.apenergy.2013.08.061>.
- [25] W. Feist, Thermische Gebäudesimulation: kritische Prüfung unterschiedlicher Modellansätze, Müller Jur Verl, Heidelberg, 1994.
- [26] R. Judkoff, D. Wortman, J. Burch, Empirical Validation of Building Analysis Simulation Programs: A Status Report, Solar Energy Research Institute, Golden, Colorado, 1982.
- [27] R. Judkoff, D. Wortman, B. O'Doherty, J. Burch, A methodology for validating building energy analysis simulations, *NREL Technical Rep.* 550–42059 (2008) 1–192.

- [28] Paul Strachan, Katalin Svehla, Ingo Heusler, Matthias Kersken, Whole model empirical validation on a full-scale building, *J. Build. Perform. Simul.* 9 (4) (2016) 331–350, <https://doi.org/10.1080/19401493.2015.1064480>.
- [29] P. Nageler, G. Schweiger, M. Pichler, D. Brandl, T. Mach, R. Heimrath, H. Schranzhofer, C. Hochenauer, Validation of dynamic building energy simulation tools based on a real test-box with thermally activated building systems (TABS), *Energy Build.* 168 (2018) 42–55, <https://doi.org/10.1016/j.enbuild.2018.03.025>.
- [30] G. Dermentzis, F. Ochs, M. Gustafsson, T. Calabrese, D. Siegele, W. Feist, C. Dipasquale, R. Fedrizzi, C. Bales, A comprehensive evaluation of a monthly-based energy auditing tool through dynamic simulations, and monitoring in a renovation case study, *Energy Build.* 183 (2019) 713–726, <https://doi.org/10.1016/j.enbuild.2018.11.046>.
- [31] Y.J. Kim, S.H. Yoon, C.S. Park, Stochastic comparison between simplified energy calculation and dynamic simulation, *Energy Build.* 64 (2013) 332–342, <https://doi.org/10.1016/j.enbuild.2013.05.026>.
- [32] S.A. Brideau, I. Beausoleil-Morrison, M. Kummert, A. Wills, Inter-model comparison of embedded-tube radiant floor models in BPS tools, *J. Build. Perform. Simul.* 9 (2) (2016) 190–209, <https://doi.org/10.1080/19401493.2015.1027065>.
- [33] ANSI/ASHRAE 140–2017, Standard method of test for the evaluation of building energy analysis computer programs, American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, 2017.
- [34] J. Neymark, R. Judkoff, M. Kummert, R. Muehleisen, A. Johannsen, N. Kruijs, J. Glazer, R. Henninger, M. Witte, E. Ono, H. Yoshida, Y. Jiang, X. Zhou, T. McDowell, M. Hiller, J. An, D. Yan, J. Allison, P. Strachan, Update of ASHRAE Standard 140 Section 5.2 and Related Sections (BESTEST Building Thermal Fabric Test Cases), Argonne National Laboratory, ANL-20/26 158451, Argonne, 2020. <https://doi.org/https://doi.org/10.2172/1643690>.
- [35] EN, 15225:2007, Energy performance of buildings - Sensible room cooling load calculation - General criteria and validation procedures 2007 European Committee for standardization Brussels
- [36] EN 15265:2007, Thermal performance of buildings - Calculation of energy needs for space heating and cooling using dynamic methods - General criteria and validation procedures, European Committee for standardization, Brussels, 2007.c
- [37] VDI-Richtlinien 6020:2016-09, Anforderungen an thermisch-energetische Rechenverfahren zur Gebäude- und Anlagensimulation, VDI-Gesellschaft Bauen und Gebäudetechnik, Düsseldorf, 2016.
- [38] VDI-Richtlinien 2078:2015-06, Berechnung der thermischen Lasten und Raumtemperaturen, VDI-Gesellschaft Bauen und Gebäudetechnik, Düsseldorf, 2015.
- [39] VDI-Richtlinien 6007:2015, Berechnung des instationären thermischen Verhaltens von Räumen und Gebäuden, VDI-Gesellschaft Bauen und Gebäudetechnik, Düsseldorf, 2015.
- [40] R. Judkoff, J. Neymark, Model validation and testing: the methodological foundation of ASHRAE standard 140, *ASHRAE Trans.* 112 (2006) 367–376.
- [41] K. Agnew, M. Goldberg, Chapter 8: Whole-Building Retrofit with Consumption Data Analysis Evaluation Protocol, The Uniform Methods Project: Methods for Determining Energy-Efficiency Savings for Specific Measures, National Renewable Energy Laboratory, NREL/SR-7A40-68564, Golden, CO, 2017.
- [42] ASHRAE Guidelines 14–2014, Measurement of energy and demand savings, American Society of Heating Refrigerating and Air-Conditioning Engineers, Atlanta, 2014.
- [43] U.S. Department of Energy, M&V Guidelines: Measurement and Verification for Performance-Based Contracts Version 4.0, Federal Energy Management Program, 2015.
- [44] International Performance Measurement & Verification Protocol Committee, International Performance Measurement & Verification Protocol (IPMVP): Concepts and Options for Determining Energy and Water Savings Volume I, (2002).
- [45] E. Fabrizio, V. Monetti, Methodologies and advancements in the calibration of building energy models, *Energies* 8 (2015) 2548–2574, <https://doi.org/10.3390/en8042548>.
- [46] M. Royapoor, T. Roskilly, Building model calibration using energy and environmental data, *Energy Build.* 94 (2015) 109–120, <https://doi.org/10.1016/j.enbuild.2015.02.050>.
- [47] G.R. Ruiz, C.F. Bandera, Validation of calibrated energy models: common errors, *Energies* 10 (2017) 1587, <https://doi.org/10.3390/en10101587>.
- [48] M.J. Witte, J. Glazer, D.B. Crawley, R.H. Henninger, J. Glazer GARD Analytics, Testing and Validation of a New Building Energy Simulation Program, in Seventh Int. IBPSA Conf., Rio de Janeiro 2001 9
- [49] TRNSYS 18: A Transient System Simulation Program, Solar Energy Laboratory - University of Wisconsin, (2021). <http://sel.me.wisc.edu/trnsys>. (accessed April 7, 2021).
- [50] M. Wetter, W. Zuo, T.S. Nouidui, X. Pang, Modelica buildings library, *J. Build. Perform. Simul.* 7 (2014) 253–270, <https://doi.org/10.1080/19401493.2013.765506>.
- [51] IDA ICE - Simulation Software | EQUA, (2020). <https://www.equa.se/en/ida-ice> (accessed July 20, 2020).
- [52] Testing and Validation | EnergyPlus, (2020). <https://energyplus.net/testing> (accessed August 3, 2020).
- [53] T. Stephane Nouidui, K. Phalak, W. Zuo, M. Wetter, Validation and Application of the Room Model of the Modelica Buildings Library, in: 9th Int. Model. Conf., Munich, 2012: p. 9.
- [54] EQUA Simulation AB, Technical report: Validation of IDA Indoor Climate and Energy 4.0 build 4 with respect to ANSI/ASHRAE Standard 140–2004, 2010.
- [55] P. Voit, T. Lechner, M. Schuler, Common EC validation procedure for dynamic building simulation programs - application with TRNSYS, in: Conf. Int. Simul. Soc., TRANSOLAR GmbH, Zurich, 1994.
- [56] M. Kottek, J. Grieser, C. Beck, B. Rudolf, F. Rubel, World Map of the köppen-geiger climate classification updated, *Meteorol. Zeitschrift.* 15 (2006) 259–263, <https://doi.org/10.1127/0941-2948/2006/0130>.
- [57] F. Ochs, M. Magni, P. Bonato, M. D'Antoni, D. Geisler-Moroder, M. Hauer, S. de Vries, R. Loonen, IEA SHC TASK 56 | Building Integrated Solar Envelope Systems for HVAC and Lighting: System Simulation Models, Solar Heating & Cooling Programme International Energy Agency, 2020. <https://doi.org/10.18777/ieashc-task56-2020-0004>.
- [58] M. Werner Dissertation: Gekoppelte lichttechnische und thermische Methoden zur Ganzjahresbewertung von Fassadensystemen für die Planungspraxis 2017 Universität Innsbruck
- [59] R. Perez, P. Ineichen, R. Seals, J. Michalsky, R. Stewart, Modeling daylight availability and irradiance components from direct and global irradiance, *Sol. Energy* 44 (1990) 271–289, [https://doi.org/10.1016/0038-092X\(90\)90055-H](https://doi.org/10.1016/0038-092X(90)90055-H).
- [60] TRNSYS 18: Programmer's Guide, Solar Energy Laboratory, University of Wisconsin-Madison, 2017.
- [61] SIA 2024:2015 Bauwesen: Raumnutzungsdaten für die Energie- und Gebäudetechnik, Schweizerischer Ingenieur- und Architektenverein, Zürich, 2015.
- [62] U.S. Department of Energy's (DOE) Building Technologies Office, EnergyPlus, (2020). <https://energyplus.net/> (accessed November 30, 2020).
- [63] EQUA, IDA ICE, (2020). <https://www.equa.se/en/ida-ice> (accessed November 30, 2020).
- [64] P. Sahlin, A. Bring, E.F. Sowell, The Neutral Model Format for Building Simulation, in: IBPSA Build. Simul., Vancouver, 1989: pp. 1–11.
- [65] S. Lohmann, Einführung in die Software MATLAB® - Simulink® und die Toolboxes CARNOT und Stateflow® zur Simulation von Gebäude- und Heizungstechnik, Düsseldorf, Germany, 2013.
- [66] K.N. Rhee, K.W. Kim, A 50 year review of basic and applied research in radiant heating and cooling systems for the built environment, *Build. Environ.* 91 (2015) 166–190, <https://doi.org/10.1016/j.buildenv.2015.03.040>.
- [67] W. Zuo, Advanced simulations of air distributions in buildings, *Purdue University*, 2010.
- [68] L. Eriksson, G. Grozman, P. Grozman, M. Havgaard Vorre, L. Ålenius, CFD-Free, Efficient, Micro Indoor Climate Prediction in Buildings, in: *Build. Simul. Optim. Conf.* 12, Loughborough, UK, 2012: p. 8.
- [69] D. Morries Grenfell, Building Heat Transfer, John Wiley & Sons, Ltd, 2004.
- [70] TRNSYS 18 User Manual: Multizone Building modeling with Type56 and TRNBuild, Solar Energy Laboratory, University of Wisconsin-Madison, 2017.
- [71] ISO 13790:2008, Energy performance of buildings - Calculation of energy use for space heating and cooling, 2008.
- [72] D. K. Arasteh, E. U. Finlayson, C. Huizenga, WINDOW 4.1: Program Description, Lawrence Berkeley Laboratory, 1994.
- [73] C. Curcija, S. Vidanovic, R. Hart, J. Jonsson, R. Mitchell, WINDOW Technical Documentation, Windows and Envelope Materials Group, Lawrence Berkeley National Laboratory; Windows and Envelope Materials Group, Berkeley, 2018.
- [74] E.U. Finlayson, D.K. Arasteh, C. Huizenga, M. Rubin, S.M. Reilly, Window 4.0: Documentation of Calculation Procedures, Lawrence Berkeley National Laboratory, Berkeley, 1993.
- [75] D.K. Arasteh, S.M. Reilly, M.D. Rubin, A versatile procedure for calculating heat transfer through windows, *ASHRAE Trans.* 95 (1989).
- [76] L. Mazzarella, M. Pasini, CTF vs FD based numerical methods: accuracy, stability and computational time's comparison, *Energy Procedia* 78 (2015) 2620–2625, <https://doi.org/10.1016/j.egypro.2015.11.324>.
- [77] M. Magni, F. Ochs, P. Bonato, M. D'Antoni, D. Geisler-moroder, S. De Vries, Comparison of Simulation Results for an Office Building Between Different BES Tools - The Challenge of Getting Rid of Modeller Influence and Identifying Reasons for Deviations, in: Proc. 16th IBPSA Conf., Rome, 2019: pp. 1475–1482. <https://doi.org/https://doi.org/10.26868/25222708.2019.210834>.
- [78] R.J. Hyndman, A.B. Koehler, Another look at measures of forecast accuracy, *Int. J. Forecast.* 22 (2006) 679–688, <https://doi.org/10.1016/j.ijforecast.2006.03.001>.
- [79] P. Sahlin, P. Grozman, IDA Simulation Environment a tool for Modelica based end-user application deployment, in: 3rd Int. Model. Conf., Linköping, Sweden, 2003: p. 11.
- [80] M. Wetter, C. Van Treeck, L. Helsen, A. Maccarini, D. Saelens, D. Robinson, G. Schweiger, IBPSA Project 1: BIM/GIS and Modelica framework for building and community energy system design and operation - Ongoing developments, lessons learned and challenges, in: IOP Conf. Ser. Earth Environ. Sci., Graz, 2019: p. 9. <https://doi.org/10.1088/1755-1315/323/1/012114>.
- [81] G. Serale, M. Fiorentini, A. Capozzoli, D. Bernardini, A. Bemporad, Model predictive control (MPC) for enhancing building and HVAC system energy efficiency: Problem formulation, applications and opportunities, *Energies* 11 (2018) 631, <https://doi.org/10.3390/en11030631>.
- [82] F. Ochs, M. Magni, M. Hauer, D. Geisler-Moroder, P. Bonato, S. de Vries, R. Loonen, D. Venus, N. Abdelnour, T. Calabrese, E. Venturi, A. Maccarini, S. Häringer, B. Bueno, Z. Ioannidis, R. Efstratios, IEA SHC T56 - System Simulation Models System Simulation Models, Solar Heating & Cooling Programme International Energy Agency, 2020. <https://doi.org/10.18777/ieashc-task56-2020-0005>.
- [83] M. Wetter, W. Zuo, T. Stephane Nouidui, Modeling of Heat Transfer in Rooms in the MODELICA "Buildings" Library, in: Proc. Build. Simul., 2011: pp. 1096–1103.