Building and Environment 180 (2020) 107081

Contents lists available at ScienceDirect

Building and Environment

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

Validation of dynamic hygrothermal simulation models for historical buildings: State of the art, research challenges and recommendations

H.E. Huerto-Cardenas [a,*], F. Leonforte [a], N. Aste [a], C. Del Pero [a], G. Evola [b], V. Costanzo [b], E. Lucchi [c]

[a] Architecture, Built Environment and Construction Engineering Department, Politecnico di Milano, Via Bonardi 9, 20133, Milano, Italy
[b] Department of Electric, Electronic and Computer Engineering, University of Catania, Viale Andrea Doria 6, 95125, Catania, Italy
[c] Institute for Renewable Energy, Eurac Research, Viale Druso 1, 39100, Bolzano-Bozen, Italy

ARTICLE INFO

Keywords:
Historical buildings
Building performance simulation
Microclimatic parameters
Model calibration
Model validation
Statistical indices

ABSTRACT

The proper simulation of the hygrothermal behaviour of historical buildings is a challenging task with several implications regarding the evaluation of indoor thermal comfort and the suitability of retrofit strategies that comply with the conservation of cultural heritage. An inaccurate simulation may lead to inadequate conclusions, which could result in inappropriate and dangerous actions for the preservation of the heritage buildings.

Then calibration and validation of hygrothermal simulation models are essential steps to achieve more accurate and reliable results.

Now, although some agencies have developed guidelines and methodologies to carry out the validation of building performance models, all of them are based on energy consumption only. However, since in some buildings the energy consumption data are not always available especially when no operating heating, ventilation and air conditioning system is installed, which is the case of many historical buildings, the microclimatic parameters are usually adopted in the validation process. In this case, neither protocols nor specific parameters have been officially recognised to perform the model validation. The present work reviewed the main approaches used by researchers for building performance model validation with special reference to historical buildings based on microclimatic parameters, highlighting the main advantages and drawbacks of the different methods reviewed. Finally, recommendations to properly carry out the model validation based on microclimatic parameters have been provided. The collected information may be useful to different subjects (e.g. designers, energy auditors, researchers, conservators, buildings' owners and policy makers) and can drive suitable and reliable retrofit and maintenance interventions.

1. Introduction

The energy refurbishment of historical buildings should be considered a pivotal issue to achieve the overall decarbonisation target by the mid-century [1,2]. Indeed, the existing building stock is extremely energy-consuming because of the age of the buildings, the type of materials used for their construction, the absence of planned maintenance, and the poor efficiency of their energy systems [3]. The high running costs may determine a progressive abandonment of historical buildings, causing their severe decay and, often, the choice to invest in more energy efficient and modern buildings [4]. Of course, thanks to their cultural value, historical buildings (as defined in Ref. [1]) are important pieces of evidence from the past that contribute to the creation of the identity of Countries and must be safeguarded for the next generations through compatible and sustainable strategies and solutions [5,6]. Nowadays, each choice of the interventions on historical buildings faces even the challenge of optimizing and balancing energy efficiency and indoor microclimate, thus ensuring their full preservation [3,6].

In such framework, building performance simulation (BPS) software tools have acquired a paramount role since they can be used both for accurate performance predictions and to investigate suitable scenarios of intervention, aimed at decreasing the energy consumption, improving thermal comfort and reducing the damage risks for building materials and artworks [7].

In this context, the accuracy of BPS models is an important issue, which can be dealt with through the so-called calibration process. ASHRAE Guidelines 14-2014 define calibration as the *"[…] process of*

* Corresponding author. Via Giuseppe Ponzio 31, 20133, Milano, Italy.
E-mail address: haroldenrique.huerto@polimi.it (H.E. Huerto-Cardenas).

https://doi.org/10.1016/j.buildenv.2020.107081
Received 16 March 2020; Received in revised form 16 June 2020; Accepted 16 June 2020
Available online 6 July 2020
0360-1323/

Nomenclature

T_a	dry bulb air temperature [°C]
T_s	surface temperature [°C]
RH	air relative humidity [%]
AH	absolute humidity [g/m^3]
SH	specific humidity [g/kg]
MR	mixing ratio or humidity ratio [g/kg]
P_v	water vapour pressure [Pa]
E_c	energy consumption [kWh]
n	number of data samples
m_i	measured data
$\overline{m}$	mean value of measured data
m_{max}	maximum value of measured data
m_{min}	minimum value of measured data
s_i	simulated data
$\overline{s}$	mean value of simulated data
MBE	mean bias error
MAE	mean absolute error
RMSE	root mean square error
NMBE	normalized mean bias error [%]
CVRMSE	coefficient of variation of the RMSE [%]
RN_RMSE	range normalization of the RMSE [%]
r	Pearson correlation coefficient
R^2	coefficient of determination
IC	Inequality Coefficient
σ	standard deviation
μ	mean value
F_i	frequency of residuals
Q_v	prediction rate
BPS	building performance simulation
HVAC	heating ventilation and air conditioning
ECM	energy conservation measure

reducing the uncertainty of a model by comparing the predicted output of the model under a specific set of conditions to the actual measured data for the same set of conditions" [8]. As stated by De Wilde, calibration includes *"[…] the process of adjusting physical modelling parameters in the computational model to improve agreement with experimental data"* [9,10]. During a calibration process, suitable statistical indices have to be calculated and compared with corresponding acceptance criteria, which allow considering the model validated and, thus, reliable [11]. The validation is *"[…] the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model"* [9,10]. Thus, validation should be considered a pivotal part of the calibration process, as stated in the ASHRAE guidelines [8,12].

The results obtained from a validated model can be useful to different subjects (e.g. designers, energy auditors, conservators, etc.), and can drive suitable maintenance interventions to prevent possible damage problems. Then, the improvement of the simulation accuracy allows decreasing the uncertainty about the outcomes and can offer a significant contribution to the development of new rules and strategies, innovating the traditional conservative approaches (i.e. restoration, preventive and planned conservation). Now, although some agencies have developed guidelines and methodologies to carry out the validation of a building performance model [12–14], all of them are based on energy consumption figures only. However, model calibration and validation can in principle be based also on microclimatic parameters such as indoor air temperature (T_a) and relative humidity (RH), especially when no operating HVAC system is present [15], which is exactly the case of many historical buildings.

Even if microclimatic parameters have been largely exploited in the field of simulation of historical buildings (see the appendix), as witnessed by a high number of studies reported in Table 5, no officially recognised method or standard is available to carry out the model validation on these bases [15–17].

2. Aims & methodology

The general goal of the present work is to describe the state of the art of the validation approaches adopted by the researchers in the domain of the dynamic energy simulation of historical buildings. The investigation is limited to those approaches that rely on the comparison between measured and simulated values of microclimatic parameters. More in detail, the work focuses on the selection of the most suitable control parameters and statistical indices required for the evaluation of model uncertainty, and on the suitability of the corresponding thresholds. The paper discusses some common trends, advantages, drawbacks, and key points emerging from the reviewed literature, highlighting potentialities and critical issues. On the other hand, it does not discuss the calibration approaches available to fine-tune the simulation models, since this topic is already addressed in other research works [18–20].

The literature background of the present work contains academic studies (i.e. scientific papers, conference proceedings, books, etc.) and "grey literature" (i.e. technical reports, and governmental guidelines), to consider both scientific aspects and empirical approaches. Finally, the references included in the scientific papers have been also considered, to ensure that all relevant published papers are covered in the study.

The outcomes of this work should be regarded as the first step in order to define a common methodology for the validation of BPS models based on microclimatic parameters. The set-up of this common methodology is pivotal to evaluate properly the reliability of BPS models for any buildings not equipped with HVAC, and not only to the historical or heritage ones.

3. Critical review on validation methodologies of dynamic hygrothermal simulation models for historical buildings

The following discussion is divided in four parts: (i) identification of the most common microclimatic parameters used in model validation (Section 3.1); (ii) selection of the uncertainty statistical indices able to evaluate the model accuracy (Section 3.2); (iii) threshold values generally adopted to consider a model validated (Section 3.3); and (iv) application in the literature of the different approaches (Section 3.4). Finally, a list of recommendations to properly carry out the model validation through microclimatic parameters is summarized.

3.1. Selection of the validation parameters

The increasing use of microclimatic parameters for calibration and validation purposes in heritage BPS is mainly related to the availability of environmental data acquired through high-accuracy measurement devices [21,22], often already installed for several purposes such as risks assessment of building materials and objects [23–38], and to the evaluation of building thermal performance [23–25,33–35,39–54]. One more reason for using microclimatic parameters is the lack of energy consumption data, generally adopted in the model validation. Of course, this latter issue can be due to the absence of heating/cooling systems or to difficulties in retrieving the energy consumption data.

In such conditions, the microclimatic variables involved in model calibration are: indoor dry-bulb air temperature (T_a), surface temperature (T_s), relative humidity (RH), absolute humidity (AH), specific humidity (SH), mixing ratio (MR) and vapour pressure (P_v), as shown in

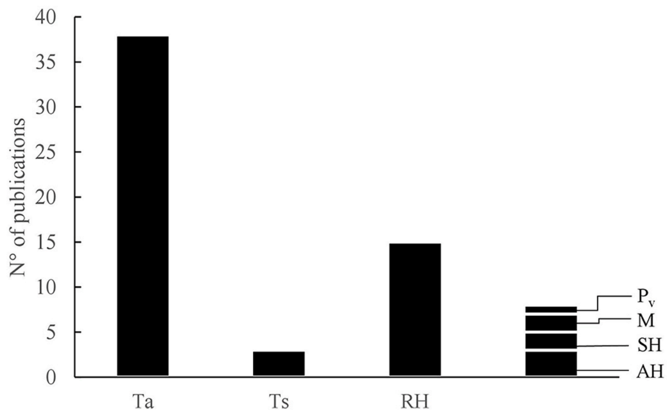

Fig. 1. Control parameters used in historical building simulations.

Fig. 1.

Amongst them, T_a is the one most used by researchers for the calibration and validation of virtual models [15,23–26,31,35,36,45,55–64]: indeed, it is the main parameter involved in human comfort assessment and for this reason it is commonly measured. Furthermore, dynamic simulation models often do not take into account the hygrometric behaviour of environments and structures, especially when the research does not intend to investigate possible problems arising from vapour condensation or humidity content in the structures.

Since the probes used in the monitoring of indoor microclimate in historical buildings can measure both T_a and RH, the latter parameter is sometimes integrated in the validation process [24–26,31,35,36,63]. However, by definition RH depends on the air temperature and this might generate an increase in the error when the simulated temperature is not properly validated. In such respect, it is recommended verifying the matching of both parameters, since a small deviation in temperature might highly affect the relative humidity output of the simulation model. Parameters independent from temperature such as AH, SH, and MR, are often used together with T_a and RH [25,26,31,35,63] as a further verification step of the reliability of a virtual model. In any case, the choice of the most appropriate parameters can be also related to the invasiveness of the sensors needed to measure them while ensuring the smooth running of the activities [30]. For example, Coelho et al. validated a virtual model of a historic church in Lisbon using only P_v to analyse the hygrometric state of the air [61]. They did not rely on RH because this depends on P_v and T_a. According to the authors, this approach avoids error replication in the validation process [61]. T_s is used in a few works as a further verification means of the model reliability, or to improve the accuracy in the estimation of the thermal properties of a building component. In such respect, Roberti et al. validated a model of an unused historic palace of the XIII century in Bolzano (Italy) by using indoor T_a and T_s in several points collected in more than 100 points through sensors located inside and outside the building [56].

In conclusion, it should be stressed that using just the T_a as the control parameter allows achieving a partial model validation. More in detail, this implies that the (vapour) mass exchanges are not taken into account [65] and some humidity-dependent phenomena, such as moisture buffering effect of materials, cannot be correctly evaluated [66]. The use of at least one humidity control parameter should lead to a more reliable model and predictions closer to a real building's response [61].

Regarding the uncertainty of the measurement devices, it should also be taken into account during the model validation process. However, such uncertainty is small compared to the one introduced by the building simulation model.

3.2. Main uncertainty indices used to assess the model accuracy

In the energy simulation of historical buildings, several indices are used to assess the discrepancy between simulated and measured values, which is a key step in the calibration and validation process. However, the nomenclature of these indices can be affected by some common errors as explained by Ruiz and Bandera [67]. At the origin, the mistakes spread through different documents (journals, thesis, reports, etc.) due to unclear existing references. The lack of unified criteria, as well as the tendency of using different methods to evaluate the model accuracy depending on the available sources, are the cause of this growing misunderstanding. For this reason, Ruiz and Bandera stressed the necessity of unifying the validation criteria defined in the ASHRAE Guideline 14 [12], FEMP [13,68] and IPMVP [14].

This review adopts the corrected nomenclature and formulas used in the publication mentioned above [67], as listed in Table 1. Fig. 2 shows which indices are mostly used by researchers in the literature.

In the early years of the building simulation practice, the gap between simulations and experimental measurements was commonly assessed just based on a simple difference between measured and

Table 1
Main uncertainty indices used to evaluate the accuracy of BPS model.

Index	Name	Formula	U.M.
% error	Percent error/difference	$\%\,error = \left(\frac{m-s}{m}\right)\times 100 = \left(1-\frac{s}{m}\right)\times 100$	%
MBE	Mean bias error	$MBE = \frac{\sum_{i=1}^{n}(m_i - s_i)}{n}$	data-dependent
MAE	Mean absolute error	$MAE = \frac{\sum_{i=1}^{n}\lvert m_i - s_i\rvert}{n}$	data-dependent
RMSE	Root mean square error	$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(m_i - s_i)^2}{n}}$	data-dependent
NMBE	Normalized mean bias error	$NMBE = \frac{1}{\overline{m}}\times\frac{\sum_{i=1}^{n}(m_i - s_i)}{n}\times 100$	%
CVRMSE	Coefficient of variation of the RMSE	$CVRMSE = \frac{1}{\overline{m}}\times\sqrt{\frac{\sum_{i=1}^{n}(m_i - s_i)^2}{n}}\times 100$	%
RN_RMSE or NRMSE	Range normalized RMSE or normalized RMSE	$RN\_RMSE = \frac{1}{(max_m - min_m)}\times\sqrt{\frac{\sum_{i=1}^{n}(m_i - s_i)^2}{n}}\times 100$	%
r	Pearson correlation coefficient	$r = \frac{\sum_{i=1}^{n}(m_i-\overline{m})\times(s_i-\overline{s})}{\sqrt{\sum_{i=1}^{n}(m_i-\overline{m})^2}\times\sqrt{\sum_{i=1}^{n}(s_i-\overline{s})^2}}$	–
R^2	Coefficient of determination	$R^2 = 1-\frac{\sum_{i=1}^{n}(m_i - s_i)^2}{\sum_{i=1}^{n}(m_i-\overline{m})^2}$	–
IC	Inequality coefficient	$IC = \frac{\sqrt{\frac{1}{n}\times\sum_{i=1}^{n}(m_i - s_i)^2}}{\sqrt{\frac{1}{n}\times\sum_{i=1}^{n}s_i^2}+\sqrt{\frac{1}{n}\times\sum_{i=1}^{n}m_i^2}}$	–

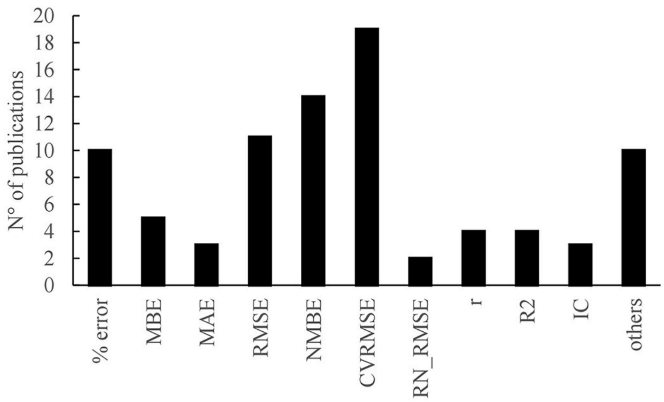

Fig. 2. Uncertainty indices used in historical building simulations.

simulated values in percentage [24,42,44,47,51,52,54,59,69]. Generally, this index applies to instantaneous values. However, for validation purposes, an average or integral value is required. The *% error* is most used to compare measured and simulated values of the total energy consumption [24,42,44,47,51,52,54,59,69], and only in a single case, it is also used with temperature data [23]. Such index is called in the literature in different ways: ERR, percentage difference, relative error and BE. Due to the simplicity and the ease of calculation, it is commonly used for a very first evaluation of the model accuracy.

The *mean bias error* (MBE) is calculated from the sum of the deviations between the measured and simulated data and dividing this number by the number of data points. The result has the same unit of measurement as the data used for its calculation. Positive and negative values for MBE indicate respectively an under and overestimation of the real data by the model. The MBE can be subject to the phenomenon of error cancellation [67], which occurs when the sum of positive and negative values produces a low result, thus it is advisable to use it together with another index, such as the *root mean square error* (RMSE) or the *mean absolute error* (MAE). The first one represents the standard deviation of the differences between measured and simulated data, while the second one takes into account the average absolute error of the differences between measured and predicted values. Both RMSE and MAE provide a result with the same unit of measurement as the data used; compared to the MBE, they are always positive and are not subject to error cancellation. The higher is their value, the lower is the reliability of the model. The RMSE is more sensitive than the MAE to the amplitude of the residuals, because of the square of the residuals in its calculation.

An application of such indices is shown by Roberti et al. [56] who calibrated and validated a BPS model of a XIII century building using EnergyPlus, by minimizing the RMSE referred to indoor T_a and T_s through the particle swarm optimization algorithm (PSO) implemented in the software GenOpt.

Other widely used uncertainty indices are the *normalized mean bias error* (NMBE) and the *coefficient of variation of the root mean square error* (CVRMSE), which are also adopted by the main validation protocols [12–14,68]. The first one is the normalization of MBE and is calculated by dividing this index by the average of the measured data (Table 1). Positive and negative values indicate respectively an underestimation and an overestimation of the real data by the simulated values. The NMBE is subject to error cancellation [67], whose effect is usually emphasized when the model is close to being validated. Instead, the CVRMSE is a normalized form of the RMSE through the average of the measured data. This index is always positive: the higher the result of its calculation, the lower the reliability of the model.

The so-called *range normalized root mean square error* (RN_RMSE or NRMSE) is a normalized form of RMSE that uses the range of the measured data (i.e. the difference between the maximum and the minimum value in the dataset) in the normalization process instead of the average value (Table 1). The RN_RMSE provides a more reliable estimation of the model accuracy compared to the CVRMSE, because it is not affected by shifting and scaling operations [70]. As an example of the drawbacks occurring with these statistical indices, Fig. 3 shows three datasets of measured energy consumption and their residuals from the results of a prediction model plotted against the outside air temperature. In detail, in the graphs a) and b) the data are expressed in kWh, with the difference that in b) each value comes from adding 100 kW h to the corresponding value in a). On the other hand, in graph c) the data are expressed in MJ, hence with a multiplicative factor of 3.6 with respect to the data illustrated in a). The graphs a) and b) show the same standard deviation (σ), but different mean values (μ): this is a case where an additive difference exists between the datasets. Data represented in graphs a) and c) have different σ and μ: this is a case where a multiplicative difference exists between the datasets. As demonstrated in Fig. 3, the CVRMSE changes where an additive difference holds between datasets, while the RN_RMSE remains the same in the three cases, confirming its higher reliability [70]. Cornaro et al. adopted this index for the validation of a historic palace of the XVI century and an office building of 1960 located in Rome (Italy) using the indoor T_a as main control parameter [11,45].

Since the normalized indices use the average (or the range in case of RN_RMSE) of the measured data in their calculation, the results can vary consistently depending on the scale of the dataset adopted [70], even if the gap between measurement and simulation is the same. Hence, defining a general threshold to be used in the validation of simulation models through the normalized indices (with particular reference to microclimatic parameters) can be challenging. Moreover, about temperatures, the use of percentage could lead to misunderstanding. Considering the same temperature discrepancy (e.g. 1 °C or 1 K), it has a completely different percentage impact if one works with Celsius, Fahrenheit or Kelvin scales (and 1 °C discrepancy on a 100 °C value is not, a priori, less important than the same discrepancy on a 50 °C value, what would emerge by using percentages). Anyway, the relevance of a discrepancy in the temperature value depends on the physical process that is being studied.

The *Pearson correlation coefficient* (*r*) and the *coefficient of determination* (R^2) are often used for the assessment of the model accuracy in combination with the other indices previously discussed. While all the other indices quantify the relative closeness of the predictions to the actual values, *r* indicates the linear relationship between two variables and R^2 measures the proportion of the variance in a dependent variable that is predictable from an independent variable. These two coefficients are calculated using measured and simulated data.

In detail, *r* is a dimensionless number ranging between −1 and 1. If *r* is less than 0, the correlation is negative, so an increase in the first parameter (in this case the simulated value) corresponds to a decrease in the second one (the measured value) and vice versa: therefore, the simulated values are not representative of the real behaviour of the building. If *r* is equal to 0 there is no correlation between the variables. When *r* is greater than 0 a positive correlation holds, then the two parameters vary similarly; the model can be considered as representative of the real behaviour of the building when $r > 0.5$, as suggested in Ref. [71]. In some works, *r* is used in the form of r^2 [11,61,72], which range between 0 and 1.

The R^2 is also a dimensionless parameter ranging between 0 and 1: a value close to 0 means that there is no dependence between the data, while a value close to 1 indicates that there is a strict dependence. It is not a mandatory parameter to evaluate model validation, however, both the ASHRAE [12] and IPVMP [14] guidelines recommend that this index must be higher than 0.75.

Another uncertainty index used by researchers is the *Inequality Coefficient* (IC), that represents the magnitude domain due to unequal tendency (mean), unequal variation (variance), and imperfect co-variation (co-variance) [37,64,73]. The IC can range between 0 and 1 (IC = 0 indicates a perfect match, while IC = 1 indicates no match) [37].

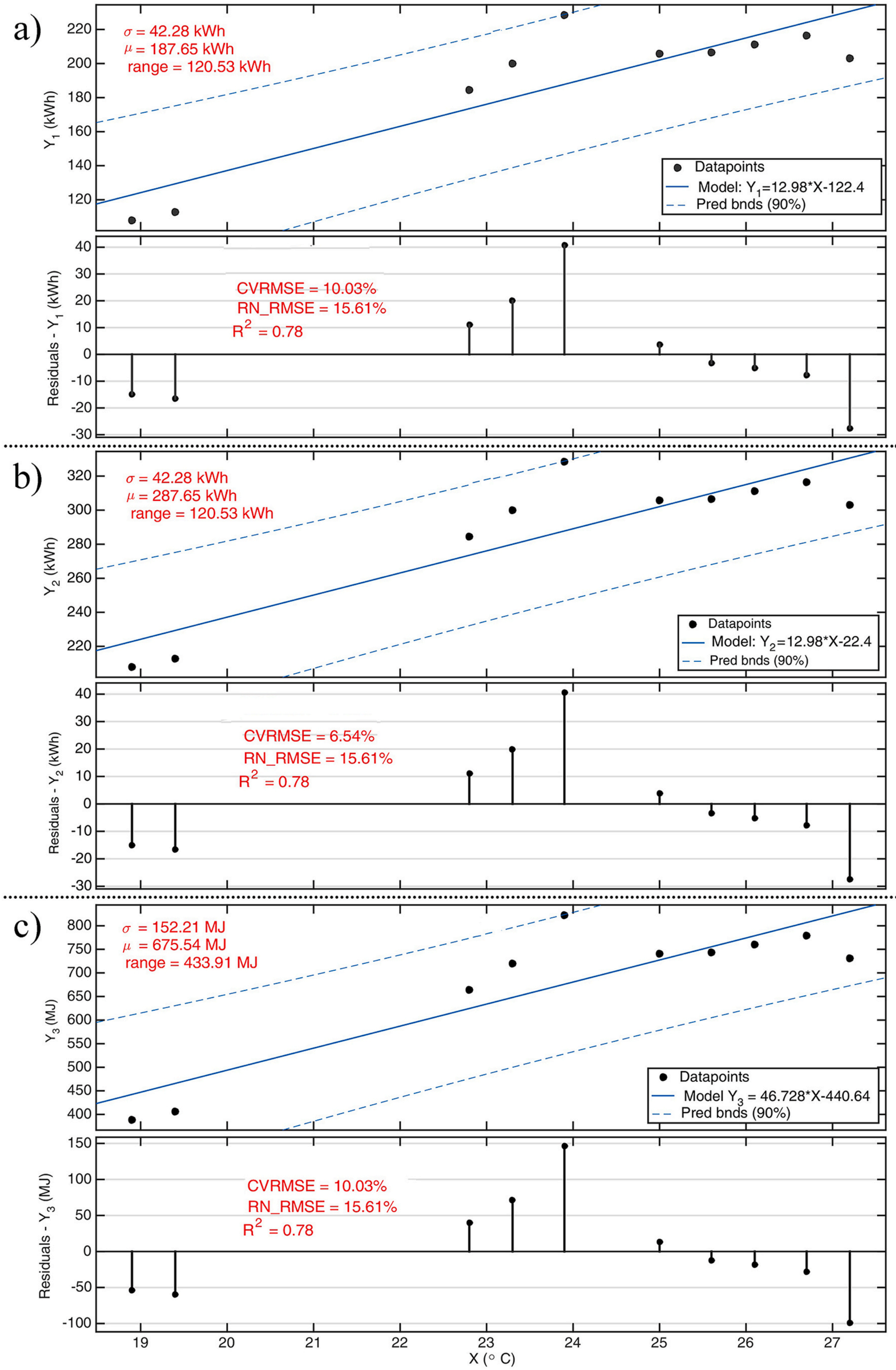

Fig. 3. Analysis of three datasets differing for additive or multiplicative factor. In the three conditions, the RN_RMSE and R^2 do not change, whereas the CVRMSE and μ give different results. Adapted from [70].

A further approach for validation consists in the *analysis of the frequency of residuals* (F_i), that is to say the differences between measured and simulated data for each time step, and in verifying that a sufficiently high amount of residuals are below a suitable threshold.

Such analysis can be more effective than statistical indices: indeed, the results of a simulation are better described through the data distribution within certain intervals and by the count of particular events, especially in the field of cultural heritage conservation [27,74]. Frequency analysis also allows not to exclude those simulations showing an overall good agreement with the measurements, but with some occasional data that fall out of the admissible ranges.

Indeed, some statistical indices (e.g. RMSE, CVRMSE and RN_RMSE) could be strongly influenced by the occurrence of rare but large differences between simulation and real values, which can instead be identified through the frequency distribution of the residuals.

Kilian et al. [74] proposed a different approach in the context of preventive conservation of cultural heritage, where the main variables affecting the physical and chemical mechanisms of decay are identified in T_a and T_s, RH, ventilation rate and air velocity [75]. Particularly, the presence of thermal fluctuations (daily or seasonally) or high thermal levels compared to the conservation standard may cause expansion, acceleration of natural damage, chemical processes, partial drying up, and increase of the fragility of organic artworks (e.g. wood, paper, and textile objects). The thermal damage is often reversible, thus not highly dangerous for the objects. On the contrary, RH has a pivotal role for the good preservation of several materials [30] because hygrometric fluctuations may produce irreversible damage and generate modifications in the size and shape of artefacts, as well as chemical reactions and biological deteriorations (e.g. degradation of materials due to salt crystallisation, frost and defrost, dimensional changes of organic hygroscopic materials, etc.).

This suggests that, in order to verify if the numerical model reproduces with adequate accuracy the relevant damage risks for hygroscopic materials [27,74], the daily fluctuations of RH should be selected as a comparative parameter. In such respect, one can calculate the *prediction rate* Q_v [74]:

$$Q_v = \frac{N_s}{N_m}$$

Here, N_s and N_m are the number of times when the simulated and measured daily RH cycles exceed a certain daily fluctuation threshold (an example is shown in Table 4 in Section 3.4.6). Basically, the method consists in the evaluation of the number of daily RH fluctuation cycles beyond a certain acceptability threshold, and then in the calculation of the ratio between the number of predicted and measured cycles. If this ratio deviates too far from 1, the reliability of the model in predicting the risks of degradation is not sufficient. As proposed by Kilian et al. [74] Q_v in the range of $0.95 \leq Q_v \leq 1.10$ is defined as “excellent”, while the ranges of $0.75 \leq Q_v < 0.95$ and $1.10 < Q_v \leq 1.5$ are “acceptable”.

The great variety of existing uncertainty indices demonstrates the lack of a unique and officially recognised methodology for the assessment of the model accuracy; thus, each researcher uses different metrics depending on the sources consulted. However, a further step to evaluate properly the level of accuracy of a BPS model should consist in the definition of a tolerance range for each index. Currently, the existing protocols define some thresholds for the NMBE and the CVRMSE, with recommended value for the R^2 (shown in Table 2 in the next section). Nevertheless, these thresholds refer to energy consumption as a control parameter.

Table 2
Existing validation criteria based on energy consumption, and their corresponding thresholds.

Data type	Index	FEMP [13,68]	ASHRAE Guideline 14 [12]	IPMVP [14]
Monthly	NMBE (%)	±5	±5	±20
	CVRMSE (%)	15	15	–
Hourly	NMBE (%)	±10	±10	±5
	CVRMSE (%)	30	30	20
Recommendations	R^2	–	>0.75	>0.75

3.3. Thresholds for model validation

As a general rule, BPS models are considered validated if they comply with the criteria set out by one of the following main protocols: ASHRAE Guideline 14 [12], FEMP [13,68] and IPMVP [14].

The origin of these three protocols and guidelines dates back to 1996 when the North American Measurement and Verification Protocol (NEMVP) was published [67]. This protocol aimed to estimate, for a given Energy Conservation Measure (ECM), the energy savings that this introduces starting from an assigned baseline condition. The first version of both the IPMVP protocol and the FEMP was published in 1997 and 2000, respectively.

In 2002, ASHRAE published its Guideline 14 regarding “Measurement of Energy and Demand Savings”. Its intention is “*to provide guidance on minimum acceptable levels of performance for determining energy and demand savings, using measurements*”. This guideline has a more technical approach than the other documents and, as a result, the majority of the scientific community refers to it [67]. Such guideline has defined how energy savings should be assessed, by following any of the four analysis methods, one of which relies on validated simulations. Unfortunately, although Guideline 14 provides procedures for using validated simulations, it does not provide a methodology to calibrate a simulation against measured conditions [18].

The criteria introduced by the previously discussed protocols identify some simple formulas to quantify the uncertainty, anchored in basic statistics; furthermore, they propose the thresholds to consider the model validated according to hourly or monthly measured data (Table 2).

Although some authors adopt such methods and values also to validate simulation models based on microclimatic variables, it should be noted that current criteria relate solely to predicted energy consumption, and do not account for uncertainty or inaccuracies of input parameters, or the accuracy of the simulated environment (as T_a, T_s, RH, etc.) [19].

Some validation thresholds based on microclimate parameters were proposed by Rajčića et al. [27] in the field of cultural heritage conservation. In order to describe the degree of compliance between simulation and measurement, a classification of the validation criteria was defined for T_a and RH through three categories: excellent, acceptable and low. In particular, the difference between simulated and measured data is interpreted as “excellent” when it lies within ±1 °C and ±5% from the median, “acceptable” when values fall within ±3 °C and ±10% from the median for temperature and relative humidity respectively, while “low” when both values are out of these ranges.

3.4. Application of the validation approaches in the literature

Hereafter, the paper will show the different approaches adopted by researchers to validate a BPS model and will present the results achieved according to the metrics used.

In the appendix all the references reviewed have been summarized and reported as a table, providing some useful information about the method adopted for the model validation. These case studies show a great variety in the building typology (churches, palaces, museums, houses, etc.), location (Italy, France, Spain, Netherlands, Sweden, etc.) and construction period (from the XIII to the XX century). Considering that none of the existing standards deals with model calibration and validation based on microclimatic parameters, the trends found in this section will be useful to highlight the most effective approaches and to find suitable recommended thresholds, which is particularly useful in the field of cultural heritage.

3.4.1. ERR (% error)

As already stated in Section 3.2, the % error is mostly used to compare measured and simulated values of total energy consumption and only in a single case it is used with along temperature data [23]. In such work, for the model validation of a Museum built between 1929 and 1930 located in Bytom (Poland) [23], the % error index was used with dry-bulb indoor air temperature. The model validation shows a good agreement between simulations and measurements: indeed, for 99% of the time, the difference between measured and simulated hourly data does not exceed 10%, while for over 85% of the time the difference is below 5%. Further verification was based on the calculation of the correlation coefficient *r*, which was high (between 0.93 and 0.98). A threshold for such index is suggested by Petrelli and Fabbri [65] in the field of historical building simulation, particularly for microclimate assessment: they recommended that virtual models validated through T_a and RH should have errors lower than 5% and 10% respectively. Here, it should be observed that using percentage statistical indices with temperature could lead to misunderstanding, as discussed in Section 3.2.

3.4.2. CVRMSE and NMBE

As stated in the previous section, the CVRMSE and the NMBE are the uncertainty indices most frequently adopted in the existing validation criteria [12–14,68] that are based on the comparison of energy data. However, in the simulation of cultural heritage buildings, these indices are often adopted also for the calibration through microclimatic data (see the Appendix).

In such respect, Dogan Sahin et al. [57] have validated the virtual model of a historic palace of the XIX century located in Izmir (Turkey) using the air temperatures collected in several rooms of the building and adopting the validation criteria defined in ASHRAE Guideline 14 through the calculation of the NMBE and the CVRMSE. The simulations were run for a period of 96 h and the results were compared to the measured temperatures in each room. The resulting values for the NMBE and the CVRMSE were between 0.5%-5.8% and 2.2%–6.7% respectively, that is to say particularly small compared to the limits for validation defined by the Guidelines (Table 2) [12].

A further case study where T_a is used as a control parameter is proposed by Ogando et al. for a school located in Galicia (Spain) built in 1970 [59]. Here, the average NMBE and CVRMSE for several rooms of the building are 2.73% and 11.52% respectively. These values are within the range proposed by the existing protocols/standard (Table 2), however, since these thresholds were defined for energy consumption, their application to microclimatic parameters does not seem the most suitable choice. In such respect, some authors [76,77] recommend to use the narrowest hourly thresholds: $\pm$ 5% and $\leq$20% for NMBE and CVRMSE respectively (Table 2).

A more accurate approach to validate BPS models, which uses the ASHRAE Guideline14 criterion but integrates it with further verification strategies, was followed by Muñoz-González et al. adopting T_a and RH as control parameters for a historic Church of the XVI century located in Seville (Spain) [35]. The calibration was based on three different indicators, namely NMBE, CVRMSE and ε (error values). Furthermore, the analysis was extended by calculating the residuals and determining the number of values that are inside a suitable range for T, RH and AH, as defined by the authors.

Since no officially recognised thresholds are defined for validation through hourly T_a and RH data with the statistical indices adopted in the main protocols, Frasca et al. [17] carried out the model validation of a historical museum built between the XIII-XVI centuries located close to Rome, assuming the measurements' uncertainty of T_a and RH as a rejection criterion. In such regards, the CVRMSE values should not exceed 2% for T_a and 5% for RH respectively, to be considered satisfactory.

The use of these indices is widespread in the literature. However, as discussed in Section 3.2, the result provided by the normalized indices depends on the scale of the dataset adopted. This aspect, particularly with microclimatic parameters could be lead to misinterpretations of the model accuracy.

3.4.3. MBE, MAE and RMSE

Since the thresholds defined by the existing criteria were not specifically introduced to validate BPS models through microclimatic parameters, in some works further verifications strategies were adopted. For example, in the model calibration of a palace of the XIII century located in Bologna (Italy), the validation results with the NMBE and CVRMSE were inside the tolerance range proposed by the standard/protocols (shown in Table 2) and further verification of the model was carried out calculating the RMSE, which was between 0.37 °C and 4.3 °C [58].

A similar approach was used in works that validate virtual models through the MBE, MAE and RMSE. Among them, RMSE is the most used parameter: five works achieved a result lower than 1 °C and considered the model validated [11,17,39,40,56,78], while only two works achieved results between 1 °C and 2 °C [15,29] (see the Appendix). Roberti et al. [56] validated a BPS through T_a and T_s by minimizing the RMSE and showed the results of the validated model in a carpet plot as a function of the hour of the day and month (see Fig. 4).

In the same work, the authors also compared the simulated and monitored time trends for the indoor air temperature (see Fig. 5). In the measured temperature trend, the band of the measurement uncertainty is also represented. A reference threshold for MBE and RMSE based on the ASHRAE Guideline [12] was defined by Pisello et al. [39] to validate a BPS model of a historic palace located in the city centre of Perugia, using sub-hourly indoor temperature as a control parameter. In such regard, in order to reduce the model uncertainty, iterative modifications of the thermal properties of the building envelope components were carried out, looking for the simulation that minimizes the error in terms of MBE and RMSE. The validation tolerance adopted were ±0.5 °C and 1 °C for MBE and RMSE respectively.

3.4.4. Pearson correlation coefficient

The *r* is used for the assessment of the model accuracy in combination with other indices. Pernetti et al. used MBE, RMSE and *r* to assess the error and the correlation between predicted and actual temperature [15], obtaining values of *r* between 0.979 and 0.992 for the calibrated model. Cornaro et al. calibrate a BPS model of a historical church in Palestrina (Italy) built before 1354 using hourly indoor T_a in different periods of the monitoring campaign. Taylor's diagram was adopted to evaluate the model accuracy, using centred RMSE in combination with the standard deviation and *r* (with $r > 0.97$). The statistical parameters evaluated the goodness of the calibration, which are iteratively applied to the model according to the experience of the operator. In such respect, the strength of using different statistical indices together has been already stressed by some authors [79,80], due to their ability to highlight different aspects of the results.

3.4.5. Inequality Coefficient

The IC indicates the degree of agreement between measured and simulated data. As stated in Section 3.2, it can range between 0 and 1, where 0 represents a strong correlation between measured and simulated data.

In the work by Cardinale et al. [64], the accuracy of a BPS model of a vernacular architecture located in Southern Italy was verified through the calculation of the IC based on temperature: the results lie between 0.13 and 0.15, thus the model is considered as reliable. Similarly, Thravalou et al. [73] considered a traditional building in the lowlands of Cyprus, reporting an IC between 0.15 and 0.19 for two different rooms.

In the study presented by Rospi et al. [37], the IC value was close to 0.25; in any case, its computation for different rooms, seasons and S/V

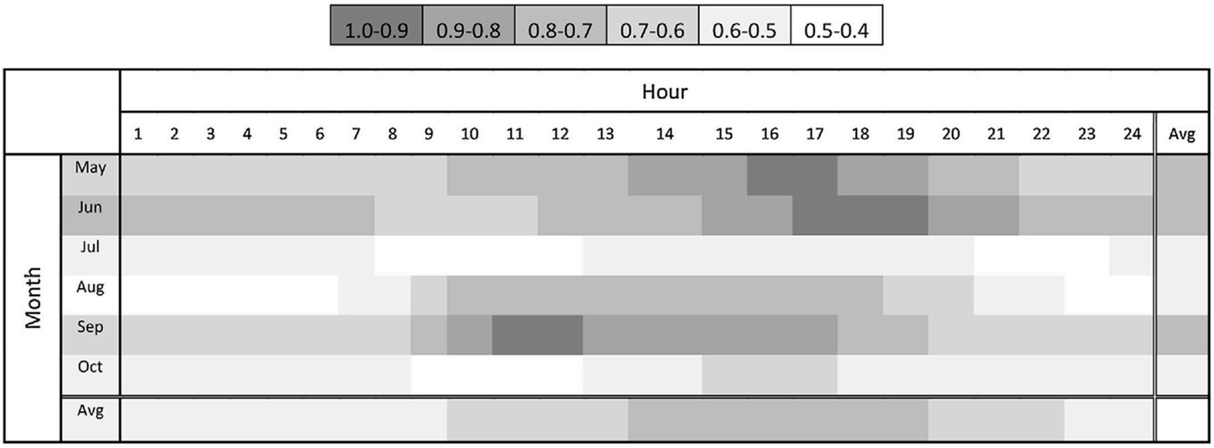

Fig. 4. Carpet plot of the RMSE depending on the hour of the day and month [56].

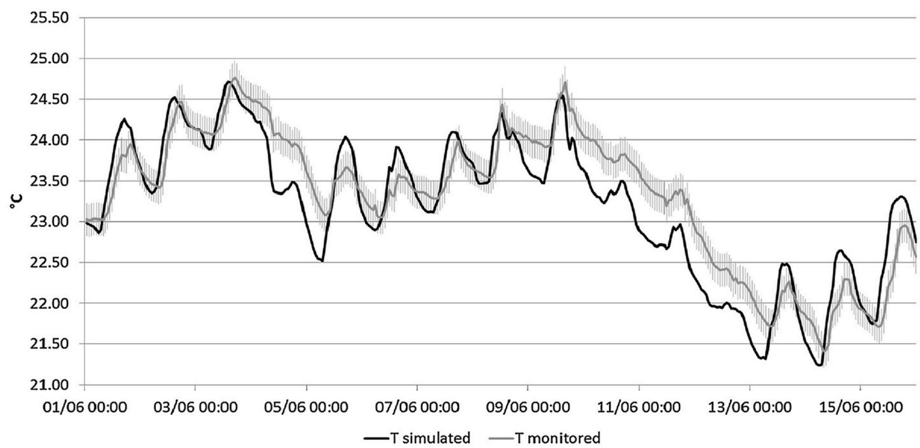

Fig. 5. Simulated and monitored air temperatures averaged over the whole building. The shaded area represents the measurement uncertainty of ±0.2 °C [56].

ratios (i.e. the ratio of the envelope surfaces to the total volume) allowed for estimating the reliability of the numerical model.

3.4.6. Other approaches

Regarding the frequency of residuals (F_i) described in Section 3.2, Kramer et al. validated the model of a museum built in the XVII century and located in Amsterdam (Netherlands) [24] by showing that for approximately 90% of the time the residuals for T_a and RH are between −0.5 °C and 2 °C, and ±4% respectively. Huijbregts et al. validated the simulation model of a small church near Eindhoven built in the XX century [32]: here, the authors have shown that the simulated T, RH and MR generally differed by no more than ±2 °C, ±10% and ±2 g/kg from the measurements, respectively.

In another case study, Muñoz-González et al. [63] validated the virtual model of a church built in the XVII century in Spain by just adopting the frequency analysis, with very good results [35]. Indeed,

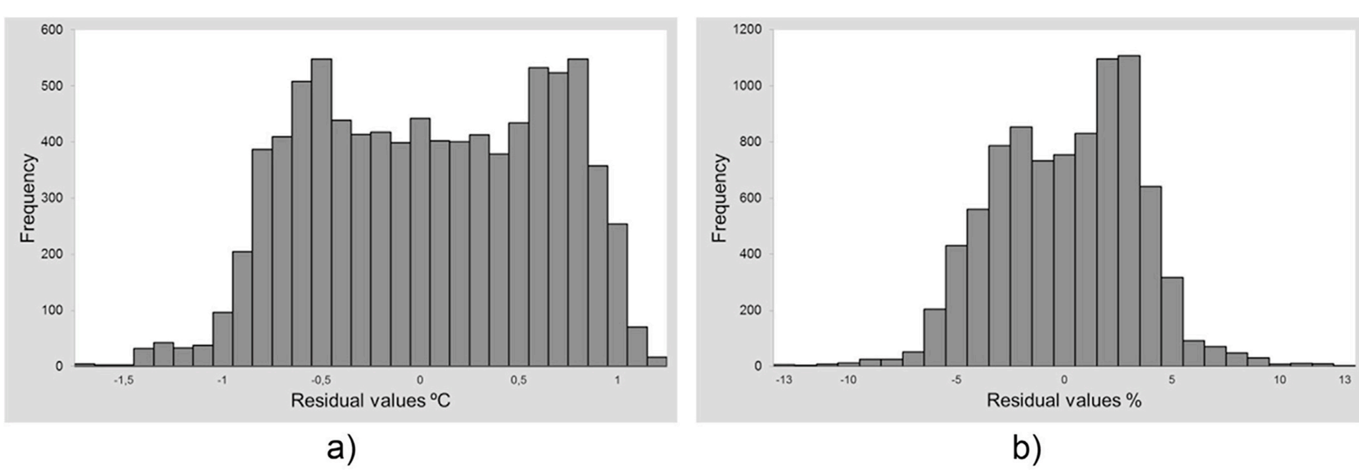

Fig. 6. Analysis of the frequency residuals distribution for T_a *(a) and RH (b)* [35].

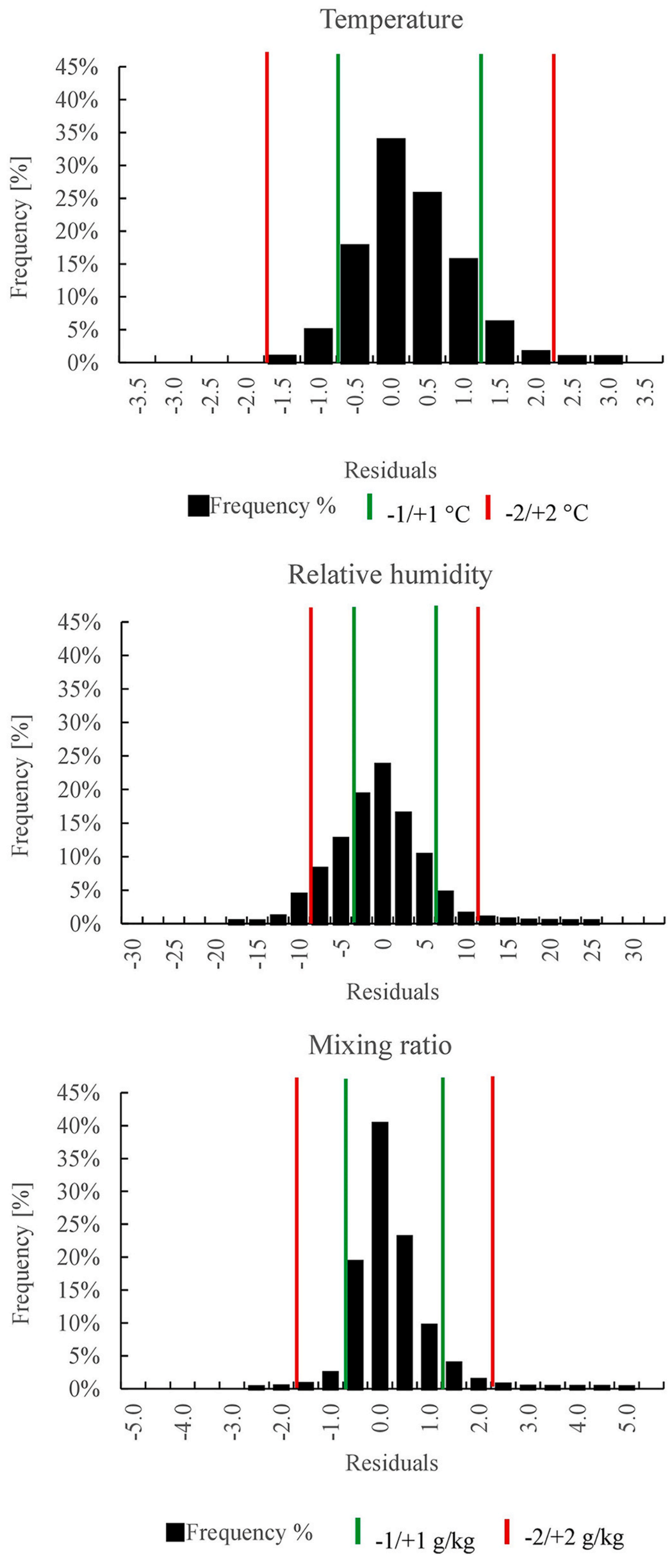

Fig. 7. Residuals frequency distribution for temperature, relative humidity and mixing ratio with two accuracy thresholds (the green and red lines represent the narrowest and widest range respectively). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 3
Comparison between temperature, relative humidity and mixing ratio residual frequency with respect to the thresholds adopted for the model validation.

Validation thresholds for T, RH and MR	Residuals frequency		
	Temperature	Relative humidity	Mixing ratio
±1 °C, ±5%, ±1 g/kg	90%	70%	91%
±2 °C, ±10%, ±2 g/kg	100%	94%	99%

Table 4
Assessment of the prediction rate Q_v for RH cycles.

Daily RH fluctuation	Q_v	N_s	N_m
$(RH_{max} - RH_{min}) \leq 5\%$	0.77	77	100
$(RH_{max} - RH_{min}) > 5\%$	1.11	288	259
$(RH_{max} - RH_{min}) > 10\%$	0.95	91	96
$(RH_{max} - RH_{min}) > 15\%$	1.15	38	33
$(RH_{max} - RH_{min}) > 20\%$	0.89	16	18

95% of the hourly residuals were between ±1 °C for T_a, while more than 90% of RH residuals was below ±5%, as shown in Fig. 6.

Fig. 7 shows the frequency of residuals for temperature, relative humidity and mixing ratio in the validation of the simulation model for the Milan Cathedral [81]. In these charts, two different ranges of accuracy are considered: between ±1 °C and ±2 °C for T_a, ±5% and ±10 for RH and ± 1 g/kg and ±2 g/kg for MR. Table 3 reports the overall residuals frequency with respect to the thresholds used for the model validation.

In detail, about 90% of the residuals are within the range of ±1 °C for temperature, with most of the residuals distributed around 0 °C as shown in Fig. 7, while the remaining 10% is within the range of ±2 °C. A similar distribution can be observed for mixing ratio, which has 91% of residuals within ±1 g/kg and 99% within ±2 g/kg. Relative humidity residuals have a more flattened distribution compared to T and MR, where about 70% of residuals are within the first range of ±5% and 94% within ±10%. In such regard, the higher amplitude of residuals is due to the dependency of relative humidity from the simulated temperature, as discussed in Section 3.1. However, these results confirm the good agreement between simulated and measured data for the three microclimatic parameters analysed.

In conclusion, according to most of the literature (see the Appendix), a model is considered validated when data very high share of residuals are within intervals ranging from ±1 °C to ±2 °C for temperature, from ±5% to ±10% for relative humidity and from ±1 g/kg to ±2 g/kg for mixing ratio and specific humidity. There is a good agreement that this should occur for at least 90%–95% of the overall dataset.

Finally, the prediction rate Q_v [27,74] is adopted as further verification of the model [81] reliability in the context of preventive conservation of cultural heritage. The results fall in the "acceptable" range, since Q_v is between 0.77 and 1.15 for all the intervals of fluctuation here considered (and reported in Table 4). It should be observed that the RH daily fluctuations higher than 10% are those more relevant for the development of mechanical damage.

4. Recommendations

The present paper gathered the main scientific research regarding the validation of hygrothermal simulation models of historical buildings based on microclimate parameters, with a focus on the selection of the most suitable variables, the uncertainty indices and methods, and their thresholds.

In order to provide a clear overview of the connection among these factors, and to outline a roadmap for a proper validation procedure, a flowchart has been arranged in Fig. 8. Here, the orange colour represents the parameters and indices that should be adopted in the validation

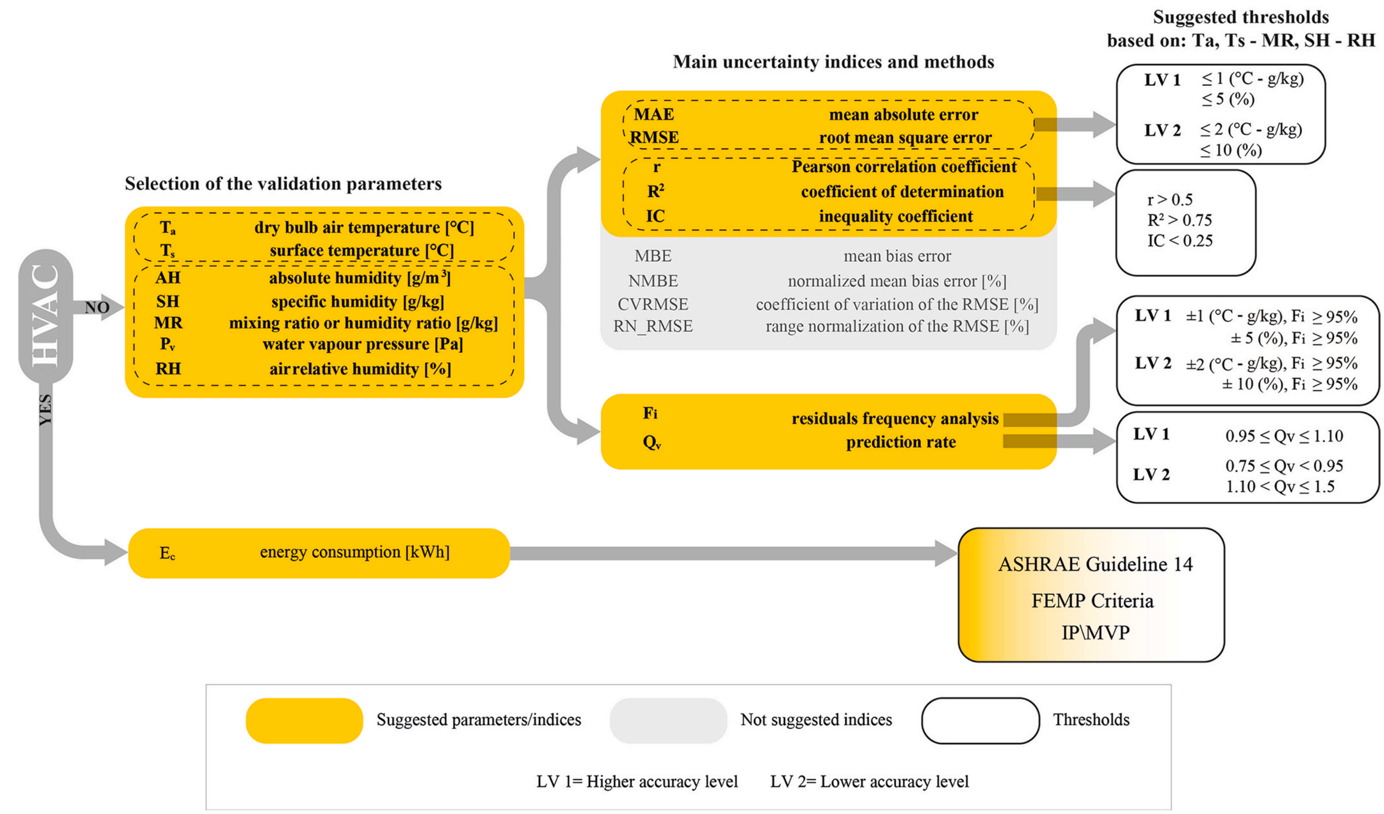

Fig. 8. Proposed validation process with microclimatic parameters. The orange and grey colours indicate respectively the parameters and indices suggested and not suggested. While the white colour indicates the thresholds for each index and method, provided by two different accuracy levels. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

approach, while the grey ones should always be used along with the recommended indices and not alone in order to avoid misunderstandings in the interpretation of the results.

Finally, the white colour shows the thresholds for each index and method according to two different accuracy levels LV 1 (high accuracy) and LV 2 (low accuracy). Of course, the required accuracy level depends on the simulation objectives and goals.

In detail, the advantages and drawbacks of using the different indices, resulting from the literature review, are summarized hereafter:

- To properly validate a hygrothermal building simulation model, T_a should be used in combination with a humidity parameter (e.g. MR, SH, etc.), particularly when dealing with historical buildings;
- Using RH only as a control parameter for validation may lead to misinterpretation of the results; in such respect, it is recommended always verifying the matching of both RH and T_a parameters, since a small deviation in the latter one might highly affect the relative humidity output of the simulation model;
- The MBE and NMBE are affected by the error cancellation effect; then, it is advisable that they are used in combination with other more reliable metrics. However, MBE and NMBE can still be useful to verify if simulated data generally over or underestimate the measured ones;
- Since normalized indices such as NMBE, CVRMSE and RN_RMSE use the average and the range of the measured data in their calculation, the results can vary consistently depending on the scale of the dataset, thus leading to possible misinterpretation. Therefore, users should be very cautious when adopting them. In such regard, defining a single general threshold to be used in the validation of simulation models through the normalized indices (with particular reference to microclimatic parameters) can be challenging;
- Statistical indices such as the MAE and the RMSE can avoid misinterpretation due to the normalization process. In particular, the RMSE is very widespread in the literature and due to its higher sensitivity to the amplitude of the residuals, it can be more cautionary and suitable concerning historical buildings. Appropriate thresholds for such indices are proposed in Fig. 8;
- The approach based on the frequency analysis of the residuals can carefully evaluate the model accuracy. As reported in Fig. 8, the literature suggests two different ranges of accuracy: a narrower one between ±1 °C, ±5% and ±1 g/kg (Lv. 1) and a wider one between ±2 °C, ±10% and ± 2 g/kg (Lv. 2) for temperature, relative humidity and mixing ratio (or specific humidity) respectively. The simulations are considered reliable when residuals are within the defined range for more than 95% of the overall data. Of course, the required accuracy level depends also on the simulation objective;
- In the field of cultural heritage conservation, taking into account also the RH daily cycles is a pivotal step, since this variable has a strong correlation with the decay of building materials. Given this, further verification of the simulation results through the prediction rate Q_v should be carried out.

5. Conclusions

The proper simulation of the hygrothermal behaviour of historical buildings is a challenging task, which may have influence the evaluation of indoor thermal comfort and the suitability of retrofit strategies complying with the conservation of cultural heritage. Indeed, an inaccurate simulation may cause inadequate conclusions, which could lead to inappropriate and dangerous actions for the preservation of the heritage building. Then, testing the predictive performance of building hygrothermal simulation models is an important step that requires as much attention as the definition of reliable calibration and validation methodologies.

Model validation is commonly based on the use of uncertainty metrics to measure the discrepancy between measured and simulated performance. If the model validation relies on the comparison of energy data (e.g. seasonal energy needs for space heating and cooling), suitable protocols and metrics have already been identified almost two decades ago and are nowadays implemented in several standards.

Now, since in some buildings the energy consumption data are not available or important, the microclimatic parameters need to be adopted in the calibration and validation process. However, in this case, neither protocols nor specific parameters have been officially recognised to perform the model validation.

Under these premises, the present work reviewed the main approaches used by researchers for BPS model validation with special reference to historical buildings through microclimatic parameters, highlighting the main issues and advantages of the different methods reviewed and defining suitable validation thresholds.

In conclusion, the information collected in this work provides guidance to different subjects (e.g. designers, energy auditors, researchers and conservators) involved in building simulations and can drive suitable maintenance interventions to prevent possible damage problems. Moreover, the improvement of the validation results allows decreasing the uncertainty about the outcomes and can offer a significant contribution to the development of new rules and strategies.

Declaration of competing interest

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

Acknowledgements

The research presented in this paper has been developed in the framework of Task 59 of the International Energy Agency – Solar Heating and Cooling Programme (IEA-SHC) "Renovating Historic Buildings Towards Zero Energy" and more in particular in the subtask B.5 "Characterization and simulation of historical buildings". This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix

Table 5
BPS models validation approaches adopted in the literature

Ref.	Index used	Building information	Validation period	Control parameter	Validation results	Validation threshold adopted	Simulation purposes
[51]	Percent difference	Palace (XIX century), Benevento (Italy)	Year	E_c	3%–3.9%	–	Energy retrofit assessment
[52]	Average difference, absolute average deviation	Town Hall (1910), Norrköping (Sweden)	Year	E_c, T_a	Average difference = 0.7% (E_c) Average difference = 0.2°C–0.5 °C (T_a)	–	Energy retrofit assessment
[53]	NMBE, CVRMSE	House (XVI -XIX centuries), France	Months	E_c	NMBE = 0.6%–2.8% CVRMSE = 0.8%–4.2%	ASHRAE guideline 14 [12], IPMVP [14]	Energy retrofit assessment
[64]	IC	House, Matera (Italy)	Year	T_a	0.13	–	Comfort assessment
[82]	NMBE, CVRMSE	Storey office building (1901), Illinois (USA)	Months	E_c	NMBE = −2.31% CVRMSE = 2.8%	ASHRAE guideline 14 [12]	Sensitivity analysis
[15, 55]	MBE, RMSE, r	Palace (1854), Trento (Italy)	From March to June	T_a		MBE* = −1.4°C-0.8 °C RMSE* = 1°C–1.7 °C r = 0.979–0.992	–
	Sensitivity analysis						
[23]	Percent difference, r	Museum (1929-1930), Bytom (Poland)	From mid-June to the end of August	T_a	Percent difference < 10% for the 99% of the time, < 5% for the 85% of the time, r = 0.93–0.98	–	Energy retrofit assessment, risk analysis for materials
[54]	Average difference, Percent difference	Houses (1700-1940), Estonia, Finland and Sweden	Various	E_c, T_a	Average difference = 3% (E_c) Average difference = 1.2 °C (T_a)	–	Energy retrofit assessment
[69]	Percent difference	3 churches (XIV -XVI centuries), Milan (Italy)	During the heating period	E_c	10%–24%	–	Simulation software comparison
[62, 78]	RMSE, CVRMSE, R^2, residuals	School (1950), Vicence (Italy)	4 periods of 2 weeks each one with different conditions	T_a		RMSE = 0.17°C–0.67 °C CVRMSE = 1.24%–3.53% R^2 = 0.92–0.99 residuals = between −2.1°C and 1.5 °C for the 100% of the time and −0.3°C-	–

(continued on next page)

Table 5 (*continued*)

Ref.	Index used	Building information	Validation period	Control parameter	Validation results	Validation threshold adopted	Simulation purposes
	Sensitivity analysis					0.4 °C for the 50% of the time	
[42]	ERR**, CVRMSE	Palace (1927), Benevento (Italy)	Year	E_c	ERR = 6.5% CVRMSE = 9%	M&V guideline [13]	Energy retrofit assessment
[24]	Relative error (E), residuals (T and RH)	Museum (XVII century), Amsterdam (Netherlands)	Year	E_c, T_a, RH	Relative error = −9%-2% (E_c) residuals: T_a between −0.5°C and 2°C for the 98% of the time RH between −4% and 4% for the 90% of the time	–	Energy retrofit assessment, risk analysis for materials
[56]	MAE, RMSE	Palace (XIII century), Bolzano (Italy)	From May to October (model calibration) and from 8th to 15th of January (model validation)	T_a, T_s	MAE = 0.38–0.67 K RMSE = 0.48–0.8 K	–	Sensitivity analysis
[31]	residuals	Castle (XVII century), Amerongen (Netherlands)	Year	T_a, RH, SH	Room: T_a ±3 °C RH ±10% SH ±2 g/kg Cabinet: T_a ±1.5 °C RH ±3% SH ±1 g/kg	–	Risk analysis for materials
[57]	NMBE, CVRMSE	Palace (XIX century), Izmir (Turkey)	96 h period (with temperature) and one year (with energy consumption)	E_c, T_a	T_a: NMBE = 0.5%–5.8% CVRMSE = 2.2%–6.7%	ASHRAE guideline 14 [12]	Energy retrofit assessment
[39]	MBE, RMSE	Palace (XVI century), Perugia (Italy)	Summer period	T_a	MBE = −0.16°C-0.28 °C RMSE = 0.66°C-0.93 °C	ASHRAE guideline 14 [12]; −1 °C ≤ MBE≤+1 °C; RMSE≤1 °C	Energy retrofit assessment, sensitivity analysis
[40]	MBE, RMSE	Residential building (XVI century), Perugia (Italy)	Two months	T_a	MBE = 0.02–0.17 °C RMSE = 0.79–0.82 °C	ASHRAE guideline 14 [12]	Energy retrofit assessment
[43]	Average difference	Houses (1650-1938), Baltic region	Various	T_a	Average difference = 0.1°C–0.6 °C	–	Energy retrofit assessment
[32]	residuals	Church, (XIX century), Netherlands	Years	T_a, RH, MR	T_a between ±2 °C RH between ±10% MR between ±2 g/kg	–	Risk analysis for materials
[33]	NMBE, CVRMSE	Fortress (1540 -1543), Perugia (Italy)	About 7-8 months	T_a	Vibio room: NMBE = 0.1%, CVRMSE = 0.2% Passage area: NMBE = 0.8%, CVRMSE = 1%	ASHRAE guideline 14 [12]	Energy retrofit assessment, risk analysis for materials, comfort assessment
[34]	residuals	Castle (XVII century), Amerongen (Netherlands)	Year	T_a, RH, MR	T_a < 2 °C RH ±10% MR ±1 g/kg	–	Energy retrofit assessment, risk analysis for materials
[35]	NMBE, CVRMSE, ε***, residuals	Church (XVI century), Seville (Spain)	Year	T_a, RH, AH	T_a: NMBE = −0.18% CVRMSE = 17.3% ε = −332 residuals ±1 °C for the 95% of the time RH: NMBE = −0.14% CVRMSE = 13.5% ε = −786 residuals ±5% for the 90% of the time AH: residuals ±1.5 g/m^3 for the 95% of the time	ASHRAE guideline 14 [12]	Energy retrofit assessment, risk analysis for materials, comfort assessment
[45]	RN_RMSE	Palace (XVI century), Rome (Italy)	1 week and 1 month	T_a	RN_RMSE = 0.18%–0.26%	–	Energy retrofit assessment
[58]	NMBE, CVRMSE, RMSE	Palace (XIII century), Bologna (Italy)	2 different periods	T_a, T_s	NMBE = 1.4%–7% CVRMSE = 1.7%–17% RMSE = 0.37°C–4.3 °C	ASHRAE guideline 14 [12]	Sensitivity analysis
[83]	RMSE, σ, r	Church (1358), Palestrina (Italy)	From 12th June to 30th November	T_a, RH	T_a: RMSE = 0.8 °C r = 0.97	–	Risk analysis for materials
[84]	Percent difference		Year	E_c	3%	–	Energy retrofit assessment

(continued on next page)

Table 5 (*continued*)

Ref.	Index used	Building information	Validation period	Control parameter	Validation results	Validation threshold adopted	Simulation purposes
		Palace (early XX century), Rome (Italy)					
[44]	Percent difference	Palace (XVIII century), Perugia (Italy)	July, January, February and December	E_c	0,1%–9%	–	Energy retrofit assessment
[59]	Percent difference, NMBE, CVRMSE	School (1970), Galicia (Spain)	2 months	E_c, T_a	E_c: Percent difference = 10.3% T_a: NMBE = 2.73% CVRMSE = 11.52%	ASHRAE guideline 14 [12]	Calibration assessment
[46]	NMBE, CVRMSE	Palace (XIX century), Benevento (Italy)	Year	E_c	NMBE = 3.2%–3.9% CVRMSE = 5.4%–13%	ASHRAE guideline 14 [12]	Energy retrofit assessment
[47]	ERR, CVRMSE	Palace (1513), Naples (Italy)	Year	E_c	ERR = −1.22%-0.03% CVRMSE = 5.85%	M&V guideline [13]	Energy retrofit assessment
[36]	MBE, RMSE, r	Palace (1565), Florence (Italy)	About 6 months	T_a, RH	–	Simulazione energetica degli edifici esistenti [71]	Risk analysis for materials
[25]	Mean deviation, maximum deviation	Museum (XV century), Amsterdam (Netherlands)	About 4 months	T_a, RH, AH	T_a: Mean dev. = 0.02 °C max dev. = 2.7 °C. RH: Mean dev. = 1.3% max dev. = 10%	No	Energy retrofit assessment, risk analysis for materials, comfort assessment
[26]	NMBE, CVRMSE	Library (1827), İzmir (Turkey)	Year	T_a, RH	T_a: NMBE = −0.71% CVRMSE = 7.16% RH: NMBE = 4.63% CVRMSE = 15.24%	ASHRAE guideline 14 [12]	Risk analysis for materials
[38]	residuals	Palace (XIV century), Florence (Italy)	Year	T_a, RH	T_a ±3 °C	–	Risk analysis for materials
[37]	IC	3 palaces (XVI-XIX centuries), Matera (Italy)	Year	T_a	0.01–0.49	–	Energy retrofit assessment
[85]	NMBE, CVRMSE	Mosque (1906), Izmir (Turkey)	Year	T_a	NMBE = −9.26-5.48% CVRMSE = 3.52–14.10%	ASHRAE guideline 14 [12]	Comfort assessment
[86]	NMBE, CVRMSE	Palace (XVI century), Modena (Italy)	2 weeks in October, January and 1 week when the building is unoccupied	T_a	NMBE = 2% CVRMSE = 5%	ASHRAE guideline 14 [12]	Comfort assessment
[87]	residuals	Church (XIII century), Bologna (Italy)	2 days	T_a, RH	T_a < 0.5 °C RH< 5%	–	Microclimate analysis
[48]	Deviation	House, Agrigento (Italy)	Year	E_c	Deviation = 9%	ASHRAE guideline 14 [8]	Energy retrofit assessment, economic analysis
[63]	residuals	Church (XVII century), Seville (Spain)	Year	T_a, RH, AH	T_a < 1 °C for the 95% of the time RH< 5% for the 90% of the time AH< 1.5 g/kg for the 95% of the time	–	Energy retrofit assessment, risk analysis for materials, comfort assessment
[73]	IC	Houses, Cyprus	From July to November	T_a	IC = 0.15–0.19	–	Comfort assessment,
[61]	NMBE, CVRMSE, r^2, fit	Church (XIII century), Lisbon (Spain)	Year	T_a, P_v	T_a: NMBE = 2.7% CVRMSE = 3.2% r^2 = 0.99 fit = 84.8%. P_v: NMBE = 3.4% CVRMSE = 4.4% r^2 = 0.97 fit = 81.7% T_a, P_v: fit = 83.2%	IPMVP [14]	Sensitivity analysis
[29]	MBE, RMSE, CVRMSE	Castle (XIV century), Perugia (Italy)	Year	T_a	MBE = −0.13°C-0.13 °C RMSE = 0.72°C-1.45 CVRMSE = 9.28%–22.4%	ASHRAE guideline 14 [12]	Risk analysis for materials, comfort assessment,
[50]	MAE, r_s****	Market (XIII century), Krakow (Poland)	Year	T_a, RH	T_a: MAE = 0.7 K r_s = 0.6–0.8 RH: MAE = 2.8% r_s = 0.6–0.8	–	Energy retrofit assessment, microclimate analysis
[27]	r, r_s****, residuals, Q_v	Church (1642), Velika Mlaka (Croatia)	Year	T_a, RH	T_a: r = 0.909 r_s = 0.855 residuals = −2.2°C-1.6 °C RH: r = 0.877	Excellent: T ±1 °C RH ±5% 0.95 ≤ Q_v ≤ 1.10 Acceptable:	Risk on materials

(continued on next page)

Table 5 (*continued*)

Ref.	Index used	Building information	Validation period	Control parameter	Validation results	Validation threshold adopted	Simulation purposes
					r_s = 0.886 residuals = −8.5%-6.2%	T ±3 °C RH ±10% $0.75 \leq Q_v < 0.95$ and $1.10 < Q_v \leq 1.5$	
[88]	Percent difference	School (1876), Florence (Italy)	Year	E_c	0–9%	–	Energy retrofit assessment, microclimate analysis
[11]	RMSE, NMBE, CVRMSE, RN_RMSE, BE, r^2	Office building (1960), Rome (Italy)	21st-26th of June, 28th July-1st August, 13th-18th December	T_a, E_c	T_a: RMSE = 0.27°C–0.624 °C CVRMSE = 1.016%–2.619% RN_RMSE = 9.058%–21.602% E_c: NMBE = −7.69%-2.11% CVRMSE = 8.37%–9.66% BE = 2.11% r^2 = 0.5433–0.8617	ASHRAE guideline 14 [12], IPMVP [14], M&V guideline [13]	Calibration assessment
[16]	NMBE, CVRMSE, residuals	Library, The Hague (Netherlands)	–	T_a, RH, SH	NMBE = −7.97%-6.91% CVRMSE = 0.01%–7.67%	ASHRAE guideline 14 [12]	Risk on materials
[89]	Percent difference	Houses, Algiers (Algeria)	From August to January	T_a, RH	–	–	Comfort assessment
[17]	MAE, RMSE, CVRMSE, r_s****	Museum (XIIIth-XVI centuries), Priverno (Italy)	From 24th September to 8th October	T_a, T_s,RH	T_a: MAE = 0.2–0.4 °C RMSE = 0.3–0.4 °C CVRMSE = 1.2–1.7% r_s = 1 RH: MAE = 0.8–1.6% RMSE = 1-2% CVRMSE = 2–3.9% r_s = 0.7–0.9	CVRMSE for T<2% and RH<5%	Risk on materials, comfort and energy assessment,
[72]	NMBE, CVRMSE, r^2	Church (XII century), L'Aquila (Italy)	From 28th September to 13th October (calibration), from 31st October to 23rd November (validation)	T_a	NMBE = 1.49–9.76% CVRMSE = 17.71–18.86% r^2 = 0.78–0.83	IPMVP [14]	Risk on materials, comfort and energy assessment,

* The simulations with the "test reference year" are not considered because not relevant.
** Calculated as the percent difference.
***ε: error values.
**** r_s: Spearman's rank correlation coefficient.

References

[1] L. Mazzarella, Energy retrofit of historic and existing buildings. the legislative and regulatory point of view, Energy Build. 95 (2015) 23–31, https://doi.org/10.1016/j.enbuild.2014.10.073.
[2] Sdsn, Feem, Roadmap to 2050: A Manual for Nations to Decarbonize by Mid-Century, 2019.
[3] M. Filippi, Remarks on the green retrofitting of historic buildings in Italy, Energy Build. 95 (2015) 15–22, https://doi.org/10.1016/j.enbuild.2014.11.001.
[4] A. Magrini, Thermal behaviour of historical buildings, materials and components: methodological framework, calculation, results, in: Hist. Build. Energy, Springer International Publishing, Cham, 2017, pp. 55–81, https://doi.org/10.1007/978-3-319-52615-7_4.
[5] A.L. Webb, Energy retrofits in historic and traditional buildings: a review of problems and methods, Renew. Sustain. Energy Rev. 77 (2017) 748–759, https://doi.org/10.1016/j.rser.2017.01.145.
[6] L. de Santoli, Guidelines on energy efficiency of cultural heritage, Energy Build. 86 (2015) 534–540, https://doi.org/10.1016/J.ENBUILD.2014.10.050.
[7] G.G. Akkurt, N. Aste, J. Borderon, A. Buda, M. Calzolari, D. Chung, V. Costanzo, C. Del Pero, G. Evola, H.E. Huerto-Cardenas, F. Leonforte, A. Lo Faro, E. Lucchi, L. Marletta, F. Nocera, V. Pracchi, C. Turhan, Dynamic thermal and hygrometric simulation of historical buildings: critical factors and possible solutions, Renew. Sustain. Energy Rev. 118 (2020) 109509, https://doi.org/10.1016/j.rser.2019.109509.
[8] ANSI/ASHRAE, Guideline 14-2014 – Measurement of Energy, Demand, and Water Savings, 2014.
[9] P. De Wilde, Building Performance Analysis, John Wiley & Sons Ltd, 2018.
[10] W.L. Oberkampf, C.J. Roy, Verification and Validation in Scientific Computing, Cambridge University Press, 2010, https://doi.org/10.1017/CBO9780511760396.
[11] C. Cornaro, F. Bosco, M. Lauria, V.A. Puggioni, L. De Santoli, Effectiveness of automatic and manual calibration of an office building energy model, Appl. Sci. 9 (2019) 1985, https://doi.org/10.3390/app9101985.
[12] ANSI/ASHRAE, ASHRAE Guideline 14-2002 Measurement of Energy and Demand Savings, 2002.
[13] FEMP, M&V Guidelines : Measurement and Verifi Cation for Federal Energy Projects- Version 3.0, Washington, DC, USA, 2008.
[14] Efficiency Valuation Organization, International Performance Measurement and Verification Protocol: Concepts and Options for Determining Energy and Water Savings, vol. 1, 2012, https://doi.org/10.1126/science.15.390.961. Washington, DC, USA.
[15] R. Pernetti, A. Prada, P. Baggio, On the Influence of Several Parameters in Energy Model Calibration : the Case of a Historical Building, 2013, pp. 263–272. IBPSA Italy.
[16] K. Kompatscher, R. Kramer, B. Ankersmit, H.L. Schellen, Intermittent conditioning of library archives: microclimate analysis and energy impact, Build, Environ. Times 147 (2019) 50–66.
[17] F. Frasca, C. Cornaro, A.M. Siani, A method based on environmental monitoring and building dynamic simulation to assess indoor climate control strategies in the preventive conservation within historical buildings, Sci. Technol. Built Environ. 25 (2019) 1253–1268, https://doi.org/10.1080/23744731.2019.1642093.
[18] T.A. Reddy, Literature review on calibration of building energy simulation programs: uses, problems, procedures, uncertainty, and tools, Build. Eng. 112 (2006) 226–240, doi:Article.
[19] D. Coakley, P. Raftery, M. Keane, A review of methods to match building energy simulation models to measured data, Renew. Sustain. Energy Rev. 37 (2014) 123–141, https://doi.org/10.1016/j.rser.2014.05.007.
[20] 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.
[21] CEN, EN 15758: 2010 – Conservation of Cultural Property. Procedures and Instruments for Measuring Temperatures of the Air and the Surfaces of Objects, 2010.
[22] CEN, EN 16242: 2012 – Conservation of Cultural Heritage – Procedures and Instruments for Measuring Humidity in the Air and Moisture Exchanges between Air and Cultural Property, 2012.
[23] J. Ferdyn-Grygierek, Indoor environment quality in the museum building and its effect on heating and cooling demand, Energy Build. 85 (2014) 32–44, https://doi.org/10.1016/j.enbuild.2014.09.014.
[24] R.P. Kramer, M.P.E. Maas, M.H.J. Martens, A.W.M. van Schijndel, H.L. Schellen, Energy conservation in museums using different setpoint stretegies: a case study for a state-of-the-art museum using building simulations, Appl. Energy 158 (2015) 446–458, https://doi.org/10.1016/j.apenergy.2015.08.044.

[25] K. Kompatscher, S. Seuren, R. Kramer, J. Van Schijndel, H. Schellen, Energy efficient HVAC control in historical buildings: a case study for the Amsterdam Museum, Energy Procedia 132 (2017) 891–896, https://doi.org/10.1016/j.egypro.2017.09.703.

[26] T. Coşkun, Ö. Gülhan, C.D. Şahin, Z.D. Arsan, G.G. Akkurt, The effect of spatial interventions on historic buildings' indoor climate (Case Study: tire Necip Paşa Library, Izmir-Turkey), Energy Procedia 133 (2017) 358–366, https://doi.org/10.1016/j.egypro.2017.09.362.

[27] V. Rajčić, A. Skender, D. Damjanović, An innovative methodology of assessing the climate change impact on cultural heritage, Int. J. Architect. Herit. 12 (2018) 21–35, https://doi.org/10.1080/15583058.2017.1354094.

[28] C. Muñoz-González, Á.L. León-Rodríguez, M.Á. Campano, C. Teeling, R. Baglioni, The assessment of environmental conditing techniques and their energy performance in historic churches located in Mediterranean climate, J. Cult. Herit. 34 (2018) 74–82.

[29] I. Pigliautile, V.L. Lucia, N. Makaremi, A.L. Laura, L.F. Cabeza, F. Cotana, On an innovative approach for microclimate enhancement and retrofit of historic buildings and artworks preservation by means of innovative thin envelope materials, J. Cult. Herit. 36 (2019) 222–231, https://doi.org/10.1016/j.culher.2018.04.017.

[30] N. Aste, R.S. Adhikari, M. Buzzetti, S. Della Torre, C. Del Pero, H.E. Huerto C, F. Leonforte, Microclimatic monitoring of the duomo (milan cathedral): risks-based analysis for the conservation of its cultural heritage, build, Environ. Times 148 (2019) 240–257, https://doi.org/10.1016/j.buildenv.2018.11.015.

[31] Z. Huijbregts, H. Schellen, J. van Schijndel, B. Ankersmit, Modelling of heat and moisture induced strain to assess the impact of present and historical indoor climate conditions on mechanical degradation of a wooden cabinet, J. Cult. Herit. 16 (2015) 419–427, https://doi.org/10.1016/j.culher.2014.11.001.

[32] Z. Huijbregts, H. Schellen, M. Martens, S.J. Van, Object damage risk evaluation in the European project Cimate for Culture, Energy Procedia 78 (2015) 1341–1346, https://doi.org/10.1016/j.egypro.2015.11.151, in: https://search.crossref.org/?q=Z.+Huijbregts%2C+H.+Schellen%2C+M.+Martens%2C+S.J.+Van%2C+Object+damage+risk+evaluation+in+the+European+project+Cimate+for+Culture%2C+Energy+Procedia.

[33] A.L. Pisello, V.L. Castaldo, G. Pignatta, F. Cotana, Integrated numerical and experimental methodology for thermal-energy analysis and optimization of heritage museum buildings, Build. Serv. Eng. Technol. 37 (2015) 334–354, https://doi.org/10.1177/0143624415609910.

[34] J. Leissner, R. Kilian, L. Kotova, D. Jacob, U. Mikolajewicz, T. Broström, J. Ashley-Smith, H.L. Schellen, M. Martens, J. Van Schijndel, F. Antretter, M. Winkler, C. Bertolin, D. Camuffo, G. Simeunovic, T. Vyhlídal, Climate for culture: assessing the impact of climate change on the future indoor climate in historic buildings using simulations, Herit. Sci. 3 (2015) 1–15, https://doi.org/10.1186/s40494-015-0067-9.

[35] C.M. Muñoz-González, A.L. León-Rodríguez, J. Navarro-Casas, Air conditioning and passive environmental techniques in historic churches in Mediterranean climate. A proposed method to assess damage risk and thermal comfort pre-intervention, simulation-based, Energy Build. 130 (2016) 567–577, https://doi.org/10.1016/j.enbuild.2016.08.078.

[36] F. Sciurpi, C. Carletti, G. Cellai, V. Muratore, A. Orsi, L. Pierangioli, G. Russo, E. D. Schmidt, Environmental monitoring and building simulation application to Vasari Corridor: preliminary results, Energy Procedia 133 (2017) 219–230, https://doi.org/10.1016/j.egypro.2017.09.393.

[37] G. Rospi, N. Cardinale, E. Negro, Energy performance and economic feasibility study of historical building in the city of Matera, Southern Italy, Energies. 10 (2017), https://doi.org/10.3390/en10122009.

[38] K. Fabbri, L. Signorelli, M. Pretelli, C. Magnani, Villa La petraia (florence) UNESCO world heritage, in: Hist. Indoor Microclim. Herit. Build., 2017, pp. 185–222, https://doi.org/10.1007/978-3-319-60343-8.

[39] A.L. Pisello, C. Piselli, F. Cotana, Influence of human behavior on cool roof effect for summer cooling, Build. Environ. 88 (2015) 116–128, https://doi.org/10.1016/j.buildenv.2014.09.025. https://search.crossref.org/?q=A.L.+Pisello%2C+C.+Piselli%2C+F.+Cotana%2C+Influence+of+human+behavior+on+cool+roof+effect+for+summer+cooling%2C+Build.+Environ.

[40] A.L. Pisello, C. Piselli, F. Cotana, Thermal-physics and energy performance of an innovative green roof system: the Cool-Green Roof, Sol. Energy 116 (2015) 337–356, https://doi.org/10.1016/j.solener.2015.03.049. https://search.crossref.org/?q=A.L.+Pisello%2C+C.+Piselli%2C+F.+Cotana%2C+Thermal-physics+and+energy+performance+of+an+innovative+green+roof+system%3A+the+Cool-Green+Roof%2C+Sol.+Energy.

[41] Ş. Doğan, A transdisciplinary approach on the energy efficient retrofitting of a historic building in the Aegean, Region of Turkey 96 (2015) 128–139, https://doi.org/10.1016/j.enbuild.2015.03.018.

[42] F. Ascione, N. Bianco, R.F. De Masi, F. De'Rossi, G.P. Vanoli, Energy retrofit of an educational building in the ancient center of Benevento. Feasibility study of energy savings and respect of the historical value, Energy Build. 95 (2015) 172–183, https://doi.org/10.1016/j.enbuild.2014.10.072.

[43] L. Eskola, Û. Alev, E. Arumägi, J. Jokisalo, A. Donarelli, T. Kalamees, L. Eskola, Û. Alev, E. Arumägi, J. Jokisalo, A. Donarelli, Airtightness , Air Exchange and Energy Performance in Historic Residential Buildings with Different Structures, 2016, p. 3315, https://doi.org/10.1080/14733315.2015.11684066.

[44] A.L. Pisello, A. Petrozzi, V.L. Castaldo, F. Cotana, On an innovative integrated technique for energy refurbishment of historical buildings: thermal-energy, economic and environmental analysis of a case study, Appl. Energy 162 (2014) 1313–1322, https://doi.org/10.1016/j.apenergy.2015.05.061.

[45] C. Cornaro, V.A. Puggioni, R.M. Strollo, Dynamic simulation and on-site measurements for energy retrofit of complex historic buildings: villa Mondragone case study, J. Build. Eng. 6 (2016) 17–28, https://doi.org/10.1016/j.jobe.2016.02.001.

[46] F. Ascione, F. Ceroni, R.F. De Masi, F. de' Rossi, M.R. Pecce, Historical buildings: multidisciplinary approach to structural/energy diagnosis and performance assessment, Appl. Energy 185 (2017) 1517–1528, https://doi.org/10.1016/j.apenergy.2015.11.089.

[47] F. Ascione, R.F. De Masi, F. De Rossi, S. Ruggiero, G.P. Vanoli, NZEB target for existing buildings: case study of historical educational building in Mediterranean climate, Energy Procedia 140 (2017) 194–206, https://doi.org/10.1016/j.egypro.2017.11.135.

[48] M. Cellura, G. Ciulla, F. Guarino, S. Longo, Redesign of a rural building in a heritage site in Italy: towards the net Zero energy target, Buildings 7 (2017) 68, https://doi.org/10.3390/buildings7030068.

[49] V. D'Agostino, F.R. D'Ambrosio Alfano, B.I. Palella, G. Riccio, The museum environment: a protocol for evaluation of microclimatic conditions, Energy Build. 95 (2015) 124–129, https://doi.org/10.1016/j.enbuild.2014.11.009.

[50] A. Sadłowska-Sałęga, J. Radoń, J. Sobczyk, K. Wąs, Influence of microclimate control scenarios on energy consumption in the gallery of the 19th-century polish art in the sukiennice (the former cloth Hall) of the national museum in krakow, IOP Conf. Ser. Mater. Sci. Eng. 415 (2018), https://doi.org/10.1088/1757-899X/415/1/012026.

[51] F. Ascione, F. De Rossi, G.P. Vanoli, Energy retrofit of historical buildings: theoretical and experimental investigations for the modelling of reliable performance scenarios, Energy Build. 43 (2011) 1925–1936, https://doi.org/10.1016/j.enbuild.2011.03.040.

[52] P. Rohdin, M. Dalewski, B. Moshfegh, Indoor environment and energy use in historic buildings - comparing survey results with measurements and simulations, Int. J. Vent. 10 (2012) 371–382, https://doi.org/10.1080/14733315.2012.11683962.

[53] A. Caucheteux, E. Stephan, C. Ouest, Transient simulation calibration of an old building using an experimental design: evaluating uncertainty results, in: 13th Conf. Int. Build. Perform. Simul. Assoc., 2013, pp. 677–684.

[54] A. Üllar, E. Lari, A. Endrik, J. Juha, D. Anna, S. Kai, B. Tor, K. Targo, Renovation alternatives to improve energy performance of historic rural houses in Baltic Sea Region, Energy Build. 77 (2014) 58–66.

[55] R. Pernetti, Problematiche relative alla definizione e all ' utilizzo di modelli delle prestazioni energetiche di edifici esistenti, Università di Pavia, 2012.

[56] F. Roberti, U.F. Oberegger, A. Gasparella, Calibrating historic building energy models to hourly indoor air and surface temperatures: methodology and case study, Energy Build. 108 (2015) 236–243, https://doi.org/10.1016/j.enbuild.2015.09.010.

[57] C.D. Şahin, Z.D. Arsan, S.S. Tunçoku, T. Broström, G.G. Akkurt, A transdisciplinary approach on the energy efficient retrofitting of a historic building in the Aegean Region of Turkey, Energy Build. 96 (2015) 128–139, https://doi.org/10.1016/j.enbuild.2015.03.018.

[58] M. Giuliani, G.P. Henze, A.R. Florita, Modelling and calibration of a high-mass historic building for reducing the prebound effect in energy assessment, Energy Build. 116 (2016) 434–448, https://doi.org/10.1016/j.enbuild.2016.01.034.

[59] A. Ogando, N. Cid, M. Fernández, Energy modelling and automated calibrations of ancient building simulations: a case study of a school in the northwest of Spain, Energies 10 (2017), https://doi.org/10.3390/en10060807.

[60] L. Tronchin, K. Fabbri, Energy and microclimate simulation in a heritage building: further studies on the Malatestiana Library, Energies 10 (2017), https://doi.org/10.3390/en10101621.

[61] G.B.A. Coelho, H.E. Silva, F.M.A. Henriques, Calibrated hygrothermal simulation models for historical buildings, Build. Environ. 142 (2018) 439–450, https://doi.org/10.1016/j.buildenv.2018.06.034.

[62] P. Penna, A. Gasparella, F. Cappelletti, F. Tahmasebi, Mahdavi, Optimization-based calibration of a school building based on short-term monitoring data, in: EWork Ebus. Archit. Eng. Constr. - Proc. 10th Eur. Conf. Prod. Process Model. ECPPM 2014, 2015, pp. 259–264, https://doi.org/10.1201/b17396-46.

[63] C.M. Muñoz-González, A.L. León-Rodríguez, M. Campano-Laborda, C. Teeling, R. Baglioni, The assessment of enviromental conditioning techniques and their energy performance in historic churches located in Mediterranean climate, J. Cult. Herit. 34 (2018) 74–82.

[64] N. Cardinale, G. Rospi, P. Stefanizzi, Energy and microclimate performance of Mediterranean vernacular building: the Sassi district of Matera and Trulli district of Alberobello, Build. Environ. 59 (2013) 590–598.

[65] M. Pretelli, K. Fabbri, Historic Indoor Microclimate of the Heritage Buildings - A | Marco Pretelli | Springer, Springer, Swizerland, 2018. http://www.springer.com/it/book/9783319603414.

[66] M. Woloszyn, T. Kalamees, M.O. Abadie, M. Steeman, A. Sasic Kalagasidis, The effect of combining a relative-humidity-sensitive ventilation system with the moisture-buffering capacity of materials on indoor climate and energy efficiency of buildings, Build. Environ. 44 (2009) 515–524, https://doi.org/10.1016/j.buildenv.2008.04.017.

[67] G. Ruiz, C. Bandera, Validation of calibrated energy models: common errors, Energies 10 (2017), https://doi.org/10.3390/en10101587.

[68] FEMP, M&V guidelines: measurement and verification for performance-based contracts - version 4.0, Washington, DC, USA. http://www1.eere.energy.gov/, 2015.

[69] V. Pracchi, Historic buildings and energy efficiency, Hist. Environ. Policy Pract. 5 (2014) 210–225, https://doi.org/10.1179/1756750514z.00000000052.

[70] D. Chakraborty, H. Elzarka, Performance testing of energy models: are we using the right statistical metrics? J. Build. Perform. Simul. 11 (2018) 433–448, https://doi.org/10.1080/19401493.2017.1387607.
[71] P. Baggio, Simulazione energetica degli edifici esistenti : guida alla definizione di modelli calibrati, EPC Editore, 2013. https://www.epc.it/Prodotto/Editoria/Libri/libro-Simulazione-energetica-degli-edifici-esistenti/2003. (Accessed 6 July 2018).
[72] D.T. Rubeis, I. Nardi, M. Muttillo, D. Paoletti, The restoration of severely damaged churches – implications and opportunities on cultural heritage conservation, thermal comfort and energy efficiency, J. Cult. Herit. (2019), https://doi.org/10.1016/j.culher.2019.11.008.
[73] S. Thravalou, M. Philokyprou, A. Michael, The impact of window control on thermal performance. investigating adaptable interventions in vernacular mediterranean heritage, J. Architect. Conserv. 24 (2018) 41–59, https://doi.org/10.1080/13556207.2018.1456058.
[74] R. Kilian, Klimastabilität Historischer Gebäude, Bewertung hygrothermischer Simulationen im Kontext der Präventiven Konservierung, 2013.
[75] E. Lucchi, Review of preventive conservation in museum buildings, J. Cult. Herit. 29 (2018) 180–193, https://doi.org/10.1016/j.culher.2017.09.003.
[76] P. Raftery, M. Keane, A. Costa, Calibrating whole building energy models : detailed case study using hourly measured data, Energy Build. 43 (2011) 3666–3679, https://doi.org/10.1016/j.enbuild.2011.09.039.
[77] 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.
[78] P. Penna, F. Cappelletti, A. Gasparella, F. Tahmasebi, A. Mahdavi, Multi-stage calibration of the simulation model of a school building through short-term monitoring, J. Inf. Technol. Construct. 20 (2015) 132–145. http://www.itcon.org/2015/10.
[79] T. Yan, Z. Sun, X. Xu, H. Wan, G. Huang, Development of a simplified dynamic moisture transfer model of building wall layer of hygroscopic material, Energy 183 (2019) 1278–1294, https://doi.org/10.1016/j.energy.2019.07.033.
[80] G. Mustafaraj, J. Chen, G. Lowry, Development of room temperature and relative humidity linear parametric models for an open office using BMS data, Energy Build. 42 (2010) 348–356, https://doi.org/10.1016/j.enbuild.2009.10.001.
[81] H.E. Huerto-Cardenas, Validation of Historical Buildings Energy Models : A Method Based on Microclimatic Control Parameters, Politecnico di Milano, 2020. PhD. Thesis.
[82] Z. O'Neill, B. Eisenhower, Leveraging the analysis of parametric uncertainty for building energy model calibration, Build. Simul. 6 (2013) 365–377, https://doi.org/10.1007/s12273-013-0125-8.
[83] C. Cornaro, D. Sapori, N. Marconi, Sustainability aspects related to preservation of historical buildings: a case of intervention, in: Proc. 16 Th CIRIAF Natl. Congr., 2016.
[84] F. Mancini, M. Cecconi, F. De Sanctis, A. Beltotto, Energy retrofit of a historic building using simplified dynamic energy modeling, Energy Procedia 101 (2016) 1119–1126, https://doi.org/10.1016/J.EGYPRO.2016.11.152.
[85] K.S.M. Bughrara, D.Z. Arsan, G.G. Akkurtc, Applying underfloor heating system for improvement of thermal comfort in historic mosques: the case study of Salepçioğlu Mosque, Izmir, Turkey, Energy Procedia 133 (2017) 290–299, https://doi.org/10.1016/j.egypro.2017.09.390, in: https://search.crossref.org/?q=K.S.M.+Bughrara%2C+D.Z.+Arsan%2C+G.G.+Akkurtc%2C+Applying+underfloor+heating+system+for+improvement+of+thermal+comfort+in+historic+mosques%3A+the+case+study+of+Salep%C3%A7io%C4%9Flu+Mosque%2C+Izmir%2C+Turkey%2C+Energy+Procedia.
[86] C. Lodi, S. Magli, F. Maria Contini, A. Muscio, P. Tartarini, Improvement of thermal comfort and energy efficiency in historical and monumental buildings by means of localized heating based on non-invasive electric radiant panels, Appl. Therm. Eng. 126 (2017) 276–289.
[87] G. Semprini, C. Galli, S. Farina, Reuse of a ancient church: thermal aspect for integrated solutions, Energy Procedia 133 (2017) 327–335.
[88] C. Balocco, A. Colaianni, Assessment of energy sustainable operations on a historical building. The Dante Alighieri high school in Florence, Sustain. Times 10 (2018), https://doi.org/10.3390/su10062054.
[89] M. Benchekroun, S. Chergui, F. Ruggiero, S. Di Turi, Improving the indoor climate of the traditional Ottoman houses in the Medina of Algiers, in: IOP Conf. Ser. Mater. Sci. Eng., 2019, https://doi.org/10.1088/1757-899X/609/4/042073.