Calibrated BEMs and LSTM Neural Networks for Indoor Temperature Prediction: A Comparative Analysis in Pre- and Post-Retrofit Scenarios

Gianluca Maracchini – University of Trento, Italy – gianluca.maracchini@unitn.it
Nicola Callegaro – University of Trento, Italy – nicola.callegaro@unitn.it
Rossano Albatici – University of Trento, Italy – rossano.albatici@unitn.it

Abstract

The need to mitigate the risks of overheating in buildings due to climate change has highlighted the importance of accurate models for predicting indoor temperatures and thermal comfort, particularly after retrofitting. To this end, white-box models, such as Building Energy Models (BEMs), and black-box models, such as Long Short-Term Memory (LSTM) neural networks, have been extensively used in recent decades. While BEMs provide detailed insights through physically-based simulations, requiring calibration for enhanced accuracy, LSTMs provide a data-driven approach that captures complex thermal dynamics with greater simplicity, albeit with less interpretability. Few studies have undertaken a comparative analysis of these models in terms of prediction accuracy, especially across pre- and post-retrofit conditions and different lengths of training periods. Thus, in this study, a comparison between the predicting capabilities of calibrated BEMs and LSTM in summer was carried out using two real monitored mock-ups in Northern Italy representing both pre- and post-retrofit conditions. The results show that, for the considered limited training periods (8 and 3 days), the dataset size does not significantly influence BEM accuracy, while LSTM accuracy is more affected. Moreover, BEMs show higher prediction accuracy in scenarios with higher indoor air temperature (IAT) variability, i.e. where unseen data could be less predictable, such as in pre-retrofit conditions. LSTMs, however, excel in low-variability scenarios, such as the post-retrofit conditions in this case. This study highlights the critical need for careful model selection and calibration based on the data availability and building typology to ensure prediction reliability.

1. Introduction

The need to mitigate overheating risks in buildings under climate change scenarios, coupled with the rise in the adoption of Model Predictive Control and fault detection and diagnosis (FDD) systems, led to the need for accurate indoor temperature and thermal comfort prediction models, especially for post-retrofit conditions. To achieve this aim, white-, grey- and black-box models have been widely adopted in recent decades (Shahcheraghian et al., 2024). White boxes, such as Building Energy Models (BEMs), use physically based simulations of building dynamics, providing detailed insights and good prediction, especially if a calibration procedure is undertaken. Back-box models, such as Long Short-Term Memory neural networks (LSTM), employ a fully data-driven approach to capture complex thermal dynamics, offering simplicity and adaptability but at the cost of reduced interpretability (Mtibaa et al., 2020; Lu et al., 2022; Cui et al., 2023). Given that each methodology presents its advantages and drawbacks, and that various models have shown differing performance in different scenarios, it is important to undertake a thorough comparison to guide model selection. However, to the authors' knowledge, few studies have aimed to compare the accuracy of these modeling approaches under different conditions. In Arendt et al. (2018), white-, grey- and black-box models are compared in terms of indoor air temperature (IAT) predictability, finding that black-box models outperform grey- and white-box models in quite almost the considered scenario, with grey-box models needing shorter training periods for good accuracy. In Afram and Janabi-Sharifi (2015), Cui et al. (2023), and Vivian

Part of
Pernigotto, G., Ballarini, I., Patuzzi, F., Prada, A., Corrado, V., & Gasparella, A. (Eds.). 2025. Building simulation applications BSA 2024. bu,press.
https://doi.org/10.13124/9788860462022

et al. (2024), grey- and black-box models are compared, finding that, on average, LSTM outperforms grey-box models albeit grey-box models remain a valid alternative, especially in case of low data availability. In Hauge Broholt et al. (2022), the robustness of black and grey-box models of thermal building behavior against weather changes is evaluated. The authors found that the predictive performance of the grey-box models was slightly better compared to the black-box model in this case. However, not all these studies include white-box models, building envelopes, and the indoor environment, as well as different scenarios such as pre- and post-retrofit conditions.

One of the main advantages of white-box models when compared to black-box ones is that once calibrated they can be modified to reflect optimization changes in the represented object. Several Standards suggest the adoption of calibrated simulations to estimate the energy saving achievable through energy retrofit measures (EVO, 2012; ASHRAE, 2014). However, a few studies verified the accuracy of a BEM calibrated in pre-retrofit conditions in predicting the thermal and energy response of a retrofitted one (Chong et al., 2021).

For these reasons, this study has two main objectives, i.e.

i) to compare the accuracy of BEMs and LSTM models in predicting the indoor thermal response of buildings considering both pre- and post-retrofit conditions and different training periods;

ii) to assess the ability of BEMs, calibrated in pre-retrofit conditions, to reproduce the post-retrofit indoor thermal response, also considering post-retrofit, i.e. second-stage, calibration.

The findings of this study can help researchers and engineers select the best model for IAT prediction based on data availability and building typology.

The paper is organized as follows: Section 2 describes the methodological approach, the experimental setup, and the adopted modeling methods. Section 3 reports the comparison between models and a critical discussion of the results. Finally, Section 5 summarizes the key findings of the research.

2. Phases, Materials and Methods

2.1 Phases

The work is subdivided into the following three phases:

- first, two identical pre-retrofit mock-ups (Cell A and Cell B) representative of Italian buildings from the 1960s were built and their thermal response was monitored and compared in pre-retrofit conditions to provide proof that the construction process led to the same thermal response;
- then, BEMs and LSTM models were created and calibrated, and then compared in pre-retrofit conditions in terms of model accuracy also considering different training periods, to identify the best solution and training period in this case;
- finally, a comparison between BEM and LSTM models in the post-retrofit scenarios was made (Cell A was retrofitted in the second year). In particular, the LSTM was recreated using the post-retrofit data. Conversely, to assess the capacity of BEMs calibrated in pre-retrofit conditions to predict the post-retrofit IAT after the appropriate upgrades, a post-retrofit BEM was created by modifying the pre-retrofit calibrated one and then recalibrating it only by tuning the properties of the new layer (e.g. external insulation) and modified building characteristics (e.g. infiltration rate).

2.2 Experimental Setup

Two identical free-running experimental test cells (Cell A and Cell B) were constructed in Malosco, Italy (1041 m a.s.l.), to allow an experimental comparison between pre- and post-retrofit scenarios considering the same outdoor conditions (see Fig. 1a and b), which is a quite rare comparison in the existing literature. The cells, designed to be representative of the construction type and size of a typical Italian room in a pre-retrofit condition, were monitored for two consecutive years. After the first year, Cell A was retrofitted with an innovative, non-intrusive, and modular timber façade system, enabling an experimental comparison between pre- and

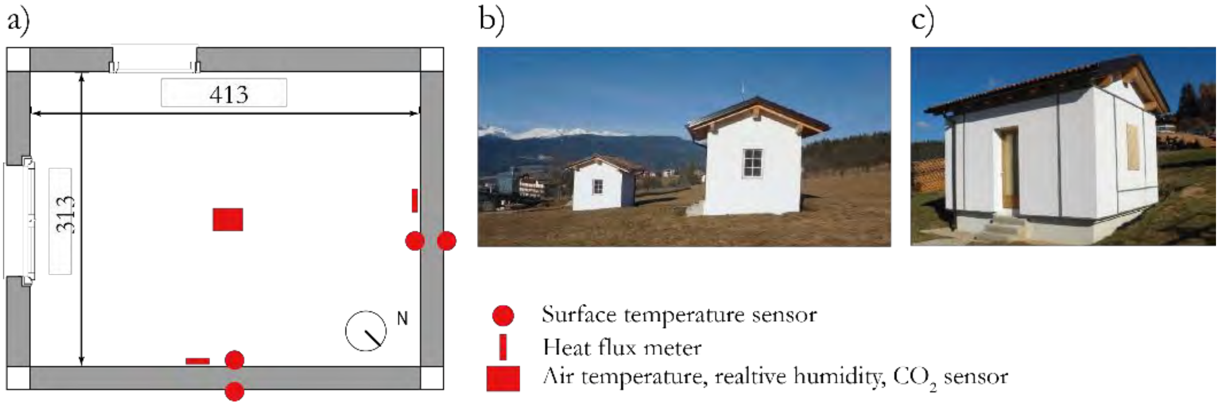
a)
413
313
N
b)
c)
● Surface temperature sensor
▮ Heat flux meter
🟥 Air temperature, realtive humidity, CO_2 sensor

Fig. 1 – a) Geometrical description of the experimental units and sensors placement (dimensions in centimeters). b) The two test cells before retrofit; c) Cell A after retrofit

post-retrofit scenarios under similar external conditions (Fig. 1c). In particular, a novel, prefabricated, multi-layered insulation panel made of a self-sustained wood frame filled with mineral wool and enclosed with OSB panels was applied to the walls providing an additional, nominal wall thermal resistance of 7.81 m^{2}K/W (Callegaro & Albatici, 2023). No internal loads were present during the periods considered to solely assess environmental influences on indoor conditions, providing ground truth data for model calibration. More information on construction features, building characteristics, and monitoring systems can be found in (Callegaro & Albatici, 2023).

2.3 White-Box Model

2.3.1 Building energy modeling

The Sketchup v. 2013 was used as a graphical interface of the EnergyPlus v.22.2.0 simulation engine (DOE 2017) to create the BEM of the test cells, following modeling methods described in (Maracchini & D'Orazio, 2022). In particular, the Conduction Transfer Function (CFT) was adopted as a heat balance algorithm with 6 timesteps per hour considered for calculation. Internal and external convective heat transfer coefficients were computed by adopting an adaptive convection algorithm, since generally more reliable for calculating convective heat transfer coefficients if compared to other techniques (Costanzo et al., 2014). The Flow Coefficient model was finally implemented to model the infiltration rate (Maracchini & D'Orazio, 2022).

2.3.2 Sensitivity analyses and calibration

A software tool specifically developed by the authors was used for the model calibration (Maracchini, 2023). This tool integrates the Morris method for parameter screening (Saltelli et al., 2008; Tian and Wei, 2013) with the Non-dominated Sorting Genetic Algorithm (NSGA-II) for the optimization-based calibration (Martínez et al., 2020; Vera-Piazzini & Scarpa, 2024). In particular, the Morris method is used to identify the parameters with uncertainty that mostly impact model accuracy and then those that can be discarded from the calibration process to reduce the computational burden without reducing calibration effectiveness.

All the most relevant parameters were considered for both sensitivity analysis and calibration. The parameters considered for calibration and the related range of variations are reported in Table 1 for both pre-retrofit and post-retrofit scenarios. A multiplier approach was adopted to avoid compensation errors among layers of the same building component. For example, for each component type (walls, roof, and floor), all the conductivities were grouped with a single multiplier that was varied between a range of ±20% (WALL_COND, ROOF_COND, etc.). Similarly, for each component, the density and specific heat capacity of all the layers were grouped using a volumetric heat capacity (VHC) multiplier, used to avoid compensation errors in terms of thermal inertia effects.

Concerning the target function used in both the sensitivity and calibration processes, different error metrics can be used (Martínez et al., 2020). In this work, the Root Mean Square Error (RMSE) was used as a target function for the sensitivity analysis while, for the calibration, a single-objective optimization approach was adopted with a target function computed as the product between the RMSE, the R2, and a novel indicator introduced in this study, named Maximum Absolute Hourly Difference (MAD) computed between simulated and measured data. This approach is considered more effective than the use of multi-objective optimization since it can provide optimized solutions with a lower computational time while being sufficiently accurate in terms of both absolute errors and inertial effect.

Concerning the post-retrofit conditions, only the parameters of the added layer and the modified building characteristics (e.g. infiltration rates, i.e. flow coefficient) are varied for model calibration (see Table 1). Due to the low number of parameters considered, a sensitivity analysis was not necessary in this case.

Table 1 – Parameters and related ranges considered in sensitivity and BEM calibration processes. Parameters with * are multipliers. WIN: windows; VHC: Volumetric Heat Capacity; TA: Thermal Absorptance; SA: Solar Absorptance; COND: conductivity; SR: Solar Reflectance; T: Temperature. Parameters with X in the post-retrofit column are considered for calibration of new layers/modified building characteristics.

Parameters	Range	Post-retrofit
WALLS_COND*	[0.8, 1.2]	X
WALLS_VHC*	[0.8, 1.2]	X
WALLS_TA	[0.8, 0.95]	X
WALLS_SA	[0.1, 0.3]	X
ROOF_COND*	[0.8, 1.2]	
ROOF_VHC*	[0.8, 1.2]	
FLOOR_COND*	[0.8, 1.2]	
FLOOR_VHC*	[0.8, 1.2]	
ROOF_TA	[0.8, 0.95]	
ROOF_SA	[0.3, 0.8]	
DOOR_COND*	[0.2, 2.0]	
DOOR_VHC*	[0.2, 2.0]	
DOOR_TA	[0.8, 0.95]	
DOOR_SA	[0.3, 0.8]	
WIN_COND*	[0.2, 2.0]	X
WIN_TA	[0.8, 0.95]	X
WIN_SA	[0.3, 0.8]	X
FLOW_COEF	[0.0001, 0.007]	X
GROUND_T [°C]	[13.0, 17.0]	
GROUND_SR	[0.15, 0.25]	

2.4 Black-Box Model

Concerning black box models, pure Artificial Neural Networks such as the Long-Short-Term-Memory neural networks (LSTM) were adopted using their capacity to learn long-term dependencies in dynamic systems like buildings (Lu et al., 2022). LSTM models consist of chains of neural network modules, focusing on a cell state mechanism that manages information flow through gates, by adopting the following workflow:

1. First, a forget gate determines what information to discard from the cell state.
2. Then, an input gate decides which new information to add to the cell state. This involves a gate that selects values to update and a *tanh* layer that generates a vector of new candidate values that could be added to the state;
3. Thirdly, the old cell state is updated with the new values identified in the previous step;
4. Finally, the final output is generated based on the updated cell state, which is modified by an output gate that applies a *tanh* function to scale the values between -1 and 1.

In this study, LSTM was developed with the Python TensorFlow library (TensorFlow Developers, 2024) and calibrated through different steps involving generating the network, optimizing hyperparameters, and training it to assimilate system behavior. This process includes layering LSTM with a fully connected layer and a sigmoid function, normalizing input data, and adjusting hyperparameters like learning rate, hidden layer size, and optimization algorithms (see Table 2) to minimize loss, ensuring the model avoids underfitting or overfitting by regulating training iterations (epochs) (Vivian et al., 2024).

Table 2 – Hyper-parameters optimization.

Hyper-parameters	Options
Learning rate	[0.01, 0.001, 0.0001]
LSTM hidden layer size	[16, 32, 64, 128]
Optimization algorithm	Adam, RMSprop

2.5 Model Accuracy Comparison

The evaluation and comparison of the performance of the models was carried out both qualitatively (graphical comparison) and quantitatively. In the latter case, reference is first made to common error indicators computed between predicted and observed data, such as RMSE, R2, and the additional MAD indicator introduced in this study.

The period considered for model calibration and comparison goes from the 21st of June to the 1st of July 2021 for both Cell A (post-retrofit condition) and Cell B (pre-retrofit condition). The hourly dataset was subdivided into training and testing datasets. For the training one, different lengths were considered for comparison purposes equal to 3 and 8 days, respectively. The test datasets instead referred to the last 3 days of the period considered.

3. Results and Discussion

3.1 Experimental Comparison

In this section, a comparison between the experimental IAT profiles obtained for the two mock-ups in pre-retrofit conditions (first summer) is reported. As can be seen from Fig. 2, a very good agreement is obtained in terms of IAT. In particular, RMSE, R2, and MAD values computed between the two datasets are equal to 0.14 °C, 1.00, and 0.20 °C, respectively. The R2 value denotes a perfect alignment of the two curves in terms of trend and inertial effect, while the two error indicators (RMSE and MAD) denote a negligible difference between the two mockups, even lower than the instrument accuracy, thus denoting a complete overlapping between the two IAT profiles. Thus, the two cells showed the same thermal response, and the two buildings can be correctly used for comparative purposes between pre- and post-retrofit conditions.

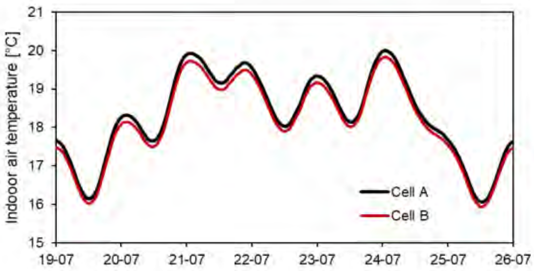

Fig. 2 – Comparison between test cells in pre-retrofit conditions

3.2 Pre-Retrofit Comparison

In this section, the results of the pre-retrofit comparison are reported and discussed. Fig. 3 reports the results of the sensitivity analyses in terms of the mean value of the absolute values of the elementary effect μ*, which is used to rank the parameters from the most to the least important in calibration processes (Tian & Wei, 2013).

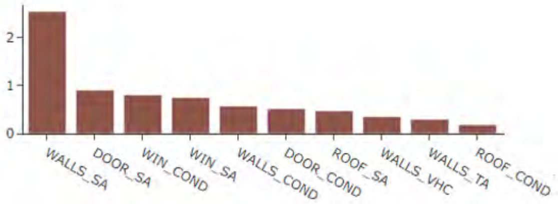

Fig. 3 – Sensitivity analyses results in terms of μ* computed on RMSE values [°C]. Only the most parameters are reported for the sake of brevity

As can be seen, in this case, the solar absorptance of the walls (WALLS_SA) is the most important parameter for IAT prediction, followed by door SA and window thermal conductivity. Parameters not reported in Fig. 3 have a negligible effect on the IAT, and therefore are not considered for model calibration.

In Fig. 4a, the comparison between BEM and LSTM output is provided in terms of IAT profiles, while Table 3 reports the values of the accuracy indicator. Both the BEMs calibrated with 3 and 8 training days (BEM B3 and BEM B8, respectively) show a good agreement with the experimental data in both the training and testing phases, with a strong reduction of the initial RMSE, R2, and MAD (equal to 2.74 °C, 0.59 and 3.6 °C, respectively, for the training period). The overlapping between the BEM B3 and BEM B8 model outputs indicates that, in this case, BEM models seem not to be affected by the variation in training length, thus the shorter length is considered sufficient for predicting testing data.

Conversely, as expected, LSTM models (LSTM B3 and LSTM B8) seem to be more affected by the training dataset length (see Fig. 4a), with increasing accuracy when longer periods are considered. In general, however, LSTMs outperform BEMs in the training phase, while the inverse is observed in the testing phases. This denotes the difficulty of LSTMs to predict new data points with the considered training periods and highlights the need for a larger dataset for training.

Table 3 – Accuracy indicators in pre- and post-retrofit conditions with 3 and 8 training days. RMSE and MD values in °C. tr: training; t: testing. Best results are underlined

Model	RMSE$_{tr}$	R2$_{tr}$	MAD$_{tr}$	RMSE$_{t}$	R2$_{t}$	MAD$_{t}$
BEM B3	0.28	0.93	0.44	0.21	0.99	0.50
BEM B8	0.36	0.91	0.69	0.24	0.98	0.38
LSTM B3	0.27	0.98	0.52	0.56	0.99	1.28
LSTM B8	0.03	1.00	0.20	0.49	0.91	0.95
BEMB3up	0.44	0.74	0.82	0.70	0.80	1.26
BEMB8up	1.24	0.56	2.13	1.24	0.88	1.78
BEM A3	0.30	0.82	0.43	0.40	0.91	0.90
BEM A8	0.41	0.74	0.84	0.42	0.94	0.74
LSTM A3	0.01	1.00	0.03	0.10	0.94	0.30
LSTM A8	0.02	1.00	0.14	0.09	0.94	0.23

3.3 Post-Retrofit Comparison

One of the main advantages of BEMs when compared to LSTM is that they can be modified to reflect changes in the building over its lifetime and that were not considered in the training phase. To understand the capability of BEMs calibrated on pre-retrofit conditions to predict the thermal response of a retrofitted building, BEM B3 and BEM B8 were updated to reflect the energy retrofit modifications that occurred in Cell A.

In Fig. 4b, the comparison between BEM and LSTM output is provided in terms of IAT profiles, while Table 3 reports the values of the accuracy indicator. As expected, despite the improvement, the obtained model (BEM B3up and BEM B8up in Fig. 2b) provides accuracy indicators lower than that previously achieved for BEM B3 and BEM B8 (see Table 3).

This can be traced back to:

i) the different construction features that are not considered in the model upgrade (e.g. infiltration);

ii) the uncertainty related to the thermal properties of the newly added materials;

iii) overfitting/compensation errors that may have occurred during the calibration phase of B3 and B8 models.

To reduce the impact of the two first error categories, a recalibration of B3up and B8up was carried out by fine-tuning building infiltration parameters

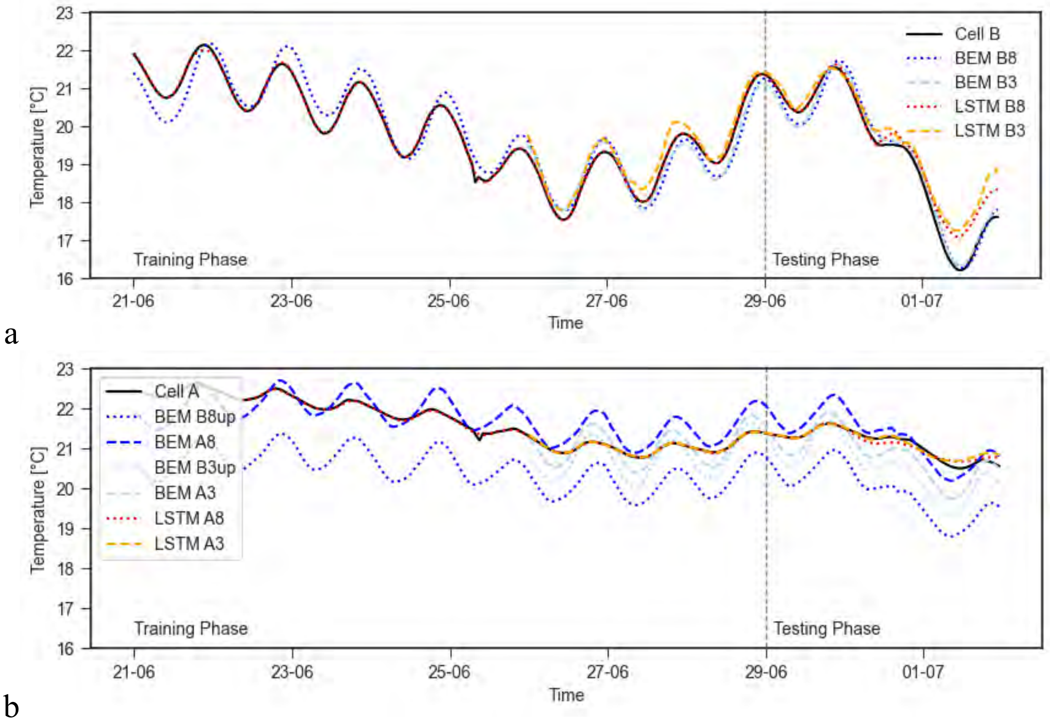

Fig. 4 – Comparison between predicted and observed IAT in both training and testing phases considering (a) pre- and (b) post-retrofit conditions

and thermal properties of the added wall layers, obtaining two "two-staged" calibrated BEMs BEM A3 and BEM A8. The obtained profiles show an accuracy higher than that previously observed (Fig. 4b) but still lower than that obtained through a full calibration with all relevant parameters involved (as in BEM B3 and BEM B8 cases, see Table 3). This may be caused by the presence of overfitting/compensation errors in the pre-retrofit calibration phase, which is also stressed by the different output profiles obtained for B3up and B8up (which were identical in Fig. 4a). Concerning LSTMs (LSTM A3 and LSTM A8), better performance is obtained, even better than that achieved in the pre-retrofit phase (LSTM B3 and LSTM B8) in both training and testing phase. Moreover, a lower dependence of LSTM on the training period can be observed. This can be due to the lower variability and dependence of observed indoor data on outdoor predictors in this case (post-retrofitted Cell A), which makes this scenario easier to predict than a pre-retrofit one, even with shorter training datasets.

4. Conclusions

This study investigated the accuracy of calibrated BEMs and LSTM models in predicting IATs in summer, in both pre- and post-retrofit scenarios. The results showed that in pre-retrofit conditions BEMs have consistent accuracy regardless of the training dataset size. Post-retrofit updates to pre-retrofit calibrated BEMs decreased their accuracy due to unaccounted changes and uncertainties. Recalibration of new parameters improved the performance, although it did not reach pre-retrofit accuracy levels, probably due to overfitting/compensation errors in the pre-retrofit phase. Therefore, particular attention should be paid to this aspect or compensation errors when using calibrated simulations should be reduced. LSTMs increase accuracy with longer datasets in pre-retrofit conditions while performing better in post-retrofit, benefiting from reduced data variability and external dependencies, indicating shorter training datasets could be sufficient in this scenario.

The main limitation of this study lies in the use of: a) a single and limited monitoring period, b) a single construction system, and c) a specific building geometry and location for carrying out the comparisons. Therefore, future studies should be carried out to extend this work and make the results more generalizable. Further studies will also investigate the accuracy of calibrated BEMs in predicting other important output variables not considered in this study, such as relative humidity and heat fluxes.

Acknowledgement

This work is part of the Renew-wall project, sponsored by the Fanti Legnami company based in Malosco (IT), and supported by Provincial Law No. 6/1999 (Autonomous Province of Trento).

References

Afram, A., Janabi-Sharifi F. 2015. "Black-box modeling of residential HVAC system and comparison of gray-box and black-box modeling methods". Energy Build 94:121–149. https://doi.org/10.1016/j.enbuild.2015.02.045

Arendt, K., Jradi M., Shaker H.R., Veje C. 2018. "Comparative analysis of white-, gray- and black-box models for thermal simulation of indoor environment: teaching building case study".

ASHRAE 2014. ASHRAE "Guideline 14 - Measurement of Energy, Demand, and Water Savings".

Callegaro, N., Albatici R. 2023. "Energy retrofit with prefabricated timber-based façade modules: Pre- and post-comparison between two identical buildings". J Facade Des Eng, 11(1), 001–018. https://doi.org/10.47982/jfde.2023.1.01

Chong, A., Gu Y., Jia H. 2021. "Calibrating building energy simulation models: A review of the basics to guide future work". Energy Build 253:111533. https://doi.org/10.1016/j.enbuild.2021.111533

Costanzo, V., Evola G., Marletta L., Gagliano A. 2014. "Proper evaluation of the external convective heat transfer for the thermal analysis of cool roofs". Energy Build 77:467–477. https://doi.org/10.1016/j.enbuild.2014.03.064

Cui, B., Im P., Bhandari M., Lee S. 2023. "Performance analysis and comparison of data-driven models for predicting indoor temperature in multi-zone commercial buildings". Energy Build 298:113499. https://doi.org/10.1016/j.enbuild.2023.113499

DOE (2017) EnergyPlus. US Dep. Energy's

EVO (2012) IPMVP - International Performance Measurement and Verification Protocol - Concepts and Options for Determining Energy and Water Savings Volume 1

Hauge Broholt, T., Dahl Knudsen M., Petersen S. 2022. "The robustness of black and grey-box models of thermal building behaviour against weather changes". Energy Build 275:112460. https://doi.org/10.1016/j.enbuild.2022.112460

Lu, C., Li S., Lu Z. 2022. "Building energy prediction using artificial neural networks: A literature survey". Energy Build 262:111718. https://doi.org/10.1016/j.enbuild.2021.111718

Maracchini G. 2023. "A set of calibrated BEMs for real demonstration cases and proposed standardization". H2020 BIMSPEED Deliverable D3.4

Maracchini, G., D'Orazio M. 2022. "Improving the livability of lightweight emergency architectures: A numerical investigation on a novel reinforced-EPS based construction system". Build Environ 208(September 2021):108601. https://doi.org/10.1016/j.buildenv.2021.108601

Martínez, S., Eguía P., Granada E., Moazami A., Hamdy M. 2020. "A performance comparison of multi-objective optimization-based approaches for calibrating white-box building energy models". Energy Build 216:109942. https://doi.org/10.1016/j.enbuild.2020.109942

Mtibaa, F., Nguyen K-K., Azam M., Papachristou A., Venne J-S., Cheriet M. 2020. "LSTM-based indoor air temperature prediction framework for HVAC systems in smart buildings". Neural Comput Appl 32(23):17569–17585. https://doi.org/10.1007/s00521-020-04926-3

Saltelli, A., Ratto M., Andres T., Campolongo F., Cariboni J., Gatelli D., Saisana M., Tarantola S. 2008. "Global Sensitivity Analysis. The Primer". Wiley.

Shahcheraghian, A., Madani H., Ilinca A. 2024. "From White to Black-Box Models: A Review of Simulation Tools for Building Energy Management and Their Application in Consulting Practices". Energies 17(2):376. https://doi.org/10.3390/en17020376

TensorFlow Developers (2024) TensorFlow

Tian, W., Wei T. 2013. "A review of sensitivity analysis methods in building energy analysis". Renew Sustain Energy Rev 20:411–419. https://doi.org/10.1016/j.rser.2012.12.014

Vera-Piazzini, O., Scarpa M. 2024. "Building energy model calibration: A review of the state of the art in approaches, methods, and tools". J Build Eng 86:108287. https://doi.org/10.1016/j.jobe.2023.108287

Vivian, J., Prataviera E., Gastaldello N., Zarrella A. 2024. "A comparison between grey-box models and neural networks for indoor air temperature prediction in buildings". J Build Eng 84:108583. https://doi.org/10.1016/j.jobe.2024.108583