SEEDS International Conference

Responsible retrofit measures for traditional listed dwellings: An energy simulation validation strategy

Michela Menconi[1], Noel Painting[1], Poorang Piroozfar[1,2]

[1] School of Environment and Technology, University of Brighton, Brighton, BN2 4GA, UK
[2] Digital Construction Lab, University of Brighton, Brighton, BN2 4GA, UK

Keywords: Building Energy Simulation, Calibration, Cultural Heritage dwellings, Energy retrofit.

Abstract

Energy and carbon retrofitting of traditional listed dwellings (TLDs) in the South-East England is much required but faces a multi-faceted and complex suite of issues and problems. A research project has been designed to specifically address those problems which utilises a mixed methods approach centred on multi-staged dynamic Building Energy Simulations (BES). Representative case studies of TLDs in Brighton and Hove have been selected, surveyed, modelled and simulated to assess their current energy performance and thermal behaviour as well as potential benefits of responsive and effective retrofit interventions.

The use of simulation implies the need for a thorough validation strategy to ensure that the data generation and analysis tool is reliable, valid and replicable in similar or identical contexts. Case studies research allows for an empirical validation, based on the calibration of simulated models with monitored data.

This paper describes the calibration process specifically devised for this on-going research project. Based on the findings of a critical review of literature, it utilizes energy consumption data as well as temperature and relative humidity data for each case study. Providing a brief overview of the methodological framework of the research, the paper describes in detail the approach utilised to ensure that the datasets collected and generated using different sources corroborate each other. The models created therefore accurately represent the real case studies in their status-quo and can be reliably used during the following stages of analysis when the impact of selected retrofit interventions is to be evaluated.

INTRODUCTION

Research background

The use of dynamic Building Energy Simulation (BES) tools, as a means to optimise the energy performance of buildings, has sensibly increased since they first emerged in the 70s. This has been driven by more stringent requirements for tighter energy conservation measures. However, several studies to date (BRE, 2014; Dall'O, G. et al., 2012; Heat et al., 2010; Hubbard, 2011; Ingram & Jenkins, 2013; Jenkins, 2008; Moran, 2013; STBA, 2012; Thompson & Bootland, 2011; Wingfield et al., 2011) have highlighted significant discrepancies between simulated performance and measured data when BES tools were deployed to assess the energy behaviour of existing buildings or to foresee the energy performance of new ones. Therefore, such instruments are yet frequently criticised for their lack of precision (Coakly et al. 2014). Concerns are even more profound when it comes to simulation of traditional (and more so for traditional listed) buildings because of the complexities of processes and synergies that characterise this part of the stock and increase the challenges in simulating the thermal behaviour of traditional buildings correctly (BRE, 2014; STBA, 2012).

Therefore, the need for thoroughly devised strategies for validating the output of energy simulations is evident, to ensure that the models created can generate realistic results. When BES is deployed for the performance analysis of the status-quo of a building in use and/or for testing possible solutions to improve such performance, the comparison between the results generated by simulation and empirical data - calibration - provides a powerful validation tool (Baranowski & Ferdin-Grygierek, 2009; Maile et al., 2012; Raftery et al., 2011; Ryan & Sansquist, 2011).

Aim of this paper

This paper sets out to describe the validation strategy specifically devised for an ongoing study that aims to propose responsive and effective retrofit interventions for C19th traditional listed dwellings (TLDs) in South-East England.

Overview of the research methodology

The research deploys BES, performed using IES-VE, for nine representative case studies (CSs) carefully selected in the city of Brighton and Hove (UK), using stringent inclusion/exclusion criteria set in this project. The study stems from two phases of data collection (as indicated in green in Figure 1), aimed to ensure that the simulated models accurately represent the thermal behaviour and energy performance of the real CS dwellings.

The first phase of data collection deployed multiple methods to gather a wide range of input data, necessary for the creation of the energy models, as follows:

- A critical review of literature established the gaps in knowledge with reference to: methodology, methods, data collection/generation and analysis instruments. This was followed by secondary data collection and expert consultation with local conservation experts, to help with necessary assumptions concerning the construction methods and materials build-up of the envelope as well as air leakage values;
- Visual and measured surveys provided data concerning location, typology, orientation, layout and measures, openings size and typology, traditional features, appliances, lighting fixtures;
- Questionnaires and interviews complemented the data previously gathered and added data concerning occupancy profiles, pattern of use of the heating and domestic hot water systems and appliances as well as window operation;
- Thermographic surveys enriched the understanding of the composition of the thermal envelope and aided in identifying possible thermal bridges, or areas of air leakage.

The second phase of data collection started in parallel and was carried out over a period of one year, aimed at gathering energy consumption and indoor conditions data for each CS, to be used in the following stage of calibration of the results of energy simulations, involving the following steps:

- Utility bills relative to the previous year's energy consumption were collected from the participants;
- Gas/LPG and electricity meter readings were performed as spot measurements over one full year;
- Temperature and relative humidity (RH) data were collected using sensors, over at least three months (during heated and unheated periods), in two rooms (living area and bedroom) for each period.

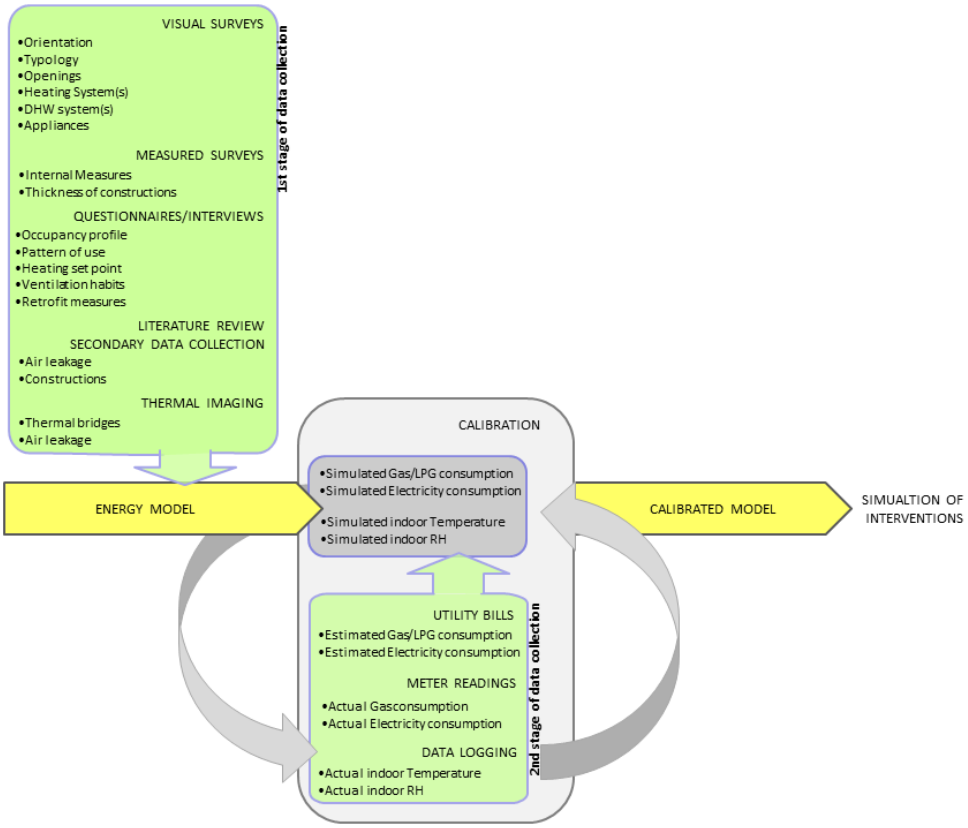

Figure 1. The methodological framework

As explained in detail in the following section, the comparison of simulated results and measured data allowed for calibration of the energy models, to ensure that the models created, virtual simplifications of the CS dwellings, accurately reproduce the energy performance and thermal behavior of the real dwellings. The calibrated models can then be used in the next stages of the research to assess the benefits or disadvantages, of potential retrofit interventions, one at a time and in combination.

MODELS CALIBRATION

The validation procedure specifically devised for this study, based on the critical literature review, develops further the approach already taken by other studies such as Bertagnolo (2012), Mustafarj et al. (2014), and Raftery et al. (2011). It constitutes of three subsequent stages of calibration, as follows:

- Stage one, based on input data obtained from visual and measured surveys, questionnaires, interviews, thermo-graphic surveys, literature review and secondary data collection;
- Stage two, where monitored energy consumption data is compared to simulated results and the input data are fine-tuned accordingly;
- Stage three, where monitored both temperature and RH data are compared to the simulation results from the model to complete the calibration.

A detailed description of the three subsequent stages of calibration follows:

Stage 1

At this stage, the energy models were created using the input data gathered during the first phase of data collection (see Figure 1). The energy simulations were then run using the average weather file for Brighton, as provided by Meteo-Norm[1].

Stage 2

An initial screening of the validity of the results generated by the energy simulations, was carried out using the estimated annual energy consumption data, as evinced from the energy bills, in order to assess the capability of the models to predict the energy consumption of the dwellings as investigated in their status-quo, as well as energy and CO_2 emission savings in the later design stages.

Meanwhile, the energy meter readings of each CS were carried out. This allowed for a calibration to be performed using actual figures where a whole year of energy consumption data was accounted for. The simulation results were compared to actual monitored data in two main categories of building energy usage: electricity and gas (or LPG wherever applicable, which was limited to one CS). Percentage Differences (PD) between simulation results and measured data were calculated for each energy consumption category over the annual period of investigation. This was done using the following equation, based on the work of Reeves et al. (2012):

PD (%) = [(Simulated Results – Measured Results) / Measured Results] x 100

Where positive values of the PD show that the simulation overestimated the annual energy consumption, and negative values indicate that it underestimated such consumption. Values of the PD in the range of ±10% were considered acceptable. This is in line with previous research (Ogando et al., 2017) and more challenging than the tolerances adopted by other researches (Reeves et al., 2012, following Maamari et al., 2006). The input values used in the first set of simulations for each CS were therefore fine-tuned, where needed, to calibrate the simulation outcome with the metered data and obtain results within the acceptable range.

At this stage, to increase the reliability of the models, a further calibration was performed using sub-annual monitored and simulated energy data. This calibration phase was more challenging than the one operated on annual data. In fact, the achievement of similar values for simulated output and monitored data relative to a shorter period, was much more affected by changes in a multitude of behavioural

---

[1] IES-VE uses site data that contain values for latitude, longitude and altitude of a wide range of sites throughout the world, taken from standard tables published by CIBSE and ASHRAE. Brighton is not included in such locations; therefore, local weather data were requested to Meteo-Norm.

factors and occupancy patterns. In the case of dwellings in use, such patterns may not be consistent, during shorter periods of investigation, with the general profiles relative to the whole year, produced according to the questionnaires and interviews with the occupants. Therefore, on the day of each meter reading, a few more questions were asked from the occupants and was taken note of any changes, that may have happened in the general profiles of use generated from the first interviews, during the specific sub-period of investigation. The energy consumption figures produced by the simulations, were estimated by the software given the input data concerning envelope constructions, heating schedule(s), heating set-point(s), appliances use pattern provided by the surveys and interviews. Such data were updated at this stage, in each sub-annual period between the meter readings, according to any special condition that the occupants were aware of.

At this stage, the applicability of ASHRAE (2002) criteria for calibration was considered as frequently adopted in previous research (Pan et al., 2007; Parker et al., 2012; Raftery et al., 2011; Yang & Becerik-Gerber, 2015; Yoon et al., 2003) when validating BES models using energy data. ASHRAE (2002) recommends the use of statistical indices, namely the Normalised Mean Bias Error [2] (NMBE) and the Coefficient of Variation of the Root Mean Square Error[3] (CV(RMSE)), applying them to hourly or monthly energy use data collected over one year. However, due to the constraints of this research (which makes use of CS dwellings with all the implications linked to the occupants' actual involvement) energy use data for a year was collected at longer intervals that spanned from 2 to 5 months. Nevertheless, it needs to be noted that the ASHRAE procedure (2002) only aims at achieving an acceptable correspondence between predicted and actual energy consumption; hence, the statistical indices proposed by ASHRAE only relate to energy in the context of the Guidelines. This validation approach, as noted by Coakley (2014) and Fabrizio and Monetti (2015) does not take into account other influential parameters, such as indoor condition, i.e. temperature and RH patterns, at the risk of producing models that do not really correspond to the actual building or of potentially producing more than one namely calibrated solution.

The strategy developed for this research instead, aims at balancing accuracy in energy consumption outputs as well as thermal behaviour data; this was considered of uttermost importance when studying traditional dwellings, whose materials and constructions are different from modern ones, hence it is challenging to reproduce accurately their complex thermal behaviour adopting the same methods utilized for modern buildings (BRE, 2014). Therefore, a more flexible approach than the one suggested by ASHRAE was taken for the calibration, when considering energy data. This stage in fact, was aimed at achieving ±15% PD between each sub-annual metered and simulated energy consumption. Because of the challenges imposed by the implications of varied patterns of use during short periods of time, such limits were not possible to achieve for all the CSs resulting, in a few cases, in values slightly

---

[2] NMBE (%) = [Σ(aD-sD) / ((n-1) x μ)] x 100 where: aD = actual data; sD= simulated data; n= number of actual/simulated data; μ= mean of actual data.

[3] CV(RMSE) (%) = (RMSE/ μ) x 100 = [√ (Σ(aD-sD)^2 /(n-1)) / μ] x 100 where: aD = actual data; sD= simulated data; n= number of actual/simulated data; μ= mean of actual data.

in excess of ±15% for one or more periods of investigation. Therefore, NMBE and CV(RMSE) were finally also calculated over one year on all the sub-annual periods; these indexes, for the reasons explained before, were allowed more flexible limits than the ones imposed by ASHRAE for monthly data.

Therefore, the models were iteratively refined, and the outputs were compared to measured data until:

- the PD calculated for annual energy data were within the limits of ±10%;

- the PD calculated for sub-annual energy data were within the limits of ±15% wherever possible and/or the NMBE and CV(RMSE) calculated over one year of sub-annual data, were respectively within ±10% and below 30%.

Such level of accuracy concerning energy data, was considered enough at this stage to proceed with the following stage of calibration. It has to be stressed in fact, that the final purpose of the calibration process described in this paper, is to serve the final aim of this research, which is to test the relative effects of potential retrofit interventions on the CS selected, in order to propose responsive and effective combinations of retrofit measures applicable to TLDs. This is done assessing the change in energy consumption between the base case dwellings and the versions with interventions. Therefore, whereas it is desirable to produce models capable of predicting the actual energy consumption of the dwellings investigated as accurately as possible, it is the change in energy consumption between the base case dwellings and the versions with interventions what this research aims to investigate.

Stage 3

When the calibration process was successful to this point, a further stage of calibration was performed using indoor temperature and RH data. This was done initially using graphic analysis (see Figure 2 and 3 for an example), when the winter and summer cycles of data logging were completed. The sub-hourly temperature and RH data acquired by the sensors were compared with the ones outputted from the dynamic simulations for the same periods and for the same rooms. At this stage, the objective was to validate the thermal behaviour of the models, establishing if the graphs presented remarkable discrepancies or were otherwise reliable, while also aiding in the understanding of the building envelope characteristics and the behaviour of its thermal mass. Furthermore, the graphic analysis provides a straightforward visual comparison of simulated and monitored data, aiding in identifying where the most evident discrepancies between such data exist, therefore where it is most likely that errors occur. Hence, complementary to the manual iterative calibration method - used in stage 1 and 2 - this further stage of calibration was conducted using visual analysis of the graphs to facilitate the decision concerning the specific periods of investigation needing further checks.

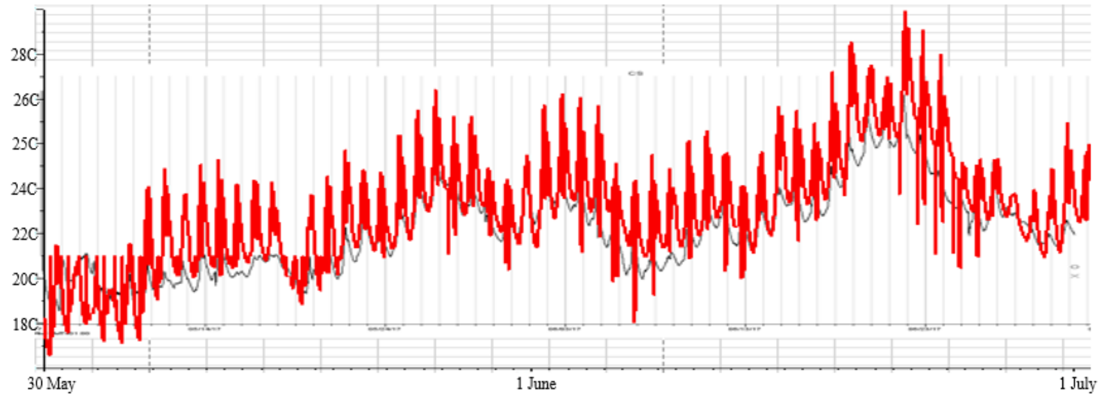

Figure 2. CS2: Graphic analysis of simulated (in red) and monitored (in black) temperature data for the living room over two summer months.

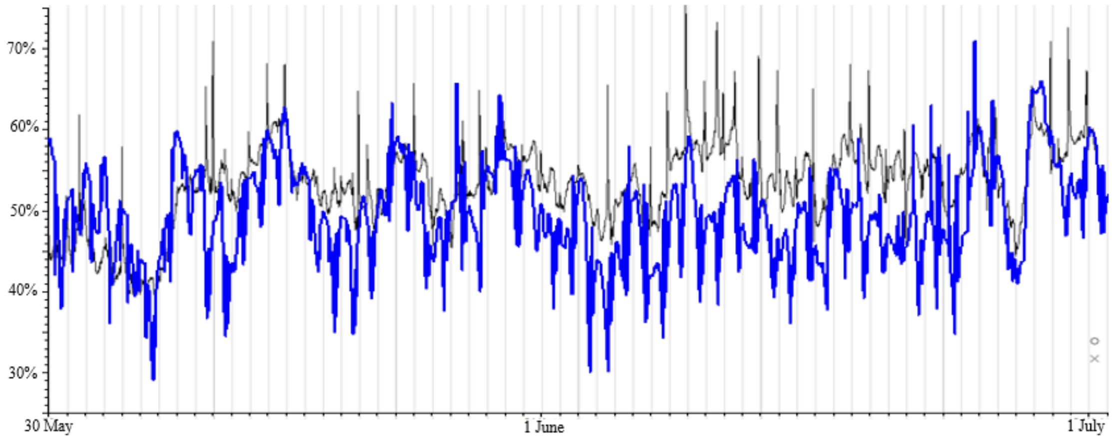

Figure 3. CS2: Graphic analysis of simulated (in blue) and monitored (in black) RH data for the living room over two summer months.

This stage of calibration was performed using the average Brighton weather file for all the data collected during 2018. For the data collected during 2017, a parallel simulation was run of the same models using the weather file generated via Weather Analytics for Brighton, relative to that specific year (as recommended by ASHRAE Guideline, 2002) acquired from IES. This further simulation was aimed at excluding the variables potentially causing discrepancies, as a result of weather, from the fine-tuning process. Having excluded differences between the weather file used in simulation and the actual weather, other factors - mainly pertaining to the pattern of use, heating system, building fabric, as indicated in the literature as the principal sources of errors in simulation - were examined more confidently to find the source of discrepancies.

When such analysis produced satisfactory results, simulated and monitored sub-hourly (with time intervals of 10 minutes and 30 minutes) recorded temperature and RH data, were used in the final stage of calibration, comparing them according to ASHRAE (2002) statistical indices for hourly data. All the CSs were calibrated calculating NMBE and CV(RMSE) for temperature and RH data for at least two (heated and unheated) periods, of 1 to 2 months each, in two different rooms (in living areas and bedrooms respectively).

The iterative calibration process, illustrated in Figure 4, was repeated for each CS, until:

- PD between simulated and measured annual energy data was within 10%;
- PD between simulated and measured sub-annual energy data was within 15% and/or NMBE and CV(RMSE) between simulated and measured sub-annual energy data was within ±10% and <30% respectively;
- A good similarity between the graphs of simulated and monitored temperature and RH data was reached for all the periods of data collection and all the rooms monitored;
- NMBE and CV(RMSE) between simulated and monitored temperature and RH sub-hourly data achieved values within ±10% and <30% respectively, for all the periods of data collection and all the rooms.

When all these criteria were fulfilled, the models were considered calibrated.

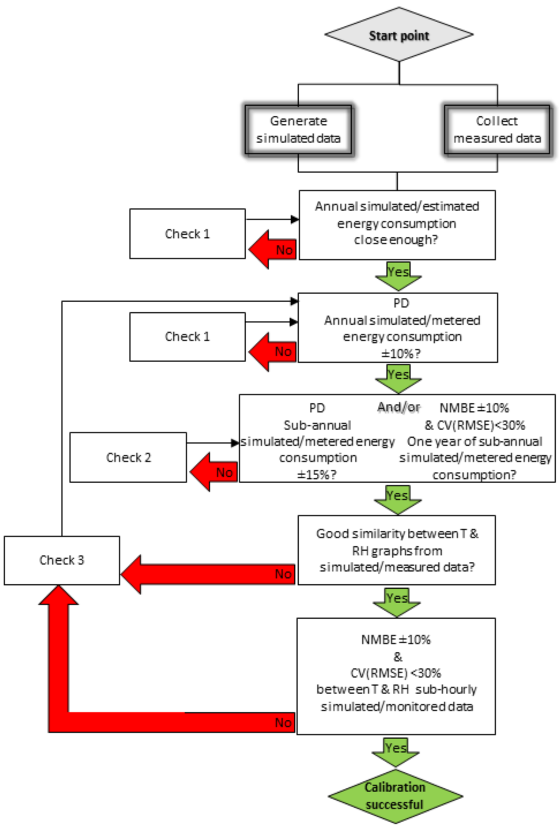

Figure 4. The calibration process.

INPUT DATA FINE TUNING

Hierarchy of input data

Previous research has pointed out how the calibration process can be extremely dependent on the researcher's personal judgment of the individual relevance of the multitude of parameters that need to be inputted in the simulation software (Maile et al., 2012; Raftery et al., 2011). Therefore, to avoid such subjective approach and to improve the validity and reliability of the calibration and the study, a strategy has been applied, developing the one used in previous studies (Gines Cooke, 2018; Parker et al., 2012; Raftery et al., 2011). First of all, a hierarchy was devised for the wide range of input data, depending on the source used to obtain them. The data sourcing strategies capable of providing on-site measured data concerning each specific CS under investigation were considered the most reliable ones in the context of this study. Therefore, to ensure consistency throughout the calibration process, input data have been sourced and checked according to the following hierarchy:

- Data from direct observation (mainly concerning envelope constructions -when visible-, heating/DHW systems, appliances, lights) and measured data, recorded during the site surveys;
- Data from questionnaires and interviews with the occupants concerning occupancy profile and pattern of use (related to heating schedules and temperature set-points, lighting schedules, DHW use, natural ventilation habits and frequency of use of the equipment);
- Data from benchmark studies, codes and legislations, best practice guides, standards and guidelines, operation manuals (CIBSE, 2015; EST, 2008; IES, 2009; Wood et al., 2009) to cover areas where it was not possible to secure data from other sources.

Data quality checks

During the subsequent stages of calibration, three levels of data quality checks (see Figure 4), were carried out, based on the hierarchy of sources devised, and on the literature reviewed concerning the sources of uncertainty in calibration (including, inter-alia, Blecich et al., 2016; De Wit & Augenbroe, 2002; Gucyeter, 2018; Hoes et al., 2016; Marini, 2014; Ogando et al., 2017; Parker et al., 2012; Reeves et al., 2012; Ryan & Sanquist, 2012).

CONCLUSION

The process of calibration specifically devised for this study was carried out until the results achieved were within the specified error margin. This iterative process can be described as:

- Running the simulations of the models created;
- Comparing their outputs with the monitored data;
- Identifying discrepancies and potential and relevant source(s) of error;
- Fine-tuning relevant parameters;
- Running the simulations of the newly modified models.

Numerous simulation runs were sometimes needed to obtain acceptable calibration levels. The calibration performed was successful for all the CSs investigated, adding to the validity, reliability and hence generalizability of future findings of the research project and, equally importantly, to its novelty. In fact, most of the precedent studies about traditional dwellings that aspired to complement the use of BES with a calibration strategy (Ingram, 2013; Mohammadpourkarbasi, 2015; Moran, 2013), were only based on a limited - sometimes one - CSs. This study by contrast uses a multiple CS approach to be able to increase its reach and validity.

All the models were calibrated using a full year record of energy consumption data as well as heated and unheated periods of temperature and RH data. It needs to be noted that the periods of data logging and meter readings are slightly different for each CS, but this is not affecting the quality of the calibration as each CS has been validated independently and individually within the period of data monitoring associated with it.

The calibration process devised this way, therefore, moves one steps further from previous research that deployed both energy and indoor conditions data for calibration of BES using CSs of public buildings with heritage values (Ogando et al., 2017; Sahin et al., 2015), dwellings (Georgiou, 2015) or traditional dwellings (Flores, 2013).

In fact, in this study, temperature data was monitored over an overall longer period of 3 to 6 months (depending on the actual participation of the occupants and ease of access to the dwellings). Notably, the use of two monitoring stages allowed for considering heated and unheated periods for calibration instead of just one, which is important in a temperate climate. Furthermore, the monitored temperature data were collected at intervals shorter than one hour (half hour and 10 minutes) and calibrated using graphic analysis as well as ASHRAE statistical indexes for all the data (vs exclusive use of graphic analysis, or satisfactory PD for most of data, as done by previous researches concerning dwellings). Importantly, this research also makes use of RH data for calibration adopting ASHRAE criteria, adding to the novelty of the study. In fact, the few precedent calibrations of BES applied to dwellings, that attempted the use of temperature as well as RH data (Bozonnet et al., 2011; Drissi Lamrhari and Benhamou, 2018), adopted PD and graphic analysis or maximum deviation, on indoor conditions data only, and did not involve energy consumption data at all. The novelty of this calibration approach is even more evident when compared to what has previously been done in the studies concerning traditional dwellings in the UK (Ingram, 2013; Mohammadpourkarbasi, 2015; Moran, 2013), where the calibration was only performed on energy (electricity and/or gas) data comparing annual total or monthly individual data - simulated and measured - using graphic analysis and/or PD.

The unique way in which the calibration process was devised in this study contributes to the novelty of the study and aims to ensure that the models created are accurate virtual reproductions of the real dwellings. To date, all the models are fully calibrated and ready to be used in the following stages of research to simulate the effect of potential retrofit interventions on the selected TLDs.

REFERENCES

ASHRAE 2002. ASHRAE Guideline 14-2002 for Measurement of Energy and Demand Savings. Atlanta, GA.: American Society of Heating, Refrigerating and Air-conditioning Engineers (ASHRAE).

BARANOWSKI, A. & FERDYN-GRYGIEREK, J. 2009. Heat demand and air exchange in a multifamily building — simulation with elements of validation. *Building Services Engineering Research & Technology,* 30**,** 227-240.

BERTAGNOLIO, S. 2012. *Evidence-based model calibration for efficient building energy services.* PhD, University of Liège.

BLECICH, P., FRANKOVIĆ, M. & KRISTL, Ž. 2016. Energy retrofit of the Krsan Castle: From sustainable to responsible design—A case study. *Energy and Buildings,* 122**,** 23-33.

BOZONNET, E., DOYA, M. & ALLARD, F. 2011. Cool roofs impact on building thermal response: A French case study. *Energy & Buildings,* 43**,** 3006-3012.

BRE 2014. Solid walls heat losses and the potential for energy saving.

CIBSE 2015. *CIBSE guide A: Environmental Design*
London, Chartered Institution of Building Services Engineers.
COAKLEY, D. 2014. *Calibration of detailed building energy simulation models to measured data using uncertainty analysis.* NUI Galway.
DALL'O, G., SARTO, L., SANNA, N., MARTUCCI, A. 2012. Comparison between predicted and actual energy performance for summer cooling in high-performance residential buildings in the Lombardy region. *Energy & Buildings,* 54**,** 234.
DE WIT, S. & AUGENBROE, G. 2002. Analysis of uncertainty in building design evaluations and its implications. *Energy & Buildings,* 34**,** 951-958.
DRISSI LAMRHARI, E.-H. & BENHAMOU, B. 2018. Thermal behavior and energy saving analysis of a flat with different energy efficiency measures in six climates. *Building Simulation,* 11**,** 1123-1144.
EST 2008. Measurement of domestic hot water consumption in dwellings. Energy Saving Trust.
FABRIZIO, E., MONETTI, V. 2015. Methodologies and Advancements in the Calibration of Building Energy Models. *Energies,* 8**,** 2548-2574.
FLORES, J. A. M. 2013. *The investigation of energy efficiency measures in traditional buildings in the Oporto World Heritage Site.* PhD, Oxford Brookes University.
GEORGIOU, G. 2015. *Assessing energy and thermal comfort of domestic buildings in the mediterranean region.* PhD, Loughborough University.
GINES COOKE, M. 2018. Analysis of implication and value of the process of calibration of Building Energy Models (BEMs). IES.
GUCYETER, B. Calibration of a building energy performance simulation model via monitoring data. Building Performance Analysis Conference and SimBuild co-organized by ASHRAE and IBPSA-USA. 2018 Chicago, IL.
HEATH, N., PEARSON, G., BARNHAM, B., ATKINS, R 2010. Historic Scotland Technical Paper 8: Energy modelling of the Garden Bothy, Dumfries House. Edinburgh: Historic Scotland.
HOES, P., HENSEN, J. L. M., LOOMANS, M. G. L. C., DE VRIES, B. & BOURGEOIS, D. 2009. User behavior in whole building simulation. *Energy and Buildings,* 41**,** 295-302.
HUBBARD, D. 2011. Ventilation, Infiltration and Air Permeability of Traditional UK Dwellings. *Journal of Architectural Conservation,* 17**,** 59-74.
IES 2009. Historic Scotland Technical Paper 5: Energy modelling of a mid 19th century villa - Baseline performance and improvement options. Edinburgh: Historic Scotland.
INGRAM, V. 2013. *Energy performance of traditionally constructed dwellings in Scotland.* PhD, Heriot-Watt University.
INGRAM, V., JENKINS, D. 2013. Historic Scotland Technical Paper 18: Evaluating energy modelling for traditionally constructed dwellings. Edinburgh: Historic Scotland.
MAAMARI, F., ANDERSEN, M., DE BOER, J., CARROLL, W. L., DUMORTIER, D. & GREENUP, P. 2006. Experimental validation of simulation methods for bi-directional transmission properties at the daylighting performance level. *Energy and Buildings,* 38**,** 878-889.
MAILE, T., BAZJANAC, V. & FISCHER, M. 2012. A method to compare simulated and measured data to assess building energy performance. *Building and Environment,* 56**,** 241.
MARINI, D., WEBB, LYNDA H., DIAMANTIS, G., BUSWELL, RICHARD A. Exploring the impact of model calibration on estimating energy savings through better space heating control. 2014 Building Simulation and Optimization Conference. 2014 London.
MOHAMMADPOURKARBASI, M. 2015. *The Eco-Refurbishment of a 19th Century Terraced House.* PhD, University of Liverpool
MORAN, F. 2013. *Benchmarking the energy use of historic dwellings in Bath and the role for retrofit and LZC technologies to reduce CO2 emissions.* Doctor of Philosophy, University of Bath.

MUSTAFARAJ, G., MARINI, D., COSTA, A. & KEANE, M. 2014. Model calibration for building energy efficiency simulation. *Applied Energy,* 130**,** 72-85.
OGANDO, A., CID, N. & FERNÁNDEZ, M. 2017. Energy Modelling and Automated Calibrations of Ancient Building Simulations: A Case Study of a School in the Northwest of Spain. *Energies,* 10**,** 807.
PAN, Y., HUANG, Z. & WU, G. 2007. Calibrated building energy simulation and its application in a high-rise commercial building in Shanghai. *Energy & Buildings,* 39**,** 651-657.
PARKER, J., CROPPER, P. AND SHAO, L. 2012. A calibrated whole building simulation approach to assessing retrofit options for Birmingham Airport. *First Building Simulation and Optimization Conference.* Loughborough, UK.
RAFTERY, P., KEANE, M. & O'DONNELL, J. 2011. Calibrating whole building energy models: An evidence-based methodology. *Energy & Buildings,* 43**,** 2356-2364.
REEVES, T., OLBINA, S. & ISSA, R. Validation of building energy modeling tools: Ecotect, Green Building Studio and IES. Winter Simulation Conference, 2012. IEEE, 1-12.
RYAN, E. M. & SANQUIST, T. F. 2012. Validation of building energy modeling tools under idealized and realistic conditions. *Energy & Buildings,* 47**,** 375-382.
ŞAHIN, C. D., ARSAN, Z. D., TUNÇOKU, S. S., BROSTRÖM, T., AKKURT, G. G., HUMANISTISK-SAMHÄLLSVETENSKAPLIGA, V., UPPSALA, U., HISTORISK-FILOSOFISKA, F. & KONSTVETENSKAPLIGA, I. 2015. A transdisciplinary approach on the energy efficient retrofitting of a historic building in the Aegean Region of Turkey. *Energy and Buildings,* 96**,** 128-139.
STBA 2012. Performance and Energy Efficiency of Traditional Buildings: Gap Analysis Study. London: STBA.
THOMPSON, P., BOOTLAND, J. 2011. Good Home Alliance Monitoring Programme 2009-11: Technical Report. *In:* ALLIANCE, G. H. (ed.).
WINGFIELD, J., BELL, M., MILES-SHENTON, D., SEAVERS, J. 2011. Elm Tree Mews Field Trial – Evaluation and Monitoring of Dwellings Performance: Final Technical Report *In:* UNIVERSITY, L. M. (ed.).
WOOD, C., BORDASS, B. AND BAKER, P. 2009. Research into the thermal performance of traditional windows: timber sash windows. English Heritage.
YANG, Z. & BECERIK-GERBER, B. 2015. A model calibration framework for simultaneous multi-level building energy simulation. *Applied Energy,* 149**,** 415-431.
YOON, J., LEE, E. J., CLARIDGE, D. E. 2003. Calibration Procedure for Energy Performance Simulation of a Commercial Building. *Journal of Solar Energy Engineering,* 125**,** 251-257.