

Domestic building fabric performance: Closing the gap between the *in situ* measured and modelled performance

A Marshall^{a*}, R Fitton^b, W Swan^b, D Farmer^c, D Johnston^c, M Benjaber^b & Y Ji^b

^a School of Computing Science and Engineering, University of Salford, Salford, Manchester, M5 4WT

^b School of the Built Environment, University of Salford, Salford, Manchester, M5 4WT

^c Centre for the Built Environment Group, Leeds Sustainability Institute Leeds Beckett University, Leeds, LS1 3HE

* Corresponding Author: a.s.marshall1@salford.ac.uk

Abstract

There is a growing body of evidence available to indicate that there is often a discrepancy between the *in situ* measured thermal performance of a building fabric and the steady-state predicted performance of that fabric, even when the building fabric has been modelled based upon what was actually built. However, much of the work that has been published to date does not fully investigate the validity of the assumptions within the model and whether they fully characterise the building. To investigate this issue, a typical pre-1920's UK house is modelled in Designbuilder in order to recognise and reduce the gap between modelled and measured energy performance. A model was first built to the specifications of a measured survey of the Salford Energy House, a facility which is housed in a climate controlled chamber. Electric coheating tests were performed to calculate the building's heat transfer coefficient; a difference of 18.5% was demonstrated between the modelled and measured data, indicating a significant '*prediction gap*'. Accurate measurements of air permeability and U-value were made *in-situ*; these were found to differ considerably from the standard values used in the initial model. The standard values in the model were modified to reflect these *in-situ* measurements, resulting in a reduction of the performance gap to 2.4%. This suggests that a better alignment between the modelling and measurement research communities could lead to more accurate models and a better understanding of performance gap issues.

Keywords: performance gap, domestic, energy modelling, measurement, building energy, energy performance, U-value

Definitions

ATTMA	Air Tightness Testing and Measurement Association
CIBSE	Chartered Institution of Building Services Engineers
DEAP	Dwellings Energy Assessment Procedure
EPC	Energy Performance Certificate
GCH	Gas Central Heating
HTC	Heat Transfer Coefficient
HVAC	Heating, Ventilation and Air Conditioning
INSHQ	Irish National Survey of Housing Quality
RDSAP	Reduced Data Standard Assessment Procedure
SAP	Standard Assessment Procedure
U-value	Thermal transmittance ($W m^{-2} K^{-1}$)

1. Introduction

There is a growing body of evidence available to indicate that there is often a discrepancy between the *in situ* measured thermal performance of the building fabric and the steady-state predicted performance of that fabric, even when the building fabric has been modelled based upon what was actually built, rather than the original design intent. In some cases, the measured *in situ* performance of the building fabric has exceeded the as-built predicted performance of the building fabric by more than 100%. This discrepancy or gap in the thermal performance of the building fabric is commonly referred to as the building fabric thermal '*performance gap*'. If the issues surrounding the '*performance gap*' are to be addressed, then it is not only imperative that the *in situ* measurements are accurate, but that the models that are used to predict the performance of the building fabric are also correct and accurate (Swan, Fitton, & Brown, 2015). Otherwise, there is a risk that one of the contributors to the '*performance gap*' is poor modelling predictions, resulting in a '*prediction gap*'.

Understanding the energy performance of buildings has become an important factor in design and retrofit. With a global aim of reducing the anthropological impact on climate change, recognising how buildings behave is crucial in the analysis of energy and fuel consumption; methods can then be developed in order to mitigate the enormous carbon footprint we have already established. Among international efforts targeting CO₂, the UK has pledged to achieve an 80% reduction in CO₂ emissions by 2050 against the 1990 baseline (DECC and HM Treasury, 2011). While commercial and industrial sectors are under scrutiny, the domestic sector accounts for 27% of total UK energy consumption (ECUK, 2015). It is in this area that a strong focus is being directed.

UK government legislation addresses decarbonisation by setting out requirements for all new domestic buildings; for instance the introduction of mandatory EPCs in the sale and rental of property, influencing the market towards the occupancy of more energy efficient buildings. Condensing boilers are also a necessity in new buildings (DECC and HM Treasury, 2010), which supplement a reduction in heating inefficiency. As dwellings in the UK have traditionally had very slow replacement cycles

and long physical lifetimes, attention must also be placed upon the existing housing stock, which is expected to make up at least 70% of the UK's total housing by 2050 (SDC, 2006). Realising how these buildings perform for the purpose of retrofitting can provide a pathway to reducing domestic energy consumption.

While the focus of this study will be based upon a typical UK home, it is important to recognise the presence of regulations on a broader scale and the incumbent issues found therein. Casals (2006) shows a good example of this by reviewing Directive 2002/91/CE, a European-wide effort to provide the following: a framework for calculating a building's integrated energy performance, a minimum energy performance requirements for new buildings, minimum energy performance requirements for large scale renovations, energy certification of buildings and a regular assessment of HVAC systems. Casals reveals a stark difference between EU members and their approach to building energy performance, identifying members such as Germany and Denmark who already impose strict measures on the energy performance of their buildings, whilst members such as Spain and France have a much more relaxed attitude to building energy performance. Realised with this directive, as has been found in the UK's approach to building energy regulations, is the limitation of enforcement for existing buildings, where the focus of energy performance assessment is reserved for new constructions. As has already been established, new constructions will make up a small proportion of the overall building stock over the coming decades, and so the scope for impact of this directive and implementation of any consequential energy regulations is particularly narrow.

1.1 The Performance Gap between Measurement & Modelling

Energy use in the home is dominated by space heating, which is the largest single end-use category in the domestic sector, accounting for just over 60% of all of the energy delivered to the existing housing stock in 2011 (Palmer and Cooper, 2013). A key approach in reducing energy consumption would, therefore, be to look at how heat is lost from the home. The U-value of an element is a measure of the rate of heat transfer through a particular construction. This factor is used to predict the quantity of plane element heat loss through the external elements of a building. U-values are typically calculated using surface composition and depend on a material's thickness and its thermal conductivity. Discrepancies are often noted in works that compare U-values measured *in situ* to their theoretical calculated equivalents (Bankvall, 1978; Lecompte, 1990, Doran, 2001; Hens *et al.*, 2001 & 2007; Doran and Carr, 2008). Consequently, these discrepancies result in a deviation between energy consumptions predicted by the models and the energy consumption measured *in situ*. Li *et al.* (2015) discuss the standard assumptions made when using U-values to evaluate the energy performance of buildings. They review work by Baker (2011) and Rye & Scott (2010), which propose that these U-values are often greatly overestimated. Baker suggests that the limitations associated with using theoretically calculated U-values are typically due to misinterpretation of the true building structure. Environmental factors, such as of wind, moisture and ventilation, coupled with the age of the structure, can also lead to the over-estimation of the U-value. Using a large sample of measured versus modelled U-values, Baker points out that buildings are typically able to perform better thermally *in situ* than predicted.

Further works compare the calculated U-value of a building element to its operational U-value, where the scale of discrepancy was found to vary across a wide range. Gaspar *et al.* (2016) for instance investigate the theoretical and actual thermal performance of three façades of different composition. In using both the average method and the dynamic method of ISO 9869-1:2014, they found that

results for two of their façades gave a reasonable difference between the theoretical and actual U-value (not greater than $\pm 5\%$ for the average method and not greater than $\pm 1\%$ for the dynamic method. A third façade tested however, demonstrated a difference of more than $\pm 20\%$ when using the average method, which was reduced to -9.6% when using the dynamic method. The larger discrepancy was largely accredited to poor environmental conditions, and although discrepancy was highlighted in each case, the remedy of applying the dynamic method over the average method is also highlighted in their work.

Evangelisti et al. (2015) had previously investigated the difference between calculated and measured U-values in a similar study. Again placing three different façade permutations under scrutiny, thermal transmittance was measured using the average method of ISO 9869. EN ISO 6946 was used to calculate the theoretical transmittance in each case and then used for comparison against the measured value. Their work revealed a staggering discrepancy, in particular for the older building of the three which demonstrated a difference of $+153\%$; this goes to show how older buildings could be earmarked for extensive retrofit based upon the assumptions of poor thermal performance, when in reality they massively outperform expectations. In a reverse of this, a different façade (and its composite layers) are found to perform considerably worse than predicted, with an average difference of -37% , thus revealing the benefit of overly pessimistic assumed U-values. Evangelisti et al. identify that a key issue in calculating the U-values for comparison is the non-invasive nature of their approach, in that they cannot identify homogeneity of their wall structures; assumptions of the façade composition therefore, could be incorrect. This demonstrates how – in the field – it is possible to be faced with the kind of problem where non-invasive measurement is required and where little or no information is available regarding wall structure and composition. Assumptions therefore have to be made when there is every possibility that the structure will differ physically (perhaps an additional material is inserted within), which could be detrimental to the results. In this case, modelling the older building from this study using the U-values calculated from the assumptions made would generate a grossly inaccurate review of the building's thermal performance.

To overcome the downfalls of non-invasive investigations of a building's envelope fabric, the 'destructive method' can be used in which a hole is bored from the outer envelope using a hollow drill; the wall composition is then analysed for accurate measurement of its thermal properties. Desogus et al. (2011) used such a technique in a comparison of non-invasive and destructive methods for analysing the thermal properties of a particular building's envelope. Their study discovered that although the benefit of knowing the full composition of the wall in question, greater uncertainty was attributed to the destructive method than with the non-invasive method, which used the average method of ISO 9869. The higher uncertainty was accredited to the inability to fully recognise the construction materials taken from the destructive method, with further difficulty arising when analysing material density and moisture content. An additional revelation from the study concerned the temperature gradient across the element. Desogus et al. chose to establish two distinct temperature differences of $7\text{ }^{\circ}\text{C}$ and $10\text{ }^{\circ}\text{C}$, finding that the higher temperature difference gave a lower uncertainty, and is most likely due to the increase in mono-directional heat flow.

Situational discrepancy in U-values can be considered in the modelling of existing retrofit buildings. Fitton (2013) looks at energy monitoring in retrofit projects, identifying a performance gap between the measured building performance and its modelled counterpart. High levels of discrepancies are seen in a significant proportion of dwellings when considering post-retrofit energy performance against the modelled predictions. Tronchin & Fabbri (2008) previously comment on this energy performance gap, accrediting inconsistencies to varying archetypes and unpredictability in Mediterranean climate conditions. As similar unpredictability is expected in the UK climate due to

global warming, this additional effect may be another cause for concern. Therefore, a high probability for misinterpreting the building fabric thermal performance in retrofit buildings can be expected, leading to an under or over estimation of its benefit.

The impacts of using overly pessimistic U-values can be far reaching, as discussed by Ahern, Norton, & Enright (2016); primarily that buildings are assigned an energy performance rating lower than they deserve and, potentially, buildings could be earmarked for costly and unnecessary retrofit. In their study, they investigate how buildings in Ireland are identified for performance assessment. Without additional data, the archetype and construction period alone are used to determine a default U-value. EPCs for these buildings are then automatically generated in software – DEAP (SEAI, 2012). Incorrect EPCs would lead to poor predictions of building performance, potentially acting as a trigger to instigate unneeded retrofit. The pessimistic rating of dwellings did in fact lead to a rapid period of retrofit in Ireland (INSHQ, 2002). The knock-on effect of this is that the validity of EPC ratings in Ireland has been questioned, and as a consequence Ahern, Norton and Enright propose that a reconsideration of EPCs is required. While the UK method of approach for EPC accreditation of RDSAP employs more data, such as structure, dimensions, heating system type, levels of insulation and fenestration/lighting types, it can be expected that the default U-values incorporated within the software, though likely to be more realistic, will still incline towards the pessimistic.

Housez et al. (2014) are among authors of works which explore the projected and actual energy performance of buildings post-retrofit. Their study looks at seven buildings situated in Austria which have received varying depths of retrofit in an attempt to improve on thermal performance and reduce energy consumption. Varied methods were found to be used in the preparation of the original energy performance certificates, which presented enough uncertainty to warrant the authors generating their own predictions for energy consumption. In all but one case, a huge discrepancy was found between the predicted energy consumption of each building vs the actual energy consumption of the same buildings (by up to a factor of 6). The authors, recognising this performance gap, sought to investigate the variance of three factors within their predictive calculations: localised weather profiles to more accurately model for the realistic weather conditions of the building; air change rates due to increased ventilation behaviour; and the maintenance of internal air temperatures, which were measured on site. By applying corrections to their predictions using these factors a much reduced performance gap was observed, and indeed a review of occupant behaviour did show that increased ventilation accounted for a majority of the original discrepancy.

While Housez et al. show how discrepancy in predicted performance of retrofit buildings can be effected by occupant behaviour and the presence of excess ventilation, Gupta & Gregg (2015) demonstrate how, by investigating two different buildings post-retrofit, discrepancies in the anticipated fabric performance contributes to a performance gap not only due to a varied thermal performance, but due to the occupant behaviour also. Their study took two archetypes into account – an older pre 1920's Victorian house and a more modern house. Discrepancies in the post-retrofit U-values were found on individual building elements (some higher than expected, some lower), however this meant the actual global building U-value met the as predicted value. In contrast to this, the actual U-values of the post-retrofit modern house were found to be higher than anticipated. In both cases however, the predicted (target) consumption of energy was found to be lower than the actual value. Some explanation as to why this is lies with the discrepancy of building fabric thermal values, however the impact of occupant behaviour (specifically the 'pre-bound' effect) is also explored – where occupants consume less energy for space heating than predicted pre-retrofit, anticipating established poor thermal performance and the inability of the building to retain heat. The two studies demonstrate how both occupant behaviour and inaccurate assumed U-values and air changes can lead to considerable

performance gaps and poor predictions of energy consumption with retrofit projects in particular; calling for some means of correction the predictive models, such as the steps taken by Housez et al.

These observations are summarised well by de Wilde (2014) in a review of the core reasons behind the performance gap and the implications thereof. Three main types of performance gap are identified in this work, which are predictions vs. measurement, machine learning vs. measurement, and prediction vs. displayed energy performance. Root causes of the gap are attributed to a large number of factors: misalignment of design for purpose and actual purpose, efficient yet complex devices, poor construction practice, a lack of constructive feedback, and general human behaviour such as the building culture and lack of client education on building performance. Conversely, de Wilde discusses an effort to try and close this gap; education of performance gap issues, improved monitoring and measurement techniques, adapting the construction culture, and the improvement of predictive models are all ways in which researchers are actively attempting to close the gap. This review supports the recommendation made in this paper that closer collaboration, not only between measurement and modelling communities, but also within construction is needed.

1.2 Model Calibration Techniques

The method of more accurately modelling the building fabric thermal performance of a building is designated as 'calibration' in this paper. Manfren, Aste, & Moshksar (2013) suggest that energy performance modelling is merely the simplification of complex physical procedures resulting from the periodic replenishment of energy. They note that as simulation tools become more sophisticated, additional system parameters may be modelled; however, as each parameter is attributed to a specific uncertainty, combined uncertainty grows with model intricacy. They go on to introduce regression techniques in order to improve accuracy in model predictions. In particular, Bayesian analysis is applied as a probabilistic method for anticipating uncertainty of inputs and outputs. The framework as developed by Kennedy & O'Hagan (2001) looks at likelihoods of subsequent probabilities in order to reduce these uncertainties. Bayesian analysis is supported as a calibration method in a number of other works, for example Heo *et al.* (2012) and Tian *et al.* (2014 & 2016).

A similar method for reducing uncertainties in energy performance modelling looks at deterministic calibration. Pan, Huang, & Wu (2007) use this method, refining their model by repeatedly varying factors of uncertainty until a closer match is ascertained. While this method demonstrates superiority in uncertainty modelling, it proves to be time consuming in nature.

An alternative calibration method involves manual adaptation of the created software model until the output metrics match those of the *in situ* measured data as seen in work by Marini *et al.* (2016). In their work, the Designbuilder software was used to develop a base model from a real domestic building located in Loughborough in the UK. Using a network of sensors, *in-situ* data was collected in order to alter their model over several stages of calibration. Weather data, HVAC operation, infiltration rates and heat flow were measured and used to replace the built-in data contained within the software and a comparison of energy consumption was made against the original model at each calibration step. The study found that significant reductions in prediction error gap from could be achieved by applying these calibrations. The remaining error is attributed to the model's inability to accurately replicate the true behaviour of the heating system, for example the dynamic behaviour of supply heat timing and the system's response to peaks in consumption.

Work comparing the process of both automatically calibrating a building energy model with that of manual calibration is explored by Chaudhary et al. (2016). In their study, numerous parameters within the model – classified under material, people, lighting, electrics and HVAC – are purposefully sabotaged; the authors are then tasked with regenerating an accurate model. A trade-off between accuracy of the final models and the time taken to achieve these models is noted. Importantly that the automatic process of calibration took much less time, but the manual process delivered a marginally better accuracy.

The work in this paper uses the manual adaptation method of calibration using the unique Energy House facility at the University of Salford. While techniques using automatic calibration with some function for optimisation are a more efficient way of closing the performance gap, as shown by Mihai and Zmeureanu (2013), the logical correction of input parameters more closely matches the model to the building with a successful reduction in the performance gap.

In this facility, a typical pre-1920's UK home has been built using largely reclaimed materials inside a climate controlled chamber. Details of this test facility are contained within Ji *et al.* (2014). Instead of simply measuring impact factors that might attribute variation in energy consumption, such as weather, temperature and system performance, it is proposed that these are all controlled. By creating a steady-state environment, it is possible to accurately measure important building characteristics such as U-value and infiltration rates and compare these to the steady-state default values that are used as input into the software models. In accordance with the work undertaken by Marini *et al.* (2016), the measured *in situ* data can then be used to modify the software model of the Energy House (EH) and thus produce a model that, in theory, more accurately reflects reality. By controlling all variables, this study also eliminates uncertainty in system performance as observed in the Marini study, moving for more accurate predictions of energy consumption. The intended outcome of this study therefore, will be a reduction of the '*prediction gap*', and ultimately in the building fabric thermal '*performance gap*'.

2. Method

A measured building survey of the Energy House was carried out to produce a scaled representation of its structural elements. Floorplans of each level of the building are shown in figure 1a and 1b, with an external shot of the building in figure 2. Using the floorplans from this survey, an accurate model was built in Designbuilder. This software performs dynamic energy simulations and is a front end user interface for EnergyPlus. It has been chosen for this study due to its prolific use in simulating for energy performance analysis and as it was developed specifically for this purpose by the US Department of Energy (DOE) (Crawley *et al.*, 2000).

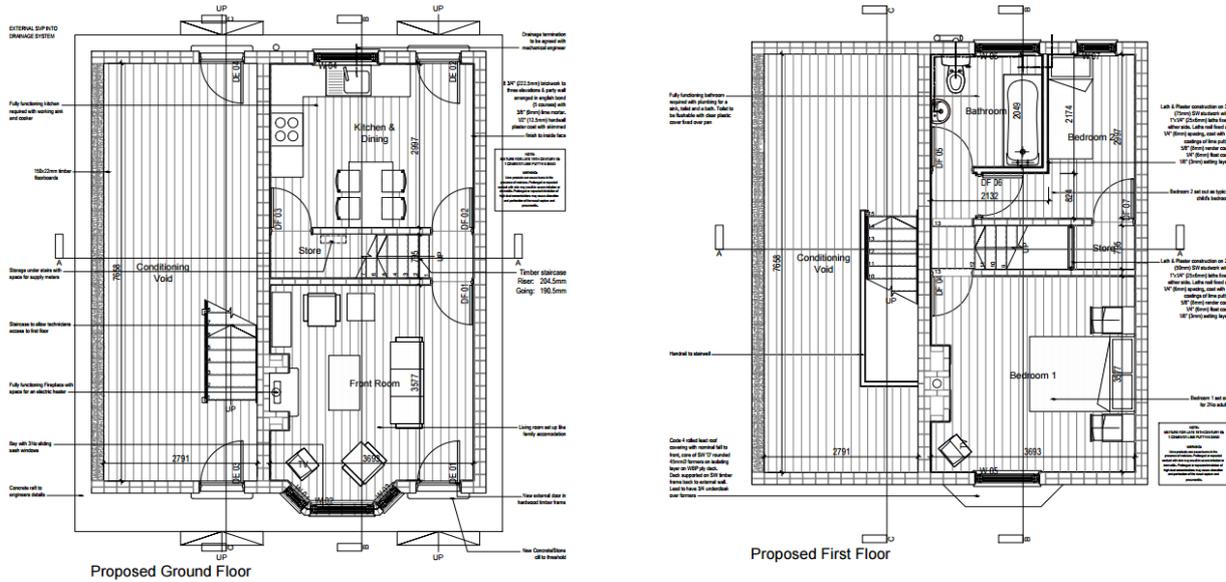


Figure 1a –Energy House Ground Floor Plans, Figure 1b – Energy House First Floor Plans



Figure 2 – The Energy House Facility at the University of Salford

The model calibration work in this study is accomplished in three key stages. The first stage is to build the ‘standard model’ of the energy house using a measured survey and the standard values offered by the software. Next is to perform building fabric tests to obtain *in situ* data for U-values and air permeability. Finally, the standard model will be calibrated over a number of stages, by adjusting input parameters to match the *in situ* data. Models at each calibration stage will provide an indication of the global Heat Transfer Coefficient (HTC) for the building. Further comparison of these HTC data would then allow for the quantification the prediction gap in each case.

2.1 The Building Fabric Tests

In order to be able to characterise the performance of the Energy House building fabric, a series of building fabric tests were undertaken on the Energy House. These tests comprised the following:

- Air pressure tests.
- Electric coheating test.
- Heat flux density measurements.

Table 1 lists the main technical information for each instrument used in these tests.

2.1.1 Air pressure tests

An air pressure test was undertaken on the Energy House both prior to and immediately following completion of the electric coheating test as laid out in section 2.1.2. The mean air leakage rate obtained from both tests was then averaged to obtain a mean air leakage rate over the coheating test period. These tests were undertaken to determine the airtightness of the Energy House and to establish whether the airtightness of the house varied over the duration of the electric coheating test. All of the pressure tests were performed in accordance with ATTMA Technical Standard L1 testing protocol for dwellings (ATTMA, 2010) and were undertaken using an Energy Conservatory Model 3 Blower Door with a DG700 dual-channel pressure gauge. During the tests, both a pressurisation and a depressurisation test were undertaken and the results were averaged. This procedure is consistent with the advice given in CIBSE TM 23 (CIBSE, 2000).

2.1.2 An Electric Coheating Test

An electric coheating test is an *in situ* test method that is capable of measuring the aggregate whole dwelling heat loss (both transmission and infiltration) of an unoccupied dwelling. It is described as a 'quasi-steady-state' or a 'quasi-static' test method as the internal temperature within the test dwelling is artificially held at a predetermined static state, whilst the external conditions vary dynamically in response to the external climatic conditions. The test method was originally devised in North America in the late 1970's to determine the net efficiency of domestic heating and cooling systems (see Socolow, 1977; Sonderegger & Modera, 1979 and Sonderegger *et al.*, 1979). At this time, the method involved the use of the buildings installed space heating system, in addition to supplementary portable electric resistance heaters, to simultaneously heat the building. Hence, the use of the term "*coheating*".

In the UK, the earliest documented use of the test method was in the mid 1980's, where a modified version of the "coheating" test method was used to measure the aggregate heat loss from a dwelling (see Siviour, 1985 and Everett, 1985). The modified version removed the requirement for simultaneous "coheating" of the dwelling, and instead, used electric resistance point heaters only to measure the aggregate heat loss of a dwelling. Following sporadic use of the test method in the 1990's, it wasn't until the Stamford Brook project in mid-2000 (Wingfield *et al.*, 2011) that the test method was applied in the field in any significant number of instances, During this project, the test

method was used by Leeds Metropolitan University (now Leeds Beckett University) to not only identify a significant performance gap in new build housing, but it was also used to help identify and quantify the party wall bypass heat loss mechanism (see Lowe *et al.*, 2007). The method has since undergone further refinement and development by Leeds Metropolitan University (for instance see Wingfield *et al.*, 2010) and in 2010 it became recognised as an established test method in the UK when it was incorporated within the Post Construction and Initial Occupation studies undertaken under the Technology Strategy Boards (now Innovate UK's) Building Performance Evaluation Programme (Technology Strategy Board, 2010).

The latest version of the test method described by Johnston *et al.* (2013) involves using strategically positioned thermostatically controlled electric resistance point heaters to artificially heat the interior spaces of an unoccupied building to a mean elevated and homogeneous internal temperature. Electrically powered air circulation fans are also strategically positioned within the building to ensure that the internal air is adequately mixed and to minimise any potential air stratification. Once the building is in thermal equilibrium internally and the thermal mass within the building is fully charged, the mean elevated temperature is maintained constant for a specified period of time, typically between 7 and 21 days. The daily heat input to the building (in Watts) can then be established by measuring the total amount of electrical energy that is required to maintain the artificially elevated mean internal temperature. If the daily heat input in Watts is plotted against the daily difference in temperature (ΔT) in Kelvin, then the uncorrected raw heat transfer coefficient (HTC) for the building can be determined in W/K. This raw uncorrected figure can then be corrected to account for external environmental effects, such as solar radiation, using linear regression analysis techniques. Further details of the analysis techniques and the test method can be found within Johnston *et al.* (2013).

An electric coheating test was carried out on the Energy House in order to determine the *in situ* aggregate (both fabric and ventilation) heat loss of the dwelling. This test was undertaken in accordance with the latest version of the Leeds Beckett University test method (see Johnston *et al.*, 2013). However, as the Energy House is located within an environmental chamber, so is not subjected to any changes in external environmental conditions, the temperature set-point of the test chamber HVAC was reduced on three separate occasions from 15°C to 10°C and then to 5°C to determine whether the comparison would be influenced by changes in the temperature differential between the internal and external (chamber) environments (ΔT). Internal temperatures within the Energy House were maintained at a constant mean elevated temperature of 25°C throughout the electric coheating test period, by thermostatically controlled electric resistance point heaters. These heaters were installed in the living room, kitchen, bedroom 1, bedroom 2 and bathroom of the dwelling and air circulation fans were deployed to minimise any air stratification.

Inter dwelling heat transfer between the Energy House and the adjacent space (the conditioning void) was mitigated against by installing additional thermostatically controlled heaters and air circulation fans in this space. In addition, the electric resistance heaters within the conditioning void were also set to the same mean elevated temperature as the Energy House. The test was carried out over a total of 9 days, with data being recorded at 10 minute intervals.

2.1.3 Heat Flux Density Measurements

A number of Hukseflux HFP-01 heat flux plates (HFPs) were attached to the internal surfaces of the external wall, party walls, windows and doors of the Energy House to enable the heat flux density

through each of these elements of the fabric to be measured. The heat flux density (measured in W/m^2) were then used in conjunction with corresponding internal/external air temperature measurements to calculate an *in situ* air-to-air U-value for each HFP location using the average method contained within ISO 9869:1994 (ISO, 1994). The positioning of the heat flux plates on the fabric was informed by a thermographic survey.

Measurement of internal/external air temperatures and heat flow were taken over a period of 7 days. Heat was administered into each room with electric radiators separate to the building HVAC, using a set point of 25°C; the external chamber temperature was maintained at 5°C throughout using a HVAC system. Data is used from the period where steady state conditions had been established.

Table 1 – Main technical information for equipment used in air pressure, coheating and heat flux tests

Equipment	Model	Measurement Tolerance
Blower Door Fan	Energy Conservatory Model 3 Fan	-
Pressure / Flow Gauge	Energy Conservatory DG-700	± 1%
Temperature Sensor	Shielded 4-wire PT100 RTD	± 0.1 °C
kWh meter	Elster A100C single phase meter	± 1%
Heat Flux Plates	Hukseflux HFP-01	± 3%
Data Logger	dataTaker DT80	± 0.1%

2.2 Designbuilder Model

The Energy House model was built and developed in Designbuilder over a number of stages. An initial model was built based upon a measured survey of the facility; floorplans of which have been given in figure 1. Building layers within the model use standards provided by CIBSE Guide A (CIBSE, 2006) in order to calculate U-values of each element according to the method set out in BS EN ISO 6946:2007 (British Standards Institution, 2007). A rendered model of the house is shown in figure 3. It should be noted that as with the Standard Assessment Procedure (SAP), Designbuilder follows the standard of ISO 13790:2008 (International Organisation for Standardization, 2008) for its calculation of energy for space heating and cooling. It also agrees with EN 15265:2007 (CEN, 2007). Air permeability is determined by infiltration of the exposed surface, considered where a pressure difference of 50 Pa is observed between the inside and outside. Hourly air change at the standard pressure is calculated in Designbuilder using the method in BS EN 12831:2003 (E) (British Standards Institute, 2003).



Figure 3 – The Energy House model rendered in Designbuilder

The first step in conditioning the model was to remove occupancy and scheduling, factors automatically assumed within the model. Each habitable zone was declared as having no activity, with occupancy density at zero and occupancy scheduling set to off. Lighting (and the associated gains), which is controlled by occupancy schedules within the model, was also changed to be permanently off.

Next, the assigned weather data is modified to reflect the conditions monitored within the Energy House chamber. All variables, such as precipitation and solar radiation were reduced to zero. Dry bulb temperatures were fixed at three separate temperatures to reflect the different ΔT 's experienced during the electric coheating test. During the test, the chamber temperature was measured at three locations of each face of the building; an average of these was used for the external temperature in the model. The recorded average wind speed in the chamber was 0.35 m s^{-1} and is also incorporated within the simulation.

The standard model was finally conditioned by matching set points for the coheating test in each heated zone. It should be noted that the adjoining building (conditioning void space) was heated at a constant 25°C , therefore all connecting surfaces in the model were considered adiabatic. This standard model was then simulated for direct comparison to the data obtained in the coheating test.

2.3 Model Calibration

The model was calibrated using a 2 stage process. Firstly the air permeability measured during the *in-situ* test was inserted into the model. The HTC value calculated using data from this model is then compared to the measured value and the previously modelled value. Secondly the standard U-values contained within the model were substituted for those measured *in-situ*. A final comparison is made between the HTC found in this model to those previously, where HTC differences are analysed.

3. Results and Discussion

Data will be initially compared between the standard model and the coheating test. To understand any gaps in performance, the standard values of air permeability and U-value will then be compared to those measured in-situ. After adjusting these factors, the measured data is to be compared with the calibrated model and implications discussed.

3.1 Coheating Test Results

Figure 4 illustrates the measured power input into the Energy House and the mean air temperatures within the Energy House, conditioning void, and chamber throughout the coheating test. The internal temperature of the Energy House and conditioning void remained reasonably constant throughout most of the coheating test period. The reduction in internal temperature of the Energy House following the final change in chamber temperature was due to a temporary loss of power input to the kitchen. The ΔT between the Energy House and conditioning void was typically $< 0.5K$ throughout the test period, which allows the party wall to be considered an adiabatic boundary in this instance. The HVAC cycle was responsible for the variation in chamber temperature measured. The measured power input into the Energy House stabilised towards the end of each phase, as can be seen in the moving average of power in figure 4; this suggests that the Energy House had reached steady-state.

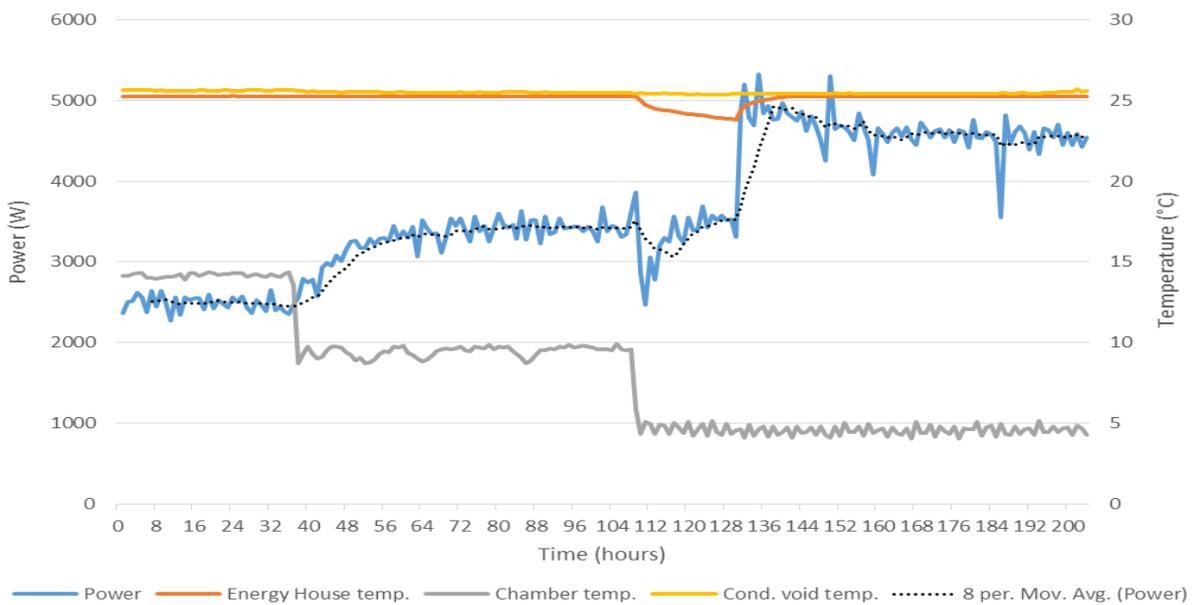


Figure 4 – Hourly mean power input and temperatures measured throughout the coheating test (colour)

The Energy House HTC measurement during each test phase was calculated by dividing the mean power input by the ΔT measured during the final eight hours of each coheating test phase. These periods were selected as the measurements suggested that the Energy House was at a steady-state for this period in each phase. The eight hour duration also smoothed out variation in power input caused by the thermostatic controllers and variation in ΔT caused by the HVAC system. Table 2 provides the ΔT , power input, and HTC for each test phase.

Table 2 – Measured ΔT and power input and HTC in each coheating test phase

ΔT (K)	Power (W)	HTC ($W K^{-1}$)
11.2	2447	219.5 (± 3.3)
15.6	3432	219.5 (± 3.3)
20.7	4538	219.7 (± 1.8)
Mean		219.6

There was no statistically significant difference between the HTC measured during each coheating test phase. The mean of the three measurements is $219.6 W K^{-1}$. Figure 5 shows a perfect correlation between ΔT and power input for the three phases of the coheating test. The HTC derived from the gradient of $219.9 (\pm 0.2) W K^{-1}$ is in very good agreement with the mean from each test phase. If the regression line is forced through the origin (as is typical in coheating analysis) then the slope of the regression line gives a HTC of $219.6 (\pm 0.1) W K^{-1}$.

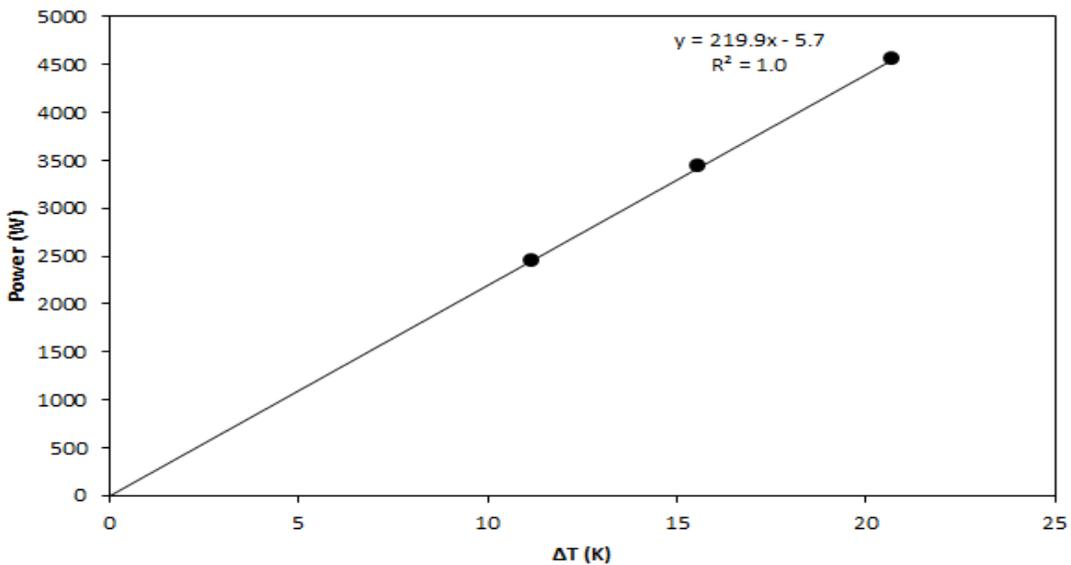


Figure 5 – Regression plot for the measured coheating data

From Table 3 it can be seen that the regression statistics show that there is a high degree of confidence with the HTC derived from the regression analysis. Note that in table 2, B indicates the unstandardized regression coefficients, Std Error is the standard error associated with each unstandardized regression coefficient. The constant of $-5.7 W$ suggests that there was a small power input into the Energy House that was not measured (e.g. party wall heat gain, the Energy House's internal sensors) or a systematic error in the ΔT measurement. However, it must be noted that the constant is not statistically significant. The good agreement between the HTC derived from the two calculation methods provides confidence that the coheating test accurately measured the Energy

House HTC. As only three data points present themselves for each regression, the most accurate way of interpreting the HTC value would be to take the mean. The HTC of 219.6 W K⁻¹ will therefore be used for the comparison with modelled data.

Table 3 – Regression statistics for the measured coheating data

	Coefficients		95% Confidence Interval for B	
	B	Std Error	Lower Bound	Upper Bound
Constant (W)	-5.7	3.3	-50	36.7
Power / ΔT (W K ⁻¹)	219.9	0.2	217.3	222.5

3.2 Standard Model

The standard model, i.e. the model using the software's standard values, is compared to the measured data in order to assess how well it imitates the true thermodynamic behaviour of the Energy House. Table 4 compares the calculated HTC values for both measured and modelled data at each ΔT, with an associated difference in each case.

Table 4 - Measured and modelled HTC values for heated zones at ΔTs of 11.2K, 15.6K and 20.7K.

ΔT (K)	Measured HTC (W K ⁻¹)	Modelled HTC (W K ⁻¹)	Difference
11.2	219.5 (±3.3)	274.1	23.2%
15.6	219.5 (±3.3)	257.7	17.4%
20.7	219.7 (±1.8)	248.8	13.4%

The coheating test, as stated in the guidance by Johnston *et al.* (2013), benefits from large differences in temperature, specifying a minimum ΔT of 10 K in order to ensure monodirectional heat flow. The data shows an observable decrease in difference with an increase in ΔT, supporting the benefit of a larger difference in temperature. This uncalibrated model produces a mean HTC value of 260.2 W K⁻¹, differing from the measured mean by 18.5%.

3.3 In-Situ Data Correction

It is important to show how the performance gap in the modelling and measured data arises so that it may be quantified and resolved. Firstly, an assumed air permeability value of 16 m³ h⁻¹ m⁻² is corrected with that found *in situ*. As discussed in section 2.1.1, the air permeability test was performed both before and after the coheating test to obtain a mean value of pressurisation and depressurisation.

The values from the test are given in Table 5. $13.95 \text{ m}^3 \text{ h}^{-1} \text{ m}^{-2}$ is used in place of the model's standard value.

Table 5 – In-Situ air permeability results

Air Permeability ($\text{m}^3 \text{ h}^{-1} \text{ m}^{-2}$ @50 Pa)		
Pressurisation	Depressurisation	Mean
14.04	13.85	13.95

Literature proposes that the application of pessimistic U-values plays a large part in erroneous results. For this study, U-values were measured at a number of locations on various surface types within the Energy House, then compared to the standard value in the model; these are shown in Table 6.

Table 6 - Comparison between calculated and measured U-values, with percentage difference

Surface	Standard U-value ($\text{W m}^{-2} \text{ K}^{-1}$) Using (CIBSE, 2006)	Measured U-value ($\text{W m}^{-2} \text{ K}^{-1}$)	Difference between U-values
External Wall: Living Room	2.243	1.596 (± 0.086)	28.85%
External Wall: Kitchen	2.243	1.530 (± 0.087)	31.79%
External Wall: Bedroom 2	2.243	1.567 (± 0.091)	30.14%
Ceiling: Bedroom 1	0.400	0.462 (± 0.026)	15.50%
Window: Living Room	3.746	3.835 (± 0.257)	-2.376%
Window: Kitchen	3.746	3.670 (± 0.246)	2.029%
Door: Living Room	2.995	2.697 (± 0.178)	9.950%

Clear discrepancies are seen for each of the surfaces in the table. The greatest difference however, is seen for external wall U-values with only minor differences reported for the windows and door. Accurate representation of the building's envelope U-value is crucial for obtaining both a value for the HTC and subsequently the energy consumption for heating. External walls dominate the surface area of the building; the window area to wall ratio for the Energy House is 0.13. Given this, it is important that the reflective properties in the model are accurate. With significant difference in U-value for the walls, considerable differences would be expected in any standard model of the Energy House.

3.4 A Calibrated Model

Dynamic energy simulations were performed with two stages of calibration: modifications of air permeability and U-value. Table 7 now compares the measured HTC values with the newly obtained values from each step of calibration. Associated differences are also displayed.

Table 7 – Measured vs modelled HTC values for comparison of calibration stages

ΔT (K)	HTC (W/K) [Initial Model]	HTC (W/K) [Modified Air Permeability]	HTC (W/K) [Modified Air Permeability & U-value]	HTC (W/K) [Measured]
11.2	274.1	271.4	238.4	219.5
15.6	257.7	254.5	222.4	219.5
20.7	248.8	244.9	214.0	219.7

Observed with each stage in calibration is a definite reduction in the differences between measured and modelled HTC values. Though it can be assumed from these data that identifying the correct air permeability for modelling energy performance only goes a small way reduce inaccuracy; correcting the U-value is a crucial factor for improving it.

In correcting the model, all HTC values fall within a closer range of the accepted mean measured HTC of 219.6 W K^{-1} . By adjusting the air permeability of the model a mean HTC value of 256.9 W K^{-1} is found, with a difference of 17.0% from the mean measured value. With the additional adjustment of the U-value of the external elements of the model a mean HTC value of 224.9 W K^{-1} is found, reducing the difference between measured and modelled mean HTC value to 2.4%.

It is noted that the mean measured HTC values demonstrate considerable similarity and that this is not seen for the modelled data in each stage of calibration. Additional analysis shows that some additional heat flux is present in all cases. To further investigate as to why this is, the model was simulated with a ΔT of 0, where it was assumed that there would be zero net heat flux. Some excess was discovered in the simulation results and was traced back to heat flow through partition elements of unconditioned spaces. Successive reductions in chamber temperature were used experimentally to vary the ΔT of the system for measured HTC data. This means that at each successive ΔT , unconditioned spaces were pre-heated to a higher than expected temperature before steady state was achieved, a factor not afforded to the individual simulations. This resulted in an unanticipated quantity of heat exchange within the overall system.

The effects of this led to a higher than expected power at larger ΔT s and a lower than expected power at smaller ΔT s. Although this discrepancy has been identified as a potential problem when interpreting individual simulation results, this problem is mitigated when repeating simulations at various ΔT s and by taking the mean HTC value.

4. Conclusion

By building a model to typical standards, a performance gap has been shown to exist between measured and modelled energy performance. From a standard model it has been possible to demonstrate that, through several stages of *in situ* measurement calibration, it is possible to close the performance gap between measured and modelled data. Substitution of in-situ measurements of air permeability and U-value have been used to reduce this gap considerably. A discrepancy in the overall heat flux for unconditioned spaces has been identified when comparing simulated results against those from the coheating test, however effects of this have been mitigated by taking the mean HTC value of a number of simulations. A reduction in the difference between measured and modelled HTC value of 18.5% to 2.4% has been achieved in this study.

The limiting factors of this study are presented by the fact that measured data is obtained under controlled conditions. This means any complexity offered by weather effects, moisture, ventilation etc. and occupancy, which exist in the field, do not exist here. Evidence of these limiting factors are visible when comparing standard U-values with in-situ measurements; considerable variance is seen for the walls of the building, with conditions necessitating a dry, static environment. While this may present limitations, it has been possible to collect extremely accurate measurement data within the facility; this led to the development of a robust model which accurately depicts the Salford Energy House.

Returning to the initial problem of the performance gap, it is clear that the differences between what is measured and what is modelled are at the heart of the problem. However, while some of this can be attributable to defects in construction, some may be attributed to both the modelling and measurement process. This study has identified a greater need for the measurement and modelling communities to better engage to ensure that performance gap is more effectively explored as a problem.

Funding

This research was funded by the School of Built Environment at the University of Salford.

References

- Ahern, C., Norton, B., & Enright, B. (2016). The Statistical Relevance and Effect of Assuming Pessimistic Default Overall Thermal Transmittance Coefficients on Dwelling Energy Performance Certification Quality in Ireland. *Energy and Buildings*, 268-278.
- ATTMA (2010) ATTMA Technical Standard L1. Measuring the Air Permeability of Building Envelopes (Dwellings). October 2010 Issue. Northampton, UK, Air Tightness Testing and Measurement Association.
- Baker, P. (2011). *U-values and Traditional Buildings*. Glasgow: Historic Scotland Conservation Group.
- Bankvall, C.G. (1978) Forced Convection: Practical Thermal Conductivity in an Insulated Structure Under the Influence of Workmanship and Wind. ASTM STP 660, pp. 409-425.
- British Standards Institution. (2007). BS EN ISO 6946:2007 Building Components and Building Elements - Thermal Resistance and Thermal Transmittance - Calculation Method.

- British Standards Institution. (2003). BS EN 12831: Heating Systems in Buildings - Method for Calculation of the Design Heat Load.
- Casals, X.G., 2006. Analysis of building energy regulation and certification in Europe: Their role, limitations and differences. *Energy and Buildings*, 38(5), pp.381-392.
- CEN. (2007). 15265-2007. *Thermal Performance of Buildings - Calculation of Energy Needs for Space Heating and Cooling using Dynamic Methods - General Criteria and Validation Methods*. European Committee for Standardisation.
- Chaudhary, G., New, J., Sanyal, J., Im, P., O'Neill, Z., Garg, V., 2016. Evaluation of "Autotune" calibration against manual calibration of building energy models. *Applied Energy*, 182, pp. 115-134
- CIBSE. (2006). *Guide A: Environmental Design*. London: The Chartered Institution of Building Services Engineers.
- CIBSE (2000) Testing buildings for air leakage: CIBSE Technical Memoranda TM23: 2000. London, The Chartered Institute of Building Services Engineers.
- Crawley DB, Pedersen CO, Lawrie LK, Winkelmann FC. EnergyPlus: Energy Simulation Program. *ASHRAE Journal* 2000; 42: 49–56.
- de Wilde, P., 2014. The gap between predicted and measured energy performance of buildings: A framework for investigation. *Automation in Construction*, 41, pp. 40-49.
- DECC and HM Treasury. (2010). *Valuation of Energy Use and Greenhouse Gas (GHG) Emissions*. London: Department of Energy & Climate Change & HM Treasury.
- DECC and HM Treasury. (2011). *The Carbon Plan: Delivering our Low Carbon Future*. London: HM Government.
- Desogus, G., Mura, S. and Ricciu, R., 2011. Comparing different approaches to in situ measurement of building components thermal resistance. *Energy and Buildings*, 43(10), pp.2613-2620.
- Directive 2002/91/CE of the European Parliament and of the Council of 16 December 2002 on the energy performance of buildings, 2002.
- Doran, S. (2001) Field Investigations of the Thermal Performance of Construction Elements As Built. BRE Client Report No. 78132. A DETR Framework Project Report. East Kilbride, UK, Building Research Establishment (BRE).
- Doran, S. and Carr, B. (2008) Thermal Transmittance of Walls of Dwellings Before and After Application of Cavity Wall Insulation. A Report Prepared by BRE Scotland for the Energy Saving Trust, Report Number 222077, East Kilbride, UK, BRE Scotland.
- ECUK. (2015). *Energy Consumption in the UK (2015)*. Department of Energy & Climate Change.
- Evangelisti, L., Guattari, C., Gori, P. and Vollaro, R.D.L., 2015. In situ thermal transmittance measurements for investigating differences between wall models and actual building performance. *Sustainability*, 7(8), pp.10388-10398.
- Everett, R. (1985) Rapid Thermal Calibration of Houses. Technical Report ERG 055 for the Science and Engineering Research Council, Milton Keynes, UK.
- Fitton, R. (2013). Energy Monitoring in Retrofit Projects: Strategies, Tools and Practices. In W. Swan, & P. Brown, *Retrofitting the Built Environment* (pp. 141 - 153). Manchester: John Wiley & Sons.

- Gaspar, K., Casals, M. and Gangoellés, M., 2016. A comparison of standardized calculation methods for in situ measurements of façades U-value. *Energy and Buildings*, 130, pp.592-599.
- Gupta, R. and Gregg, M., 2015. Do Deep Low Carbon Retrofits Actually Work?. *Energy Procedia*, 78, pp.919-924.
- Hens, H. Janssens, A. and Depraetere, W. (2001) Hygrothermal Performance of Masonry Cavity Walls with Very Low U-factor: A Test House Evaluation. Proceedings of the Performances of Envelopes of Whole Buildings VIII Conference, Clearwater Beach, Florida, pp. 1-13.
- Hens, H. Janssens, A. Depraetere, W. Carmeliet, J. and Lecompte, J. (2007) Brick Cavity Walls: A Performance Analysis Based on Measurements and Simulations, *Journal of Building Physics*, Vol. 31, No. 2, pp. 95-124.
- Heo, Y., Choudhary, R., & Augenbroe, G. A. (2012). Calibration of Building Energy Models for Retrofit Analysis under Uncertainty. *Energy and Buildings*, 550-560.
- Housez, P.P., Pont, U. and Mahdavi, A., 2014. A comparison of projected and actual energy performance of buildings after thermal retrofit measures. *Journal of Building Physics*, 38(2), pp.138-155.
- INSHQ, Irish National Housing Survey of Ireland, in: E.a.S.R. Institute (Ed.), 2001–2002.
- ISO (2008). 13790: 2008 . *Energy Performance of Buildings - Calculation of Energy Use for Space Heating and Cooling*. Geneva, Switzerland: International Organization for Standardization.
- ISO (1994) ISO 9869: Thermal insulation – Building elements - In situ measurement of thermal resistance and thermal transmittance. Geneva, Switzerland, International Organization for Standardisation.
- Ji, Y., Fitton, R., Swan, W., & Webster, P. (2014). Assessing Overheating of the UK Existing Dwellings - A Case Study of Replica Victorian End Terrace House. *Building and Environment*, 1-11.
- Johnston, D. Miles-Shenton, D. Farmer, D. and Wingfield, J. (2013) Whole House Heat Loss Test Method (Coheating). June 2013 [Internet] Leeds, Leeds Metropolitan University. Available from: <www.leedsbeckett.ac.uk/as/cebe/projects/cebe_coheating_test_method_june2013.pdf> [Accessed 22nd September 2016].
- Kennedy, M. C., & O'Hagan, A. (2001). Bayesian Calibration of Computer Models. *Journal of the Royal Statistical Society*, 425-464.
- Lecompte, J. (1990) Energy – Insulation: The Influence of Natural Convection on the Thermal Quality of Insulated Cavity Construction, Results of Experimental Research and Computer Simulation. *Building Research and Practice*, Number 6, pp. 349-354.
- Li, F. G., Smith, A. Z., Biddulph, P., Hamilton, I. G., Lowe, R., Mavrogianni, A., . . . Oreszczyn, T. (2015). Solid-Wall U-values: Heat Flux Measurements Compared with Standard Assumption. *Building Research & Information*, 238-252.
- Lowe, R., Wingfield, J. Bell, M. and Bell, J. (2007) Evidence for heat losses via party wall cavities in masonry construction. *Building Services Engineering Research and Technology*, 28 (2) pp. 161-181.
- Manfredi, M., Aste, N., & Moshksar, R. (2013). Calibration and Uncertainty Analysis for Computer Models - A Meta-Model based Approach for Integrated Building Energy Simulation. *Applied Energy*, 627-641.
- Marini, D., He, C., Buswell, R., Hopfe, C., & Crawley, D. (2016). Modelling and Calibration of a Domestic Building using High-Resolution Monitoring Data. *Proceedings of 2016 3rd Conference of IBPSA-England*:

Building Simulation and Optimization (BSO16). Newcastle: International Building Performance Simulation Association (IBPSA).

- Mihai, A. and Zmeureanu, R., 2013, August. Calibration of an energy model of a new research center building. In *Proceedings of the 13th Conference of International Building Performance Simulation Association, Chambéry, France* (pp. 25-28).
- Palmer, J. and Cooper, I. (2013) United Kingdom housing energy fact file 2013. A report prepared under contract to DECC by Cambridge Architectural Research, Eclipse Research Consultants and Cambridge Energy. London, UK, Department of Energy & Climate Change.
- Pan, Y., Huang, Z., & Wu, G. (2007). Calibrated Building Energy Simulation and its Application in a High-Rise Commercial Building in Shanghai. *Energy and Buildings*, 651-657.
- Rye, C., & Scott, C. (2010). *The SPAB Research Report 1. U-value Report*. Society for the Protection of Ancient Buildings.
- SDC. (2006). *'Stock Take': Delivering Improvements in Existing Housing*. Sustainable Development Commission.
- SEAI, Sustainable Energy Authority of Ireland, Irish Official Method for Calculating and Rating Energy Performance of Dwellings, 2012
- Siviour, J. (1985) Experimental Thermal Calibration of Houses. In: Everett R Rapid Thermal Calibration of Houses. Technical Report ERG 055 for the Science and Engineering Research Council, Milton Keynes, UK.
- Socolow, R. H. (1977) The Twin Rivers Program on Energy Conservation in Housing: Highlights and Conclusions. *Energy and Buildings*, 1 (1977/78) 207 – 242.
- Sonderegger, R. C. Condon, P. E. and Modera, M. P. (1979) In-situ Measurements of Residential Energy Performance Using Electric Co-heating. *ASHRAE Transactions*, 1979; Vol. 86 (I).
- Sonderegger, R. C. and Modera, M. P. (1979) Electric Co-Heating: A Method for Evaluating Seasonal Heating Efficiencies and Heat Loss Rates in Dwellings, Presented at the Second International CIB Symposium on Energy Conservation in the Built Environment, Copenhagen, Denmark, May 28 - June 1, LBL-8949, 1979.
- Swan, W., Fitton, R., & Brown, P. (2015). A UK Practitioner View of Domestic Energy Performance Measurement. *Institution of Civil Engineering*, 1-9.
- The Technology Strategy Board (2010) Building Performance Evaluation, Domestic Buildings - Guidance for Project Execution. Swindon, The Technology Strategy Board.
- Tian, W., Wang, Q., Song, J., & Wei, S. (2014). Calibrating Dynamic Building Energy Models using Regression Model and Bayesian Analysis in Building Retrofit Projects. *International Building Performance Simulation Association*. Ottawa: IBPSA.
- Tian, W., Yang, S., Li, Z., Wei, S., Pan, W., & Liu, Y. (2016). Identifying Informative Energy Data in Bayesian Calibration of Building Energy Models. *Energy and Buildings*, 119, 363-376.
- Tronchin, L., & Fabbri, K. (2008). Energy Performance Building Evaluation in Mediterranean Countries: Comparison between Software Simulations and Operating Rating Simulation. *Energy and Buildings*, 1176-1187.

Wingfield, J. Bell, M. Miles-Shenton, D., South, T. and Lowe, R (2011) Evaluating the impact of an enhanced energy performance standard on load-bearing masonry domestic construction - Understanding the gap between designed and real performance: lessons from Stamford Brook. London, DLCG Publications.