# Measuring and modelling retrofit fabric performance in solid wall conjoined dwellings

James Parker[1]*, David Farmer[1], David Johnston[1], Martin Fletcher[1], Felix Thomas[1], Christopher Gorse[1] and Steven Stenlund[2]

(1) Leeds Beckett University, Leeds, UK

(2) BRE, Watford, UK

**Corresponding author*

## ABSTRACT

There remains a significant number of occupied and uninsulated solid wall dwellings in the UK. Deep retrofit is often required for these buildings to become energy efficient but it is difficult to determine how these buildings will respond to retrofit without a detailed understanding of their fabric thermal performance Greater certainty can however be achieved by combining theoretical models and practical field tests, prior to the design of retrofit programmes. This type of approach can then be used to inform and optimize the design of retrofit interventions. This paper presents results from a series of *in situ* fabric performance tests undertaken on two no-fines concrete, conjoined dwellings pre- and post-retrofit and demonstrates how empirical data can be used to inform and calibrate the thermal performance of dynamic simulation models (DSMs). This is a particularly pragmatic calibration method as it eliminates the need for actual weather data, which is expensive and prohibitive to collect and collate. The DSM inputs and outputs were compared with those obtained from Standard Assessment Procedure (SAP) calculations. The results illustrate how the fabric performance of no-fines concrete can vary between similar house types within the same development. This research also validates the effectiveness of the calibration methodology that uses the whole house Heat Transfer Coefficient (HTC) as the qualifying metric. Furthermore, results also emphasize the importance of appropriately characterizing the physical properties of existing buildings before designing retrofit strategies. This paper contributes to the growing knowledge base concerned with the energy performance gap. In this instance, SAP predicts higher absolute savings then measured in situ which is problematic when assessing the financial viability of retrofits.

**Keywords:** Calibration, Coheating, Domestic, Retrofit, Thermal Modelling

**Highlights:**

- Calibrated models used to improve accuracy of predicted absolute savings from deep retrofit solutions.
- Calibration method using *in situ* test results validated against pre- and post-retrofit measurements.
- *In situ* test results presented for fabric performance of solid wall no-fines concrete dwellings and for traditional and innovative external wall insulation systems.
- Co-pressurisation test results shown to have a significant impact on modelled energy performance.

## 1. INTRODUCTION

In the UK, most dwellings have long physical lifetimes and there is a slow stock turnover, with estimates suggesting that the number of dwellings in the UK increases by less than 1% per year [1]. Consequently, the vast majority of the country's current domestic stock of circa twenty-eight million dwellings [2] are existing rather than new build, a trend that will continue for the foreseeable future. Therefore, it has been projected that most of the existing housing stock will remain in use by the year 2050 [3-6]. Dwellings also account for approximately 30% of Carbon Dioxide ($CO_2$) emissions in the UK [7]. Therefore, the low-carbon retrofit market in the UK is significant and is likely to remain so, despite the drive to increase new build programmes to accommodate predicted increases in the population. To comply with the European Directive 2002/91/EC, all dwellings that are available to either buy or rent must be accompanied by a prediction of their energy consumption. In the UK, legislation enforces the production of an Energy Performance Certificate (EPC) in order to comply with this directive [8]. An EPC, produced using the Government's Standard Assessment Procedure (SAP), provides a benchmark of a building's energy performance, when compared against a national standard. The SAP methodology is also approved by the UK government to demonstrate compliance with Part L: Conservation of Fuel and Power of the Building Regulations in England and Wales, with Technical Handbook 6 in Scotland and with Technical Booklet F in Northern Ireland [9-12].

Limitations in the SAP calculation method can lead to perceived gaps between the estimated and actual energy consumption of a dwelling; reasons for this and examples of how it manifests in reality are presented in section 4 of this paper. It is important to acknowledge that this is a perceived or 'predicted' gap in many cases. Regulatory compliance calculations are often assumed to be providing an accurate estimate of energy performance, although this was not their intended purpose. The performance gap is now widely discussed in the built environment sector, with a growing evidence base underpinning the phenomenon [13-23]. The contribution of SAP calculations to the performance gap should however be treated with caution. The SAP calculation is a benchmarking tool and should not be automatically assumed to be producing an accurate design intent based prediction of a building's future energy performance. It is in effect an asset rating tool that can be used to compare the relative performance of various properties. Despite its intended purpose as a benchmarking tool, in reality, SAP is often employed as a design tool and this can lead to various problems including sub-optimal design, favouring cost efficiency over energy or carbon efficiency and excessive overheating due to the highly simplified method designed to avoid excessive solar gains [24].

The thermal performance of the building fabric is particularly significant in the domestic sector, as space heating is the largest single end-use, representing approximately 60% of the total energy consumption in UK dwellings [25-27]. Consequently, the gap between the predicted and actual fabric performance has become the focus of many research projects, which have sought to quantify the discrepancies [16, 18, 19, 28-30].

The research presented in this paper considers two case study no-fines concrete dwellings, both subjected to pre- and post-retrofit *in situ* fabric performance tests, with the test data used to calibrate dynamic thermal models of the buildings. In conjunction with pre- and post- retrofit SAP calculations, the test results have been used in this research to quantify differences between the modelled estimates and the as-built fabric performance, and to explore the reasons for any discrepancies. The case study dwellings were part of a broader retrofit research project. It is important to stress that, in the wider research project, SAP was used exclusively to calculate the pre- and post-retrofit energy performance of the dwellings; all the dynamic models described in this work have been produced in the context of this paper and were not used to evaluate the retrofit measures as part of the design process.

The calibration methodology used here was first defined in previous work [31] and has been further validated by extending its use to the retrofitted case study dwellings. The methodology is particularly useful for building modellers, as its effectiveness is not limited by lack of access to actual year building simulation weather files. Local building simulation weather files for specific periods are difficult to obtain, potentially expensive to purchase or resource intensive to produce. Results demonstrate the importance of refining the model inputs wherever possible, especially when considering the type of deep retrofit project considered in this work. They also demonstrate that, whilst it is not inherently flexible and cannot be used to understand detailed aspects of performance, SAP can be useful when predicting proportionate savings from retrofit programs, but has limitations when predicting absolute values. This in turn has implications for forecasting cost savings, which are of greatest concern to the vast majority of end users.

# 2. LITERATURE REVIEW

## 2.1 Domestic Energy Modelling and the Performance Gap

The department for Business, Energy and Industrial Strategy (BEIS) require that SAP be used to produce performance estimates and, subsequently, an Energy Performance Certificate (EPC) for dwellings that are either being sold or offered for rent. Data from domestic EPCs has now been made publically available [32] and is especially useful for meta-analysis; it should however be noted that these records do not include enough information to reproduce the SAP calculations upon which they are based. Reduced data SAP (RdSAP) is used for existing buildings, which simplifies inputs and makes a number of fixed assumptions [24, 33, 34]. The SAP calculation methodology is central to UK domestic energy efficiency policy and is, primarily, used to demonstrate regulatory compliance [10, 35]. It is also linked to policy instruments, including Stamp Duty exemption for zero carbon homes, Feed in Tariffs and the Renewable Heat Incentive [24]; it was also central to the now defunct Green Deal, with RdSAP used to inform economic payback for specific retrofit measures [24, 34].

The fixed boundary conditions and restricted inputs of the SAP calculation methodology can limit the accuracy of performance estimates for specific dwellings. Standard occupancy profiles, heating/cooling set points and a central UK weather file are used as part of the calculation [24, 33]. Although the central UK weather file is used to calculate the Environmental Impact (EI) rating and Dwelling Emissions Rate (DER), predicted energy costs are based upon average regional weather conditions. Due to the fixed value of some SAP inputs, there is often a gap between predicted and actual building performance. However, this perceived gap is not limited to SAP calculations; performance gaps are also common in buildings simulated using dynamic models, especially those designed to produce non-domestic EPCs [17, 36, 37]. The greater reliance on the assumptions incorporated within RdSAP, which are intended to reduce the burden of data collection, the complexity of the assessment and, ultimately, the cost of an assessment, can serve to further widen the gap between the predicted and in use performance [24].

In cold and temperate climates, the thermal performance of the building fabric of a dwelling has the most significant impact on its energy consumption and associated $CO_2$ emissions [16, 18, 19, 29]. Several methodologies have been developed that allow the aggregate thermal performance of a whole building's fabric to be measured in the field; it is through this *in situ* testing that the building fabric's contribution to gaps in thermal performance can be quantified. Testing methodologies include the Primary and Secondary Terms Analysis and Renormalization (PSTAR) method [38, 39] and the electric coheating test method [38, 40-44]. The PSTAR method is a dynamic test method that uses short-term energy monitoring (STEM) tests, such as temperature monitoring, flux measurement and air pressurization tests, to gather performance characteristics on the building undergoing testing. The STEM results are then used as inputs to a model, which is then extrapolated to produce an annual simulation [40]. Data obtained through this dynamic method can be highly complex, limiting its application in the field; accuracy of results obtained using this method have also been questioned in recent years [45].

The electric coheating test has undergone significant development in the field. It comprises a quasi-steady-state test method which involves artificially heating the internal environment within an unoccupied building to an elevated homogeneous state, using electric resistance point heaters, whilst the external environment conditions vary naturally [46, 47]. By measuring the daily heat input in to the building in Watts over a specified period of time, and plotting this against the daily difference in internal/external air temperature (ΔT) in Kelvin, the uncorrected raw heat transfer coefficient (HTC) for the building can be determined in Watts per Kelvin (W/K). The electric coheating approach has been used repeatedly to help identify discrepancies between the predicted and actual fabric performance of dwellings in the UK [18, 29, 30]. Although the coheating test in isolation cannot explain reasons for gaps in performance, it does provide an opportunity to carry out additional fabric tests, such as heat flux measurement and air pressurization, which can provide information to identify areas of poor performance [48]. Other dynamic testing methods are available, which offer much shorter test durations than the coheating methodology, such as the Quick U-value of Buildings (QUB) method [49] and ISABELE (In Situ Assessment of the Building EnveLope pErformances) [50]. However, the analysis of test data is significantly more complex and, consequently, their use has so far been primarily limited to research and development purposes.

The same type of models described in this work have been used to investigate the effect that thermal mass and solar radiation has on electric coheating test results [51]. Stamp et al. [51] compared outputs from Energy Plus models with a small sample of empirical coheating data and found that for heavyweight, highly glazed dwellings, the heat loss may be underestimated when there are high levels of incident solar radiation. These results are pertinent to the method described in this work, but did not have an impact on the test results reported here. The case study dwellings do not incorporate especially high levels of thermal mass, and the tests were carried out during periods with very low solar radiation. As Stamp et al. note, electric coheating tests are time consuming, therefore modelling electric coheating conditions offers an efficient means of understanding a wider sample set of building thermal performance [51]. This represents another potential benefit of validating the type of model calibration method described in this paper.

In warmer climates, the ability to limit solar gains and ventilate dwellings can often play a more dominant role than the fabric insulation and air tightness of dwelling. However, recent work has aimed to parametrise performance of fabric insulation in the context of overheating [52]. A large-scale modelling exercise was combined with data mining techniques, and validated against monitored data, to evaluate the impact of key variables, including insulation, thermal mass, glazing ratio, shading, occupancy, infiltration and ventilation. This work concluded that, although fabric insulation can contribute to both increased and decreased overheating, it only accounted for 5% of overall overheating response when ranked against other influencing variables [52].

Recently published work has investigated the performance of housing stock energy models, including SAP, and identified a number of issues that have an impact on their efficacy [53]. Sousa et al. evaluated 29 different models and found that they were collectively over-simplified and non-transparent, thus it was difficult to understand the algorithms and methods used to predict energy performance [53]. It is not however relevant to describe the SAP methodology as ‘non-transparent’ as the algorithms used in the calculation engine are published as part of the SAP methodology document [33]. Sousa et al. [53] call for a means of improving model predictions and for the calibration of models to be implemented long-term. As already noted, much of the published work related to the performance gap has been produced in the UK. There are, however, a number of international examples of simplified energy models contributing to gaps in performance. A wide number of these focus on the non-domestic sector and many are summarised by Zou et al. [54] Specific to domestic models, there are examples from European countries where simplified energy modelling has contributed to the performance gap, including multiple examples from Germany, the Netherlands, Italy and Switzerland. [20, 21, 55-57]

DesignBuilder software was used in the research presented in this paper. Although it is itself a propriety piece of software, the building physics engine that it uses, Energy Plus, is entirely open-source and could therefore offer a potential means of addressing some of the issues identified by Sousa et al. [53] if it were applied at scale. The calibration methodology described in this paper could also help address this issue, building upon the wide-spread application of the modelling techniques identified by Stamp et al [51]. In the context of the calibration method used in this work, and in contrast to simplified steady-state models, Energy Plus is particularly useful as it allows analysis to be completed at hourly time steps, allowing the dynamic nature of heat loss to be simulated fully. Energy Plus also allows data outputs to be aggregated at daily time steps, which means they can be directly compared with the measured coheating test data as part of the calibration process.

## 2.2 Building Simulation Calibration Techniques

Simulation calibration is defined as a process through which the inputs for an existing building simulation model are fine-tuned so that a specific output (or multiple outputs) closely match observed performance in the real building [58]. A research project commissioned by the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) in 2003, aimed to develop a 'coherent and systematic calibration methodology' [58]. Subsequently, this field has become a broad area of research, with more recent papers reviewing developments in calibration techniques [59, 60]. Coakley et al. [59] published a comprehensive review of calibration techniques which divides approaches to calibration into two core categories, "manual" and "automated" methods. The method described in this paper can be described as a manual approach, as the specific inputs are systematically updated using *in situ* measured data. Coakley et al. further sub-categorize calibration techniques and the approach used here fits into the sub-category of a "procedural extension" that utilizes an "evidence-based development" method [59]. The method used in this work has similarities, in terms of procedure, to an 'evidence-based' methodology developed by Raftery et al. [61, 62] A simple iterative approach would not necessarily use *in situ* measured data to refine inputs and record evidence of changes in model inputs as it is updated [59]. Coakley et al. [59] describe the problem of accuracy within complex building simulation as one that is “over-parameterized” and “under-validated.” The method described in this paper partly mitigates against these issues, as the key model input parameters are measured *in situ* and the results can be validated directly against the measured aggregate whole house thermal performance.

The strength of a calibrated model relies on the input data available for calibration, which can be reported at different levels (see Table 1). The additional data and resource requirement for higher calibration levels often presents an obstacle, with calibration often being a compromise between data availability and accuracy. To obtain the range of data required for levels 4 and 5 of calibration, there is no prescriptive number of parameters measured or frequency of measurement; this will be decided on an ad-hoc basis. For example, in a dwelling, it may be both practical and useful to measure air temperature and relative humidity in every room, whereas, in a very large non-domestic facility, this may not be viable. For specific fabric elements, and particularly U-value measurement, there is ongoing work that is exploring the requirement for more detailed in situ measurement of specific plane elements. [63] The creation of accurate models of existing buildings allows building stakeholders to evaluate major retrofit upgrades thoroughly at an early stage. Informing ‘investment grade’ decisions relating to low-energy retrofits is a broadly acknowledged application of calibrated simulation; it is also particularly important when retrofits are multiple and interactive [61, 64-66]. It is important to note that the technique described in this paper is particularly useful for the analysis of retrofit measures that increase the efficiency of space heating, which places less emphasis on occupancy, equipment use and associated internal heat gains. If retrofit measures associated with cooling or energy management are being evaluated, it is more important to calibrate models against data recorded during periods of occupancy.

Table 1: Calibration levels based on the building information available [58, 60]

| Calibration Levels | Building Input Data Available | | | | | |
|---|---|---|---|---|---|---|
| | **Utility Bills** | **As-Built Data** | **Site Visit or Inspection** | **Detailed Audit** | **Short-Term Monitoring** | **Long-Term Monitoring** |
| Level 1 | X | X | | | | |
| Level 2 | X | X | X | | | |
| Level 3 | X | X | X | X | | |
| Level 4 | X | X | X | X | X | |
| Level 5 | X | X | X | X | X | X |

Reddy [58], originally used four main categories to characterize calibration techniques: (a) manual, iterative and pragmatic interventions; (b) informative graphical comparative displays; (c) special tests/analytical procedures; and (d) analytical and mathematical methods [58]. Some limitations can restrict the accuracy of iterative approaches (a) although the process can be refined [62, 67, 68]. There is often an over-reliance upon an analyst's own knowledge and experience, which is sometimes applied in an unstructured way [59, 60, 64]. It is also possible to structure iterative approaches as described in multiple publications [59-61]; one example uses the concept of 'calibration signatures' and 'characteristic signatures' [69]. Calibration signatures can be identified from consumption data and plotted against an external variable, for example, external dry-bulb temperature, and then used to compare with model outputs [70]. Characteristic signatures provide a baseline profile from a similar reference building type to compare with model outputs. This approach is best suited to the calibration of heating and cooling system energy consumption where signature values are plotted against a temperature range to allow errors to be visualized. The approach described in this paper employs a similar method of comparing output data by plotting modelled and measured power demand against the temperature differential ($\Delta T$) between the internal and external environments.

Standard systematic methodologies for calibrated simulation based upon the iterative approach have been developed. They incorporate elements of the four types previously defined and a higher resolution and range of real data from the subject building; most accurately hourly end-use records [61, 62, 64, 65]. These do however rely upon actual weather data and detailed information recorded within existing buildings. Invasive 'blink-tests', where different end-use loads are activated and deactivated, can also be used to quantify consumption and improve model inputs [71, 72].

Although the accuracy of DSM software has been questioned in the past, inaccuracies are most often associated with a misunderstanding between using DSM for regulatory compliance purposes and using it as a low-energy design tool. It has however been demonstrated that using refined data inputs reduces this element of the performance gap significantly [19, 67, 73-75]. There has also been criticism of the accuracy achievable by simulation software [76, 77], but results from case studies [62, 73, 74] have demonstrated for some time that DSM calculations can be calibrated to closely simulate actual performance at different temporal resolutions, when inputs are refined and a systematic approach to calibration is employed. The accuracy of the physical-law driven software used to produce building energy models is also independently validated by professional bodies [78, 79]. Although the competency of modellers can affect the accuracy of predictions as highlighted in recent work [77], it is mainly the assumptions that are made in regulatory compliance calculations that account for the modelling related performance gap identified in published research [17, 36, 37, 74]. Ultimately, as identified by de Wilde [23], a co-ordinated effort and change of industry practice to address the range of issues that contribute to the performance gap. Pertinent to this paper, de Wilde notes that further work is required to validate energy models outputs and to improve the quality of model input data [23].

## 3. METHODOLOGY

### 3.1 Case Study Dwellings

As part of a larger development, the two case study dwellings are examples of the ‘no-fines’ construction technique used during the 1960s to help meet the rapidly growing demand for housing throughout the UK. The two case study houses are referred to in the rest of this paper as ‘Dwelling A’ and ‘Dwelling B’ to ensure anonymity. The term ‘no-fines’ is used to describe concrete mixed using only large aggregate and is considered to be a non-traditional method of construction. Moss [80] describes no-fines concrete as consisting of Portland cement and course aggregate, in ratios of 1:8 to 1:10. The concrete mix does not include any fine material, such as sand or gravel. Instead, it has an open textured cellular structure that creates air voids which can act as insulation, depending upon the particular density of mixed concrete, giving solid no-fines concrete walls U-values which are considered comparable to uninsulated masonry cavity walls [81]. Images of the no-fines material in Dwelling A are shown in Fig. 1.

Fig. 1. *In situ* no-fine concrete wall constructions.

Although proposed as an energy saving construction in the 1970s [80], when compared with modern building regulations, no-fines walls have very poor thermal performance. A 250mm thick solid no-fines wall finished internally with plaster (the density of which is not specified) has been found to have a U-value of 1.71 $W/m^2{\cdot}K$, whilst the same wall finished with plasterboard on timber battens has been found to have a U-value of 1.23 $W/m^2{\cdot}K$ [82, 83]. These were however taken from a very small sample of three dwellings. The thermal performance and quality of no-fines concrete walls varies due to the lack of a standard approach to construction and variance of the size and shape of aggregates used, which can affect the proportion and nature of the voids within the structure. Sommerville et al. [82] investigated the nature of the voids within no-fines concrete. Sample cores, removed from no-fines concrete homes were found to consist of 66% voids, consisting of a mix of isolated closed voids, and interconnected voids. In contrast, laboratory analysis of cast samples found 37% voids, although more compaction on site reduced the proportion of voids to 25%. The differences in aggregate between lab samples and actual drilled cores, as well as different practices on site, are likely to be responsible for the large difference in void proportion.

This range of measured values is not unique to no-fines, but is also evident in other solid wall dwellings. A large-scale survey in the UK found that the mean value for standard solid wall constructions was 1.57 $W/m^2{\cdot}K$ (with a standard deviation of 0.32), with non-standard solid wall constructions having a mean value of 1.28 $W/m^2K$ (with a standard deviation of 0.42) [84]. For the purposes of the cited report, ‘standard’ solid walls were of brick construction with a thickness < 330mm; ‘non-standard’ cases were brick walls with a thickness >= 330mm or solid walls not constructed of brick (stone or concrete walls for example). It has been established that measured U-values in the UK are often higher than calculated values for the same construction [22, 84, 85]. However, in some instances, the in situ U-values for solid walls have been found to be lower than the default values used in energy calculations [86, 87]. Recent work has also emphasised the limitations of point measurements in determining whole element U-values, particular to solid wall construction. Marshall et al employed three different techniques to measure in situ U-values in controlled conditions and found there to be a significant difference between the plane element U-values derived using point measurements when compared to those derived with the aid of high-resolution thermal imaging [88]. Akin to the work presented in this paper, the range of U-values estimated by Marshall et al were used to calibrate energy models that predicted HTC values between 183 – 285 W/K [88].

Dwelling A is a mid-terrace property and Dwelling B is an end-terrace property. The houses are owned by the Northern Ireland Housing Executive and formed part of a retrofit research project funded through the Innovate UK 'Scaling Up Retrofit' programme of funding; this specific project aimed to investigate the retrofit of hard-to-treat solid wall homes (www.s-impler.com). Dwelling A was retrofitted with traditional closed cell external wall insulation (EWI). Dwelling B was retrofitted with a novel dynamic external wall insulation system that aimed to recover heat by drawing external ambient air through a thin cavity within the insulation panels. The heat that is harvested is then returned to the dwelling through a low-speed positive input ventilation system. Section drawings for the retrofitted walls of Dwelling A and B are shown in Fig. 2.

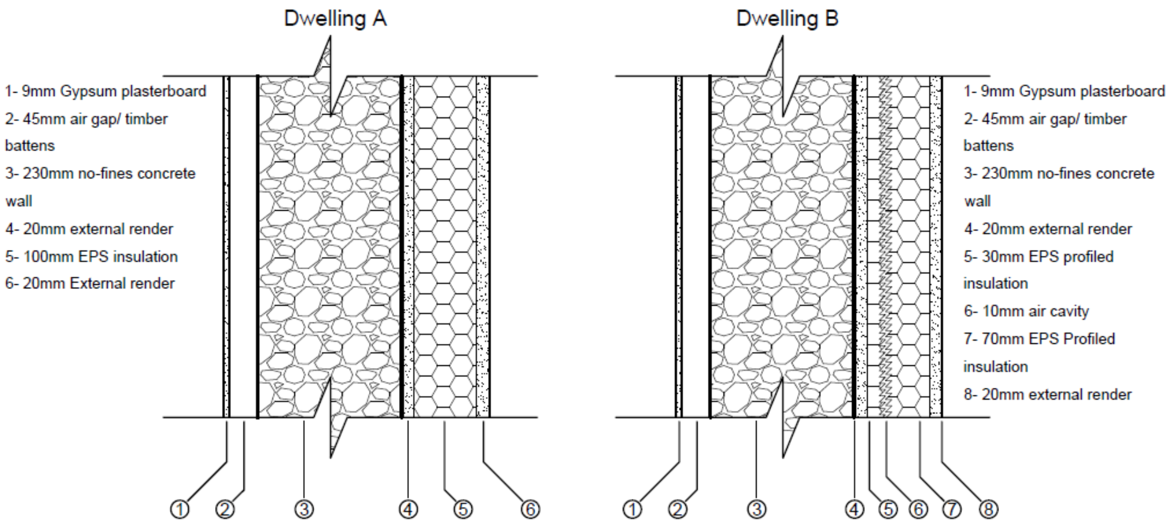


Fig. 2. Retrofitted wall section drawings.

Other upgrades to both case study dwellings included increased loft insulation, double-glazed windows, upgraded doors and zonal heating controls. Technical data relating to the pre- and post-retrofit construction details are presented in Table 2 later in this paper. The retrofit materials were specified by designers as part of a competitive tender process and the overall retrofit proposals were evaluated by third parties in the wider research project. Analysis of retrofit proposals were evaluated using pre- and post-retrofit SAP calculations.

The approximate total floor areas of Dwelling A and Dwelling B are 98m$^2$ and 100m$^2$ respectively. The retrofit of both dwellings increased the floor areas slightly, as the front doors were moved forward to create a single front façade and reduce conductive heat loss areas. Old boiler cupboards were integrated into the rear porch areas, which were also externally insulated. Floor plans of the pre-retrofit properties and visualizations of the building models both before and after retrofit are shown in Fig. 3 and 4.

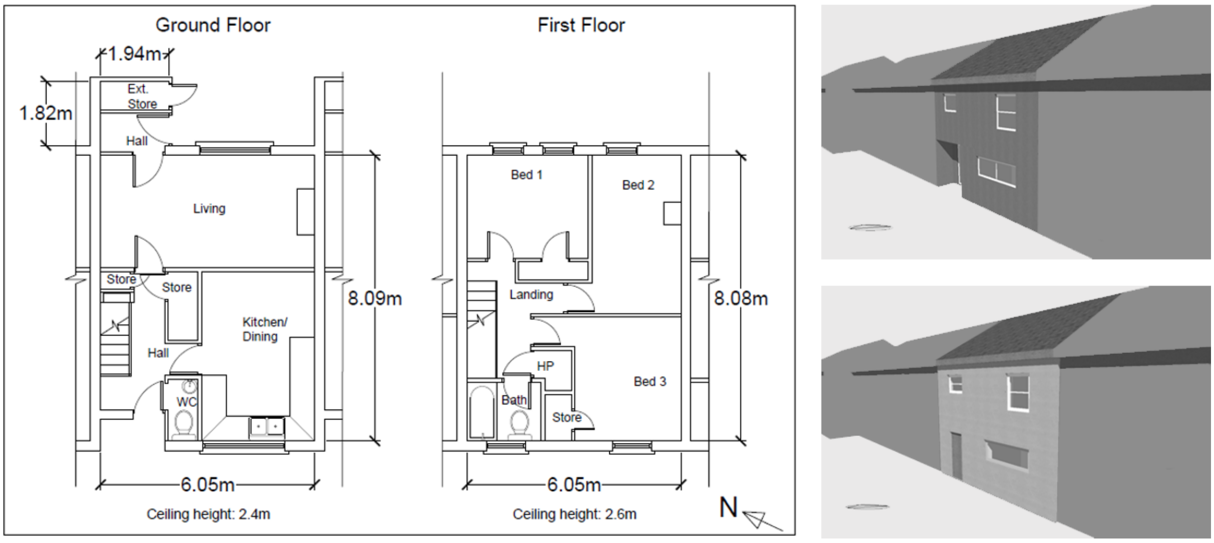


Fig. 3. Floor plans and visualisation of Dwelling A before and after retrofit.

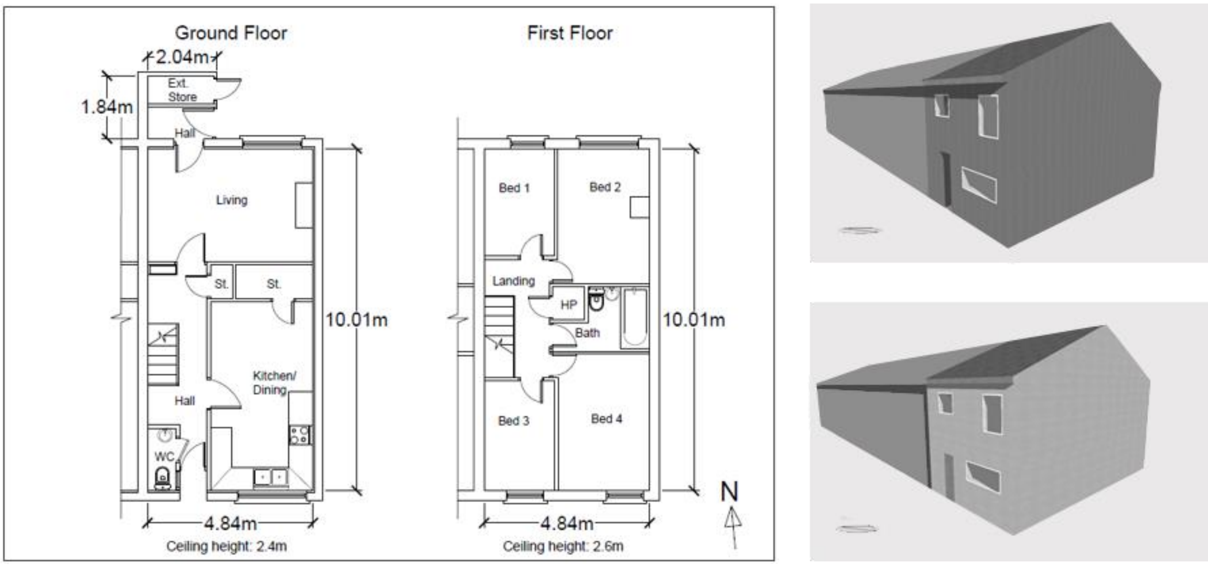


Fig. 4. Floor plans and visualisation of Dwelling B before and after retrofit.

## 3.2 Coheating and *in situ* fabric performance measurement

The two case study dwellings (Dwelling A and B) were tested both pre- and post-retrofit. The Heat Transfer Coefficient (HTC) was calculated for each dwelling through the completion of electric coheating tests both before and after retrofit. Tests were completed following the protocol defined by Leeds Beckett University [47]. The uncertainty associated with this test method is between ±8 to 10% [89]. Pre-retrofit testing was completed between the 6$^{th}$ and 24$^{th}$ February 2014. The post-retrofit testing was completed for Dwelling A between 18$^{th}$ December 2016 and 6$^{th}$ January 2017 and the post-retrofit testing for Dwelling B was completed between 10$^{th}$ and 25$^{th}$ April 2016.

In accordance with the electric coheating test protocol, a range of parameters were monitored in detail during the test, including total electrical energy input to the dwelling, internal air temperatures and relative humidity, and local weather conditions. Sensor type RTDPT-100 (accuracy or ±0.1C) were used to monitor the internal air temperature and proportional, integral, derivative (PID) temperature controllers were used to control a series of electric resistance point heaters, such that the internal temperature throughout the dwelling was maintained at a stable internal temperature for the duration of the tests. Total electrical power input was measured using a series of Elster A100C kWh meters (accuracy of ± 1%). A Vaisala WXT520 weather transmitter was used to measure external air temperature (accuracy of ± 0.3°C), relative humidity (±3% RH at 0 to 90 % RH), wind speed (accuracy of ±3% at 10ms$^{-1}$) and wind direction throughout all test periods. This was originally attached to the gable wall of Dwelling B during the pre-retrofit testing and was relocated to a nearby gable wall to enable the external wall insulation (EWI) to be installed. A south facing vertically orientated pyranometer (Kipp & Zonen CMP3) was also attached to the weather station installation to measure solar insolation (typical accuracy of ±5%). All measurements were logged at ten-minute intervals using an Eltek Squirrel RX250AL data logger, with missing data being corrected using linear interpolation.

Elevated steady internal temperatures are a feature of the electric coheating test environment and provide optimized conditions for determining *in situ* U-values. The *in situ* U-values were determined in accordance with ISO 9869 [90] by measuring the heat flux density using Hukseflux HFP01 plates. The uncertainty associated with the in situ U-values measurements was calculated to be ±10% using the quadrature sum of the uncertainties listed in the ISO 9869. Either Thermo Fisher Scientific DataTaker DT80 data loggers or Eltek SQ851 loggers were used to record voltage at one minute and two minute intervals respectively. Point *in situ* U-values were then derived from the heat flux density measurements. In both dwellings, multiple heat flux plates (HFPs) were placed in areas that were considered representative of the relevant plane element, with particular care taken to ensure they avoided thermal bridges at elemental junctions. A FLIR T620bx thermal imaging camera was also used to ensure that HFPs were not placed over floor and ceiling joists and timber studwork. The plates were secured using thermal compound paste and adhesive tape around their edges. The diagrams in Fig. 5 illustrate the positioning of the HFPs in Dwelling A as an example of the careful placing of these; these locations were recorded and the HFPs were placed in identical locations for both the pre- and post-retrofit tests. In total, 49 HFPs were used in each of the case study dwellings for both the pre- and post-retrofit tests; this includes HFPs used to measure heat flux across party walls also.

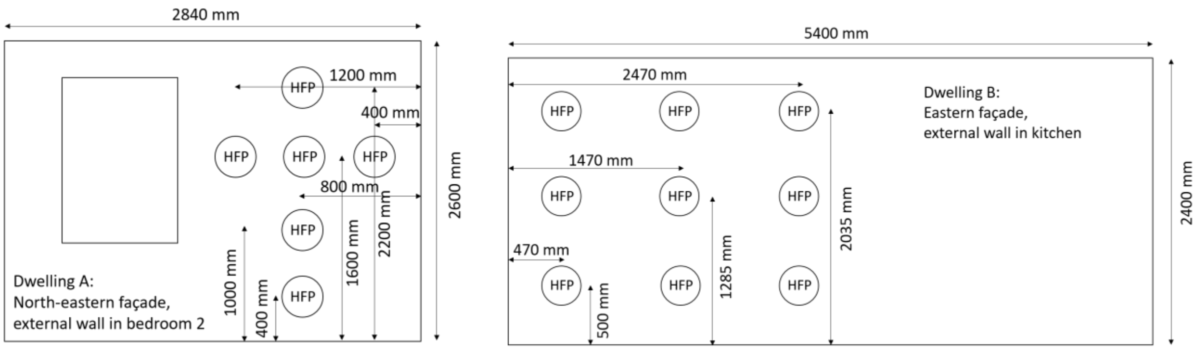


Fig. 5. Example placement of HFPs in Dwellings A and B (not to scale).

During the electric coheating tests, electrically driven air circulation fans were used to ensure that the internal air was thoroughly mixed and distributed evenly throughout the dwellings and to minimise any potential temperature stratification. The fans were also positioned to avoid air being directly blown onto the HFPs, as this can potentially influence readings. To mitigate against the thermal mass of the building compromising the overall results, steps were taken to ensure that the building mass was fully charged before the commencement of the electric coheating tests. In addition to this, the 'Average Method' defined in ISO 9896 also avoids the influence of thermal inertia on the results.

Air tightness of the dwellings was assessed using a blower door to conduct a series of building pressurization tests consistent with the ATTMA Technical Standard L1 [91]. Based upon the results of the pressurization tests, background ventilation rates were approximated using the $n_{50}/20$ 'rule of thumb' [92] The $n_{50}/20$ rule uses a correction factor for local shelter and adheres to the SAP 2012 methodology [33]. Local environmental conditions during the time of each air pressurization test dictate the level of associated uncertainty, which is normally less than 10% [93], so they were conducted during days with relatively low wind speed.

## 3.3 Calibration methodology

As with the *in situ* testing , the model calibration has previously been described in detail [31] and is paraphrased here. As part of the calibration process, results from the pre-retrofit and post-retrofit tests were used to refine the model inputs. The HTC value was used to calculate the total difference between the modelled and measured fabric performance. A simple four stage iterative process was followed to calibrate the models: a baseline model was first created based upon digital drawings and site surveys (stage 1); the model was then refined to account for calculated thermal bridging (stage 2); the air changes per hour were approximated based upon the measured air leakage rate (stage 3); and measured *in situ* U-values were used to update the model constructions (stage 4). These updates were introduced chronologically in terms of when the data was available (a scenario that is likely to be repeated in other projects).

Case study models were produced using DesignBuilder software version 5.0.2.003 [94], which uses Energy Plus as its physics engine. The CIBSE Test Reference Year 2016 (TRY) building simulation weather file for Belfast [95] was used in all models, as this was the closest available file for the building location. Model geometry inputs were based upon digital plan and elevation drawings, and these were validated using on-site measurements. In this instance, the geometry modelling convention in DesignBuilder was set to calculate fabric heat loss based upon internal dimensions, as opposed to external dimensions, in order to aid comparison with the SAP methodology. This method was also used as the thermal bridging Ψ values were calculated to create a y-value suitable for use in SAP. When using internal dimensions, the calculated Ψ will differ from those calculated using external dimensions. This is most pronounced when calculating Ψ values for linear bridges at the external corner wall-wall junctions, which will have a negative value under the default software modelling convention. The orientation of the buildings was measured using plan drawings of the full estate area. As indicated in Fig. 3 and 4, neighbouring dwellings were included to represent local shading and the adiabatic spaces directly adjacent to the case study dwellings.

A summary of the main thermal plane elements is presented in Table 2; these values were used in the baseline models; the U-values used in the SAP calculations are presented in Table 3 of the results section for comparison. The U-values used in the DesignBuilder software are calculated using the construction build-up in adherence to BS EN ISO 6946 [96]. The U-values for all constructions were adjusted to account for repeat bridging from details such as timber stud frames; this adjustment is performed within the software. Before calibration, a default value for a leaky dwelling of 17 $m^3/(h.m^2)$ @ 50Pa was used for the air permeability, which, when using the $n_{50}/20$ rule of thumb, approximates to 0.85 air changes per hour [97].

Table 2: Main fabric elements in 'Surveyed' iteration of models

| Element: | Construction pre-retrofit (outside – inside): | U-value ($W/m^2 \cdot K$) | Construction post retrofit | U-value ($W/m^2 \cdot K$) |
|---|---|---|---|---|
| External wall | External render 20mm; No-fines concrete 215mm; Air void 25mm (bridged by softwood battens); Plasterboard 12.5mm | 1.418 | Dwelling A: as pre-retrofit + dynamic insulation 100mm (65mm/35mm); render 10mm. Dwelling B: as pre-retrofit + dense EPS insulation 100mm; render 10mm. | A = 0.445<br>B = 0.444 |
| Roof | Concrete roof tiles 12mm;<br>Air void (various depth due to pitched roof);<br>Mineral wool insulation (Dwelling A 150mm; Dwelling B 250mm); Plasterboard 12.5mm | A = 0.162<br>B = 0.416 | As pre-retrofit but mineral wool layer increased by 300mm (A increased to 450mm and B increased to 550mm overall) | A = 0.068<br>B = 0.086 |
| Ground floor | Lightweight concrete 200mm; Floor screed 10mm; Linoleum floor tiles 5mm. | 1.286 | As pre-retrofit, with an additional 100mm perimeter insulation in both cases. | A = 0.59 |
| Windows | Softwood frame, single glazing 6mm | 5.778 | PVC frame, double glazed low-E windows 10mm/12mm/10mm | 1.400 |

The electric coheating test is modelled by adjusting inputs to reflect the real test conditions, with no internal heat gains from people, lighting and equipment included in the simulation. A convective electric heat source with a performance coefficient of 1.0 was specified to model the required heat input. Heating set points were specified as constantly running at 22°C and 23°C in dwellings A and B respectively, for the pre- and post-retrofit models as per the mean on-site test conditions.

Analysis of the real electric coheating test data removes the influence of solar gains through linear regression and this technique has been replicated using the model outputs. The power output is used as the dependant variable and results are regressed using the mean daily global solar and the difference between indoor and outdoor dry bulb temperature ($\Delta T$) as the independent variables. Through Analysis of Variance (ANOVA), a coefficient is calculated and used to multiply the mean daily solar values before they are summed to calculate the power used to maintain the internal temperature that simulates the coheating test conditions. This is in contrast to previous work [31], where solar gain in the model was minimised by using permanently closed blinds in all of the windows and placing a large continuous local shading device around the east, south and west façades of the dwelling. In the previous method, any remaining negligible solar gains were added to the delivered power in the analysis of the output data. However, this has the potential to impact on the results through altered exposure to weather conditions and the slight increase in thermal transmittance associated with the closed blinds. Therefore, the solar regression described above is applied in this work.

Both pre- and post-retrofit SAP calculations were prepared as part of the wider research project. The Heat Transfer Coefficients included in the SAP calculations have been used for comparison in the results section, and these have been compared to the HTCs estimated in the DSM calculations and those derived from the *in situ* electric coheating test. A full set of thermal bridging calculations were also undertaken and the resultant Ψ-values from each junction present in each dwelling were used to calculate a whole house y-value to include in the SAP and DSM calculations. For the pre-retrofit buildings, y-values were calculated to be 0.133 $W/m^2 \cdot K$ and 0.121 $W/m^2 \cdot K$ for dwellings A and B respectively, with the post-retrofit values being calculated as 0.064 $W/m^2 \cdot K$ and 0.065 $W/m^2 \cdot K$ respectively.

Although quantitative data that is mainly analysed in this work, it is important to draw attention to the limitations of it being presented in a case study context. Whilst the calibration methodology itself is robust and transferable, the implications of the overall results are specifically relevant to this case study. Existing work related to the performance gap, particularly for retrofit projects, suggests that the implications will be relevant to other house and construction types. However, it should be understood that the results are directly attributable to this specific case study, despite their potential relevance to the wider field of research.

### 3.4 Modelling conventions in SAP and DSM calculation methods

There are integral differences between the SAP and DSM calculation methods and describing these provides some context to the comparison. The most obvious difference is that SAP uses steady-state monthly heat balance calculations, whereas the DSM calculates heat balances dynamically at an hourly temporal resolution. Geometry inputs are essentially the same and the geometry modelling convention in the DSM software was configured to calculate fabric heat loss based upon internal dimensions, as this is consistent with the SAP methodology (most DSM software default conventions use external dimensions). The means of accounting for solar heat gain are more sophisticated in the DSM models, as the dwellings are orientated as they are in reality and solar gain is calculated at an hourly time-step, dependent upon local weather conditions at each step. The SAP methodology does make some allowance for orientation, but only in a simplified way, by designating the orientation of the glazing at steps of 45°.

Zone types in DSM models include spaces for circulation, lounge, kitchen, bathrooms, toilets and bedrooms; this contrasts with the SAP model, which includes only two zones, living and bedroom space. The heating set points are 21°C for the main living spaces (lounge and kitchen) and 18°C for all other rooms. This differs slightly from SAP, which uses 21°C and either 18°C, or 19°C depending upon the dwelling's heat loss parameter (HLP); SAP has two zones, zone 1 is main living space, and zone 2 is the rest of the dwelling. The temperature set point in zone 1 is 21°C, while zone 2 is dependent upon the HLP and heating control type. Therefore, in poorly insulated dwellings zone 2 will use a set point of 18°C, while in better insulated dwellings this will be above 19°C. The set points in the DSM were set up to match these inputs. The DSM also inherently accounts for savings assumed in SAP from improved control systems, as individual heating set points control the heat input in each specific zone. A gas-fired boiler with the same specification of that used in the SAP calculations was also used in the DSM calculations.

In the SAP calculation outputs, only electricity consumed through lighting and any HVAC associated activities are reported. Heat gains from equipment are however included in the SAP heat balance calculation, along with heat gains from occupants. The DSM models use operating schedules for different zones (rooms) of the dwelling that are based upon the NCM thermal templates used in SBEM calculations [98] (the templates are used for larger buildings that include residential spaces). Although the NCM and SBEM are similar in nature to SAP, SAP was originally developed from the BRE Domestic Energy Model. The NCM templates do however offer a simple means of defining the daily occupancy and operating patterns required for the DSM calculations; the SAP calculations do not include any hourly patterns of this nature.

# 4. RESULTS

## 4.1 Pre-retrofit model calibration

The pre-retrofit measured and modelled HTCs for Dwellings A and B are illustrated in Fig. 6 and 7, respectively. Although the measured HTC values have changed only marginally from those presented in the authors' previous paper [31], measured HTC values have been updated as further analysis has refined the results, which are presented here for clarity. The modelled HTC values represent the final calibrated iteration of each model that have been updated to include thermal bridging losses, approximated air change rates and *in situ* U-values. There was also a further refinement made to these models to account for a measured phenomenon relating to the in-use air change rates and the inherent permeability of the no-fines constructions. During early site investigations, leakage detection during air pressurization tests suggested that air was moving through the no-fines walls. Thermography studies also indicated that convective heat transfer was occurring within the no-fines structure.

During the final *in situ* testing activities, it was possible to co-pressurize Dwelling A and the neighbouring dwelling. This test found that the measured air change rate reduced by approximately 25% when the co-pressurization tests were completed. This reduction is thought to be due to the inherently leaky, honeycomb structure of the no-fines concrete party walls, enabling unintentional air exchange between neighbouring dwellings when one of these dwellings is artificially pressurized in comparison to the other. This is in keeping with existing work reviewed and validated by Jones et al [99]. There is very little published research that identifies this, although Jones et al [99] cite three studies from three different countries that identified air flows through adjoining walls in terraced houses and apartment blocks.

This phenomenon will be explored in more detail in future work. During normal operation (and during coheating test conditions) neighbouring properties are unlikely to be operating under different internal air pressure levels. Therefore, the air change in this archetype between conditioned properties would account for very little, if any, heat loss in reality. It was also observed that all of the neighbouring properties were conditioned during all testing periods. To account for this effect, the final calibrated models included an adjustment to the final air change rate. A 25% reduction was measured for Dwelling A, which was a mid-terraced dwelling. A lower reduction of 10% was estimated for Dwelling B due to the lack of co-pressurization test results and it being an end-terrace dwelling with less party-wall area. Therefore, there is more uncertainty associated with the adjustment in the Dwelling B model.

For consistency, in Fig. 6 and 7, and in all the remaining figures in this paper, the measured HTC values are shown in the bottom-left corner of each chart area using bold italic text.

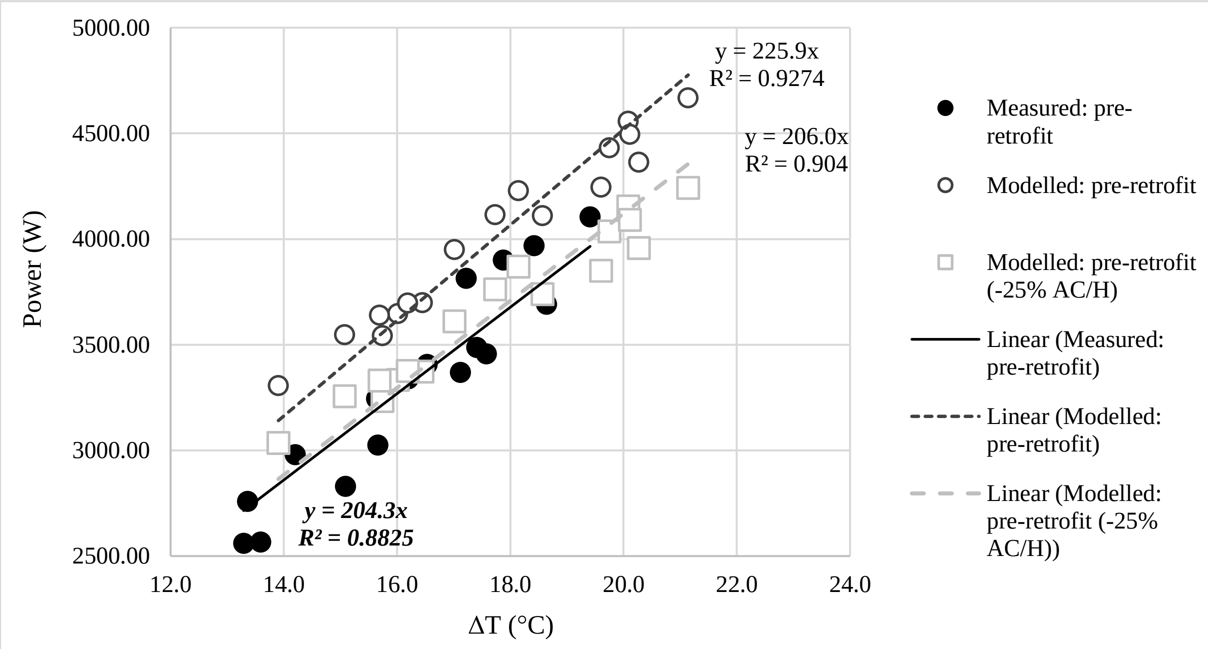


Fig. 6. Measured and modelled HTC values for Dwelling A pre-retrofit.

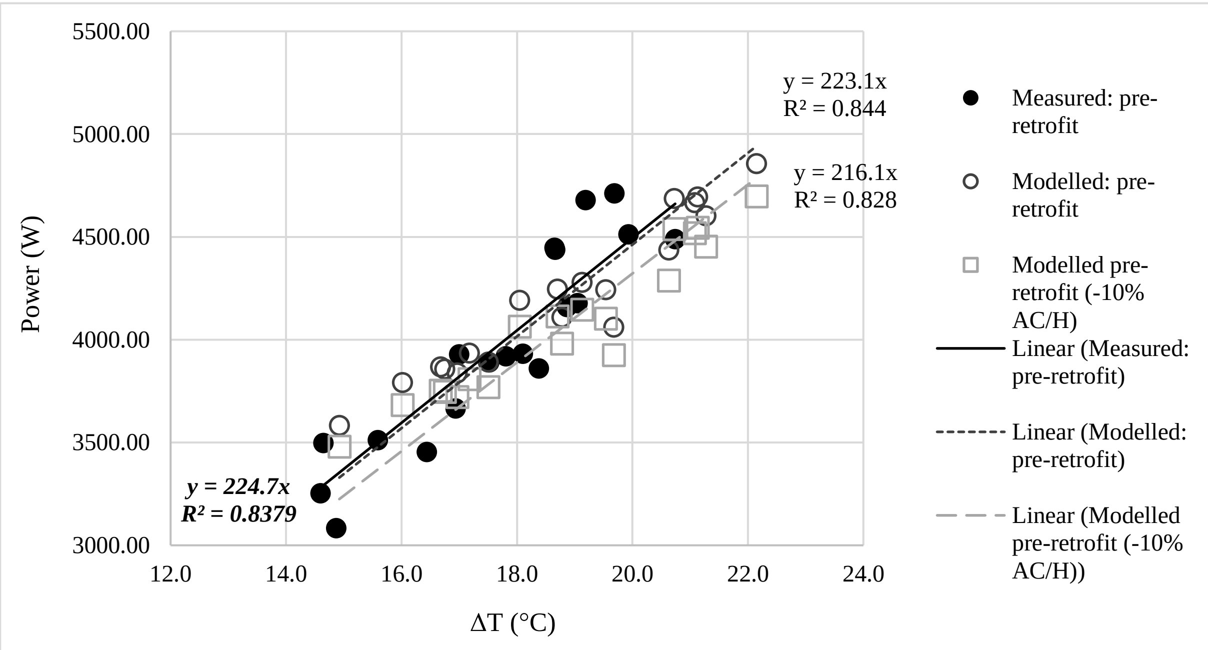


Fig. 7. Measured and modelled HTC values for Dwelling B pre-retrofit.

The results for Dwelling A indicate that the use of the air change rates calculated from the co-pressurization test results produce a more accurate prediction of heat loss than when using the results from the air pressurization test that was conducted in isolation. However, the same effect is not observed in the model outputs for Dwelling B; the version of the model using the measured air change rate predicts a HTC closer to the measured value than the adjusted model. Further testing is required to quantify this effect in an end-terrace building of this nature. The HTC for Dwelling B is only slightly under estimated by 1.6 W/K when using the measured air change rate and 8.6 W/K when including the reduced air change rate. The HTC for Dwelling A is considerably over estimated by 21.6 W/K before the air changes through the no-fines party walls are accounted for, however, once the air change rate is reduced, the model demonstrates a very close match with the measured fabric performance (1.7 W/K higher than measured). The difference between predicted HTCs for both models and their respective coheating test results are within the uncertainty associated with the test method, specifically, ±9.7 W/K for Dwelling A and ±8.8 W/K for Dwelling B. The need to adjust the air change rate in the Dwelling A model to calibrate them against *in situ* test results has potentially important implications for energy modelling, as will be discussed later in this paper.

## 4.2 Post-retrofit energy performance modelling

Results for the pre- and post-retrofit *in situ* U-value measurements and the HTCs are summarized in Table 3. The input values for the SAP calculations (referred to in the remainder of this paper as 'SAP As-built') and are presented in Table 3 for comparison. It should be noted that a more detailed data and analysis will be presented in a future paper that focuses explicitly on the *in situ* performance testing of the pre- and post-retrofit dwellings included in the S-IMPLER project. The original target of the S-IMPLER project was to achieve up to a 60%

reduction in total running costs. The data presented here is only relevant to the energy consumption associated with space heating but, as can be seen from the values in Table 3, a significant improvement in fabric performance was realized in both variants of the retrofit. It is also worth reiterating that the retrofit strategy for both properties was similar, and only differed in terms of the type of EWI system used. The proportionate improvement in the measured HTC results for both dwellings are comparable, with Dwelling A achieving an improvement of 52% and Dwelling B achieving an improvement of 43%. It is important to reiterate that there are some limitations to the purely quantitative nature of this analysis due to the case study context. Even within this small sample of seemingly similar buildings, there are significant differences in performance both before and after retrofit.

Table 3. Summary of measured pre- and post-retrofit fabric performance test results.

| | Air changes per hour (AC/H) | Thermal bridging (y-value) | Wall U-value ($W/m^2 \cdot K$) | Roof U-value ($W/m^2 \cdot K$) | Ground floor U-value ($W/m^2 \cdot K$) | HTC (W/K) |
|---|---|---|---|---|---|---|
| Dwelling A: | | | | | | |
| SAP As-built: pre-retrofit | 0.71 | 0.133 | 1.23 | 0.19 | 0.42 | 272.3 |
| SAP As-built: post-retrofit | 0.74 | 0.064 | 0.26 | 0.08 | 0.37 | 146.6 |
| Measured: pre-retrofit | 0.83 | N/A | 1.44 | 0.17 | 0.68 | 204.3 |
| Measured: post retrofit | 0.74 | N/A | 0.43 | 0.16 | 0.47 | 98.4 |
| Dwelling B: | | | | | | |
| SAP As-built: pre-retrofit | 0.71 | 0.121 | 1.23 | 0.32 | 0.59 | 286.2 |
| SAP As-built: post retrofit | 0.68 | 0.065 | 0.17 | 0.09 | 0.59 | 164.9 |
| Measured: pre-retrofit | 0.78 | N/A | 0.83 | 0.41 | 0.81 | 224.7 |
| Measured: post retrofit | 0.68 | N/A | 0.38 | 0.14 | 0.75 | 127.6 |

Although the values are described as ‘measured’ in Table 3, the air changes per hour are derived from the air tightness test results. The SAP air changes per hour value for the post-retrofit model of Dwelling A increases slightly as it uses the measured as-built value, the value included in the pre-retrofit SAP model was the default estimate for this building type. The thermal bridging y-value is calculated based upon plan and section drawings and specifications. It is not possible to measure the heat loss through linear thermal bridges in situ, which is the reason that no measured value is included. The calculated y-values were used to adjust elemental U-value inputs in the dynamic models. The U-values in the SAP calculations are considerably different to those measured in situ, most significantly for the no-fines walls. The SAP value does however align with the values stated in previously published work on no-fines thermal properties [82, 83]. In most cases, the calculated U-values in SAP are better than those measured in situ, apart from the pre-retrofit value for the no-fines walls in Dwelling B. This lower U-value will account for some of the difference between the measured and predicted HTC values. It is also important to note that the main area of the ground floor slabs were not treated as part of the retrofit but perimeter insulation was added in both cases. In the case of the dwelling using the dynamic wall insulation system (B), strips of insulation were added to the exposed perimeter of the ground floor slab. For the dwelling with EPS insulation (A), insulation boards were installed to cover the exposed perimeter of the ground floor slab.

As noted above, in the ‘SAP As-built’ calculations for the retrofitted dwellings, the air change rates derived from measured air permeability results were used. It may be that higher estimates of heat loss are, in part, indicative of the in-use air change phenomenon described earlier in this section. This would be particularly relevant for the results for Dwelling A, where the air change through party walls was directly observed. It is important to note that the HTC calculated using SAP includes heat losses through intermittent extraction fans; this heat loss is not included in the HTC calculated using the dynamic models. Therefore, the heat loss associated with the controlled ventilation has been removed from the reported HTC values for all SAP data in this paper. In the SAP calculations for Dwelling B, there is also an air exchange associated with the positive input ventilation system included in the HTC calculation. However, this is in theory offset by the U-value adjusted using the Appendix Q calculation, which takes account of the higher temperature of the incoming air that is pre-heated by the dynamic insulation system.

Despite the differences already discussed, the proportionate improvement in HTCs following retrofit that were calculated using SAP are similar to those in the measured results. Dwelling A achieves a reduction of 46% (compared to a measured reduction of 52%) and Dwelling B achieves a reduction of 42% (compared to a measured reduction of 43%). Although there are notable differences between the baseline performance of the two dwellings, and particularly the external wall U-values, it can be inferred from results, within this case study context, that this type of EWI-led deep retrofit can achieve reductions in HTC values of approximately 47.5% within an error of ± 4.5%. However, the particular limitations of the case study approach in extrapolating this quantitative analysis can be demonstrated by the fundamental differences between these similar buildings. The baseline external wall U-values in Dwelling A were considerably higher than the default values that provided more scope for overall improvement. Conversely, Dwelling B is an end-terrace building and has a larger external wall area which would intuitive suggest that a greater propionate saving could be achieved. It is also worth noting that the heat losses predicted by SAP are considerably higher than those measured in situ for both the pre- and post-retrofit scenarios. Therefore, the absolute savings predicted by SAP will be higher and would indicate a greater absolute cost saving, albeit, based upon a higher baseline cost to begin with; this is quantified and discussed further in section 4.4.

To inform an impartial comparison, the SAP calculations were also calibrated using the measured U-values that had a notable impact on estimated performance (referred to in the remainder of this paper as ‘SAP Calibrated’). The measured HTC and those estimated in the

‘SAP As-built’ and the ‘SAP Calibrated’ calculations are compared in Table 4. These results show that the SAP estimate of proportionate fabric savings is slightly lower in the calculations updated to include the measured U-values. This again has important implications for forecasts of energy savings through retrofit, also discussed in section 4.4.

Table 4. Comparison of reduction in HTC between measured values and predictions made using As-Built SAP and Calibrated SAP.

| | HTC (W/K) | Post-retrofit absolute reduction | Post-retrofit percentage reduction |
|---|---|---|---|
| Dwelling A: | | | |
| SAP As-built: pre-retrofit | 272.3 | | |
| SAP As-built: post retrofit | 146.6 | 125.72 | 46% |
| SAP calibrated: pre-retrofit | 258.2 | | |
| SAP calibrated: post-retrofit | 148.8 | 109.32 | 42% |
| Measured: pre-retrofit | 204.3 | | |
| Measured: post retrofit | 98.4 | 105.9 | 52% |
| Dwelling B: | | | |
| SAP: pre-retrofit | 286.2 | | |
| SAP: post retrofit | 164.9 | 121.27 | 42% |
| SAP calibrated: pre-retrofit | 264.9 | | |
| SAP calibrated: post-retrofit | 159.8 | 105.15 | 40% |
| Measured: pre-retrofit | 224.7 | | |
| Measured: post retrofit | 127.6 | 97.1 | 43% |

From a modelling perspective, it is important to note that the total footprint of the building DSM geometry was extended slightly when the retrofit EWI layer was added. This was done to ensure that the internal floor areas remained accurate. In the methodology used here, if EWI layers were added to the pre-retrofit models and the total floor areas were not adjusted, the internal floor areas of the models would be reduced due to the thicker walls following retrofit.

In order to produce a prediction of retrofitted performance based upon the calibrated pre-retrofit models, the improved U-values were first calculated within the DSM software using the specified construction layers for each element. In the first instance, simple updates to the constructions were made by introducing the additional layers of insulation as per the retrofit design specification for the walls, roof, doors and windows (windows were upgraded to Low-E double-glazing). None of the DSM U-value inputs, at this stage in the investigation, directly reference either the measured or SAP calculated retrofit values. They did however include the improvement in calculated thermal bridging and include the air changes per hour predicted in the design stage SAP calculations for the proposed retrofit. The modelled HTC for both dwellings is presented in Fig. 8 and 9 alongside the pre-retrofit values.

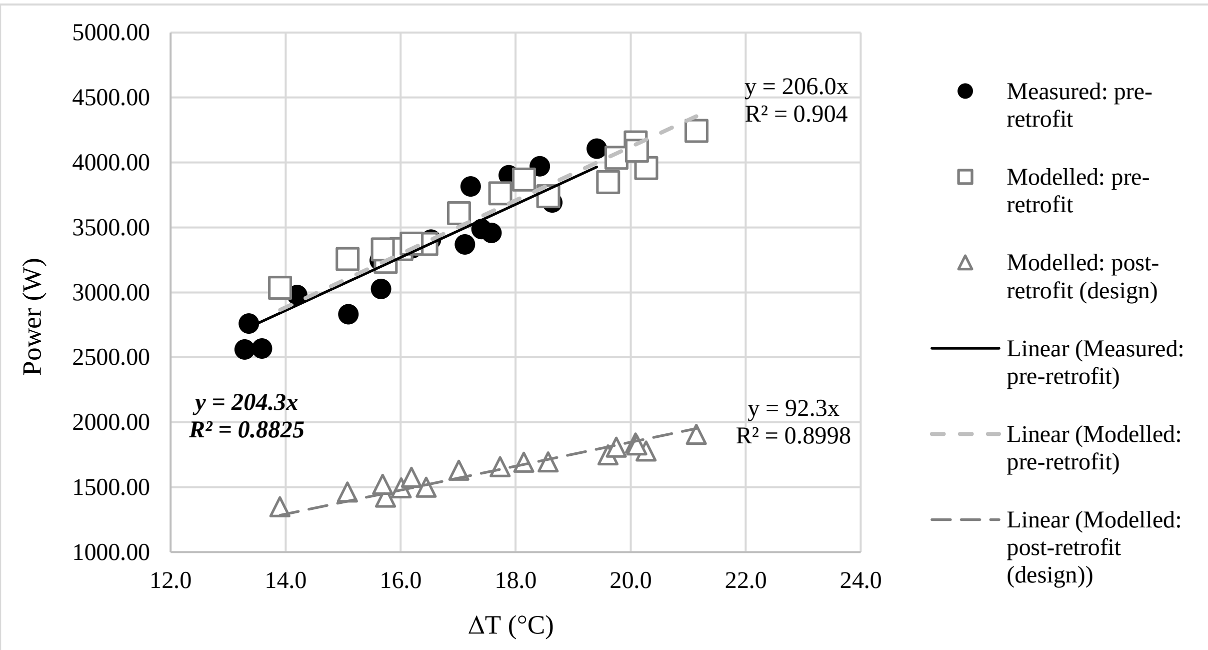


Fig. 8. DSM calculated improvement in HTC post retrofit for Dwelling A.

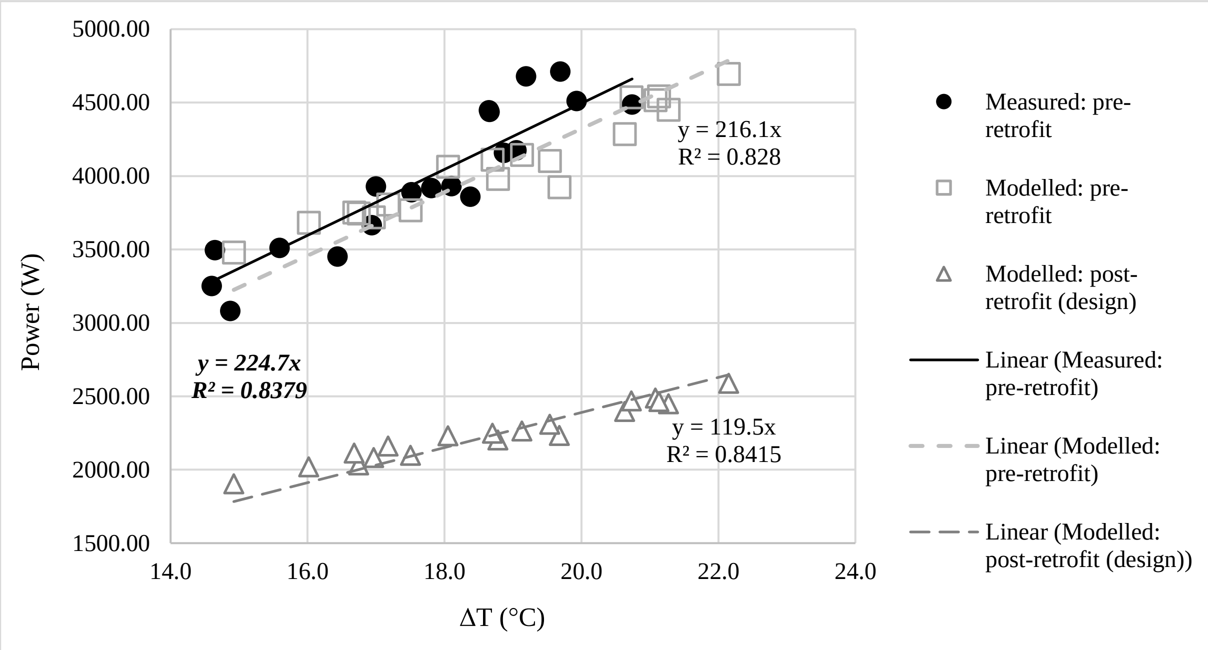


Fig. 9. DSM calculated improvement in HTC post retrofit for Dwelling B.

The estimated savings produced using the calibrated pre-retrofit dynamic models, using the design specifications for the retrofitted fabric, predict reductions in HTCs of 55% and 45% for Dwellings A and B respectively. These results are again similar to the proportionate savings calculated using SAP and measured *in situ*. The absolute savings are though closer to the measured values. Being able to use proportionate estimates of achievable improvements for this type of retrofit can be useful if the baseline fabric performance of a dwelling is known. The inferred proportionate improvement in HTC using the models in this specific case study is 50% ± 5%, which is very similar the average measured savings of 47.5% ± 4.5%. However, if the baseline performance of a building is not known, using these proportionate values to predict absolute savings can be inaccurate as noted below and described further in section 4.4.

## 4.3 Post-retrofit model calibration

In the final stage of the modelled HTC studies, the previously described model calibration methodology was applied to the post-retrofit DSMs. U-values, thermal bridging and infiltration inputs were all updated to match the *in situ* measurements of performance. The results from this exercise are shown in Fig. 10 and 11. Within Fig. 10, the descriptions of the models for Dwelling A refer to:

- the version of the model based upon the design values for the retrofit, noted as 'Modelled: post-retrofit (design)';
- the version of the model calibrated using the measured values but without an adjustment to the air change rates, noted as 'Modelled: post-retrofit (@ 100% AC/H)';
- and the version of the model that includes the adjustment to the air change rate due to leakage through the party wall, noted as 'Modelled: post-retrofit (@ 75% AC/H).'

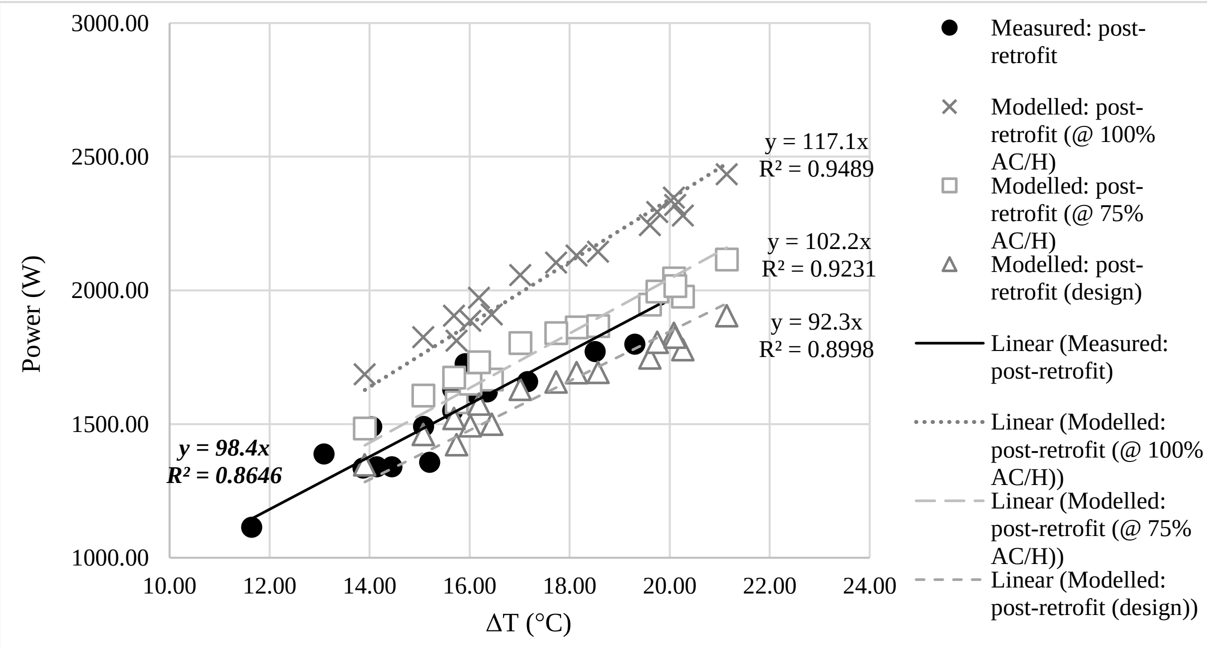


Fig. 10. Measured and modelled HTC values for Dwelling A post retrofit

Calibration of the model for Dwelling A required one further update following the adjustments for measured U-values, and thermal bridging. The air change rate was again reduced by 25% to account for the air leakage in to neighbouring dwellings through the no-fines party walls. As with the pre-retrofit calibrated model, this again produced a more accurate prediction of the HTC. Before this was adjusted, the dynamic model estimated an HTC of 117.1 W/K. Following the adjustment, an HTC 102.2 W/K was predicted, 3.7 W/K higher than the measured value.

Dwelling B was retrofitted with a novel dynamic insulation system, which recovers heat through a very small cavity. The cavity is made up from triangular grooves that are up to 5mm deep in each panel, with the grooves being perpendicular to the opposite panel once installed. The cavity is too narrow to model explicitly within the type of DSM software used in this work. The resultant recovered heat is delivered into the dwelling through a low-speed positive input ventilation system. It is therefore necessary to account for the effects of this dynamic insulation system through the adjustment of pertinent model inputs. In SAP, the wall U-value is modified using the bespoke SAP Appendix Q methodology approved to account for the effects of the dynamic insulation system. This supplementary calculation results in an adjusted U-value for the dynamic insulation system of 0.17 $W/m^2 \cdot K$ in the SAP calculation. The DSM calculations used the same means of evaluating this type of system to account for its effects on predicted thermal performance. In Fig. 11, the version of the model incorporating the effects of the dynamic insulation (DI) system is noted as ‘Modelled: post-retrofit (calibrated DI).’

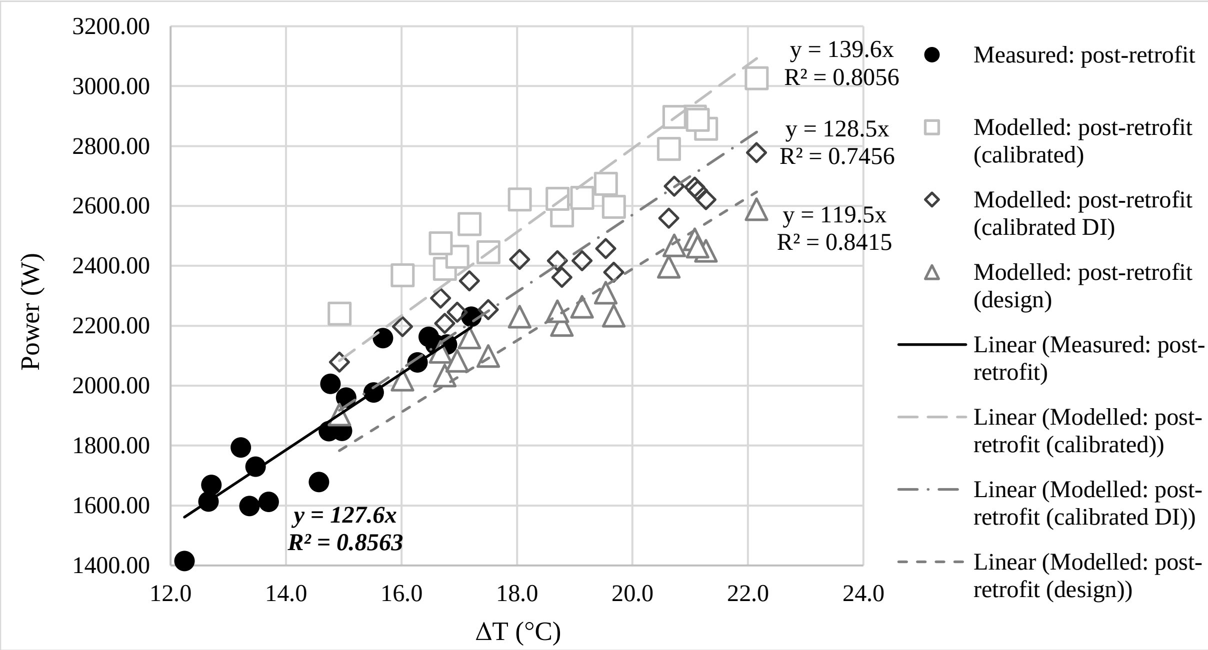


Fig. 11. Measured and modelled HTC values for Dwelling B post retrofit

The results shown in Fig. 11 also help to quantify the impact that the dynamic insulation system had under the test conditions. The 'Modelled: post-retrofit (calibrated)' model was used to calculate the HTC when all inputs are adjusted to match those measured on-site. When the model is adjusted to meet these inputs, it results in an HTC of 139.6 W/K. It was then possible to adjust model inputs further to match the measured HTC and therefore account for the dynamic insulation effects. In keeping with the SAP Appendix Q methodology, the approach involved adjusting the wall U-value to account for any recovered heat. A reduced wall U-value of 0.34 W/m$^2$·K resulted in a HTC that closely matched the HTC measured *in situ*. This adjusted U-value is marginally lower than the U-value measured *in situ* (0.38 W/m$^2$K) and is within the bounds of measurement error. However, it does not include any additional adjustment made for thermal bridging (the wall U-value used in the calibrated model was 0.445 W/m$^2$·K which is based upon the measured U-value, 0.380 W/m$^2$·K, and includes an adjustment for the calculated y-value of 0.065 W/m$^2$·K). The *in situ* physical performance of the dynamic insulation system will be subject to more detailed investigation in a future publication.

Overall, when using the calibrated DSM models to predict retrofitted performance based purely on fabric specification, the estimated proportionate improvements in HTC are marginally higher than the measured improvement and those predicted using SAP. It is, however, the magnitude of the predicted absolute energy and cost savings for this type of building and retrofit that could prove to be more problematic, as is discussed in sections 4.4 and 5, and in the conclusions of this paper.

### 4.4 Post-retrofit modelled annual energy performance

Once calibrated, the post-retrofit dynamic models were used to predict annual energy consumption and to compare predicted internal heat gains and heat losses with those estimated in the 'As-built' and 'Calibrated' SAP calculations.

The primary focus of this work is to understand the differences between the space heating energy consumption, although the electricity consumption does have an impact on the space heating demand, due to the internal heat gains associated with equipment consumption. The equipment and lighting consumption and resultant heat gains in the DSM models were adjusted to match the totals in SAP to provide a fair comparison. Heat gains predicted by both the DSM models and the SAP models from solar irradiation, occupants, lighting and equipment are compared for both Dwellings in Fig. 12.

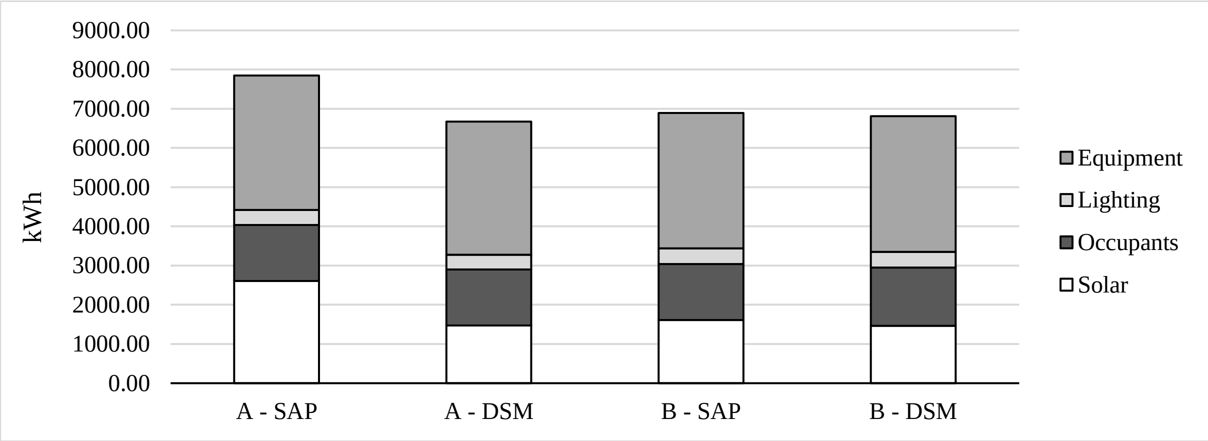


Fig. 12. Comparison of heat gains predicted in SAP and DSM models

This simple comparison of the internal heat gains shows that SAP includes higher solar heat gains than those in the DSM models. Intuitively, given that the remaining heat gains are consistent between all models, this would lead to a higher estimate of the space heating energy consumption in the DSM models than that predicted in the SAP calculations. However, this is not the case. Although modelled internal heat gains are lower in the DSM calculations, the calibrated models also incorporate lower HTC values than either the As-built or calibrated versions of the SAP calculations. It is also important to note that the hourly time-step dynamic models include internal heat gains during the hours that space heating is required, which will also help to off-set space heating energy consumption. Consequently, the predicted space heating and energy consumption are lower in the DSM models than those predicted in SAP, as can be seen in Fig. 13, which illustrates the predicted space heating demand and heat losses.

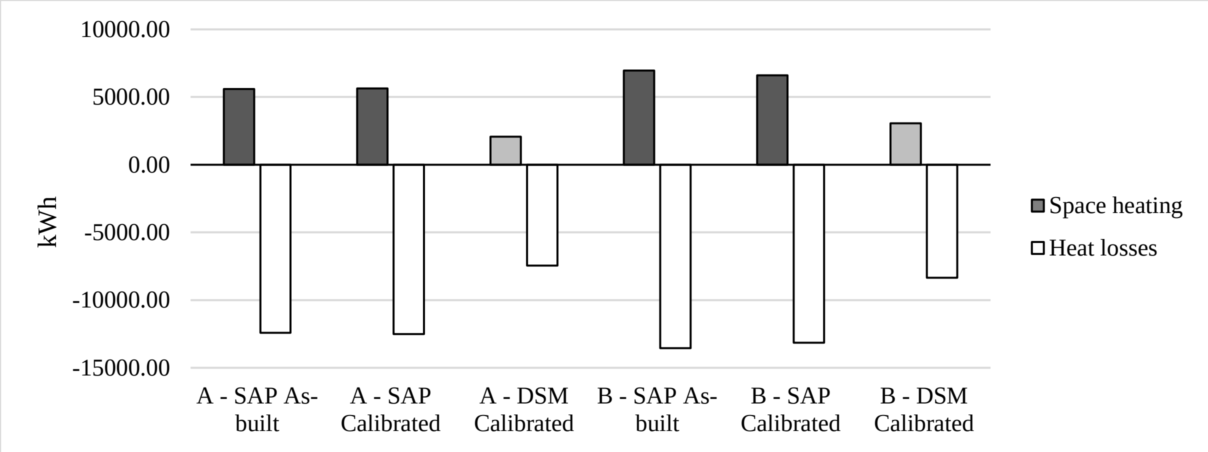


Fig. 13. Comparison of annual space heating energy consumption predicted in SAP and DSM models

As can be inferred from the solar gains shown in Fig. 12, the external environmental conditions assumed for the calculations also differ. It is worth noting that Dwelling B is orientated almost directly south, and the SAP solar gains are similar to those in the DSM model. Dwelling A is orientated at 245°, approximately south-west facing. The SAP calculation uses central UK weather data (for East Pennines) to ensure consistency between comparison for generating the EPC, but the regional weather data for Northern Ireland was used in calculating running costs. The DSM model uses a simulation weather file specific to Belfast. As a regional weather file is used in SAP, the monthly average temperatures in this instance are very similar as can be seen from the chart in Fig. 14. It is therefore unlikely that the ΔT contributes significantly to the differences in prediction.

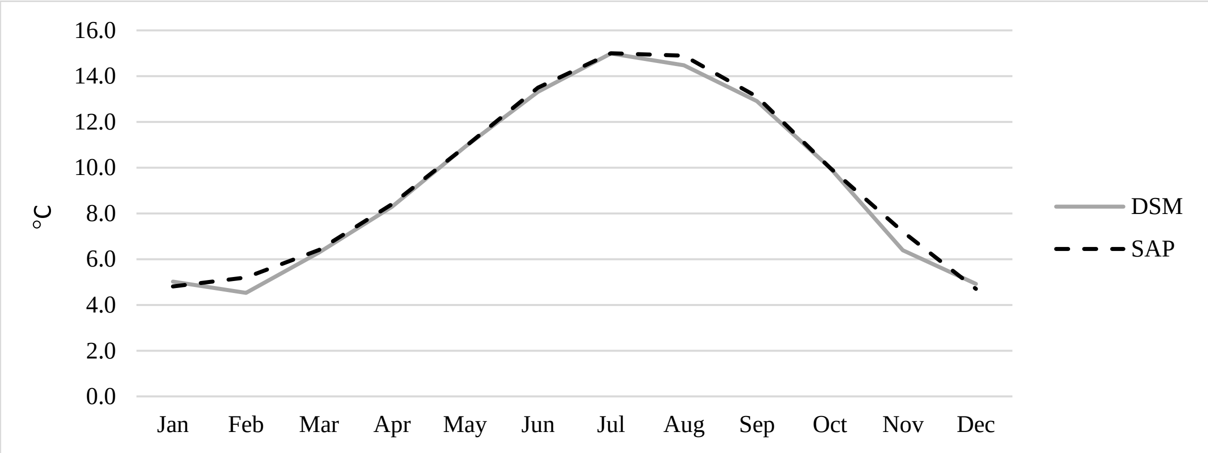


Fig. 14. Comparison of monthly average external dry-bulb temperature used in SAP and DSM models

Over-prediction of absolute energy savings, and therefore cost savings, has potential repercussions for building owners/operators. This is especially pertinent if the forecast savings have been included in any payback calculations, like those included in the now defunct Green Deal policy [34]. The chart in Fig. 15 illustrates the difference in cost savings predicted as a result of the retrofit measures applied to the two case study buildings. The cost savings for the SAP predictions are taken directly from the SAP calculation outputs, whilst the cost savings for the DSM models were calculated following the same methodology, using the same fuel cost data and standing charges [100]. Most pertinent to this work is the cost savings associated with space heating. Using a hypothetical retrofit cost of £7,000 per dwelling, and the space heating savings predicted by the As-built SAP, the simple payback period would be 22.4 years and 23.7 years for Dwellings A and B respectively; these periods increase when using the calibrated SAP calculations, to 26.1 years and 28.4 years respectively. Based upon the DSM calculations, these would be extended further to 39.6 years and 34.6 years, respectively. If the exercise is repeated based upon total energy savings, the periods for the As-built SAP would be 17.5 years and 19.9 years and the periods for the calibrated SAP would be 19.7 and 23.1 years. The periods for the

DSM calculations would be 26.6 years and 27 years. There are elemental differences between the two dwellings as have been noted throughout this paper. However, within this case study context, the difference between predicted payback periods for this type of retrofit is relatively low and within ± 0.2 years of the inferred average period of 26.8 years.

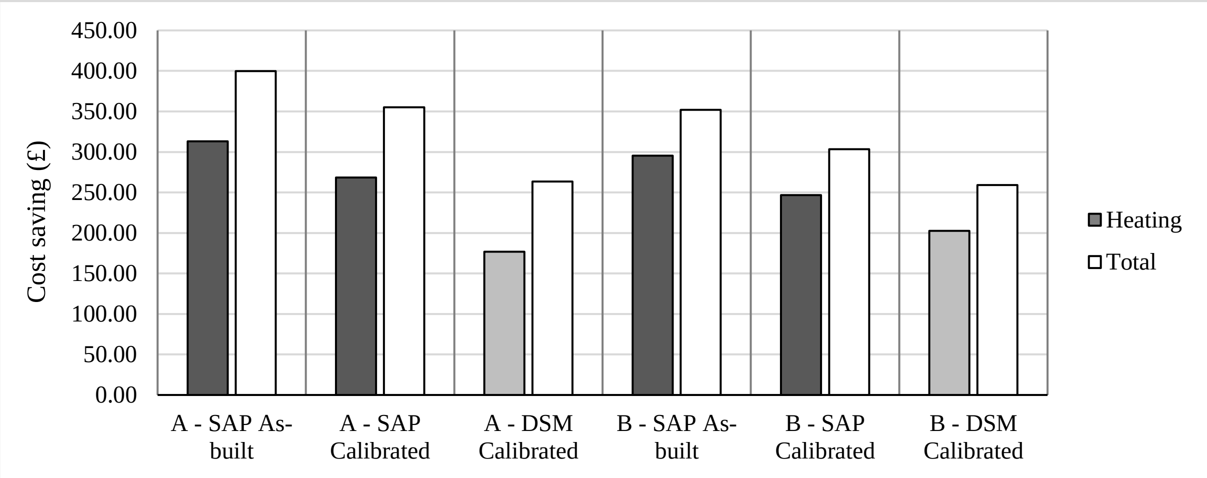


Fig. 15. Comparison of predicted annual fuel cost saving post retrofit SAP and DSM models

## 5. DISCUSSION

This paper has illustrated that by using the building HTC as the qualified metric, *in situ* measured fabric performance data can be successfully used to calibrate pre- and post-retrofit DSM models without the need for a costly site and time specific detailed simulation weather file. This means that more confidence can be placed in calibrated baseline models to predict the fabric performance of retrofitted buildings. Although this research utilises a case study methodology and a small sample, this helps to further validate this calibration method, and to consolidate previous work by the authors and other researchers who have employed a similar methodology [31, 51]. In this instance, it was also possible to calibrate these models when quantifying the improvement in performance of an innovative dynamic insulation system under real conditions. It should be noted that this followed the same method of accounting for this as is found in the SAP calculations. This does however demonstrate that this methodological approach could be useful for investigating unusual building characteristics and technologies that affect thermal performance.

The required input *in situ* data included measured U-values, calculated and measured air change rate; modelled thermal bridging values were also used. Of all of these input parameters, the U-value measurements showed the greatest deviation from the assumed designed values, in the particular case studies chosen. However, not all of the deviation in performance could be attributed to differences in the U-values. Further analysis revealed that the air change rates used by the models was also important. When revised air change rates were introduced into the models, the predicted performance of the DSM models more closely matched the *in situ* measured data. It is reasonable to draw this conclusion from the observed phenomenon and subsequent model calibration; however, it is not possible to quantify the impact of the phenomenon due to the case study context and limited sample. Results presented here suggest that retrofits of this type can achieve improvements in HTC of approximately 45%. However, rather than demonstrate the homogeny of retrofit for this type of dwelling, the results have emphasised the idiosyncratic nature of existing hard-to-treat buildings and the design, modelling and implementation of retrofit measures.

Although treated in calculations as solid walls, the no-fines structure includes a network of small voids, which supported the descriptions found in the literature. There was some indicative evidence of this phenomenon during thermography surveys. These voids collectively act as a form of cavity, therefore improving the elemental U-value. However, on exposed walls, this also aids the transfer and loss of heat from the dwelling as a whole, which will be quantified in future work. During later *in situ* testing following the retrofit of EWI, co-pressurization tests identified a significant reduction in measured air change rates, when compared to isolated single dwelling tests, which indicates a relatively large amount of inter-dwelling air movement. It has been demonstrated through the modelling exercises included in this work that the difference between the results recorded from isolated air pressurization tests and those under co-pressurization, has a significant impact on the forecast heat loss and, ultimately, the predicted energy and cost savings gained through the retrofit fabric upgrades. The calibration of the DSM model for Dwelling A was reliant on using the reduced air change rates that would be expected to occur under normal operating conditions.

The issue related to air change rates was also compounded by the high heat loss estimates included in SAP, despite the higher heat gain values assumed in the calculations. It has however been demonstrated that the *in situ* values can also be used to refine the estimates produced by SAP as well. Absolute fuel cost savings were however higher than those predicted by the calibrated DSM models for both versions of the SAP calculations. Although the results presented here are specific to Belfast, the proportionate savings are expected to be similar in other areas of the UK. This is however specific to the case study context of the work evaluating EWI retrofitted to no-fines concrete constructions. However, based upon the inferred proportionate savings, absolute savings would be likely to increase slightly in colder regions and reduce slightly in the warmer areas. It is also crucial to note that RdSAP is most commonly used for existing buildings and the limitations in this calculation method would exacerbate these issues further. In practice, this would all be problematic if retrofit measures are financed on an estimated payback period, which could become much longer in reality than that accounted for at the retrofit design stage when using any version of SAP to forecast savings for specific dwellings.

The combined methods of measurement and modelling described in this work could be employed by large stock owners to improve the efficacy of their retrofit programmes. For instance, a sample of a particular dwelling archetype would undergo *in situ* testing to understand the intricacies of their thermal performance. The results would then be used to calibrate dynamic simulation models, which in turn would be used to evaluate a wide range of retrofit options and identify the most effective measures. Modelled results could then be extrapolated across the entire stock for this archetype. There would of course be initial costs associated with this exercise, but the potential long-term savings for all stakeholders could be significant. If there were a commitment at governmental level to retrofit existing homes to make them more thermally efficient, an approach of this nature could be used to underpin policy implementation. This would be particularly significant to large social housing stocks and could help towards addressing issues of on-going fuel poverty.

The results also confirm existing understanding that internal heat gains associated with occupancy, human operation of HVAC plant performance and natural ventilation, represent much more variable parameters than the fabric elements. Although the DSM models could also have been calibrated to accurately represent the impact of these variables when in-use, this would have required further refining of the model inputs to align with a detailed set of monitored data. The main value of this type of additional exercise would be to further validate the accuracy of the DSM models and to verify their value when predicting the performance of more complex and interactive retrofit packages. This type of model calibrated against in-use data would also be useful for overheating analysis and investigations that consider the impact of behavioural change. In practice, SAP operates well as a benchmarking tool and there would be limited value in trying to refine each iteration of a SAP calculation to match the *in situ* performance, as it would not be fit for purpose.

## 6. CONCLUSION

Results presented in this paper emphasise the importance of understanding the fabric performance of dwellings when specifying deep retrofit measures. They also demonstrate that *in situ* testing can greatly improve the accuracy of the modelling used to validate these design choices. The work reported here, at this point, is based upon a small and non-random sample and more research will be required to further validate this calibration methodology, ideally including a wider range of dwelling types and forms and fabric specifications. There is further scope for comparing the SAP inputs and outputs to increase the accuracy of performance estimates, especially in relation to deep retrofit projects. It is however a reliable example of proof of concept and demonstrates the repeatable nature of the methodology.

The results are of general interest to researchers working in the field of model calibration and validation and can contribute towards the development of building energy performance assessment methods based upon *in situ* measurements. This methodology is well suited to large-scale retrofit programmes. By mimicking the coheating test conditions, the methodology could also be used to calculate HTCs for new buildings where only U-values and air tightness have been measured. This would be particularly useful if U-values could be measured more rapidly in the future and has the potential for application to sample sets of dwellings across large new-build developments, to help ensure the thermal performance of dwellings matches the design intent.

## ACKNOWLEDGEMENTS

The fabric testing element of this work was funded through the Innovate UK Scaling up Retrofit programme as part of the S-IMPLER (www.s-impler.com) project. The Author's would also like to thank Adrian Blythe and Robert Hopkin of the Northern Ireland Housing Executive and

Davey Wilton formerly of Carillion during the project for their invaluable help throughout the duration of the S-IMPLER project.

## REFERENCES

[1] J. Ravetz, State of the stock-What do we know about existing buildings and their future prospects?, Energy Policy, 36 (12) (2008) 4462-4470.

[2] HM Government, Live tables on dwelling stock including vacants. Table 101: by tenure, United Kingdom, in, HM Government,, London, 2014.

[3] M.J. Kelly, Retrofitting the existing UK building stock, Building Research and Information, 37 (2) (2009) 196-200.

[4] A. Power, Does demolition or refurbishment of old and inefficient homes help to increase our environmental, social and economic viability?, Energy Policy, 36 (12) (2008) 4487-4501.

[5] B. Boardman, Home truths: A low carbon strategy to reduce UK housing emissions by 80% by 2050, in: ECI Research Report 34, University of Oxford Environmental Change Institute, Oxford, 2007.

[6] G. Killip, Building a greener Britain: Transforming the UK's existing housing stock. A report for the Federation of Master Builders by the University of Oxford Environmental Change Institute, in, London, 2008.

[7] DECC, Energy Consumption in the UK, in, The Stationary Office, London, 2014.

[8] HM Government, Improving the energy efficiency of our buildings: A guide to Energy Performance Certificates for the construction, sale and let of dwellings., in: D.f.C.a.L. Governance (Ed.), The Staionary Office, Wetherby, 2008.

[9] HM Government, UK Building Regulations. Part L2A: Conservation of Fuel and Power in Buildings Other than Dwellings in, RIBA Publishing Ltd, London, 2014.

[10] HM Government, UK Building Regulations. Part L1B: Conservation of Fuel and Power in Existing Dwellings in, RIBA Publishing Ltd, London, 2014.

[11] Scottish Government, Technical Handbook 6 - Domestic - Energy, in, Glasgow, 2017.

[12] Northern Ireland Government, Technical Booklet F - Domestic - Energy, in, Belfast, 2017.

[13] R. Cohen, M. Standeven, B. Bordass, A. Leaman, Assessing building performance in use 1: The Probe process, Building Research and Information, 29 (2) (2001) 85-102.

[14] B. Bordass, R. Cohen, M. Standeven, A. Leaman, Assessing building performance in use 2: Technical performance of the Probe buildings, Building Research and Information, 29 (2) (2001) 103-113.

[15] B. Bordass, R. Cohen, M. Standeven, A. Leaman, Assessing building performance in use 3: Energy performance of the Probe buildings, Building Research and Information, 29 (2) (2001) 114-128.

[16] Zero Carbon Hub, Closing the performance gap: Building low carbon housing for real. Report of Topic Work Group 4, Carbon Compliance Tool Policy Assumptions Task Group. , in, Milton Keynes, 2010.

[17] R. Cohen, The performance gap in non-domestic buildings: evidence collected from the Technology Strategy Board's Building Performance Evaluation Programme, in, London, 2013.

[18] D. Johnston, D. Miles-Shenton, D. Farmer, Quantifying the domestic building fabric 'performance gap', Building Services Engineering Research and Technology, 36 (5) (2015) 614-627.

[19] A. Marshall, R. Fitton, W. Swan, D. Farmer, D. Johnston, M. Benjaber, Y. Ji, Domestic building fabric performance: Closing the gap between the in situ measured and modelled performance, Energy and Buildings, 150 (2017) 307-317.

[20] D. Calì, T. Osterhage, R. Streblow, D. Müller, Energy performance gap in refurbished German dwellings: Lesson learned from a field test, Energy and Buildings, 127 (2016) 1146-1158.

[21] N. Kampelis, K. Gobakis, V. Vagias, D. Kolokotsa, L. Standardi, D. Isidori, C. Cristalli, F.M. Montagnino, F. Paredes, P. Muratore, L. Venezia, M.K. Dracou, A. Montenon, A. Pyrgou, T. Karlessi, M. Santamouris, Evaluation of the performance gap in industrial, residential & tertiary near-Zero energy buildings, Energy and Buildings, 148 (2017) 58-73.

[22] S. Doran, B. Carr, Thermal Transmittance of Walls of Dwellings Before and After Application of Cavity Wall Insulation, in: Report prepared for the Energy Saving Trust, Report 222077, Glasgow, 2008.

[23] P. de Wilde, The gap between predicted and measured energy performance of buildings: A framework for investigation, Automation in Construction, 41 (2014) 40-49.

[24] S. Kelly, D. Crawford-Brown, M.G. Pollitt, Building performance evaluation and certification in the UK: Is SAP fit for purpose?, Renewable and Sustainable Energy Reviews, 16 (9) (2012) 6861-6878.

[25] M.F. Jentsch, G.J. Levermore, J.B. Parkinson, M.E. Eames, Limitations of the CIBSE design summer year approach for delivering representative near-extreme summer weather conditions, Building Services Engineering Research and Technology, 35 (2) (2014) 155-169.

[26] J. Palmer, I. Cooper, United Kingdom housing energy fact file 2013. A report prepared under contract to DECC by Cambridge Architectural Research, Eclipse Research Consultants and Cambridge Energy, in, London, 2013.

[27] HM Government, Digest of United Kingdom Energy Statistics (DUKES), in: D.o.E.a.C. Change (Ed.), DECC, London, 2012.

[28] D. Johnston, J. Wingfield, D. Miles-Shenton, Measuring the fabric performance of UK dwellings, in: Association of Researchers in Construction Management, ARCOM 2010 - Proceedings of the 26th Annual Conference, 2010, pp. 1371-1380.

[29] D. Johnston, D. Farmer, M. Brooke-Peat, D. Miles-Shenton, Bridging the domestic building fabric performance gap, Building Research and Information, 44 (2) (2016) 147-159.

[30] D. Johnston, M. Siddall, The building fabric thermal performance of passivhaus dwellings-Does it do what it says on the tin?, Sustainability (Switzerland), 8 (1) (2016) 1-14.

[31] J.M. Parker, D. Farmer, M. Fletcher, Calibrating whole house thermal models against a coheating test, in: System Simulation in Buildings 2014 Proceedings of the Ninth International Conference, December 10-12, 2014., Atelier des Presses, Liege, 2015, pp. 211-219.

[32] HM Government, Live tables on Energy Performance of Buildings Certificates, in, HM Government,, 2018.

[33] HM Government, The Government's Standard Assessment Procedure for Energy Rating of Dwellings, 2012 Edition., in: DECC (Ed.), Building Research Establishment Ltd, Watford, 2014.

[34] HM Government, Standard Assessment Procedure: Guidance on how buildings will be SAP energy assessed under the Green Deal and on recent changes to incentivise low carbon developments., in: DECC (Ed.), HMSO, London, 2013.

[35] HM Government, UK Building Regulations. Part L1A: Conservation of Fuel and Power in New Dwellings in, RIBA Publishing Ltd, London, 2014.

[36] The Carbon Trust, Closing the gap: lessons learned on realising the potential of low carbon building design, in, London, 2011.

[37] A.C. Menezes, A. Cripps, D. Bouchlaghem, R. Buswell, Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap, Applied Energy, 97 (2012) 355-364.

[38] K. Subbarao, PSTAR - Primary and Secondary Terms-Analysis and Renormalization: A Unified Approach to Building and Energy Simulations and Short-Term Testing – A Summary. , in: SERI/TR-254-3347, Solar Energy Research Institute, Colarado, USA, 1988.

[39] K. Subbarao, J. Burch, C.E. Hancock, How to accurately measure load coefficients of residential buildings, in, 1990, pp. 419-425.

[40] K. Subbarao, J. Burch, C.E. Hancock, A. Lekov, J.D. Balcomb, Short-Term Energy Monitoring (STEM): Application of the PSTAR Method to a Residence in Fredericksburg, Virginia., in: SERI TR-3356, Solar Energy Research Institute, Colorado, USA, 1988.

[41] R.C. Sonderegger, M.P. Modera, Electric co-heating: A method for evaluating seasonal heating efficiencies and heat loss rates in dwellings, in: Second International CIB Symposium, Energy Conservation in the Built Environment, LBL-8949, Copenhagen, 1979.

[42] R.C. Sonderegger, P.E. Condon, M.P. Modera, In-situ measurements of residential energy performance using electric co-heating, ASHRAE Transactions, 86 (Pt) (1980) 394-407.

[43] J.K.E. Ortega, J.V. Anderson, J.M. Connolly, C.E. Bingham, Electric coheating experiment to determine the heat loss coefficient of a double-envelope house, in, 1981, pp. 74-78.

[44] R.H. Socolow, The twin rivers program on energy conservation in housing: Highlights and conclusions, Energy and Buildings, 1 (3) (1978) 207-213.

[45] J. Palmer, G. Pane, M. Bell, J. Wingfield, Comparing Primary and Secondary Terms Analysis and Re-Normalisation (PSTAR) Test and Co-Heating Test Results; A Report to the Department for Communities and Local Government, in, London, 2011.

[46] D. Johnston, D. Miles-Shenton, J. Wingfield, D. Farmer, M. Bell, Whole house heat loss test method (Coheating). Part of: International Energy Agency Annex 58: Reliable building energy performance characterization based on full scale dynamic measurements. , in, Leeds Metropolitan University, Leeds, 2012.

[47] D. Johnston, D. Miles-Shenton, D. Farmer, J. Wingfield, Whole House Heat Loss Test Method (Coheating). in, Leeds Metropolitan University, Leeds, 2013.

[48] A. Stafford, D. Johnston, D. Miles-Shenton, D. Farmer, M. Brooke-Peat, C. Gorse, Adding value and meaning to coheating tests, Structural Survey, 32 (4) (2014) 331-342.

[49] E. Mangematin, G. Pandraud, D. Roux, Quick measurements of energy efficiency of buildings, Comptes Rendus Physique, 13 (4) (2012) 383-390.

[50] R. Bouchie, F. Alzetto, A. Brun, P. Boisson, S. Thebault, Short Methodologies for In-situ Assessment of the Intrinsic Thermal Performance of the Building Envelope, in: Sustainable Places, Nice, 2014.

[51] S. Stamp, R. Lowe, H. Altamirano-Medina, An investigation into the role of thermal mass on the accuracy of co-heating tests through simulations & field results, in: Proceedings of BS 2013: 13th Conference of the International Building Performance Simulation Association, 2013, pp. 39-46.

[52] D. Fosas, D.A. Coley, S. Natarajan, M. Herrera, M. Fosas de Pando, A. Ramallo-Gonzalez, Mitigation versus adaptation: Does insulating dwellings increase overheating risk?, Building and Environment, 143 (2018) 740-759.

[53] G. Sousa, B.M. Jones, P.A. Mirzaei, D. Robinson, A review and critique of UK housing stock energy models, modelling approaches and data sources, Energy and Buildings, 151 (2017) 66-80.

[54] P.X.W. Zou, X. Xu, J. Sanjayan, J. Wang, Review of 10 years research on building energy performance gap: Life-cycle and stakeholder perspectives, Energy and Buildings, 178 (2018) 165-181.

[55] F. Heesen, R. Madlener, Consumer behavior in energy-efficient homes: The limited merits of energy performance ratings as benchmarks, Energy and Buildings, 172 (2018) 405-413.

[56] B. Frei, C. Sagerschnig, D. Gyalistras, Performance gaps in Swiss buildings: An analysis of conflicting objectives and mitigation strategies, in: Energy Procedia, 2017, pp. 421-426.

[57] P. van den Brom, A. Meijer, H. Visscher, Performance gaps in energy consumption: household groups and building characteristics, Building Research and Information, 46 (1) (2018) 54-70.

[58] T. Reddy, Literature review on calibration of building energy simulation programs: uses, problems, procedures, uncertainty and tools, ASHRAE Transactions, (211) (2006) 226-240.

[59] D. Coakley, P. Raftery, M. Keane, A review of methods to match building energy simulation models to measured data, Renewable and Sustainable Energy Reviews, 37 (2014) 123-141.

[60] E. Fabrizio, V. Monetti, Methodologies and advancements in the calibration of building energy models, Energies, 8 (4) (2015) 2548-2574.

[61] P. Raftery, M. Keane, J. O'Donnell, Calibrating whole building energy models: An evidence-based methodology, Energy and Buildings, 43 (9) (2011) 2356-2364.

[62] P. Raftery, M. Keane, A. Costa, Calibrating whole building energy models: Detailed case study using hourly measured data, Energy and Buildings, 43 (12) (2011) 3666-3679.

[63] S. Pelsmakers, B. Croxford, C.A. Elwell, Suspended timber ground floors: measured heat loss compared with models, Building Research and Information, 47 (2) (2019) 127-140.

[64] T. Reddy, I. Maor, C. Panjapornpon, Calibrating detailed building energy simulation programs with measured data – Part I: General methodology, HVAC&R Research, 13 (2) (2007) 221-241.

[65] T. Reddy, I. Maor, C. Panjapornpon, Calibrating detailed building energy simulation programs with measured data – Part II: Application to three case study office buildings, HVAC&R Research, 13 (2) (2007) 243-265.

[66] Y. Heo, R. Choudhary, G.A. Augenbroe, Calibration of building energy models for retrofit analysis under uncertainty, Energy and Buildings, 47 (2012) 550-560.

[67] J. Yoon, E.J. Lee, D.E. Claridge, Calibration procedure for energy performance simulation of a commercial building, Journal of Solar Energy Engineering, Transactions of the ASME, 125 (3) (2003) 251-257.

[68] Y.S. Kim, M. Heidarinejad, M. Dahlhausen, J. Srebric, Building energy model calibration with schedules derived from electricity use data, Applied Energy, 190 (2017) 997-1007.

[69] D.E. Claridge, Building simulation for practical operational optimization, in: Building Performance Simulating for Design and Operation, 2012, pp. 365-401.

[70] A.E. Kandil, J.A. Love, Signature analysis calibration of a school energy model using hourly data, Journal of Building Performance Simulation, 7 (5) (2014) 326-345.

[71] J.A. Shonder, P.J. Hughes, J.W. Thornton, Using calibrated engineering models to predict energy savings in large-scale geothermal heat pump projects, in: ASHRAE Transactions, 1998, pp. 944-954.

[72] V.I. Soebarto, Calibration of hourly energy simulations using hourly monitored data and monthly utility records for two case study buildings, in: International Building Performance Simulation Conference, IBPSA, Prague, 1997.

[73] J.M. Parker, P. Cropper, L. Shao, A calibrated whole building simulation approach to assessing retrofit options for Birmingham Airport, in: IBPSA England First Building Simulation and Optimization Conference, IBPSA, Loughborough, 2012, pp. 49-56.

[74] E. Burman, D. Rigamonti, J. Kimpain, D. Mumovic, Performance gap and thermal modelling: a comparison of simulation results and actual energy performance for an academy in North West England, in: IBPSA England First Building Simulation and Optimization Conference, IBPSA, Loughborough, UK, 2012, pp. 35-42.

[75] Z. Mylona, M. Kolokotroni, S.A. Tassou, Frozen food retail: Measuring and modelling energy use and space environmental systems in an operational supermarket, Energy and Buildings, 144 (2017) 129-143.

[76] R. Raslan, M. Davies, Results viability in accredited building energy performance compliance demonstration software in the UK: an inter-model comparative study, Journal of Building Performance Simulation, 3 (1) (2010) 63-85.

[77] S. Imam, D.A. Coley, I. Walker, The building performance gap: Are modellers literate?, Building Services Engineering Research and Technology, 38 (3) (2017) 351-375.

[78] CIBSE, TM33: Tests for Software Accreditation & Verification, in, CIBSE, London, 2006.

[79] ASHRAE, Standard 140-2007: Standard method of test for the evaluation of building energy analysis computer programs, in, ASHRAE, Atlanta, 2007.

[80] J.K. Moss, No-fines building gives energy-conserving homes, in: International Construction, The Aberdeen Group, Sutton, 1979.

[81] A.W. Williams, G.C. Ward, The renovation of no-fines housing: A guide to the performance and rehabilitation of load bearing no-fines concrete dwellings built using the Wimpey and Scottish housing association systems, in, Building Research Establishment, Watford, 1991.

[82] J. Sommerville, N. Craig, A. Charles, No-fines concrete in the UK social housing stock: 50 years on, Structural Survey, 29 (4) (2011) 294-302.

[83] N. Craig, J. Sommerville, A. Charles, No-fines concrete homes: Atypical thermal performances, Structural Survey, 31 (1) (2013) 43-55.

[84] J. Hulme, S. Doran, In-situ measurements of wall U-values in English housing, in, Watford, 2014.

[85] S. Doran, Field Investigations of the Thermal Performance of Construction Elements As Built, in: DETR framework project report, Glasow, 2001.

[86] P. Baker, Technical Paper 10 – U-values and Traditional Buildings, in, Edinburgh, 2011.

[87] C. Ahern, B. Norton, B. Enright, The statistical relevance and effect of assuming pessimistic default overall thermal transmittance coefficients on dwelling energy performance certification quality in Ireland, Energy and Buildings, 127 (2016) 268-278.

[88] A. Marshall, J. Francou, R. Fitton, W. Swan, J. Owen, M. Benjaber, Variations in the U-value measurement of a whole dwelling using infrared thermography under controlled conditions, Buildings, 8 (3) (2018).

[89] R. Jack, D. Loveday, D. Allinson, K. Lomas, First evidence for the reliability of building co-heating tests, Building Research and Information, 46 (4) (2018) 383-401.

[90] ISO, ISO 9869: Thermal insulation – Building elements - In situ measurement of thermal resistance and thermal transmittance, in, International Organization for Standardisation, Geneva, 2014.

[91] ATTMA, ATTMA Technical Standard L1. Measuring the Air Permeability of Building envelopes (Dwellings), in, Air Tightness Testing and Measurement Association, Northampton, 2010.

[92] M.H. Sherman, Estimation of infiltration from leakage and climate indicators, Energy and Buildings, 10 (1) (1987) 81-86.

[93] BSI, BS EN ISO 9972:2015. Thermal performance of buildings – determination of air permeability of buildings – fan pressurization method, in, British Standards Institution, London, 2015.

[94] DesignBuilder Software Ltd, DesignBuilder Version 5.0.2.003, in, DesignBuilder Software Ltd, Stroud, UK, 2016.

[95] CIBSE, Belfast Test Reference Year building simulation weather file, in, CIBSE, London, 2016.

[96] BSI, BS EN ISO 6946: Building components and building elements. Thermal resistance and thermal transmittance. Calculation method., in, BSI, London, 2008.

[97] CIBSE, Guide A: Environmental Design, in, London, 2006.

[98] HM Government, National Calculation Methodology (NCM) modelling guide (for building other than dwellings in England and Wales), in, BRE Ltd, London, 2013.

[99] B. Jones, P. Das, Z. Chalabi, M. Davies, I. Hamilton, R. Lowe, J. Milner, I. Ridley, C. Shrubsole, P. Wilkinson, The effect of party wall permeability on estimations of infiltration from air leakage, International Journal of Ventilation, 12 (1) (2013) 17-29.

[100] HM Government, The Government's Standard Assessment Procedure for Energy Rating of Dwellings: 2012 Edition, in: DECC (Ed.), BRE, Watford, 2014.