UK Government

GHG SMETER Project Executive Summary

The application of SMETERs to determine the change in thermal performance of homes retrofitted through the Green Homes Grant

August 2025

Executive Summary of GHG-SMETER Project Final Report

Overview and Key Findings

- Project Scope: Assessed thermal performance changes in homes retrofitted under the Green Homes Grant (GHG) scheme using Smart Meter Enabled Thermal Efficiency Rating (SMETER) methods.
- Sample Size:
    - 2,455 homes from GHG Voucher (GHG-V) scheme.
    - 105 homes from GHG Local Authority Delivery (GHG-LAD) scheme.
    - Control group of 13,000+ homes from the Smart Energy Research Laboratory.
- Retrofit Measures Studied: Included loft, cavity wall, external wall, roof, and underfloor insulation; solar thermal; heat pumps; and secondary measures, such as glazing and door upgrades.
- SMETER Types Used:
    - Type A: Remote-only data (e.g., smart meter + weather).
    - Type B: Includes internal temperature data (more accurate but more intrusive).
- Thermal Performance Metric: In-use Heat Transfer Coefficient (HTC) used to represent real-world heat loss, incorporating building fabric, technologies, and occupant behavior.
- Plausibility Rates:
    - Type A SMETERs: 54–55%.
    - Type B SMETERs: 66–69%.
    - Higher plausibility linked to larger, older, lower EPC-rated homes with regular heating patterns.
- Change in HTC Post-Retrofit: Type A SMETERs estimated an average reduction of ~20 W/K (from ~250 W/K baseline). EPC-data Model predicted ~60 W/K reduction, likely overestimated due to assumptions and lack of occupant behaviour consideration.
- Model Comparison: SMETERs consistently estimated lower in-use HTCs than EPC-based models, which highlights EPCs' tendency to overestimate energy use, especially in lower EPC bands.
- Limitations:
    - SMETERs less accurate for underheated homes or those with unmetered heating.
    - Type A methods not suitable for evaluating individual homes but effective for cohorts (≥50 homes).
- Recommendations:
    - SMETERs are promising for large-scale retrofit evaluation and policy support.
    - Need for consistent methodology pre- and post-retrofit.
    - Development of standardised data sharing and validation protocols is essential.

Introduction

This summary provides an overview of the main findings and methods of the GHG-SMETER project, which applied Smart Meter Enabled Thermal Efficiency Rating (SMETER) methods to homes retrofitted through the Green Homes Grant (GHG) Voucher (GHG-V) and Local Area Delivery (GHG-LAD) schemes. The project investigated the performance of properties retrofitted through the GHG and the ability of SMETER methods to assess their performance and change in performance on retrofit.

In the GHG-SMETER project, the in-use Heat Transfer Coefficient (HTC) was estimated to represent the thermal performance of the homes. The HTC is the rate at which heat is lost from a building per degree temperature difference between the inside and outside spaces. The in-use HTC, assessed here with SMETER technologies, includes the thermal performance from all aspects of the occupied home, including the building fabric, installed technologies (e.g. ventilation) and how the home is operated (including occupant practices e.g. ventilation use).

The in-use HTC contrasts to thermal performance testing, which uses standardised procedures, such as the new Aggregate Heat Loss Test (BS EN 17887-2:2024) that focuses on thermal performance of the fabric without deliberate ventilation, with fans to distribute air within the home and using tester-installed electric heaters. Similarly, in-use performance measurement is expected to produce different results to incumbent survey and modelling methods (such as the Standard Assessment Procedure, SAP) because it accounts for the actual performance of the home in its measured state and occupant practices, rather than assumed thermal performance of building elements and standardised occupant factors.

Recruitment and sample

2455 homes were recruited from the GHG-V scheme, with 105 recruited from GHG-LAD. All participants completed a survey detailing the technologies within their homes, their practices and the fabric of the home. All participants also provided consent for their smart meter data to be accessed and analysed for the purpose of this research. Participants in the GHG-LAD sample additionally self-installed four temperature sensors, provided by the research team, according to supplied instructions. All research was completed according to appropriate data security and privacy arrangements, and following a process approved by the Ethics Committee at University College London (UCL).

The most commonly installed retrofit in the recruited sample from the GHG-V scheme was loft insulation, with other fabric measures including cavity wall insulation, external wall insulation, roof insulation and underfloor insulation. Solar thermal had the highest uptake of the low carbon heating measures, followed by heat pumps, with fewer secondary measures such as glazing and door upgrades in the sample.

In comparison to a representative sample of the building stock, taken to be the English Housing Survey (EHS) 2021, the recruited sample from the GHG-V and GHG-LAD schemes were both skewed older and to a lower Energy Performance Certificate (EPC) band, as expected for a retrofit scheme. The GHG-V sample was also skewed to a larger floor area than EHS 2021.

A control study of a representative sample of homes recruited to the Smart Energy Research Laboratory (>13,000) was analysed using UCL-Power Temperature Gradient (UCL-PTG) over the same time period analysed for the GHG. This analysis checked whether stock-level changes in SMETER estimated in-use thermal performance, due to other factors such as energy price changes, may have affected the results of the GHG-SMETER sample. No significant change in thermal performance was observed in the control sample and changes in thermal performance observed in this project are therefore associated with the GHG retrofit.

Methods

This project used Type A and Type B SMETERs. Type A only use data that is gathered remotely (e.g. smart meter plus weather data); UCL- Power Temperature Gradient (UCL-PTG) and EDF Deconstruct+ were applied. Type B SMETERs use all remotely available data plus internal temperature data; UCL-Multiple Linear Regression (UCL-MLR), UCL-Siviour and Build Test Solutions Smart HTC (BTS SmartHTC) were applied. The results of this work reflect the applied methods, which have been shown to perform well in the SMETER TEST project (Allinson, et al., 2022), but do not preclude the development of methods with improved performance in the future.

Type A SMETERs offer low cost and disruption to households, with no equipment or in-home measurement; they are likely to be highly scalable but depend on assumptions about the internal conditions. Type B SMETERs utilise the internal temperature of homes and therefore changes to heating setpoints, cooling rates when the heating is switched off, and schedules are incorporated into the analysis.

Different SMETERs make their own assumptions about occupants, practices and physical effects, and incorporate them, or not, in different ways. These can result in different estimates for the in-use HTC, which account for different effects. SMETERs may account for occupant take-back and increased internal temperatures following retrofit; choice of type of SMETER for a particular purpose depends on the characterisation required alongside the costs, accuracy and intrusion that is acceptable in that instance.

SMETER analysis was complemented by a simple area-based elemental model, termed the EPC-Data Model, of heat flow (based on the same physical principles and material properties as SAP and RdSAP, but with some additional simplification) to produce a model-predicted HTC comparator. This was possible for around 73% of homes across the GHG-V sample (1814 before GHG retrofit and 824 after – after was only estimated for homes with single, fabric retrofits) which had recent EPC assessments; EPC input data was used for this calculation.

The impact of GHG retrofit on in-use HTC was evaluated using the Type A methods for the GHG-V sample, for households having had GHG measures installed during the period of data availability. Single fabric measure installations were studied. However, an accurate EPC-data Model estimate of the change of HTC on retrofit was not possible in all cases with the available data, due to the lack of detailed and reliable information on the retrofits beyond the category (e.g. area of insulation fitted). Application of SMETER Type B (with internal temprature measurement) was not possible in both pre and post retrofit due to timing of access issues collecting temperature data pre –retrofit.

SMETER range of plausibility

The reliability of SMETERs was explored through investigation of their self-robustness and through comparison between different methods and to the EPC-data Model; the accuracy of the SMETER results was unable to be explored due to the lack of ground truth HTCs for comparison. Self-robustness was assessed via a “plausibility rate”, the proportion of results deemed credible according to chosen criteria for the physical plausibility of in-use HTC estimates (i.e. $50 < HTC < 1000WK^{-1}$) and giving sufficient statistical fit to the data.

Plausibility rates of 54-55% were recorded for Type A SMETERs analysing the GHG-V sample. Higher plausibility rates are recorded for Type B methods, as expected, at 66-69%. These plausibility rates reflect both the availability of suitable data for each methods’ requirements in addition to the methods’ ability to characterise the data, thus highlighting the importance of planning the data collection for future, similar projects.

The relationship between the plausibility rate and characteristics of the sample was investigated using the GHG-V and GHG-LAD samples. Homes with characteristics related to expected higher heating energy requirements were linked to better credibility rates (above 70%): larger homes, older homes, and homes with lower EPC ratings. Homes with fewer party elements also returned higher credibility rates, as did those with more occupants, up to four and lower thereafter. Households self-reporting to be “financially comfortable” also recorded higher plausibility rates. For the Type A SMETERs, a large dataset is important, with a wide range of recorded heating/energy use. For Type B SMETERs a minimum internal/external temperature difference of 7°C leads to higher plausibility rates, as do regular heating patterns.

The ideal home for SMETER evaluation has high energy consumption, a regular pattern of heating, few unmetered gains, minimal energy consumption outside the heated envelope and is monitored for a long time (Type A) or long enough to enable data selection to meet SMETER specific criteria (Type B).

The performance of SMETERs was also evaluated for homes with unmetered heating (primary or secondary from solid fuels, oil and solar thermal) or significant uncertainty in their heat output (heat pumps), in addition to those with solar PV (without measurement of self-consumption of generation). A complementary metric to the HTC was introduced to enable this: the Grid Power Loss Coefficient, measuring metered energy input into the home. No significant changes in plausibility rates were observed for heat pump-heated homes or those with solar PV; this may be due to the core inclusion of energy import (heat pumps) and of methods to account for solar gains (PV). However, reduced credibility rates were observed for homes with other unmetered heating.

For homes in the GHG-V sample, for which sufficient EPC data was available, HTC estimates from the EPC-data Model are systematically higher than the in-use HTCs of both Type A SMETERs. This aligns with recent work suggesting that EPCs significantly over-estimate the energy use of homes in lower EPC bands.

Change to in-use HTC on GHG retrofit

The change of in-use HTC on retrofit under the GHG-V scheme was studied using Type A methods. The UCL-PTG and EDF Deconstruct+ methods estimated a similar average reduction of in-use HTC of ~20W/K (from ~250W/K baseline) across the households with fabric measures installed (including loft insulation, cavity wall insulation, pitched roof insulation, external wall insulation and under-floor insulation), this is shown per measure in Figure 1. This in-use HTC reduction is much lower than that predicted using the EPC-data Model of ~60W/K; however, due to data quality issues we note that this modelled value is likely to overestimate the expected efficiency improvement of retrofit. The EPC-data Model will also be unaffected by any occupant takeback or comfort taking, unlike the in-use estimates. The systematically lower HTCs estimated using the Type A SMETERs compared to the EPC-data Model also contribute to this difference between modelled and empirically estimated change in HTC on retrofit.

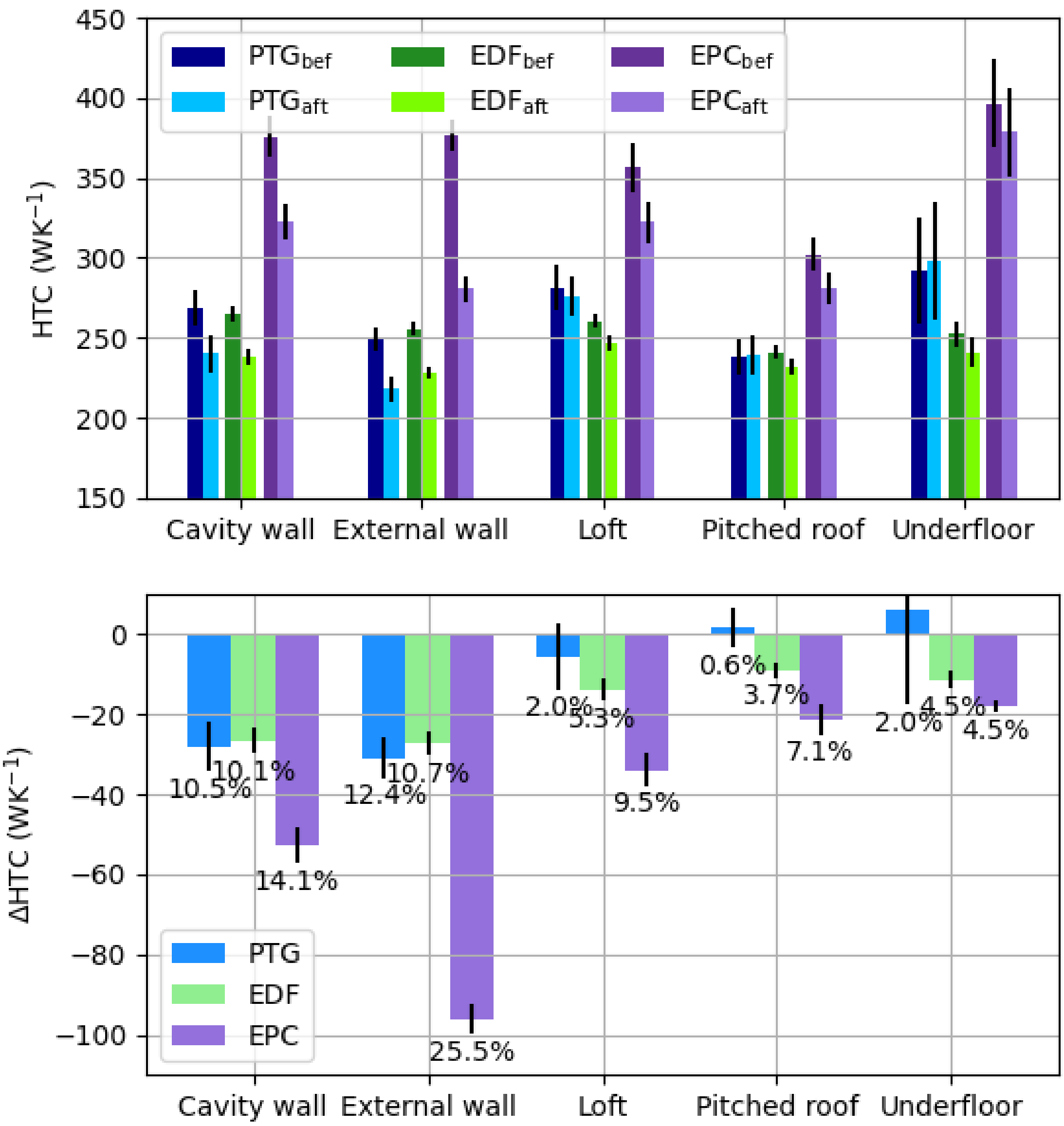

Figure 1 (Top) The HTC calculated from different methods for each insulation measure, showing the mean and standard error and (bottom) the corresponding ΔHTC for insulation measures. These results present data from all households where methods were successfully applied before and after, and the main heating fuel was metered (i.e. different homes are presented for each method).

The EPC-data Model predicts clear trends in HTC improvement on retrofit with property characteristics, such as age, floor area, EPC rating and house type due to the underlying deterministic model. The Type A SMETERs did not display the exact same clear trends, but trended in broadly the same way. It is also notable that the change to in-use HTC on retrofit associated with different measures broadly followed expectation relative to each other, but consistently lower than predicted by the EPC-data Model. The biggest difference was for external wall insulation where change of in-use HTC was around a third of that predicted by the model; data issues are likely to contribute to, but not entirely explain, this difference.

Investigation of the change to estimated in-use HTC on homes retrofitted through the GHG scheme suggests that Type A SMETERs are suitable for analysis of a cohort of homes receiving retrofit based on the analysis of combined error and cohort size. Error is reduced significantly at the size of 50 homes and above with reducing improvement above cohorts of 200 homes. However, they are not sufficiently accurate to identify the performance of works in individual homes.

The sample (cohort) size to provide consistent insights into the performance of retrofit on the mean in-use HTC was investigated by repeated random sampling of the population of participating GHG-V homes. The number of homes required depends on a range of factors including the similarity of the homes (therefore the expected variation in performance across them) and the required accuracy to meet the analysis objectives. For 200 homes, the mean in-use HTC for UCL-PTG had uncertainty (combined error of estimates) of ~4 W/K and EDF Deconstruct+ ~0.5 W/K, with the difference reflecting the much narrower distribution of results from Deconstruct+ compared to PTG and the EPC-data Model.

Applying and developing SMETERs

SMETERs may be used to evaluate the performance of individual and a stock of homes in support of a wide range of applications, such as quality assurance, retrofit performance and policy evaluation. However, their application requires care as does interpretation of the results that they deliver; in-use HTC is fundamentally different from the HTC derived from a controlled test (such as the Aggregate Heat Loss Test) and a wide-range of well-documented factors contribute to a performance gap between expected and observed energy use. As with all empirical methods, SMETERs must be selected according to the criteria for the required application.

SMETERs work best on low-efficiency homes, those that would most benefit from retrofit, but can perform poorly for underheated homes, which may affect some applications.

Internal temperature data from the GHG-LAD sample indicated that the pattern of temperatures is highly variable, with different heating patterns and setpoints. This may affect SMETER accuracy, particularly for Type A SMETERs. Temperature profiles indicated homes with low heating demand, and underheated homes, are expected to have high in-use HTC error with all studied methods; this has implications for the evaluation of thermal performance within households at risk of fuel poverty.

The difference between the SMETER and EPC-data Model results highlights the need for the application of the same methodology both before and after intervention. Some interventions, such as retrofit, may also affect internal temperatures and different SMETERs account for such effects in different ways and with differing success. Type B SMETERs, with internal temperature measurement, are better able to account for these effects than Type A (remote-only data), but at higher cost and disruption to householders.

The EPC-data Model produces an HTC based solely on the (assumed based on the RdSAP methodology) characteristics of the building fabric, whilst SMETER methods produce in-use HTC estimates, which will to a varying extent incorporate the way in which the occupants use their home, for example heating hours and window opening. The difference between these

HTCs for the GHG-V sample highlights that using heat loss parameters (HLP, HTC normalised by floor area) derived from SMETER estimates could reduce the overestimation of energy consumption by EPCs (Few, et al., 2023). However, utilising an in-use HTC, which will have some level of specificity to the present occupants, to then predict the energy consumption of a standard occupancy as in the EPC introduces a new source of uncertainty.

The HTC has natural variability from seasonal changes to the fabric and impacts of the conditions (such as higher ventilation rates through the same size vents as the temperature difference between inside and outside increases, or simply the impact of wind). Similarly, occupant factors affect the real efficiency of a home, for example their window opening practices. This presents challenges for the ability to compare in-use HTC over time that contribute to results of the GHG-SMETER project, but may be resolved in the future. It also provides the opportunity to provide better insights into the true energy use of homes and how it varies.

Schemes across large stocks of homes (such as national or local retrofit schemes) provide the opportunity to trial and develop SMETERs at low cost. Successful outcomes of these trials can be heavily dependent upon achieving timely access to high quality data and engagement with those delivering the housing scheme. Development of standard contracts, data sharing and data management systems for such schemes may facilitate better outcomes for SMETER application and development as well as other complementary evaluation activities.

Because in-use HTC is different from tested HTC, and noting the cost and disruption of standard HTC tests, cross-validation of SMETER methods may provide a valuable opportunity to develop and validate their performance. This may include the collection of and providing access to in-use energy, temperature, weather and complementary data (and also use of different products in the same homes to assess the level of variation in measurement).

This publication is available from: https://www.gov.uk/government/collections/heat-pump-research

If you need a version of this document in a more accessible format, please email alt.formats@energysecurity.gov.uk. Please tell us what format you need. It will help us if you say what assistive technology you use.