Department for
Business, Energy
& Industrial Strategy

Technical Evaluation of SMETER Technologies (TEST) Project

Loughborough University, Halton Housing, Leeds Beckett University, UCL

January 2022

Acknowledgements

Loughborough University: David Allinson, Ben Roberts, Kevin Lomas, and Dennis Loveday.

Leeds Beckett University: Chris Gorse, Adam Hardy, Felix Thomas, Dominic Miles-Shenton, David Johnston, David Glew, Kambiz Rakhshanbabanari, Fiona Fylan, and David Farmer.

UCL: Cliff Elwell, Jenny Crawley, Frances Hollick, and Jez Wingfield.

Halton Housing: Gavin Roberts and Lee Reevell.

Corresponding author:

David Allinson
d.allinson@lboro.ac.uk
Building Energy Research Group (BERG)
School of Architecture, Building and Civil Engineering
Loughborough University
LE11 3TU

© Crown copyright 2022

This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. To view this licence, visit nationalarchives.gov.uk/doc/open-government-licence/version/3 or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: psi@nationalarchives.gsi.gov.uk.

Where we have identified any third-party copyright information you will need to obtain permission from the copyright holders concerned.

Any enquiries regarding this publication should be sent to us at: Richard.fitzgerald@beis.gov.uk

Contents

Table of Figures

Table of Tables

Executive summary

This report details work carried out by the Technical Assessment Contractor for the Department of Business Energy and Industrial Strategy (BEIS) during Phase 2 of the Technical Evaluation of SMETER technologies (TEST) Project under the Smart Meter Enabled Thermal Efficiency Ratings (SMETER) Innovation Programme. The Technical Assessment Contractor, referred to here as the TEST team, comprises experts from Loughborough University, Leeds Beckett University, UCL, and Halton Housing.

Background

SMETER technologies use algorithms to calculate the Heat Transfer Coefficient (HTC) of occupied homes from smart meter data. Some SMETER technologies also have a product that is installed in the home to measure the parameters required by those algorithms, such as sensors to measure indoor air temperature.

The HTC is a widely recognised metric for describing building heat loss expressed as the rate at which heat is lost per degree Celsius air temperature difference between the inside and outside of a building in units of W/K. It includes the heat loss by conduction through the fabric and by infiltration and ventilation. A lower HTC demonstrates a lower rate of heat loss and therefore better thermal performance. The HTC is predicted as part of an Energy Performance Certificate (EPC) for new homes (using the SAP method) and for existing homes (using the RdSAP method). The HTC can be measured in unoccupied homes using well-established tests, a co-heating test to measure the fabric and infiltration heat loss, plus a blower door test to add the ventilation component. It is hypothesised that SMETER technologies may be able to calculate the HTC more accurately than predicted by RdSAP in a way that is more practical and cost-effective than measurement by the co-heating test.

Eight participating organisations (A-H, as shown in the table below) developing their own SMETER technologies, completed the Phase 1 stage gate[1] of the TEST Project and took part in the Phase 2 field trials. The participating organisations were provided with tailored support throughout – communicating key information, responding to enquiries, collating results and offering help with the development of their method and the estimation of measurement uncertainty. A ninth organisation (SMETER I, Knauf Energy Solutions) joined the project part way through Phase 2 and was evaluated separately.

[1] Phase 1 used simulated data from dynamic thermal simulation - see Department of Business, Energy and Industrial Strategy, TEST Project Phase 1 Stage Gate Report – 4th December 2019.

Table 1: Description of each of the SMETER technologies

SMETER	Participating organisation	Identifier	Brief description of each SMETER technology and product installed for this trial[1]
A	Building Research Establishment	BRE	Used only smart meter data and required no additional hardware product in the home. Required data that could be found in an EPC survey, plus: number of bedrooms.
B	Build Test Solutions	BTS	Five wireless and battery-powered sensors (temperature and relative humidity) that report to a hub. The hub was connected to the internet. Required data that could be found in an EPC survey, plus: floorplan; and type, area, and orientation of each window.
C	Cambridge Architectural Research	CAR	A proprietary heating controller with a touch screen interface (with temperature sensor), wireless boiler receiver unit, and five wirelessly controlled (battery powered) motorised TRVs (with temperature sensors) to install on radiators. Additionally, five wireless battery-powered sensors (each measuring temperature, relative humidity, light, and motion detection) report to a hub. The hub and the heating controller were connected to the internet. Required data that could be found in an EPC survey.
D	Centre for Sustainable Energy	CSE	Seven battery-powered data logging air temperature sensors, placed in different rooms, and then mailed back to the participating organisation at the end of monitoring. Included a shielded external air temperature sensor mounted to an external wall. No other information about the home was required.
E	EDF	EDF	Used only smart meter data and required no additional hardware product in the home or any other information about the home.
F	Hoare Lea	HOA	Four wireless battery-powered sensors (temperature and relative humidity) that report to a hub. The hub was connected to the internet. Required data that could be found in an EPC survey, plus: floorplan.
G	Passiv UK	PAS	Two battery-powered wireless sensors (temperature and relative humidity) connected to the local Sigfox wireless network. Required data that could be found in an EPC survey, plus: floorplan, number of bedrooms, number of occupants.
H	Switchee	SWI	A proprietary smart heating controller. The heating controller measured temperature, relative humidity, and motion detection. There were no additional sensors, and no external internet connection (the participant reported GSM cellular communications were built into the system). Required data that could be found in an EPC survey, plus: floorplan; and type, area, and orientation of each window

[1] This describes the SMETER technology and product that was installed by the TEST team, and the data that were requested during this trial. Participating organisations may not have used all of these data or sensors in their calculation of the HTC.

Methods for the technical evaluation of SMETER technologies

The field trials took place in 30 homes belonging to Halton Housing in NW England. The homes comprised two-storey houses and single-storey bungalows (one detached, 10 semi-detached and 19 end-terrace), built between approximately 1927 and 1990, with floor areas between 38m2 and 83m2, and with EPC bands of C or D. The homes were chosen because they were typical of those found in the UK, they were empty for a period before the field trials and could yield data that would provide a robust evaluation of all the SMETER technologies. The homes were not a representative sample of any particular housing stock.

Every home was carefully surveyed by experts in the TEST team, and this included the information required to calculate an EPC. The survey data were used by experts in the TEST team to predict each home's HTC using the RdSAPv12 software (called herein an expert RdSAP HTC). A commercial domestic energy assessor was also employed to produce EPCs for 22 of the 30 homes and the HTC predicted using the Elmhurst Energy RdSAP platform Version 9.94 (called herein a commercial RdSAP HTC). The assessor regularly carried out EPC assessments for Halton Housing and was unaware that the data collected for these EPCs were being used for a research project.

Smart meters were installed in every home to measure gas and electricity demand, and temperature and relative humidity sensors were installed in five to eight rooms of each home. Prior to occupancy, the HTC of every home was measured by a co-heating test and air tightness by fan pressurisation was used to account for additional ventilation heat loss. The range of measured HTCs was from 127 W/K to 269 W/K. The participating organisations were not told the results of any of these tests.

The well-established measurement methods were compared with newer alternatives. Because the co-heating test can take a week or more, the alternative QUB test, which estimates the HTC within a day, was evaluated. Fan pressurisation tests measure airtightness at an elevated pressure difference (ΔP = 50Pa), so the Pulse method of airtightness measurement at more natural pressure differences (ΔP = 4Pa) was also trialled. The QUB or Pulse measurements were not revealed to the participants or used to evaluate the SMETER technologies[2].

Ten of the homes were allocated to each participating organisation and had their SMETER product installed by the TEST team resulting in two SMETER products installed in every home[3]. The homes were then let to Halton Housing customers who gave their consent to be part of this project and moved in at various times between October 2019 and February 2020. All homes were monitored continuously until 6 August 2020.

The homes that were occupied last had shorter periods of winter data and this may have disadvantaged some SMETER technologies. To provide a longer duration data set, the TEST team gained consent from 27 households to extend monitoring over a second winter

[2] Except that the QUB result was used to produce the measured HTC in two cases where the co-heating test was unsuccessful.

[3] Two of the SMETER technologies relied only on the smart meter data and an algorithm with no associated product.

(01/08/2020 to 25/02/2021), but all the SMETER products were removed. The data needed by the SMETER technologies were measured by the TEST team. The participating organisations were invited to resubmit their results following this extended period of monitoring and five organisations accepted this opportunity: SMETER A (BRE), SMETER C (CAR), SMETER E (EDF), SMETER F (HOA) and SMETER G (PAS). These results relied on the monitored data provided by the TEST team.

To enable the participating organisations to test and, if necessary, refine the algorithms used in their SMETER technology, all eight participants were provided with the measured HTC for six homes, along with the survey information for these homes[4]. No other information, or feedback on the performance of their SMETER technologies was given. The participants were unaware of the performance of any other participant's SMETER technology.

Results of the technical evaluation of SMETER technologies

The accuracy of each SMETER product was evaluated by comparison with the measured HTC in two ways. Firstly, the SMETER result (calculated HTC, including the 95% confidence interval) was compared directly to the measured HTC. Where the confidence intervals of these two results overlapped, the SMETER technology result was deemed to be successful. Secondly, the difference between each SMETER result (central estimate, ignoring confidence intervals) and the corresponding measured HTC was analysed.

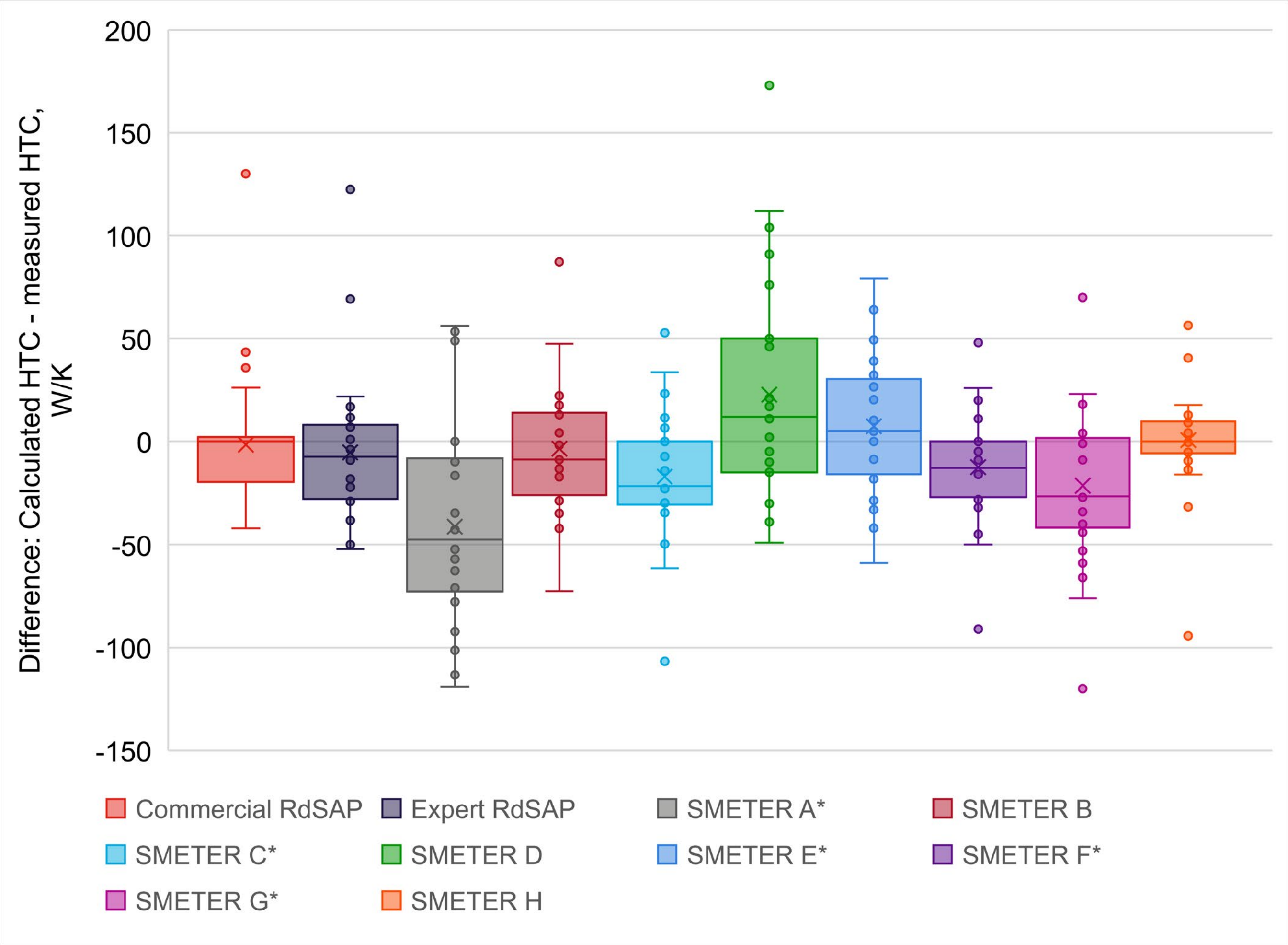
All participating organisations were able to report confidence intervals for every calculated HTC. Overall, the SMETER technologies were successful for between 70% and 97% of the homes, with average confidence intervals between 12% and 33%. Five participating organisations provided calculated HTC results that were more than 90% successful overall:

- SMETER B (BTS, 28 out of 30 homes, with an average confidence interval of +/-18%),
- SMETER E (EDF, 26 out of 27 homes, with an average confidence interval of +/-26%),
- SMETER F (HOA, 26 out of 27 homes, with an average confidence interval of +/-14%),
- SMETER G (PAS, 25 out of 27 homes, with an average confidence interval of +/-33%),
- SMETER H (SWI, 29 out of 30 homes, with an average confidence interval of +/-21%).

For each SMETER technology, the difference between each calculated HTC and the associated measured HTC was determined (see box-whisker plot below). The normalised mean bias error (NMBE) quantifies the magnitude and direction of the average bias in the calculated HTC. This is a measure of the trueness, or systematic agreement, of the measurement and would ideally be zero. The coefficient of variation of the root mean square error (CVRMSE) is a comparative measure of the precision of the calculated HTC. A lower CVRMSE is better. The NMBE ranged from -0.7% (best) to -26.9% (worst) and the CVRMSE from 13.4% to 38.9%.

[4] The results presented below include those for these six homes; excluding them made little difference to the overall results or conclusions.

Box-whisker plot of the differences between the HTC calculated by each SMETER and the corresponding measured HTC for each home



For all 30 homes, the difference between the HTC predicted from the expert RdSAP was compared with the measured HTCs, yielding an NMBE of -2.8% and CVRMSE of 18.2%. The two best-performing SMETER technologies (the lowest CVRMSE and NMBE closest to zero) were more accurate than the expert RdSAP: SMETER B (BTS, NMBE -2.1%, CVRMSE 15.2%) and SMETER H (SWI, NMBE -0.7%, CVRMSE 13.4%). However, compared with the commercial RdSAP HTC results for 22 of the homes (NMBE -1.1%, CVRMSE 19.6%) only SMETER H (SWI, NMBE -0.7%, CVRMSE 13.4%) was more accurate[5]. Two SMETER technologies had a lower CVRMSE than either of the RdSAP HTC predictions, but their NMBE was not as good: SMETER E (EDF, NMBE 4.1%, CVRMSE 17.4%) and SMETER F (HOA, NMBE -7.7%, CVRMSE 15.9%).

Previous work (e.g. Crawley, et al., 2019[6]) has revealed large discrepancies in the SAP ratings produced by different assessors. However, this is the first time that the accuracy of HTCs predicted using RdSAP survey data has been quantified. The commercial assessor was very experienced, and the homes were relatively simple to assess being small in size and without complicated features such as rooms-in-the-roof, extensions with different wall types, or

[5] The NMBE of SMETER B (BTS) being further from zero.
[6] Quantifying the Measurement Error on England and Wales EPC Ratings: https://doi.org/10.3390/en12183523

conservatories. It is hypothesised that the average commercial RdSAP HTC would not be as accurate as observed here.

The suitability of a SMETER technology for a particular application will depend on various factors, including accuracy (success rate, reported confidence interval, NMBE and CVRMSE), duration (average length of data period required for calculation[7]), and cost/convenience (number of sensors in the home, whether a professional installation or survey is required). The survey requirement may be less expensive and less inconvenient if the information can be taken from an existing EPC survey. The summary table below has been colour coded (green, amber, red) to ease interpretation: success rate >90%/>80%/<80%; average confidence interval <15%/<20%/>20%; NMBE <5%/<10%/>10%; CVRMSE <20%/<30%/>30%; average length of data period required <14days/<31 days/>31 days; and Type (which relates to increasing cost/inconvenience) T1, T2&T3 or T4. The comparison with the RdSAP predicted HTCs have been colour coded green if the calculated HTC is more accurate than the RdSAP value or red if it less accurate.

This colour coding is not indicative of fundamental problems with any SMETER technology. For example, SMETERs which use longer periods of data may be well-suited to many applications and SMETERs that are integrated in heating controllers (T4) offer little additional cost/inconvenience to a household choosing that controller. The required accuracy will depend on the application, and it may be possible to calibrate the SMETER technologies to improve the accuracy.

[7] This data period may be longer than the number of days of data used in the calculation as, for example, a SMETER algorithm might ignore days in the period when a home appears unheated and/or days when there are missing data.

Summary table (results in bold used data from the second monitoring period)

Aspect	Criterion	SMETER							
		A BRE	B BTS	**C CAR**	D CSE	**E EDF**	**F HOA**	**G PAS**	H SWI
Accuracy	Number attempted	**27**	30	**26**	30	**27**	**27**	**27**	30
	Success rate	**70%**	93%	**85%**	77%	**96%**	**96%**	**93%**	97%
	Average CI declared	**33%**	18%	**12%**	18%	**26%**	**14%**	**33%**	21%
	NMBE	**-26.9%**	-2.1%	**-10.9%**	9.8%	**4.1%**	**-7.7%**	**-13.1%**	-0.7%
	CVRMSE	**38.9%**	15.2%	**20.3%**	28.2%	**17.4%**	**15.9%**	**24.4%**	13.4%
	NMBE better than expert RdSAP (-2.8%)?	✗	✓	✗	✗	✗	✗	✗	✓
	CVRMSE better than expert RdSAP (18.2%)?	✗	✓	✗	✗	✓ [1]	✓	✗	✓
	NMBE better than commercial RdSAP (-1.1%)?	✗	✗	✗	✗	✗	✗	✗	✓
	CVRMSE better than commercial RdSAP (19.6%)?	✗	✓	✗	✗	✓	✓	✗	✓
Duration	Average length of data period used by participants (self-reported, days)	**208**	22	**101**	21	**311**	**51**	**7**	74
Cost or convenience	Total number of room sensors	**0**	5	**5**	8	**0**	**4**	**2**	0
	Heating controller included			✓					✓
	Professional install?			✓ [2]	✓ [3]				✓ [2]
	Uses data from an EPC survey	✓	✓	✓			✓	✓	✓
	Requires additional home survey	✓	✓				✓	✓	✓
	Type[4] of SMETER product	**T1b**	T3	**T4**	T2	**T1a**	**T3**	**T3**	T4

1 The CVRMSE of SMETER E (EDF) increased (from 17.4% to 19%) when the six homes with a known HTC were removed and was therefore no longer better than the CVRMSE for the expert RdSAP HTC (18.2%), but still better than the commercial RdSAP HTC (19.6%). The CVRMSE of all SMETER technology results changed (some up and some down) because of the smaller sample, but this was the only one that was so close to the RdSAP value.

2 Professional installation required as central heating controller electrically connected to boiler.

3 Professional installed deemed required as an external temperature sensor was mounted above head height on the outside of the home.

4 Type relates to increasing cost/inconvenience: T1a=only smart meter data required, T1b=only smart meter data and survey information required, T2=smart meter data and room sensors required, T3= smart meter data and room sensors and survey data required, T4= smart meter data and heating controller (with sensors) and survey required.

Other relevant findings

Households were interviewed during the field trial to understand their views on the SMETER products:

- Initially, 97% of households reported that they did not notice the SMETER products in their home.
- 93% of households said they would be happy to have a SMETER product in their home forever and the remaining 7% would be happy for a SMETER product to be installed for 6 months.
- 7% of households found the SMETER product's use of a plug socket to be inconvenient.
- 13% of the household reported that they did not like the flashing light on some of the sensors that were installed for the TEST Project monitoring.

The two best-performing SMETER technologies, SMETER B (BTS) and SMETER H (SWI), were of Type T3 and Type T4 respectively, and so required a home survey and additional installed equipment. More complex SMETER products such as these may be more vulnerable to hardware failure and all Type 3 SMETERs (SMETER B (BTS), SMETER F (HOA), and SMETER G (PAS)) experienced some problems with sensors that were not reporting as expected[8], while some households reported that they had problems using the heating controller that was part of the Type T4 SMETER H (SWI).

SMETER I (Knauf Energy Solutions) joined the project too late for their product to be installed in the field trial homes. Therefore, this participating organisation supplied two additional homes, located in Manchester (UK) and Genk (Belgium), and installed their own SMETER product. To determine the measured HTC, the TEST team carried out co-heating and blower door tests for the Manchester home and existing results for the Genk home were provided by a team at KU Leuven (Belgium). The SMETER was able to successfully predict the HTC in both of the homes with a self-reported confidence interval of +/- 8% in one home and +/-3% in the other. The SMETER used an average monitoring period of 74 days in length and was of type T4, due to the requirement to install a heat meter in the central heating pipework. There was no other evaluation of this product and direct comparison with the results from other SMETER technologies was not appropriate due to the differences in the field trial methods, sample, and sample size.

Observations can also be made about the two additional tests that were carried out in each of the field trial homes (the QUB, a shorter duration alternative to the co-heating test and Pulse, an alternative to the blower door airtightness test which is carried out at lower pressure differences using compressed air). The QUB test tended to give HTC results that were lower than the co-heating test and with a larger uncertainty; not all the tests were successful and so values were obtained for 26 of the 30 homes. The Pulse test was found to be slightly quicker than the blower door test, but results were not as reliable or repeatable in less airtight homes

[8] In at least one case, the equipment was unplugged by the household

and those with suspended timber floors[9]; overall, no simple linear conversion between the Pulse and Blower door results was identified for the tests carried out in this study.

Limitations and further work

There were some limitations to the field trials. The homes, while typical and with a diverse range of occupant types, were not representative of the UK housing stock. They had EPC ratings of C or D and there were no flats, new-build homes, or larger homes (the maximum floor area was 83m^2). All homes had the same, or very similar, gas combi, central heating, boilers, and there was little use of secondary heating.

There were some data problems, such as occasional spikes in energy data – these were identified by the QA procedures, but any data cleaning was left to participating organisations. Similar data problems should be expected when SMETER technologies are deployed at scale. Indeed, some participating organisations used indoor air temperature data from the TEST Project due to hardware problems, so their own sensor performance is not tested.

Additional field trials are recommended which should include more highly insulated homes, perhaps homes that comply with the current and future Building Regulations. In well-insulated homes[10], the proportion of all energy use that is for domestic hot water is greater, and internal heat gains from the sun and occupants' activities substantially contribute to space heating. This may lead to systematic errors, as well as greater uncertainty[11], in the SMETER-calculated HTC. The impact of party-wall heat transfer on the accuracy of SMETER-calculated HTCs is still not fully understood and mid-terraced homes, back-to-back terraces and flats should be investigated further. Homes with a wider range of energy technologies[12] should also be investigated, such as those with solar thermal and solar PV systems, heat pump heating systems, and mechanical ventilation systems, including those with heat recovery.

The repeatability of HTC calculations from SMETER technologies should be assessed. Where a home is not physically altered, SMETER technologies should be able to provide consistent calculations of HTC. The ability of SMETER technologies to calculate changes in HTC should also be assessed. The HTC will change when a home undergoes an insulation retrofit, has new windows or doors fitted, or an extension added. It is hypothesised that relatively small changes in the HTC may be identifiable using SMETER technologies as the uncertainty in the change may be smaller than the absolute uncertainty in each calculated HTC[13].Thus, the

[9] Box-whisker plot of the differences between the HTC calculated by each SMETER and the corresponding measured HTC for each home
[10] And potentially flats that are bounded on most sides by other heated spaces.
[11] To reduce uncertainty, the un-metered heat gains, such as those from the occupants and the sun, and un-metered heat losses, such as those from hot water going down the drain, should be much less than the metered heat gains from using gas and electricity to heat the home (Li et al., 2019): https://doi.org/10.1108/IJBPA-02-2019-0022.
[12] It is relatively straightforward to calculate the heat input to a home from a gas boiler using smart meter data. SMETERs may require additional measurement hardware or new assumptions when some other energy technologies are present.
[13] This could be the case if the systematic error was the same before and after a retrofit.

calculated change in energy demand might be more reliable than the change calculated by an RdSAP calculation.

Implications for the general implementation of SMETER technologies

This work has shown that the concept of using smart meter data to calculate HTCs clearly has merit. The use of SMETERs might provide a more robust procedure, with more clearly defined error characteristics, than HTCs derived by surveyors and RdSAP. The SMETER approach might also be more discriminating than RdSAP surveys, e.g., between nominally similar homes, where one was constructed with missing sections of cavity wall insulation that cannot be seen, and one that was not. SMETERs could also overcome difficulties associated with the need for presumptions in RdSAP, e.g., about loft insulation where the loft is not accessible. SMETERs could play a role, not only in the energy rating of homes, but also in quantifying the improvement to energy efficiency following refurbishment and identifying under-performance of new homes.

The SMETER technologies with no product in the home (Types T1a and T1b) did not always perform as well as those with sensors (Types T2, T3 and T4). This suggests that the measurement of internal temperatures is likely to lead to more accurate SMETER-calculated HTCs. However, the cost, intrusiveness and reliability of SMETER products must be considered. Integrating the SMETER technology into a new heating controller may offer an unobtrusive solution; but is only possible if the household want a new heating controller as the costs are relatively high. The requirement to collect survey data from the home (Types T1b, T3 and T4) is another potential barrier for some SMETER technologies, but this is eliminated if these data can be obtained from existing EPCs. SMETER technologies that use only smart meter data and no survey data (Type T1a) offer advantages in their ease of mass deployment and low costs.

The view of households is crucial to the success of SMETER deployment. Based on the response of the 30 households in this study, almost all would have no problem having SMETER products installed in their home, and especially if the use of plug sockets is minimised and sensors are unobtrusive.

If used for rating homes and other regulatory purposes, all the SMETER technologies will need protocols that define the homes to which they can be reliably applied and those to which they cannot and give guidance on the uncertainty. Such protocols would also describe how to deal with other diverse matters such as: unmetered heat sources, e.g., wood burners; large energy using appliances that are outside of the heated envelope, e.g., hot tubs and electric vehicles; and homes with an ill-defined thermal envelope, e.g., homes with conservatories, and especially those that are unheated but have internal doors that could be left open by the household[14]. Other situations in which SMETERs cannot be reliably used, or that require

[14] Conservatories could significantly increase the HTC if they are not thermally separated from the heated envelope of the home; RdSAP assumes any doors will be closed in unheated conservatories.

guidance on the interpretation of raw SMETER data, are likely to emerge once SMETER use becomes more widespread.

The findings from this work are ground-breaking given the small amount we know about the thermal performance of our housing stock – homes are very rarely measured. The SMETER approach opens up the prospect of a consistent and more reliable national database of domestic home energy efficiency. Further validation of SMETER technologies in more homes of varied types should be seen as an important short-term priority. The co-heating test, together with blower door tests to account for ventilation heat loss, should be used as the benchmark value as this remains the most accurate method for measuring the HTC.

There were some limitations to the field trial, in particular, it is not yet possible to comment on the reliability of SMETERs for more energy efficient homes (Rated A and B) or with more complex energy technologies (e.g., heat pumps). It is expected that the range of application, and the accuracy of SMETER algorithms will improve with experience and the collection of new data. Sharing the measured HTCs, dwelling characteristics, and ancillary measurement for more homes from this project, will stimulate further innovation.

While further field work is required to extend our understanding of SMETER technologies, this does not preclude their immediate use for homes of the type monitored here. In fact, some participating organisations are already offering this service and potentially we stand to learn much about the thermal performance of our housing stock this way.

1 Introduction

This report describes work carried out by the Technical Assessment Contractor for the Department of Business, Energy and Industrial Strategy (BEIS) under their Smart Meter Enabled Thermal Efficiency Ratings (SMETER) Innovation Programme. This work comprises Phase 2 of the Technical Evaluation of SMETER technologies (TEST) Project. The Technical Assessment Contractor, referred to here as the TEST team, comprises experts from Loughborough University, Leeds Beckett University, UCL, and Halton Housing.

SMETER technologies use algorithms to calculate the Heat Transfer Coefficient (HTC) of an occupied home from smart meter data. Some SMETER technologies also have a product that is installed in the home to measure the parameters required by those algorithms, such as temperature sensors to measure indoor air temperature.

The HTC is a widely recognised metric for building heat loss expressed as the rate at which heat is lost per degree Celsius air temperature difference between the inside and outside of a building in units of W/K. It includes the heat loss by conduction through the fabric and by ventilation and infiltration[15]. A lower HTC demonstrates a lower rate of heat loss and therefore better thermal performance. The HTC is predicted as part of an Energy Performance Certificate (EPC) for new homes (using the SAP method) and for existing homes (using the RdSAP method). The HTC can be measured in unoccupied homes using well-established tests, a co-heating test to measure the fabric and infiltration heat loss, plus a blower door test to add the ventilation component. It is hypothesised that SMETER technologies may be able to calculate the HTC more accurately than predicted by RdSAP and also be more practical and cost-effective than measurement by the co-heating test[16].

For clarity, the HTC values used in this report are defined as follows:

- Calculated HTC – the HTC calculated by a SMETER technology
- RdSAP HTC – the HTC predicted using the RdSAP method
- Measured HTC – the HTC measured using a modified version of the co-heating test with a blower door test to account for ventilation and additional uncertainty added to account for seasonal variation in heat loss

Eight participating organisations (A-H, Table 1) developing their own SMETER technologies were funded by BEIS under the SMETER Innovation Competition. They all completed the Phase 1 stage gate[17] of the TEST Project and took part in the Phase 2 field trials.

[15] See Appendix H for formal definition of HTC
[16] The co-heating test requires the home to be empty for two or more weeks during winter while electric heaters and fans maintain a constant elevated indoor air temperature.
[17] See Department of Business, Energy and Industrial Strategy, TEST Project Phase 1 Stage Gate Report – 4th December 2019.

Table 2: The eight participating organisations that took part in both Phase 1 and Phase 2 of the TEST Project

SMETER A	Building Research Establishment
SMETER B	Build Test Solutions
SMETER C	Cambridge Architectural Research
SMETER D	Centre for Sustainable Energy
SMETER E	EDF
SMETER F	Hoare Lea
SMETER G	Passiv UK
SMETER H	Switchee

The purpose of the TEST Project was to support and evaluate the development of SMETER technologies. Phase 1 had used simulated dwelling energy demand and temperature data to develop and evaluate the SMETER algorithms. Phase 2 was based on a field trial of 30 occupied homes and so tested the algorithms and the associated products. The evaluation considered both the intrusiveness of the SMETER technologies and the accuracy of the calculated HTC in three stages (TEST 4, TEST 5 and TEST 6). The participants were provided with tailored support throughout - communicating key information, responding to enquiries, collating results, and offering help with the development of their method and the estimation of measurement uncertainty.

This report explains the methods used by the TEST team to support the participating organisations and evaluate the SMETER technologies (Section 2), describes the homes and the households (Section 3) and describes the SMETER products (Section 4). The accuracy of the calculated HTCs is evaluated (Section 5 and Section 6), the results are compared with predictions of the HTC using RdSAP (Section 7), and the results are discussed (Section 8), before drawing conclusions (Section 9).

2 Methods for Support and Evaluation

The methods for supporting the participating organisations and for evaluating the SMETER technologies in the Phase 2 field trials are detailed below but summarised here.

1. 30 homes were identified from Halton Housing's stock, and households were recruited to live in them.
2. Each home was surveyed, to enable an expert to predict the RdSAP HTC, and a co-heating test and airtightness test carried out to produce the measured HTC.
3. The participating organisations were asked to calculate the HTC of 10 homes to which their SMETER was allocated (and their product installed if necessary) (TEST 4: blind field test of SMETER technologies).
4. The participating organisations were asked to calculate the HTC of the remaining 20 homes using the data collected by the TEST team (TEST 6: blind test of SMETER algorithms using monitored data).
5. There was an interim stage in which the measured HTC of six of the 30 homes was shared with participating organisations to help with algorithm development, before the final set of results were submitted.
6. The SMETER products were evaluated by the installer, and the occupants of the homes were asked pertinent questions about the SMETER products during an initial and follow-up interview (TEST 5: user acceptability evaluation).
7. The participating organisations were supported throughout Phase 2 with queries around the methods of the TEST team, the measurements made and properties under study, identifying issues with their SMETER product, and developing their analysis algorithms.
8. There was an optional second monitoring period in which data were collected from 27 of the homes for a second winter and SMETER participants were invited to re-submit their results.
9. A commercial home energy assessor produced EPCs and the data collected were used to predict the commercial RdSAP HTC for comparison with the results.
10. The TEST team applied a range of quality assurance (QA) processes to ensure the integrity of the monitored data collected from homes.

2.1 Recruitment of field trial homes and households

During the field trial, Halton housing identified 30 newly void homes (assigned identifiers HH01 to HH30) from their existing stock that would be suitable for the TEST Project. All field trial homes were located within Runcorn and Widnes in North West England (Table 2). The homes, described in full in Section 3.1, met the following criteria:

1. A single-family house or bungalow (i.e., not a flat/apartment or maisonette).
2. A suitably large meter cupboard in which to fit secondary energy meters.
3. Ready for occupation immediately after the co-heating test, i.e., no additional works to be done which might alter the HTC.

The homes were chosen because they were typical of those found in the UK, they were empty for a period before the field trials and could yield data that would provide a robust evaluation of all the SMETER technologies. The homes were not a representative sample of any particular housing stock.

Halton Housing let the homes to 30 households, described in full in Section 3.2. The households were invited to join the study and were free to leave at any time. Tenancies started from 06/05/2019 to 24/02/2020 but move-in dates were often later and continued until 30/03/2020 (Table 2). The household received financial incentives for participation in the project:

1. £100 was paid when the customer accepted the tenancy for the home.
2. £100 was paid after the first TEST 5 interview.
3. £100 was paid upon completion of the project once all SMETER products were removed.

The homes had monitoring equipment (Table 3) installed by Halton Housing before the household moved in and recording started after they had moved in (Table 4). Secondary ‘smart’ gas and electricity meters located near the supplier meter, measured half-hourly electricity and gas use. Combined temperature and relative humidity sensors were installed in as many rooms as possible. These were attached to an internal wall[18], at a height of 1.5-1.8 m from the floor, avoiding heat sources or direct sunlight.

Each of the homes had two different SMETER products installed (Table 5) by TEST team experts at Halton Housing before a household moved in (apart from HH02 and HH03 where households moved in before SMETERs were installed). Two of the eight participating organisations had SMETER technologies that required no product (SMETER A and SMETER E). The remaining six participating organisations each had ten of their SMETER products

[18] All sensors were originally attached (screwed) to the walls. In 2 homes (HH14 and HH18), the household removed the sensors to decorate. In HH18 the sensors were put on shelves/furniture for the duration of the field trial. HH14 sensors were lost by the household. It is not sure when, but new sensors were attached to the walls during the second interview.

installed in ten different homes (two SMETER products per home). SMETER C and SMETER H both incorporated central heating controllers and therefore had to be in different homes. SMETER G was installed only in specific postcode areas where the participating organisation believed their product would have a suitable SigFox signal (and signal strength was checked at installation). SMETER B, SMETER C, and SMETER F all required an internet connection (GSM modem/router supplied by each participating organisation) and so were not put together. For parity, SMETERS A and E were each allocated to ten homes for TEST 4.

Table 3: Tenancy and move-in date for the 30 field trial homes

Home ID	Postcode area	Tenancy start date	Move in date*
HH01	WA8	27/01/20	03/03/20
HH02	WA8	06/05/19	10/05/19
HH03	WA8	20/05/19	22/05/19
HH04	WA8	14/10/19	19/10/19
HH05	WA8	14/10/19	18/10/19
HH06	L24	21/10/19	26/10/19
HH07	WA8	21/10/19	10/12/19
HH08	WA8	28/10/19	06/12/19
HH09	WA7	28/10/19	02/11/19
HH10	WA7	04/11/19	04/11/19
HH11	WA7	18/11/19	08/02/20
HH12	WA8	18/11/19	06/12/19
HH13	WA7	25/11/19	02/12/19
HH14	WA8	25/11/19	30/03/20
HH15	WA8	25/11/19	04/11/19
HH16	WA8	02/12/19	06/01/20
HH17	WA8	02/12/19	07/12/19
HH18	WA7	09/12/19	16/12/19
HH19	WA8	09/12/19	21/12/19
HH20	WA8	16/12/19	03/01/20
HH21	WA8	16/12/19	18/12/19
HH22	WA8	16/12/19	22/02/20
HH23	WA8	06/01/20	10/01/20
HH24	WA7	06/01/20	12/02/20
HH25	WA8	13/01/20	16/01/20
HH26	WA8	13/01/20	13/01/20
HH27	WA8	20/01/20	07/02/20
HH28	WA7	20/01/20	13/02/20
HH29	WA8	24/02/20	21/02/20
HH30	WA8	24/02/20	07/03/20

* The move-in date was self-reported during TEST 5, Interview 1.

Each participating organisation was required to provide a method statement for the installation of their product. They were also required to provide their own internet connection if this was required. This was to avoid problems associated with using the household's internet, such as the routers being switched off, changes in provider causing disconnection, and homes without any internet subscription. Mains powered routers that used the mobile phone network were used in each case and housed in a tamperproof box.

Table 4: TEST team gas, electricity, and temperature measurement instruments

Variable	Measurement frequency	Device make/model	Communication protocol	Measurement uncertainty [3]
Gas used (m^3)	every 30 minutes	Honeywell BK-G4M connected to pulse reader and low power radio transmitter [1]	Low power radio connection to electricity meter	<1%
Electricity used (kWh)	every 30 minutes	Emlite EMA1 smart meter [1]	General Packet Radio Service (GPRS)	<1%
Indoor dry bulb temperature (°C)	every 30 minutes [4]	Combined T/RH sensor: Republic Of Things SmEILing [2]	Narrowband (NB) IoT	± 0.2 °C
Indoor relative humidity (%)	every 30 minutes [4]		Narrowband (NB) IoT	± 2%

[1] Gas and electricity meters were installed close to the supply meter as a secondary meter.

[2] Temperature and relative humidity sensors were mounted on an internal wall at a height of approximately 1.5-1.8m from the floor and avoiding heat sources/direct sunlight.

[3] Measurement uncertainty according to the manufacturer's specification.

[4] Temperature and relative humidity were measured every 10-minutes and post-processed by the TEST team to provide 30-minute values from the mean of the preceding measurements in each 30-minute period.

Table 5: Installation dates for monitoring equipment

Home ID	Energy meter install date	Temperature/Relative humidity sensors		Data recording started
		Install date	Number	
HH01	22/01/20	22/01/20	6	03/03/20
HH02	05/12/19	05/12/19	5	21/02/20
HH03	05/12/19	05/12/19	6	08/03/20
HH04	14/10/19	14/10/19	5	26/11/19
HH05	14/10/19	14/10/19	7	26/11/19
HH06	21/10/19	21/10/19	8	26/11/19
HH07	21/10/19	21/10/19	5	26/11/19
HH08	28/10/19	28/10/19	8	26/11/19
HH09	28/10/19	28/10/19	5	26/11/19
HH10	04/11/19	04/11/19	6	26/11/19
HH11	18/11/19	18/11/19	7	26/11/19
HH12	18/11/19	18/11/19	5	26/11/19
HH13	25/11/19	25/11/19	5	26/11/19
HH14	25/11/19	25/11/19	6	26/11/19
HH15	25/11/19	25/11/19	7	26/11/19
HH16	02/12/19	02/12/19	6	12/12/19
HH17	02/12/19	02/12/19	6	12/12/19
HH18	09/12/19	09/12/19	5	12/12/19
HH19	09/12/19	09/12/19	6	12/12/19
HH20	09/12/19	09/12/19	7	12/12/19
HH21	16/12/19	16/12/19	8	18/12/19
HH22	16/12/19	16/12/19	6	12/12/19
HH23	06/01/20	06/01/20	8	16/01/20
HH24	06/01/20	06/01/20	5	31/01/20
HH25	13/01/20	13/01/20	6	31/01/20
HH26	13/01/20	20/01/20	7	31/01/20
HH27	20/01/20	20/01/20	7	31/01/20
HH28	20/01/20	20/01/20	6	31/01/20
HH29	24/01/20	24/01/20	8	28/02/20
HH30	24/01/20	24/01/20	8	04/03/20

Table 6: SMETER product installation/allocation

Home ID	Date installed	SMETER product 1				SMETER product 2				SMETER allocation
		SMETER	No. of sensors	Heating controller?	Internet connection	SMETER	No. of sensors	Heating controller?	Internet connection	No product to install
HH01	23/01/20	D	7[1]	×	×	C	4+4[4]	✓	✓	A
HH02	05/12/19	D	6[1]	×	×	B	5[2]	×	✓	-
HH03	05/12/19	G	2	×	×	B	5[3]	×	✓	-
HH04	10/10/19	H	1	✓	×	F	4	×	✓	E
HH05	10/10/19	H	1	✓	×	F	4	×	✓	E
HH06	17/10/19	H	1	✓	×	F	4	×	✓	E
HH07	17/10/19	H	1	✓	×	F	4	×	✓	E
HH08	24/10/19	G	2	×	×	C	7+7[4]	✓	✓	A
HH09	24/10/19	G	2	×	×	C	4+4[4]	✓	✓	A
HH10	31/10/19	H	1	✓	×	F	4	×	✓	E
HH11	11/11/19	D	8[1]	×	×	F	4	×	✓	-
HH12	11/11/19	D	6[1]	×	×	F	4	×	✓	-
HH13	21/11/19	G	2	×	×	B	5	×	✓	-
HH14	21/11/19	G	2[5]	×	×	B	5[5]	×	✓	-
HH15	21/11/19	G	2	×	×	F	4[2]	×	✓	-
HH16	27/11/19	G	2[2]	×	×	C	6+6[4]	✓	✓	A
HH17	27/11/19	G	2	×	×	C	5+5[4]	✓	✓	A
HH18	05/12/19	G	2	×	×	F	4	×	✓	-
HH19	05/12/19	G	2	×	×	F	4	×	✓	-
HH20	10/12/19	D	7[1]	×	×	C	6+6[4]	✓	✓	A
HH21	10/12/19	D	8[1]	×	×	C	7+7[4]	✓	✓	A
HH22	10/12/19	D	7[1]	×	×	C	5+5[2,4]	✓	✓	A
HH23	19/12/19	H	1	✓	×	B	5	×	✓	E
HH24	19/12/19	H	1	✓	×	B	5	×	✓	E
HH25	07/01/20	H	1	✓	×	B	5	×	✓	E
HH26	07/01/20	H	1	✓	×	B	5	×	✓	E
HH27	09/01/20	D	8[1]	×	×	B	5	×	✓	-
HH28	09/01/20	H	1	✓	×	B	5	×	✓	E
HH29	19/02/20	D	9[1]	×	×	C	5+5[4]	✓	✓	A
HH30	19/02/20	D	9[1]	×	×	C	5+5[4]	✓	✓	A

[1] Total sensors for SMETER D includes one outdoor air temperature sensor.

[2] The number given is the total number of sensors that were installed, but one was lost in each of these homes.

[3] Two sensors were lost but were replaced.

[4] The second number is the number of TRVs installed (they also measure air temperature).

[5] All sensors went missing.

2.2 Survey and testing of the field trial homes

Survey and testing of the field trial homes was carried out during two heating seasons, March to May 2019 (the first three homes) and September 2019 to February 2020 (the remaining 27 homes) by TEST team experts at Leeds Beckett University (LBU) and Loughborough University (LU). This comprised:

1. Installation of the TEST Project weather station (LBU).
2. Home survey (LBU).
3. Co-heating test to measure HTC (LBU and LU).
4. Fan pressurisation test to measure airtightness (LBU and LU).

Two additional tests were carried out for comparison with the standard methods given above and all details of these are given in Appendix G:

5. Airtightness measurement by the alternative Pulse test (LBU).
6. HTC measurement by the QUB test (LBU).

The weather station (Figure 1), installed atop the Naughton Fields Care Home in Widnes and within 6 km of all 30 homes, monitored continuously throughout the field trial. It was mounted 12 m from the ground and recorded: air temperature, relative humidity, wind speed and direction, precipitation[19], atmospheric pressure, and total solar irradiance (vertical, south facing) (Table 6). Measurements were recorded at 10-minute intervals throughout the entirety of the field trial.

The expert home survey included: built form, approximate age[20], construction, insulation, dimensions, orientation, window type (e.g., double-glazed), and boiler make and model. Sufficient information was collected to predict the HTC using the RdSAPv12 software ('expert RdSAP HTC') and to produce an Energy Performance Certificate rating.

Two dwellings, HH09 and HH13, had a suspended timber floor with unfinished floorboards and no carpet or other covering. In each case, the floor was covered in craft paper to reduce the infiltration before any testing was carried out. This was because it was expected that the new tenants would install floor coverings when they moved in.

Co-heating tests were limited to a three-week period (Table 7). Pre-heating of the homes was necessary due to them being unoccupied prior to measurement (average 4 days pre-heating, range 1 to 11 days), the co-heating tests lasted on average a further 12 days (range 6 to 19 days). The indoor temperatures were maintained at a constant temperature for each test,

[19] Precipitation data only available until 09/02/20 due to sensor failure.
[20] Estimated using information from the Edina Digimap Historic mapping service.

though this varied between tests (range 20°C to 25°C) to maintain a consistent temperature difference depending on the expected outdoor air temperature. Purpose-provided ventilators (trickle vents and extractor fans) were open[21] for 10 tests and closed for 21 tests[22]. It was planned to keep them open in all tests, but this was found to increase the uncertainty of the measured HTC in homes that were not very airtight. Heat flux plates were fitted to party wall elements to measure the heat flux density into the party wall. An additional heat flux plate was fitted to solid ground floors to monitor the pre-heating period. A full explanation of the testing method is provided in Appendix C.

Fan pressurisation tests were conducted in all 30 homes to measure air tightness in accordance with the ATTMA standard (ATTMA, 2016). Three sets of fan pressurisation tests were carried out using a TEC Minneapolis Type 3 Blower Door with a DG700 gauge (Table 8): pressurisation with vents closed, de-pressurisation with vents closed, and de-pressurisation with vents open. When vents were closed, the trickle vents in the home were closed via their manual cover, but not additionally sealed; room airbricks and openings for mechanical extract ventilation were sealed with temporary sealing tape[23]. When vents were open, the trickle vents were opened via their manual cover and the airbricks and mechanical extract routes were unsealed and in their open position if applicable.

The measured HTC was calculated from the co-heating test in most cases. In two cases, the co-heating test failed and the result of the QUB test[24] was substituted (Table 7). The fan pressurisation test results were used to account for ventilation in the measured HTC (where vents were sealed). When fan pressurisation tests were repeated before and after the co-heating test, the average value was used to produce the measured HTC. The 95% confidence interval of the measured HTC was calculated from the measurement uncertainty, the difference in airtightness when trickle vents are open and closed[25], and an estimation of the seasonal variation that has been observed in modelled HTC values (Phase 1 of the TEST Project) during different parts of the heating season. A full explanation of the calculation of uncertainty is provided in Appendix C.

[21] Extractor fans were switched off.
[22] In total, 31 co-heating tests were undertaken because HH09 was co-heated twice, once with vents open and once with vents closed.
[23] https://energyconservatory.com/products/product/sixrollsductmask/
[24] In both cases the QUB test was successful but the resulting uncertainty in measured HTC was larger than for the co-heating tests.
[25] The position of trickle vents in each home during the field trial is not known and so accounted for in the uncertainty of the measured HTC.

Figure 1: TEST Project weather station mounted at 12 m above ground in Widnes

Table 6: Details of weather station

Variable	Device	Make/model	Measurement uncertainty
Dry bulb temperature	Capacitive ceramic sensor	Vaisala Weather Transmitter WXT520	± 0.3 °C
Relative humidity	Capacitive thin-film polymer sensor		± 3% at 0-90% RH ± 5% at 90-100% RH
Wind speed	Ultrasonic transducer		± 3% at 10 m/s
Wind direction			± 3.0°
Precipitation	Acoustic sensor		< 5% mm (daily accumulation)
Barometric pressure	Micromechanical pressure sensor		± 0.5hPa at 0-30 °C
Total vertical solar irradiance	Vertical south facing pyranometer	Kipp and Zonen CMP3	<10%

Table 7: Details of the co-heating tests

Home ID	Start date	End date	Pre-heat days	Analysis from	Analysis to	Days used	Set-point (°C)	Vent status	TEST team member
HH01	15/03/19	05/04/19	6	21/03/19	04/04/19	14	23	Open	LBU
HH02	15/04/19	03/05/19	6	21/04/19	03/05/19	12	25	Open	LBU
HH03	24/04/19	17/05/19	4	30/04/19	17/05/19	17	25	Open	LBU
HH04	13/09/19	03/10/19	11	24/09/19	03/10/19	9	25	Closed	LBU
HH05	13/09/19	30/09/19	NA	Failed [2]	Failed [2]	0	25	Open	LBU
HH06	30/09/19	14/10/19	3	03/10/19	14/10/19	11	22	Closed	LBU
HH07	23/09/19	04/10/19	NA	Failed [2]	Failed [2]	0	23	Open	LBU
HH08	30/09/19	21/10/19	10	10/10/19	20/10/19	10	22	Closed	LBU
HH09 [1]	01/10/19	18/10/19	3	04/10/19	11/10/19	7	22	Open	LBU
HH09 [1]	01/10/19	18/10/19	10	11/10/19	18/10/19	7	22	Closed	LBU
HH10	07/10/19	24/10/19	5	12/10/19	23/10/19	11	25	Open	LU
HH11	16/10/19	05/11/19	3	19/10/19	04/11/19	16	25	Open	LU
HH12	21/10/19	08/11/19	2	23/10/19	08/11/19	16	22	Closed	LBU
HH13	04/11/19	15/11/19	3	07/11/19	15/11/19	8	22	Open	LBU
HH14	06/11/19	22/11/19	9	15/11/19 [3]	22/11/19	7	20	Closed	LBU
HH15	05/11/19	18/11/19	2	07/11/19	18/11/19	11	22	Closed	LBU
HH16	05/11/19	25/11/19	2	07/11/19	25/11/19	18	25	Open	LU
HH17	18/11/19	28/11/19	3	21/11/19	28/11/19	7	25	Closed	LU
HH18	12/11/19	03/12/19	2	14/11/19	03/12/19	19	22	Closed	LBU
HH19	12/11/19	29/11/19	3	15/11/19	29/11/19	14	22	Closed	LBU
HH20	20/11/19	06/12/19	2	22/11/19	06/12/19	14	22	Closed	LBU
HH21	20/11/19	10/12/19	4	24/11/19	10/12/19	16	21	Closed	LBU
HH22	25/11/19	10/12/19	9	04/12/19	10/12/19	6	21	Closed	LBU
HH23	03/12/19	16/12/19	2	05/12/19	16/12/19	11	22	Closed	LBU
HH24	26/11/19	11/12/19	2	28/11/19	11/12/19	13	21	Closed	LU
HH25	06/12/19	19/12/19	3	09/12/19	19/12/19	10	21	Closed	LBU
HH26	02/12/19	18/12/19	3	05/12/19	17/12/19	12	21	Closed	LU
HH27	18/12/19	07/01/20	7	25/12/19	07/01/20	13	20	Closed	LBU
HH28	18/12/19	06/01/20	2	20/12/19	05/01/20	16	21	Closed	LU
HH29	04/02/20	18/02/20	2	06/02/20	17/02/20	11	20	Closed	LU
HH30	03/02/20	17/02/20	3	04/02/20	17/02/20	11	20	Closed	LBU

[1] HH09 was co-heated twice, once with vents open and once with vents closed.

[2] HH05 and HH07 co-heating tests failed due to warm weather during the testing so the QUB results from the given dates were used instead.

[3] HH14 lost power 12/11/19 to 15/11/19. Analysis started on 15/11/19.

Table 8: Details of fan pressurisation tests

Home ID	Pre- co-heating test				Post- co-heating test			
		Pressurisation	De-pressurisation			Pressurisation	De-pressurisation	
	Test date	Vents closed	Vents open	Vents closed	Test date	Vents closed	Vents open	Vents closed
HH01	N/A	×	×	×	05/04/19	✓	✓	✓
HH02	N/A	×	×	×	03/05/19	✓	✓	✓
HH03	N/A	×	×	×	17/05/19	✓	✓	✓
HH04	17/09/19	✓	✓	✓	N/A	×	×	×
HH05	20/09/19	✓	✓	✓	N/A	×	×	×
HH06	26/09/19	✓	✓	✓	N/A	×	×	×
HH07	02/09/19	✓	✓	✓	N/A	×	×	×
HH08	09/10/19	✓	×	✓	N/A	×	×	×
HH09	11/10/19	✓	✓	✓	24/10/19	✓	✓	×
HH10	N/A	×	×	×	24/10/19	✓	✓	✓
HH11	16/10/19	×	✓	✓	05/11/19	✓	✓	✓
HH12	N/A	×	×	×	08/11/19	✓	✓	✓
HH13	29/10/19	✓	✓	✓	N/A	×	×	×
HH14	30/10/19	✓	✓	✓	N/A	×	×	×
HH15	30/10/19	✓	✓	✓	N/A	×	×	×
HH16	05/11/19	✓	✓	✓	26/11/19	✓	✓	✓
HH17	05/11/19	✓	✓	✓	29/11/19	✓	✓	✓
HH18	N/A	×	×	×	05/12/19	✓	✓	✓
HH19	13/11/19	✓	✓	✓	N/A	×	×	×
HH20	N/A	×	×	×	13/12/19	✓	✓	✓
HH21	27/11/19	×	✓	✓	N/A	×	×	×
HH22	27/11/19	✓	✓	✓	N/A	×	×	×
HH23	N/A	×	×	×	16/12/19	✓	✓	✓
HH24	26/11/19	✓	✓	✓	12/12/19	✓	✓	✓
HH25	N/A	×	×	×	13/01/20	✓	✓	✓
HH26	02/12/19	✓	✓	✓	19/12/19	✓	✓	✓
HH27	N/A	×	×	×	10/01/20	✓	✓	✓
HH28	18/12/19	✓	✓	✓	06/01/20	✓	✓	✓
HH29	04/02/20	✓	✓	✓	18/02/20	✓	✓	✓
HH30	N/A	×	×	×	18/02/20	✓	✓	✓

2.3 TEST 4: blind field test of SMETER Technologies

In TEST 4 (blind test of SMETER technologies), the participating organisations were provided with weekly updates of half-hourly gas and electricity data for each of the ten homes in which their SMETER product was installed (or the allocated home in the case of SMETER A and SMETER E). Additionally, because their own SMETER products failed, several participants were sent 10-minutely indoor air temperature data (Table 9): SMETER B (three homes, five rooms per home), SMETER F (two homes, four rooms per home), and SMETER G (all 10 homes, two rooms per home); and SMETER D (all homes, all available rooms as interim data until their own product's data could be recovered[26]). To replicate the deployment of SMETER technologies in a real-world scenario, participants were expected to source their own weather data for TEST 4. However, some participants used data from the TEST team's weather station, for example where an outdoor sensor had failed (e.g., SMETER D, HH21) (Table 10). The TEST team's weather data were used for TEST 4 by SMETER C.

Participating organisations were supplied with additional information about each home at their request, and to mimic the data required for the normal use of their SMETER technology (Table 9). SMETER C requested information that could be obtained from a normal EPC survey. Others required information from an EPC survey plus additional information that would require a bespoke survey of the home/household: SMETER B requested the building floorplans plus window dimensions, type and orientation; SMETER F requested the building floorplans, SMETER G requested the building floorplans, number of bedrooms, and occupant numbers; and SMETER H requested the building floorplans plus window dimensions, type and orientation. SMETER A, SMETER D and SMETER E relied on monitoring data alone and required no extra information beyond the date of occupation. In some cases, the participating organisations reported that they did not use all of the additional data they requested, and this is noted in Table 9.

All participating organisations were asked to report the HTC for the 10 homes in TEST 4 along with four additional pieces of information:

1. The uncertainty in the calculated HTC (95% confidence interval).
2. The number of days of measured data used for the calculated HTCs.
3. The specific time period analysed.
4. The weather data source that was used.

The HTC calculated by each SMETER technology was compared with the measured HTC accounting for the confidence interval for each (see Section 5).

[26] SMETER D was not designed to transmit data, instead it logged to a device and the device was recovered from the home later. Ultimately, all devices were recovered, and it was understood that the TEST team data provided was not used in the final HTC prediction.

2.4 Test 6: blind test of SMETER algorithms using monitored data

In TEST 6 (blind test of SMETER algorithms using monitored data), participating organisations were sent data (Table 9) from the 20 homes in which their product was not installed, or the allocated homes in the case of SMETER A and SMETER E. Participating organisations were also supplied with weather data from the TEST team's weather station for TEST 6 but had the option to use their own data if they preferred[27] (Table 10).

All participating organisations were asked to report the HTC for the 20 homes in TEST 6 in the same way as for TEST 4 and the HTC calculated by each SMETER algorithm was compared with the measured HTC.

2.5 Optional interim assessment and HTC result sharing

Participating organisations were invited to submit interim results for all homes in TEST 4 and TEST 6 (based on data available up to 26/03/20). The purpose was to enable the participants to test and, if necessary, refine the algorithms used in their SMETER technology. All the organisations took part in this interim assessment except for the organisation behind SMETER A.

All eight participating organisations[28] were provided with the measured HTC for six homes: HH10, HH13, HH19, HH21, HH22, and HH25 (Table 11). For all organisation, this represented two TEST 4 homes and four TEST 6 homes. The six homes were chosen as they had: a reliable measured HTC; there was sufficient in-use data to calculate the HTC using the QA methods (see Section 2.10); and the QA methods and measured HTC had overlapping confidence intervals (CI).

Additional supplementary information on all six homes (Table 12) was provided to all eight organisations to further aid development of their SMETER algorithm. No other feedback on the performance of their SMETER technologies was given at this interim stage. The participants were unaware of the performance of any other participants SMETER technology.

[27] Participants were aware that TEST 4 and TEST 6 homes were in the same geographical area, thus weather data would be applicable for all homes.
[28] The measured HTC was sent to all eight participating organisations rather than just the seven that requested it. This ensured that organisations were producing results based on the same knowledge base.

Table 9: Summary of the data and information supplied for TEST 4 and TEST 6

Data set	SMETER							
	A	B[1]	C[1]	D	E	F	G	H
Weather data 10 minutes	✓	✓	✓	✓	✓	✓	✓	✓
Gas demand (kWh) 30 minutes	✓	✓	✓	✓	✓	✓	✓	✓
Electricity demand (kWh) 30 minutes	✓	✓	✓	✓	✓	✓	✓	✓
TEST 4 Air temperature 30 minutes		✓[2]		✓[3]		✓[4]	✓[5]	
TEST 6 Air temperature 30 minutes – number of rooms		5	6 [6]	6 [6]		4	2	1
TEST 6 Relative humidity 30 minutes – number of rooms								1
Tenancy commencement date	✓	✓	✓	✓	✓	✓	✓	✓
Self-reported permanent occupation date	✓	✓	✓	✓	✓	✓	✓	✓
Boiler make and model		✓	✓			✓	✓	✓
Partial postcode		✓	✓				✓	✓
Floorplan		✓				✓	✓	✓
Number of bedrooms							✓	
Total floor area (m^2)		✓						✓
Built form (e.g., house)		✓					✓	✓
Attachment (e.g., semi-detached)		✓					✓	✓
Occupant numbers: adult male, adult female, children							✓	
Window dimension/type/orientation		✓						✓
Requires product installation in the home		✓	✓	✓		✓	✓	✓
Requires information from an EPC survey		✓	✓			✓	✓	✓
Requires additional survey information [7]		✓				✓	✓	✓

[1] These participating organisations reported that they received these data and information but did not use all of them in their HTC calculation.

[2] SMETER B was provided with TEST 4 temperature data for five rooms per home in three homes (HH02, HH03, and HH14) due to problems with sensors that were not reporting as expected.

[3] SMETER D was provided with TEST 4 temperature data for all rooms per home in all 10 homes because they did not use transmitting sensors in their product. However, they reported that they used their own data to predict HTC after their sensors were removed from the homes.

[4] SMETER F was provided with TEST 4 temperature data for four rooms per home in two homes (HH18 and HH19) due to problems with sensors that were not reporting as expected.

[5] SMETER G was provided with TEST 4 temperature data (measured at 10-minute intervals) for two rooms per home in all 10 homes, due to problems with sensors that were not reporting as expected.

[6] Mean number of sensors (range 5 to 8) – data from all available rooms were provided (Table 4).

[7] Additional information that would require a bespoke survey of the home/household.

Table 10: Participating organisations use of their own weather data for HTC calculation

Stage	SMETER							
	A	B	C	D	E	F	G	H
TEST 4 (number of homes out of 10)	10	10	0	9[1]	10	9	8[3]	10
TEST 6 (number of homes out of 20)	20	20	0	0	20[2]	20	0	20

[1] One external sensor failed (HH21).

[2] Used a combination of own and TEST team weather data for all TEST 6 homes.

[3] SMETER G only submitted results for eight homes, so this value represents 100% of homes submitted.

Table 11: The six HTC measurements released to all participating organisations after the interim assessment

Home	Measured HTC (W/K)	Measured HTC 95% CI (W/K)	
		Lower	Upper
HH10	194	189	198
HH13	159	150	168
HH19	155	141	169
HH21	269	255	283
HH22	194	180	207
HH25	244	224	263

Table 12: Additional information provided to all participating organisations for the six homes after interim assessment

Data set	SMETER							
	A	B	C	D	E	F	G	H
Boiler make and model	✓	✓	✓	✓	✓	✓	✓	✓
Total floor area (m^2)	✓	✓	✓	✓	✓	✓	✓	✓
Dwelling type (e.g., house)	✓	✓	✓	✓	✓	✓	✓	✓
Attachment (e.g., semi-detached)	✓	✓	✓	✓	✓	✓	✓	✓
Roof insulation material	✓	✓	✓	✓	✓	✓	✓	✓
Roof insulation thickness	✓	✓	✓	✓	✓	✓	✓	✓
Wall construction type	✓	✓	✓	✓	✓	✓	✓	✓
Wall insulation material	✓	✓	✓	✓	✓	✓	✓	✓
Wall insulation thickness	✓	✓	✓	✓	✓	✓	✓	✓
Floor construction type	✓	✓	✓	✓	✓	✓	✓	✓
Floor insulation material	✓	✓	✓	✓	✓	✓	✓	✓
Floor insulation thickness	✓	✓	✓	✓	✓	✓	✓	✓
Window type (e.g., double glazed)	✓	✓	✓	✓	✓	✓	✓	✓

2.6 TEST 5: User acceptability evaluation

The acceptability of the SMETER products was evaluated by eliciting the opinions of the product installer using a questionnaire survey. To understand how intrusive the SMETER products were, two interviews were undertaken with one member of each household. The interviews also sought to identify if aspects of the household's behaviours during the trials could have abnormally influenced the ability of the SMETER technologies to calculate an HTC.

2.6.1 Installer Survey

The installer survey was designed to understand how difficult, time consuming and intrusive each SMETER product was to install. The same expert at Halton Housing installed all 60 SMETERs (six products in 30 homes at two per home) between 10/10/19 and 19/02/20. They completed one summary copy of the survey for each SMETER product based on contemporaneous notes from all of the installations for that product.

The survey (reproduced for each SMETER product in Appendix E) included the following details:

- Installer name
- Date
- Participating organisation
- Professional skills required e.g., electrician, heating engineer, etc
- Average time to install (the SMETER product, and separately the router if needed)
- Number of people required to install
- Details of each hardware item (name, size, room installed, and number of power sockets used)
- Total number of power sockets used
- Whether or not the SMETER used wireless communications and which wireless infrastructure
- How easy it was to set up any wireless communications
- Whether or not the gas and electricity supply had to be switched off at the meter during the install
- Details of the most-time-consuming aspect of the install
- Details of any problems encountered
- Details of any potential problems envisaged
- Details of any deviations from the participating organisation's installation method statement
- How much quicker the installation became with practice

- Example photographs of the equipment in-situ
- Any other comments.

2.6.2 Householder interviews to understand opinions of the SMETER products

The householder interview was designed to understand the occupants' perspective of the SMETER products, in terms of user experience, intrusiveness, and acceptability. The same expert at Halton Housing carried out all 30 interviews in February 2020. They also carried out a follow-up interview between 25/09/20 and 29/10/20. The purpose of the second interview was to establish if anything had changed since the first interview and to ask some new questions suggested by the participating organisations to aid their product development. In both cases, pilot interviews were carried out beforehand to refine the method.

The interviews were recorded electronically using Microsoft Forms (blank forms reproduced in Appendix F) and included the following details:

- Home ID
- Date
- Attitudes to technology and energy use
- Concerns related to the conduct of installers in their home
- Perceptions of the SMETER products in their home
- Any interactions with the SMETER products in their home
- Further perceptions of the SMETER product, including sensors during the follow-up interview.

The questions suggested by the participating organisations were:

5. Are you interested and able to improve the thermal performance of your home?

1. How much did you notice the sensors on the walls? Did they bother you?

2. Have you switched your SMETER device off at all? If so, why?

All 30 interviewees were asked the same questions, in the same order. The exception was where the answer to one question depended on another, e.g., answering “no” to the question “Do you set your central heating to heat to a particular temperature?” would mean the question “What temperature is your central heating set to?” would not be asked.

Because two SMETER products and the TEST Project monitoring equipment were installed in each home, care was taken to draw the interviewees attention to specific SMETER products in turn, noting which SMETER product was being referred to in the response form.

2.6.3 Additional questions to identify occupant behaviours that may affect the calculation of the HTC

During the interviews described above, householders were asked additional questions about how they used their home to identify any occupant behaviours or patterns that may affect the calculation of the HTC by a SMETER technology. The second interview was also an opportunity to ask if anything had changed since the first interview, the effect of the (March-May 2020 and beyond) COVID-19 restrictions on occupancy durations, and if they had turned their boiler off for any periods (to corroborate zero gas readings). Questions covered (see Appendix F for full details) were:

- Space heating schedules
- Use of secondary heating
- Hot water, laundry and washing
- Cooking
- Electricity used outside of the home
- Use of windows, other ventilation, and fans
- Periods away from home
- Changes to the household or the way the home is used/occupied.

2.7 Support to the participating organisations

The participating organisations were supported by TEST team experts at UCL and LU throughout the field trial using email and teleconferencing.

Email support included communicating key information, responding to enquiries, collating results, providing technical advice, and helping with ideas for algorithm development. Over 370 emails from individual participating organisations were received. The main topics of these enquiries were:

- Monitored data: some participating organisations enquired about the monitoring periods in different homes and asked questions about the released data. Issues with the data were highlighted by the TEST team with each release, and participating organisations were kept informed of the progress in resolving these.
- SMETER products: some participating organisations had problems communicating with their SMETER product, or the TEST team were unable to retrieve their sensors due to COVID-19 restrictions. Where it was not possible to resolve these issues, TEST 6 monitored data were substituted for TEST 4 weekly data releases. This option was offered to all participating organisations for TEST 4.

- Requests for additional information: some participating organisations requested additional information on the field trial homes, or about the modelling work that was carried out in the completed Phase 1 of the TEST Project, and this was provided whenever possible.

In addition to the opportunity to contact the TEST team themselves, participating organisations were invited to two rounds of teleconferences for detailed technical discussion. At the end of March 2020, they were invited to discuss their progress, any concerns with the field study, and the development of their algorithms. At the beginning of May 2020, they were invited to discuss the development of their algorithms with particular focus on their method for calculating confidence intervals. Only the organisation behind SMETER A declined the first round of calls. All joined the second round of calls after being reassured that details around IP did not need to be shared, and discussion should focus on the principles, rather than the details, of their methods. The TEST team at UCL provided tailored advice on how to improve the calculated HTC and to estimate confidence intervals, suggesting standard methods appropriate to each SMETER technology. This discussion resulted in most participating organisations planning changes to their method of calculating confidence intervals. Information about the measured HTC and QA HTC methods was also requested by most participating organisations, and this was supplied to all. Some participating organisations with unrealistic HTC results were advised on potential method changes, for instance altering their data cleaning process or the detail of their method. The effects these changes would have on the confidence interval calculations were also discussed. A follow-up call with one participating organisation showed that the advice given had led to a significant improvement in their results and much more appropriate and robust calculation of the confidence interval.

2.8 Optional second monitoring period

The homes that were occupied last had shorter periods of winter data and this may have disadvantaged some SMETER technologies. Participating organisations were invited to submit new results following a second period of monitoring over a second winter (01/08/2020 to 25/02/2021). The purpose was to provide a longer period of data for those participants that requested it. Five participating organisations accepted this opportunity: SMETER A, SMETER C, SMETER E, SMETER F, and SMETER G. 27 households gave consent for the TEST team to extend the monitoring in their home (HH07, HH18, and HH30 withdrew from the study). All SMETER products were removed, so these results relied on the monitored data provided by the TEST team (as in TEST 6). The five participating organisations were asked to report the HTC for the homes in the same way as for TEST 6, and the HTC calculated by each SMETER algorithm was compared with the measured HTC. As with the results from the first monitoring period, these included the six homes where the HTC was known.

The organisation behind SMETER A requested additional information about each home at this stage. In the first monitoring period (see section 2.3), SMETER A relied on monitoring data alone and they requested no extra information beyond the date of occupation. In the second monitoring period the organisation requested information from an EPC survey (total floor area, built form, attachment, and dwelling age) plus additional information that would require a bespoke survey of the home (number of bedrooms).

The results from the first monitoring period and those derived from the second monitoring period are presented in Section 5 and Section 6 respectively. Where participating organisation submitted results for both monitoring periods, it is those from the second that are used in the discussion and to draw conclusions in Section 8 and Section 9 respectively.

2.9 Commercial home energy assessment

A commercial domestic energy assessor was employed by Halton Housing to produce new EPCs[29] for the homes. The assessor regularly carried out EPC assessments for Halton Housing and was unaware that the data collected for these EPCs were being used for a research project. This work was completed for 22 of the 30 homes towards the end of the project. The RdSAP HTC of these homes was predicted using the Elmhurst Energy RdSAP platform (SAP Version 9.94). The aim was to compare the accuracy of SMETERs with the accuracy of the commercial RdSAP HTC, as well as with an expert RdSAP HTC (Section 2.2).

2.10 Quality Assurance

All measurements made by the TEST team (gas, electricity, indoor temperature, weather station air temperature and weather station solar irradiance) were verified by TEST team experts at UCL and any problems identified before release of the data to the participating organisations. Two basic types of QA were undertaken on every dataset (HH01-HH30): data quality checks, and data suitability checks.

The data quality checks were designed to spot missing or unfeasible data. They consisted of content tests, calculation of summary statistics, and plots for visual checking.

1. Content tests:
    a. Check the shape of the dataset and that the expected columns are present
    b. Check the format of each expected column
    c. Fulfilment for each variable (i.e., if < 100% then there is missing data)
    d. First and last datetime as expected
    e. Timesteps are the correct distance apart
    f. Rates of change of each variable are as expected
2. Summary statistics for gas, electricity and indoor air temperatures:

[29] A pre-existing EPC was in place for each home. It is not possible to derive the HTC from these pre-existing commercial EPCs as it is not recorded and cannot be calculated from publicly available data. That is why new commercial EPCs were procured.

   a. Minimum, maximum and ranges
   b. Mean and median
   c. Percentage of each column which is non-zero
3. Plots of gas, electricity, and indoor air temperatures for visual checking:
   a. Timeseries, individually and comparing outdoor air temperatures and power inputs

Data fulfilment in the first monitoring period was 99% for gas and electricity with at least 151 days of data except for two homes H22 and H06 for which fulfilment was around 95% (Table 13). The air temperature fulfilment was typically over 90% but there were numerous cases where the fulfilment one or more of the monitored rooms was lower (see Appendix A).

Data fulfilment for the second monitoring period was 94% for gas and electricity (Table 15) and 94% for temperature (see Appendix A).

Some persistent problems with the gas and electricity smart meter data were identified. Most notably, there were periods of missing data followed by a single large ‘catch up’ reading (spike). Similar problems have been observed elsewhere, notably in data from over 10,000 smart meters collated within the UK Smart Energy Research Lab (SERL).

The data suitability checks were designed to ensure that reasonable HTC calculations were possible using the data. They employed two well-established methods (see Appendix D) to calculate the HTC of the homes: the Siviour Method (QA1), and Multiple Linear Regression (QA2). These two methods were applied using both a simple automated data selection procedure, including all data recorded, and via expert selection of the data for analysis, providing a total of four estimates of HTC. These results also offered a benchmark against which the HTC calculated by each SMETER could be compared.

Data fulfilment in the first monitoring period was 99% for gas and electricity with at least 151 days of data except for two homes H22 and H06 for which fulfilment was around 95% (Table 13). The air temperature fulfilment was typically over 90% but there were numerous cases where the fulfilment one or more of the monitored rooms was lower (see Appendix A).

Data fulfilment for the second monitoring period was 94% for gas and electricity (Table 15) and 94% for temperature (see Appendix A).

Some persistent problems with the gas and electricity smart meter data were identified. Most notably, there were periods of missing data followed by a single large ‘catch up’ reading (spike). Similar problems have been observed elsewhere, notably in data from over 10,000 smart meters collated within the UK Smart Energy Research Lab (SERL).

The data suitability checks were designed to ensure that reasonable HTC calculations were possible using the data. They employed two well-established methods (see Appendix D) to calculate the HTC of the homes: the Siviour Method (QA1), and Multiple Linear Regression

(QA2). These two methods were applied using both a simple automated data selection procedure, including all data recorded, and via expert selection of the data for analysis, providing a total of four estimates of HTC. These results also offered a benchmark against which the HTC calculated by each SMETER could be compared.

Table 13: Monitoring data fulfilment (first monitoring period) – gas and electricity

Home ID	Electricity				Gas				Notes
	Start date	End date	Total days	Fulfilment (%)	Start date	End date	Total days	Fulfilment (%)	
HH01	03/03/20	06/08/20	156	99.2	03/03/20	06/08/20	156	99.9	
HH02	21/02/20	06/08/20	167	99.9	21/02/20	06/08/20	167	99.9	
HH03	08/03/20	06/08/20	151	99.8	08/03/20	06/08/20	151	99.8	
HH04	26/11/19	06/08/20	254	99.1	26/11/19	06/08/20	254	99.5	
HH05	26/11/19	06/08/20	254	99.9	26/11/19	06/08/20	254	99.5	
HH06	26/11/19	06/08/20	254	94.8	26/11/19	06/08/20	254	95.1	
HH07	26/11/19	06/08/20	254	99.9	26/11/19	06/08/20	254	99.5	
HH08	26/11/19	06/08/20	254	98.7	26/11/19	06/08/20	254	97.9	
HH09	26/11/19	06/08/20	254	99.5	26/11/19	06/08/20	254	99.9	
HH10	26/11/19	06/08/20	254	99.9	26/11/19	06/08/20	254	99.9	
HH11	26/11/19	06/08/20	254	99.9	26/11/19	06/08/20	254	99.9	
HH12	26/11/19	06/08/20	254	99.9	26/11/19	06/08/20	254	99.4	
HH13	26/11/19	06/08/20	254	99.5	26/11/19	06/08/20	254	99.9	
HH14	26/11/19	06/08/20	254	99.1	26/11/19	06/08/20	254	99.1	
HH15	26/11/19	06/08/20	254	99.9	26/11/19	06/08/20	254	99.5	
HH16	12/12/19	06/08/20	238	99.0	12/12/19	06/08/20	238	99.0	
HH17	12/12/19	06/08/20	238	99.5	12/12/19	06/08/20	238	99.9	
HH18	12/12/19	06/08/20	238	99.9	12/12/19	06/08/20	238	99.9	
HH19	12/12/19	06/08/20	238	99.9	12/12/19	06/08/20	238	99.9	
HH20	12/12/19	06/08/20	238	99.5	12/12/19	06/08/20	238	99.9	No gas use until 02/01/20
HH21	18/12/19	06/08/20	232	99.9	18/12/19	06/08/20	232	99.9	
HH22	12/12/19	06/08/20	238	95.7	12/12/19	06/08/20	238	95.7	
HH23	16/01/20	06/08/20	203	99.9	16/01/20	06/08/20	203	99.4	
HH24	31/01/20	06/08/20	188	99.9	31/01/20	06/08/20	188	99.5	
HH25	31/01/20	06/08/20	188	99.9	31/01/20	06/08/20	188	99.9	
HH26	31/01/20	06/08/20	188	99.3	31/01/20	06/08/20	188	98.8	
HH27	31/01/20	06/08/20	188	99.9	31/01/20	06/08/20	188	99.9	
HH28	31/01/20	06/08/20	188	99.9	31/01/20	06/08/20	188	99.9	
HH29	28/02/20	06/08/20	160	99.9	28/02/20	06/08/20	160	99.9	
HH30	04/03/20	06/08/20	155	99.2	04/03/20	06/08/20	155	99.9	

Note: "Start date" is the first day of day that was provided to participating organisations. Data were released from dates after the co-heating test was completed and the home was assumed occupied.

Table 14: Monitoring data fulfilment (second monitoring period)

Home ID	Start date	End date	Total days	Gas	Electricity
				Fulfilment (%)	Fulfilment (%)
HH01	01/08/2020	25/02/2021	209	94.2	94.2
HH02				92.3	93.8
HH03				94.2	94.7
HH04				91.5	92.3
HH05				94.2	94.7
HH06				89.8	89.8
HH08				94.7	94.7
HH09				94.7	94.7
HH10				94.2	94.7
HH11				94.7	94.7
HH12				94.2	94.2
HH13				94.7	94.7
HH14				94.7	94.7
HH15				94.2	94.7
HH16				92.7	93.3
HH17				92.8	94.2
HH19				94.2	94.7
HH20				94.7	94.7
HH21				94.7	92.8
HH22				94.7	94.2
HH23				94.7	94.2
HH24				94.7	94.7
HH25				93.7	93.7
HH26				94.7	94.7
HH27				94.7	94.7
HH28				94.2	94.2
HH29				92.3	93.8

Note: “Start date” is the first day of day that was provided to participating organisations.

3 Description of the homes and households

3.1 Field trial homes

The homes used in the trial were not a representative sample of any particular housing stock. Rather, they were selected because of their availability whilst attempting to include a consistent range of dwelling types. They comprised of two-storey houses and single-story bungalows: 22 two-storey and 8 single-storey; one detached, 10 semi-detached and 19 end-terrace (Table 15). They were built between approximately 1927 and approximately 1990 and had floor areas between 38m^2 and 83m^2. All were within 6 km of the TEST Project weather station. Compared to the English housing stock in general the homes were smaller on average with a higher proportion of semi-detached homes and end-terraces.

Most (23) of the homes had external walls of a masonry construction with an insulated cavity, but there was one with timber frame walls, one with a concrete frame construction and five with no-fines concrete construction and external wall insulation (Table 16). All had loft insulation, though in some homes the insulation was disturbed or piled up. There was a mixture of suspended timber and solid concrete floors which did not include any retrofit insulation. All windows were double glazed. All the homes had therefore received the commonly occurring and cost-effective refurbishment measures to improve their energy efficiency[30].

All homes had gas-fired combination boilers with rated efficiencies around 89% (Table 17). There were no homes with on-site renewables or secondary un-metered heating (e.g., none had a woodburning stove).

The measured HTC of the 30 homes was from 127 W/K to 269 W/K (Table 19). Measurement error in the co-heating test and ventilation measurements added an uncertainty of between ±3% and ±36% with an average of ±10% (±36% was for HH07 where the co-heating test failed and the QUB result was used instead). An additional uncertainty was added to represent the seasonal variation in the HTC. This change in the HTC during the heating season is caused by, for example, changes in the wind speed affecting infiltration and ventilation heat loss, and changes in the ground temperature affecting heat loss through the floor. This increased the uncertainty range in the measured HTC to an average of ±17% (range 11% to 37%). The full method for calculation of the HTC and its uncertainty are included in Appendix C.

The HTC of the 30 homes predicted by RdSAPv12 using the expert surveys (expert RdSAP HTC) ranged from 128 W/K to 334 W/K (Table 19). The average expert RdSAP HTC was 2.8% lower than the average measured HTC, and the difference between results ranged from 21%

---

[30] There was scope for topping up loft insulation in some homes.

lower to 58% higher. The average commercial RdSAP HTC was 1.1% lower than the average measured HTC, and the difference between results ranged from 17% lower to 61% higher.

Table 15: Description of the 30 field trial homes

Home ID	Number of storeys	Attachment	Estimated age	Floor area (m²)	Number of bedrooms	Distance from weather station (km) [1]
HH01	2	End-terrace	1980-1990	70.47	2	1-2
HH02	1	End-terrace	1950-1959	37.98	1	2-3
HH03	2	Semi-detached	1960-1969	64.33	2	1-2
HH04	1	End-terrace	1958-1968	41.90	1	3-4
HH05	2	Semi-detached	1927-1937	65.70	3	2-3
HH06	2	Semi-detached	1928-1954	83.37	3	4-5
HH07	1	End-terrace	1958-1969	43.18	1	<0.5
HH08	2	End-terrace	1937-1958	80.46	3	3-4
HH09	1	Detached	1962-1977	39.72	1	5-6
HH10	2	Semi-detached	1954-1968	61.82	2	3-4
HH11	2	End-terrace	1927-1937	68.31	3	4-5
HH12	1	Semi-detached	1937-1957	37.98	1	2-3
HH13	1	End-terrace	1954-1968	39.40	1	5-6
HH14	2	End-terrace	1967-1977	74.92	2	0.5-1
HH15	2	End-terrace	1937-1957	73.01	3	<0.5
HH16	2	End-terrace	1964-1984	73.48	2	2-3
HH17	2	End-terrace	1937-1958	62.07	2	<0.5
HH18	1	End-terrace	1958-1968	39.40	1	5-6
HH19	2	End-terrace	1937-1957	61.39	2	1-2
HH20	2	Semi-detached	1927-1937	64.97	3	2-3
HH21	2	Semi-detached	1927-1937	78.08	4	2-3
HH22	2	End-terrace	1967-1977	65.02	2	1-2
HH23	2	End-terrace	1967-1977	78.35	3	1-2
HH24	1	Semi-detached	1954-1962	38.96	1	5-6
HH25	2	Semi-detached	1927-1937	64.97	2	2-3
HH26	2	Semi-detached	1927-1937	70.17	3	2-3
HH27	2	End-terrace	1967-1977	72.38	3	1-2
HH28	2	End-terrace	1957-1968	64.18	2	5-6
HH29	2	End-terrace	1937-1958	75.88	3	0.5-1
HH30	2	End-terrace	1937-1958	77.04	3	0.5-1

[1] The weather station was located at 53.370590, -2.758850. Distance from weather station has been reported as a range to avoid identification of the homes.

Table 16: Construction and insulation of the 30 field trial homes

Home ID	External wall		Loft insulation				Ground floor		Windows	
	Type (cavity, timber frame, concrete frame, or concrete no-fines)	Insulation	Type	Thickness (mm)	Location	Condition	Solid concrete or suspended timber?	Insulated?	Double glazed?	Total window area (m^2)
HH01	Cavity	Blown	MWF	200	Joists	Even cover	Solid	×	✓	11.08
HH02	Cavity	Blown	MWF	?	Joists	Unknown	Solid	×	✓	6.02
HH03	Cavity	Blown	MWF	300	Joists	Disturbed	Solid	×	✓	12.63
HH04	Cavity	Blown	MWF	300	Joists	Disturbed	Solid	×	✓	9.74
HH05	Cavity	Blown	MWF	300	Joists	Disturbed	Solid	×	✓	13.39
HH06	Concrete	?	MWF	300	Joists	Piled up	Solid	×	✓	13.28
HH07	Cavity	Blown	MWF	300	Joists	Even cover	Solid	×	✓	7.85
HH08	Cavity	Blown	MWF	300	Joists	Disturbed	Solid	×	✓	13.24
HH09	Cavity	Blown	MWF	100	Joists	Piled up	Suspended	×	✓	7.51
HH10	Cavity	Blown	MWF	100	Joists	Disturbed	Solid	×	✓	12.38
HH11	Cavity	Blown	MWF	100	Joists	Disturbed	Suspended	×	✓	9.98
HH12	Cavity	Blown	MWF	300	Joists	Even cover	Solid	×	✓	7.24
HH13	Cavity	Blown	MWF	100	Joists	Piled up	Suspended	×	✓	6.97
HH14	Timber	Frame	MWF	100	Joists	Even cover	Solid	×	✓	8.02
HH15	No-fines	EWI	MWF	300	Joists	Disturbed	Solid	×	✓	12.87
HH16	Cavity	Blown	MWF	300	Joists+ rafters	Even cover	Solid	×	✓	10.71
HH17	No-fines	EWI	MWF	300	Joists	Even cover	Solid	×	✓	11.90
HH18	Cavity	Blown	MWF	150	Joists	Piled up	Suspended	×	✓	7.49
HH19	No-fines	EWI	MWF	100	Joists	Piled up	Solid	×	✓	11.87
HH20	Cavity	Blown	MWF	100	Joists	Even cover	Solid	×	✓	13.30
HH21	Cavity	Blown	MWF	100	Joists	Boarded	Solid	×	✓	13.12
HH22	Cavity	Blown	MWF	300	Joists	Even cover	Solid	×	✓	11.16
HH23	Cavity	Blown	MWF	150	Joists	Disturbed	Solid	×	✓	14.20
HH24	Cavity	Blown	MWF	100	Joists	Unknown	Suspended	×	✓	7.29
HH25	Cavity	Blown	MWF	300	Joists	Disturbed	Solid	×	✓	13.28
HH26	Cavity	Blown	MWF	300	Joists	Unknown	Solid	×	✓	12.35
HH27	Cavity	Blown	MWF	300	Joists	Even cover	Solid	×	✓	14.65
HH28	Cavity	Blown	MWF	300	Joists	Even cover	Solid	×	✓	9.36
HH29	No-fines	EWI	MWF	200	Joists	Disturbed	Solid	×	✓	11.17
HH30	No-fines	EWI	MWF	100	Joists	Even cover	Solid	×	✓	11.17

? = value unknown

Insulation codes: EWI = External Wall Insulation; MWF = Mineral Wool Fibre

Table 17: Heating systems in the 30 field trial homes

Home ID	Gas Boiler			Boiler efficiency (%)[1]		
	Type	Make	Model	SAP 2009/2012 annual	Winter seasonal	DHW
HH01	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH02	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH03	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH04[2]	Combi.	Vaillant	ecoMAX 824/2E	88.90	89.80	62.20
HH04[2]	Combi.	Intergas	Combi Compact ECO RF 30	88.80	89.70	76.80
HH05	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH06	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH07	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH08	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH09	Combi.	Intergas	Compact ECO RF 30	88.80	89.70	76.80
HH10	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH11	Combi.	Worcester	Greenstar 28i Junior	89.10	90.00	67.50
HH12	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH13	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH14	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH15	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH16	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH17	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH18	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH19	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH20	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH21	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH22	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH23	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH24	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH25	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH26	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH27	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH28	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH29	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10
HH30	Combi.	Vaillant	ecoTEC pro 28	88.80	89.70	61.10

[1] Boiler efficiency data taken from the Building Energy Performance Assessment support website (BRE, 2020).

[2] HH04 had a new boiler fitted on 06/11/2020

Table 18: Energy performance ratings of the 30 field trial homes

Home ID	TEST team expert survey		Pre-existing EPC[1]		Difference from expert survey	
	EPC rating	EPC band	EPC rating	EPC band	EPC rating	EPC band
HH01	72	C	70	C	-2	-
HH02	70	C	69	C	-1	-
HH03	72	C	73	C	+1	-
HH04	69	C	68	D	-1	-1
HH05	71	C	75	C	+4	-
HH06	64	D	63	D	-1	-
HH07	69	C	69	C	0	-
HH08	70	C	71	C	+1	-
HH09	66	D	68	D	+2	-
HH10	70	C	62	D	-8	-1
HH11	70	C	63	D	-7	-1
HH12	68	D	65	D	-3	-
HH13	68	D	70	C	+2	+1
HH14	72	C	76	C	+4	-
HH15	73	C	74	C	+1	-
HH16	74	C	71	C	-3	-
HH17	73	C	73	C	0	-
HH18	68	D	58	D	-10	-
HH19	72	C	73	C	+1	-
HH20	70	C	64	D	-6	-1
HH21	70	C	64	D	-6	-1
HH22	71	C	65	D	-6	-1
HH23	71	C	70	C	-1	-
HH24	68	D	68	D	0	-
HH25	70	C	64	D	-6	-1
HH26	71	C	71	C	0	-
HH27	71	C	75	C	+4	-
HH28	72	C	59	D	-13	-1
HH29	73	C	71	D	-2	-1
HH30	72	C	66	D	-6	-1
Average difference					-2.1	-0.3

[1] From the pre-existing commercial EPC generated by a commercial assessor

Table 19: Measured HTC, and the HTC predicted by RdSAP

Home ID	**Measured HTC** [1]							**Expert RdSAP HTC**			**Commercial RdSAP HTC**		
		Experimental uncertainty			Experimental uncertainty and seasonal variation				Difference from measured HTC			Difference from measured HTC	
	Central estimate (W/K)	lower bound (W/K)	upper bound (W/K)	expressed as +/-	lower bound (W/K)	upper bound (W/K)	expressed as +/-	Predicted (W/K)	W/K	Percentage	Predicted (W/K)	W/K	Percentage
HH01	181	166	195	8%	129	233	29%	193	12	7%			
HH02	131	113	148	14%	91	171	31%	128	-3	-2%	157	26	20%
HH03	155	146	164	6%	110	201	30%	177	22	14%	191	35	23%
HH04	153	124	182	19%	120	187	22%	147	-6	-4%	153	0	0%
HH05 [1]	*231*	*205*	*256*	*11%*	*195*	*266*	*16%*	*193*	*-38*	*-16%*	191	-40	-17%
HH06	212	188	236	11%	179	246	16%	334	122	58%	342	130	61%
HH07 [1]	*166*	*106*	*225*	*36%*	*104*	*228*	*37%*	*144*	*-22*	*-13%*			
HH08	213	190	237	11%	179	247	16%	230	17	8%	221	8	4%
HH09	172	160	184	7%	148	196	14%	165	-7	-4%	161	-11	-7%
HH10	194	189	198	3%	172	216	11%	186	-8	-4%	189	-4	-2%
HH11	213	199	228	7%	186	241	13%	209	-4	-2%			
HH12	127	116	137	9%	109	144	14%	141	14	11%	170	43	34%
HH13	159	150	168	6%	138	180	13%	150	-9	-6%	148	-10	-7%
HH14	137	126	147	8%	118	155	14%	206	69	50%			
HH15	208	192	223	8%	180	236	13%	180	-28	-13%	177	-31	-15%
HH16	184	171	196	7%	158	210	14%	176	-8	-4%			
HH17	191	173	209	9%	163	218	15%	164	-27	-14%	156	-34	-18%
HH18	148	136	160	8%	128	168	14%	146	-2	-1%	143	-5	-3%
HH19	155	141	169	9%	133	178	15%	172	17	11%	155	0	0%
HH20	229	210	248	8%	197	261	14%	200	-29	-13%	193	-36	-16%
HH21	269	255	283	5%	235	303	13%	234	-35	-13%	227	-42	-16%
HH22	194	180	207	7%	168	219	13%	192	-2	-1%			
HH23	237	213	260	10%	202	271	15%	219	-18	-8%			
HH24	171	160	181	6%	149	193	13%	145	-26	-15%	158	-13	-8%
HH25	244	224	263	8%	210	277	14%	192	-52	-21%	205	-39	-16%
HH26	254	235	272	7%	219	288	14%	204	-50	-20%	236	-18	-7%
HH27	201	189	212	6%	176	226	12%	208	7	3%	211	10	5%
HH28	172	154	189	10%	146	198	15%	173	1	1%	183	11	6%
HH29	226	204	247	10%	192	259	15%	189	-37	-16%	201	-25	-11%
HH30	227	205	249	10%	194	260	15%	198	-29	-13%			
Average	192							187	-5.3	-2.8%[2]	189	-2.1	-1.1%[2]

[1] The measured HTC of *HH05* and *HH07* was produced using the QUB test rather than using the co-heating test

[2] The average percentage difference is calculated from the average difference (W/K) and the average measured HTC (192 W/K)

3.2 Field trial households

The field trial households presented a diverse range of types but were not selected such that they were representative of those found in a particular housing stock or demographic. Their common characteristic is simply that they live in social housing.

There were 78 people living in the 30 homes, ranging from one to six people per household, with a mean of 2.6 people per household. This is slightly higher than the UK average of 2.4 people per household (ONS, 2021). There were seven single person households (23.3%), 20 households (66%) with children living at home (age 0-19), and an average of 1.3 children per home considering all 30 homes (1.95 children per home for the 20 households with children). 40% of homes have at least one adult in full-time or part-time employment. All adult occupants are unemployed in 40% homes, retired in 17%, and in tertiary education in 3%.

Most households paid their energy bills by monthly direct debit (electricity 57%, gas 60%), more than a third via a pre-pay meter (electricity = 37%, gas 33%) and the remaining household paid on receipt of the bill (electricity = 7%, gas 7%). More than a third of homes had a smart meter installed (37%).

Most households had broadband internet (83%). A little over half of the household respondents reported that they like new technology (57% agree or strongly agree). Most reported that they try to save energy at home (80% agree or strongly agree) and the majority reported that they know how to save energy (84% agree or strongly agree). More than half would choose not to rent a home if it was not energy efficient (60% agree or strongly agree) or if they knew the energy bills were going to be high (63% agree or strongly agree). Just over half also thought having an idea of the homes' energy bills in advance of moving in would be useful in helping them budget (53% agree or strongly agree).

All homes had a room thermostat and a timer to control the set-point and schedule the time of heating. 57% of households reported that they used the same fixed set-point temperature throughout the heating season. The heating set-point temperature chosen varied between 18°C and 24°C with a mean of 20.4°C. Interviewees in 37% of homes reported that they changed their heating set-point temperature regularly to control their comfort; the remaining 6% of homes changed the set-point occasionally. Most households (90%) used the same heating schedule for weekdays and weekends. None of the households used a bimodal (morning and evening) heating schedule, one household reported their heating was on 24/7 at a constant 21°C set-point temperature. 73% reported that the heating schedule was changed daily (by manual over-ride for example), and in one home changed "sometimes". Use of secondary heating was uncommon (13%) and from electrical heaters in all cases.

There was a washing machine in all 30 homes with an average of 5.4 loads of laundry done per home, per week (range 1 to 20). The most common method of drying laundry in the winter was on a laundry rack (50%) but many reported using a tumble drier (40%), with an average usage of 5.1 times per week (range 1 to 10). None of the homes had a dishwasher.

Households reported an average of 9.9 showers had per week, per home (range 0 to 20), 3.5 baths per week, per home (range 0 to 22).

Electricity was more common than gas as a cooking fuel both for the oven/grill[31] (63% gas, 33% electricity) and hob (53% gas, 40% electricity)[32]. Household reported cooking 6.7 times per week on average (range 0 to 14). There were no households with electric cars. Two households reported they charged an electric mobility scooter at home.

Half of households reported that they sometimes opened their windows in the autumn/winter (50%). The majority of those opened their windows ‘randomly’ in response to some stimuli (73%), rather than in any regular pattern (27%). The most common motivation for opening windows was for “fresh air”. Other reasons given were to remove cooking smells, or steam from the bathroom when bathing, and to dry clothes, or to cool the house (in winter). Households reported opening durations of 2.5 hours per day on average (range 0.5 to 12 hours). Two households which reported the highest window opening durations said they opened their bedroom windows overnight, all others reported 2 hours or less per day. The rooms with a window most likely to be open were the adult bedroom and bathroom (reported in 67% of the households that reported window opening) and also living room windows (53% of the households that reported window opening). Trickle vents were present in 73% of homes, and most of these (72%) reported that they were usually closed in all rooms. Mechanical extract ventilation was used in the kitchens (20% of households) and bathrooms (40%), with 7% of households reporting that their bathroom extractor was on constantly.

Households did not change very much during the trial. Only two households reported any change to the number of occupants during the field trial: one adult moving in, and one with a new baby. Eight (27%) reported turning their boiler off at some point during the trial due to leaks/breaking (5), to control their heating (2) or because they wanted to save money in summer (1). The COVID-19 pandemic, which meant that households were encouraged to stay at home during March-May 2020 and beyond, meant that households tended to be at home more than they would ordinarily.

[31] Two households did not have an oven/grill.
[32] One household did not have a hob.

4 SMETER products, installation, and household perspectives

From the findings of TEST 5, the SMETER products and their installation are described, and the views of the households are summarised.

4.1 SMETER products and their installation

The following summary of the SMETER products results from the documentation collected during the TEST 5 installer survey for each SMETER (see Appendix E). This describes the SMETER technology and product that was installed by the TEST team, and the data that were requested during this trial. Participating organisations may not have used all of these data or sensors in their calculation of the HTC.

All the SMETER technologies used the gas and electricity smart meter data. The additional sensors required (Table 20) and any other issues related to deployment of each (Table 21) are summarised below.

SMETER A used only smart meter data and required no additional hardware product in the home. In the first monitoring period, no other information about the home was requested (Table 9). In the second monitoring period requested data, as found in an EPC survey, comprised: partial postcode, total floor area, built form, attachment, and dwelling age. The following was also requested for each home: number of bedrooms.

The **SMETER B** product comprised five wireless and battery-powered sensors (temperature and relative humidity) that report to a hub. The hub was connected to the internet using a GSM modem. The modem and hub required plugging into mains power. Installation took less than one hour. Requested data, as found in an EPC survey, comprised: partial postcode, built form, attachment, floor area, and boiler make/model. The following was also requested for each home: floorplan; and type, area, and orientation of each window (Table 9).

The **SMETER C** product included a proprietary heating controller with a touch screen interface (with temperature sensor), wireless boiler receiver unit, and five wirelessly controlled (battery powered) motorised TRVs (with temperature sensors) to install on radiators. Additionally, five wireless battery-powered sensors (temperature, relative humidity, light, and motion detection) report to a hub. The hub and the heating controller were connected to the internet using a GSM modem. The heating controller, hub and modem required plugging into mains power. This was the longest installation of all SMETERs, taking around two hours. There was a requirement to turn off the heating system and mains electricity supply while the boiler receiver unit was connected to the boiler. This was therefore judged to be a professional installation (could not normally be carried out by a household). Requested data, as found in an EPC survey, comprised: partial postcode, boiler make/model (Table 9).

The **SMETER D** product comprised seven battery-powered data logging air temperature sensors. They were placed in different rooms, and then mailed back to the participating organisation at the end of Phase 2. There was a requirement to install a shielded external air temperature sensor. This was done above head height and required drilling an external wall to attach the shield. For this reason, it was judged to be a professional installation (could not normally be carried out by a household). The installation took around 30 minutes. No other information about the home was requested (Table 9).

SMETER E used only smart meter data and required no additional hardware in the home. No other information about the home was requested (Table 9).

The **SMETER F** product comprised four wireless battery-powered sensors (temperature and relative humidity) that report to a hub. The hub was connected to the internet using a GSM modem. The hub and modem required plugging into mains power. Installation took around 45 minutes. Requested data, as found in an EPC survey, comprised: boiler make/model. The following was also requested for each home: floorplan (Table 9).

The **SMETER G** product used two battery-powered wireless sensors (temperature and relative humidity) connected to the local Sigfox wireless network. This was the quickest SMETER to install, taking only 10 minutes. Requested data, as found in an EPC survey, comprised: partial postcode, built form, attachment, and boiler make/model. The following was also requested for each home: floorplan, number of bedrooms, number of occupants (Table 9).

The **SMETER H** product was a proprietary smart heating controller with a boiler receiver unit. The heating controller measured temperature, relative humidity, and motion detection. There were no additional sensors, and the internet modem was built into the system. The installation took around 90 minutes. There was a requirement to turn off the heating system and mains electricity supply while the boiler receiver unit was connected to the boiler. This was therefore judged to be a professional installation (could not normally be carried out by a household). Requested data, as found in an EPC survey, comprised: partial postcode, built form, attachment, floor area, and boiler make/model. The following was also requested for each home: floorplan; and type, area, and orientation of each window (Table 9).

To help with comparing the results, the SMETERs were grouped into Types based on the characteristics of their products (Table 21):

- Type T1a was SMETER E – it had no product, required no installation, no home survey and were therefore the least intrusive.
- Type T1b was SMETER A – it had no product and required no installation but did require information from a home survey (second monitoring period).

- Type T2 was SMETER D - it used simple battery-powered sensors with no communications (which were posted back to the organisation to recover data) and required no survey.
- Type T3 included[33] SMETER B, SMETER F and SMETER G - all used sensors with wireless communications, internet communications, and required information from a home survey.
- Type T4 included SMETER C and SMETER H – they required the installation of a central heating controller as well as information from a survey.

Thus, in general, the cost or inconvenience of the SMETER product and its installation increases from Type T1 to Type T4.

Table 20: Summary of SMETER product sensors and heating controllers

Product components	SMETER							
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**
Room sensors								
Outdoor air temperature				1				
Indoor air temperature				7[1]				
Indoor air temperature and relative humidity (RH)		5				4	2	
Indoor air temperature/RH, motion, and light			5[1]					
Total room sensors	**0**	**5**	**5**	**8**	**0**	**4**	**2**	**0**
Heating controllers								
Integrated controller/temperature sensor			1					
Smart radiator valves with temperature sensor			5[1]					
Integrated controller/temperature/RH/motion								1
Total sensors, including controllers with sensors	**0**	**5**	**11**	**8**	**0**	**4**	**2**	**1**

[1] This was an average value as sensors varied by number of rooms in the home.

[33] The QA methods used by the TEST team were of Type T3 as they used multiple temperature sensors and the boiler efficiency from the survey.

Table 21: Summary of SMETER installation and surveys

	SMETER							
	A	B	C	D	E	F	G	H
Professional install?			✓	✓ [1]				✓
Heating controller included			✓					✓
Install duration (minutes)	0 [2]	50	120	30	0 [2]	45	10	90
Internet connection required		✓	✓			✓		
Electricity sockets required		1	2			1		1 [3]
Ease of wireless comms set-up [4]		2	2 (5) [5]				1	1
Electricity off duration (minutes)			60					60
Uses data from an EPC survey	✓	✓	✓			✓	✓	✓
Requires additional home survey	✓	✓				✓	✓	✓
Type of SMETER product [6]	T1b	T3	T4	T2	T1a	T3	T3	T4

[1] A professional would only be required to mount the external sensor shield.

[2] The SMETER used smart meter data and no other sensors or controllers.

[3] One fewer socket is required if the thermostat is wired directly to the boiler.

[4] Where 1=easy, 2=OK, 3=acceptable, 4=difficult, 5=problematic.

[5] Usually 2, but two sets were received unconfigured and troubleshooting meant a 5 rating.

[6] T1a=only smart meter data required, T1b=smart meter data and survey required, T2= smart meter data and room sensors required, T3= smart meter data, room sensors and survey data required, T4= smart meter data, heating controller (with sensors) and survey required.

4.2 Household perspectives

The following summary of the household's views of the SMETER products results from the documentation collected during the TEST 5 householder interviews (see Appendix F).

In general:

- Households reported that they would appreciate the installer of a SMETER product to be[34]: clean and tidy (30%), friendly (24%), on time (22%), quick (11%), and polite (5%).
- Most households had no opinion on what they liked or disliked about the SMETER product (80%).
- 93% of households would be happy to have a SMETER product in their homes forever and 7% would only be happy for a SMETER product to be installed for 6 months.
- Two (7%) households reported that the SMETER product was inconvenient as it used a plug socket they required.

[34] Households gave multiple responses – these were the top 5.

- All 30 households said that having a SMETER product in their home did not make them change their behaviour.
- Initially, 97% of households reported that they did not notice the SMETER Products.
- After living with the SMETER products for more than 6 months, 93% of households reported that they did not notice the SMETER products.

Two of the SMETER products (SMETER C and SMETER H) included a central heating controller attached to the gas boiler:

- SMETER C was generally liked and the ability to control temperature with smart TRVs in every room seen as a positive benefit. However, two households with SMETER C experienced a broken smart TRV.
- Three households with SMETER H reported that this controller was difficult to use, and the heating came on at unwanted hours. Two households were still having problems after 6 months and resorted to switching their boilers off to control the heating[35].

In response to the questions suggested by the participating organisations:

1. Are you interested and able to improve the thermal performance of your home?

- 43% of households reported that they were able to improve the thermal performance of their home but were not interested in doing so.
- 40% said they were able and interested.
- 10% said they were interested, but not able.
- 7% said they were neither able nor interested in improving the thermal performance of their home.

2. How much did you notice the sensors on the walls? Did they bother you?

- Initially, most households reported no problems with the appearance or positioning of the SMETER products (93%). One household (3%) did not like the light on one of the sensors and did not like the sensors being screwed to the wall as they had recently decorated.
- After living with a SMETER product for 6 months, 13% commented that they did not like the flashing lights on some sensors.

3. Have you switched your SMETER device off at all? If so, why?

[35] The participating organisation offered their resident support line to directly assist these households, but this was not allowed for reasons of confidentiality.

- SMETER products were reported to remain plugged in and switched on in most homes (93.3%).
- Reasons for switching off SMETER products[36] in two of the homes were given as "unplugged by a child", and "needing a plug socket to charge a phone".

[36] Specifically, this meant householders unplugged the router power supply.

5 Evaluation of SMETER accuracy: first monitoring period

The results in this section are based on the first monitoring period. In Section 6, the results are all updated following the optional second monitoring period. In both cases the measured HTC for 6 homes (HH10, HH13, HH19, HH21, HH22, and HH25) in known to the participating organisations (see Section 2.5).

The measured HTC for the 30 homes ranged from 127 W/K to 269 W/K (Table 22). The results of applying the QA methods to calculate HTC from the monitored data from each home are included for completeness.

The HTC results calculated by the SMETER technologies (Table 23 and Table 24) were separated into TEST 4 (the 10 homes in which each participating organisation's SMETER product was installed) and TEST 6 (the other 20 homes with monitored data provided by the TEST team). This separation is somewhat arbitrary for SMETER A and SMETER E as they only used smart meter data for gas and electricity anyway (there was no product to install). All participating organisations were required to declare the periods of data used for each HTC calculation (Appendix B), which provided the number of days of data from within which data to calculate the HTC was taken (Table 25).

Two approaches were taken to the evaluation of accuracy. Firstly, the SMETER result (calculated HTC, including the 95% confidence interval) was compared directly to the measured HTC. Where the confidence intervals of these two results overlapped, this was deemed to be a successful SMETER result. This approach was designed to encourage participating organisations to actively engage with the realistic estimation of their products accuracy. Secondly, the difference between each SMETER result (central estimate, ignoring confidence intervals) and the corresponding measured HTC was analysed.

5.1 Comparison of calculated HTC with measured HTC using confidence intervals

Participating organisations provided HTC calculations for every home, except for SMETER G which reported results for only 17 homes[37]. All participating organisations were able to report confidence intervals for every calculated HTC. Their confidence intervals overlapped the confidence interval around the measured HTC in most cases (Figure 2, Figure 3, Figure 4, Figure 5, Figure 6, and Figure 7). Where confidence intervals overlapped, this was deemed to

[37] The participating organisation behind SMETER G commented that this was because of gaps in the data provided by the TEST Project. This was addressed in the second monitoring period (see Section 6).

be a successful HTC calculation by the SMETER i.e., the calculated HTC and the measured HTC were deemed to be the same.

Table 22: Measured and QA HTC results for all homes

Home ID	Measured HTC[1] (W/K)			QA method calculated HTC (W/K)					
				QA1 Multiple linear regression			QA2 Siviour method		
	HTC	lower	upper	HTC	Lower	Upper	HTC	Lower	Upper
HH01	181	129	233	170.9	111.0	230.8	179.8	118.1	241.5
HH02	131	91	171	147.4	110.7	184.1	151.0	114.6	187.4
HH03	155	110	201	135.4	79.2	191.6	131.8	80.1	183.5
HH04	153	120	187	157.3	115.5	199.1	160.9	119.0	202.8
HH05	231	195	266	225.7	164.0	287.4	234.2	169.0	299.4
HH06	212	179	246	274.4	210.3	338.5	269.6	206.4	332.8
HH07	166	104	228	145.2	111.1	179.3	155.0	120.1	189.9
HH08	213	179	247	245.6	183.9	307.3	258.5	195.6	321.4
HH09	172	148	196	175.0	136.6	213.4	178.2	139.5	216.9
HH10	194	172	216	207.7	165.0	250.4	205.8	163.1	248.5
HH11	213	186	241	298.7	223.9	373.5	290.3	213.3	367.3
HH12	127	109	144	147.0	109.5	184.5	149.0	111.4	186.6
HH13	159	138	180	158.2	121.7	194.7	160.1	123.2	197.0
HH14	137	118	155	148.8	117.8	179.8	67.5	41.6	93.4
HH15	208	180	236	204.9	153.3	256.5	203.7	151.6	255.8
HH16	184	158	210	178.5	137.1	219.9	178.6	137.2	220.0
HH17	191	163	218	179.0	136.6	221.4	163.2	121.1	205.3
HH18	148	128	168	170.1	124.0	216.2	175.5	128.9	222.1
HH19	155	133	178	164.0	126.0	202.0	170.3	131.5	209.1
HH20	229	197	261	238.8	185.2	292.4	237.6	184.1	291.1
HH21	269	235	303	309.9	236.8	383.0	316.1	242.4	389.8
HH22	194	168	219	185.2	130.8	239.6	186.5	126.2	246.8
HH23	237	202	271	276.8	202.0	351.6	279.3	204.7	353.9
HH24	171	149	193	205.2	146.3	264.1	211.2	152.1	270.3
HH25	244	210	277	221.3	156.8	285.8	222.4	158.2	286.6
HH26	254	219	288	263.5	175.7	351.3	253.9	161.3	346.5
HH27	201	176	226	230.8	174.7	286.9	232.8	176.5	289.1
HH28	172	146	198	162.4	119.9	204.9	161.7	119.2	204.2
HH29	226	192	259	179.5	137.1	221.9	177.5	135.6	219.4
HH30	227	194	260	213.7	145.8	281.6	222.8	156.7	288.9

1 Upper and lower limits of the measured HTC include measurement uncertainty, uncertainty to account for unknown state of trickle vents, and seasonal variation in the HTC

Table 23: HTC results (first monitoring period) for SMETERS A-D TEST 4 (bold) and TEST 6

Home ID	SMETER A			SMETER B			SMETER C			SMETER D		
	HTC	Lower	Upper	HTC	Lower	Upper	HTC	Lower	Upper	HTC	Lower	Upper
HH01	**122.7**	**95.7**	**175.4**	199.6	171.4	227.6	**184.0**	**162.0**	**207.0**	**272.0**	**236.0**	**307.0**
HH02	210.3	164.1	300.8	**135.2**	**115.5**	**153.3**	125.0	109.0	142.0	**143.0**	**107.0**	**180.0**
HH03	134.2	104.7	191.9	**120.2**	**85.3**	**157.8**	123.0	104.0	142.0	106.0	72.0	139.0
HH04	235.9	184.0	337.4	142.0	125.9	156.2	140.0	126.0	155.0	174.0	139.0	209.0
HH05	240.0	187.2	343.2	202.3	176.5	225.1	209.0	188.0	230.0	221.0	188.0	254.0
HH06	189.8	148.0	271.4	237.3	206.0	268.4	280.0	252.0	308.0	258.0	223.0	292.0
HH07	239.1	186.5	342.0	153.5	134.6	171.8	132.0	115.0	149.0	156.0	124.0	189.0
HH08	**189.8**	**148.0**	**271.4**	218.7	193.0	241.5	**197.0**	**177.0**	**217.0**	290.0	258.0	322.0
HH09	**106.8**	**83.3**	**152.8**	159.3	142.3	174.2	**152.0**	**124.0**	**181.0**	157.0	126.0	189.0
HH10[1]	*221.7*	*172.9*	*317.0*	*208.0*	*174.8*	*241.5*	*197.0*	*167.0*	*225.0*	*196.0*	*163.0*	*229.0*
HH11	194.1	151.4	277.6	300.2	264.1	332.2	260.0	232.0	288.0	**289.0**	**257.0**	**322.0**
HH12	192.7	150.3	275.5	125.2	104.5	144.4	140.0	125.0	154.0	**114.0**	**81.0**	**147.0**
HH13[1]	*140.8*	*109.8*	*201.3*	***150.2***	***124.7***	***174.8***	*178.0*	*160.0*	*196.0*	*176.0*	*143.0*	*209.0*
HH14	117.7	91.8	168.3	**126.7**	**96.4**	**220.1**	100.0	40.0	160.0	160.0	128.0	191.0
HH15	196.2	153.0	280.5	181.7	152.7	209.2	186.0	168.0	205.0	312.0	278.0	347.0
HH16	**112.2**	**87.5**	**160.4**	166.8	141.6	191.1	**153.0**	**138.0**	**168.0**	233.0	199.0	268.0
HH17	**108.0**	**84.3**	**154.5**	148.9	105.2	192.5	**102.0**	**76.0**	**127.0**	152.0	119.0	185.0
HH18	196.4	153.2	280.8	146.4	127.3	162.9	164.0	148.0	180.0	89.0	59.0	120.0
HH19[1]	*168.5*	*131.4*	*241.0*	*153.3*	*131.8*	*173.9*	*120.0*	*98.0*	*142.0*	*176.0*	*142.0*	*209.0*
HH20	**172.4**	**134.4**	**246.5**	203.0	174.7	229.8	**203.0**	**173.0**	**233.0**	**341.0**	**306.0**	**376.0**
HH21[1]	***164.8***	***128.5***	***235.6***	*282.0*	*246.2*	*313.8*	***231.0***	***206.0***	***255.0***	***442.0***	***407.0***	***478.0***
HH22[1]	***106.8***	***83.3***	***152.8***	*179.0*	*149.9*	*209.0*	***154.0***	***134.0***	***175.0***	***157.0***	***122.0***	***191.0***
HH23	200.4	156.3	286.5	**259.1**	**216.5**	**298.1**	252.0	225.0	279.0	248.0	213.0	283.0
HH24	207.9	162.1	297.2	**218.5**	**188.4**	**245.5**	185.0	166.0	205.0	141.0	110.0	172.0
HH25[1]	*208.1*	*162.4*	*297.6*	***210.0***	***178.7***	***238.9***	*167.0*	*150.0*	*183.0*	*214.0*	*180.0*	*249.0*
HH26	143.3	111.8	205.0	**271.6**	**232.2**	**308.3**	261.0	229.0	294.0	249.0	215.0	282.0
HH27	201.0	156.8	287.4	**207.0**	**176.8**	**235.9**	231.0	207.0	255.0	**251.0**	**216.0**	**286.0**
HH28	188.7	147.2	269.9	**158.6**	**127.8**	**189.3**	148.0	130.0	166.0	165.0	134.0	197.0
HH29	**106.8**	**83.3**	**152.8**	153.4	125.5	182.4	**138.0**	**116.0**	**160.0**	**190.0**	**154.0**	**225.0**
HH30	**160.4**	**125.1**	**229.4**	217.8	184.3	250.0	**202.0**	**178.0**	**226.0**	**274.0**	**239.0**	**310.0**

[1] The participating organisations knew the measured HTC for HH10, HH13, HH19, HH21, HH22, and HH25.

Table 24: HTC results (first monitoring period) for SMETERS E-H TEST 4 (bold) and TEST 6 (na = not reported)

Home ID	SMETER E			SMETER F			SMETER G			SMETER H		
	HTC	Lower	Upper	HTC	Lower	Upper	HTC	Lower	Upper	HTC	Lower	Upper
HH01	227.4	116.2	254.2	197.0	177.0	218.0	na	na	na	190.0	152.3	229.9
HH02	178.3	149.8	216.6	129.0	114.0	142.0	na	na	na	138.1	115.1	161.1
HH03	121.9	98.6	125.7	118.0	16.0	187.0	na	na	na	139.0	98.7	183.3
HH04	**201.8**	**163.3**	**265.0**	**140.0**	**135.0**	**145.0**	139.0	102.0	176.0	**153.0**	**132.6**	**173.6**
HH05	**259.6**	**170.1**	**314.9**	**219.0**	**212.0**	**226.0**	201.0	152.0	251.0	**228.8**	**193.2**	**264.4**
HH06	**197.4**	**144.7**	**322.8**	**247.0**	**230.0**	**264.0**	na	na	na	**268.3**	**224.5**	**316.5**
HH07	**199.1**	**173.2**	**258.6**	**137.0**	**129.0**	**146.0**	113.0	85.0	140.0	**132.6**	**110.0**	**157.6**
HH08	264.3	171.0	340.1	224.0	211.0	238.0	**203.0**	**162.0**	**243.0**	211.1	176.7	246.2
HH09	179.6	98.4	239.7	159.0	151.0	166.0	**163.0**	**124.0**	**202.0**	166.3	139.4	196.5
HH10[1]	***176.8***	***100.4***	***296.4***	***197.0***	***178.0***	***214.0***	*149.0*	*102.0*	*195.0*	***184.6***	***150.3***	***224.3***
HH11	179.8	128.2	340.2	**276.0**	**256.0**	**299.0**	na	na	na	255.4	215.5	296.8
HH12	177.0	95.5	250.5	**141.0**	**135.0**	**147.0**	na	na	na	133.7	113.7	154.0
HH13[1]	*179.1*	*124.3*	*269.0*	*156.0*	*146.0*	*168.0*	***148.0***	***86.0***	***210.0***	*171.8*	*142.0*	*205.9*
HH14	203.1	175.4	315.3	140.0	118.0	174.0	na	na	na	146.8	99.8	263.6
HH15	213.7	152.4	312.8	**207.0**	**193.0**	**220.0**	**199.0**	**134.0**	**263.0**	207.5	173.6	243.9
HH16	187.3	128.5	264.2	170.0	160.0	179.0	**128.0**	**70.0**	**186.0**	170.3	141.8	201.4
HH17	183.9	81.7	341.8	135.0	112.0	159.0	**178.0**	**129.0**	**226.0**	159.3	127.1	195.4
HH18	256.9	177.5	325.4	**157.0**	**147.0**	**167.0**	**126.0**	**96.0**	**156.0**	144.2	122.5	165.3
HH19[1]	*170.5*	*127.0*	*236.4*	***146.0***	***137.0***	***157.0***	***150.0***	***96.0***	***204.0***	*155.4*	*127.4*	*185.9*
HH20	202.5	145.8	287.0	209.0	195.0	222.0	190.0	148.0	233.0	233.1	194.5	276.6
HH21[1]	*277.2*	*179.0*	*348.2*	*267.0*	*252.0*	*278.0*	*292.0*	*230.0*	*354.0*	*286.6*	*244.9*	*331.1*
HH22[1]	*190.1*	*120.1*	*301.9*	*195.0*	*145.0*	*266.0*	*na*	*na*	*na*	*188.7*	*153.7*	*228.3*
HH23	**230.5**	**147.7**	**361.7**	272.0	249.0	297.0	na	na	na	**237.9**	**200.4**	**275.6**
HH24	**268.2**	**149.1**	**333.2**	205.0	179.0	232.0	na	na	na	**166.2**	**140.9**	**190.9**
HH25[1]	***252.8***	***147.5***	***330.5***	*214.0*	*196.0*	*236.0*	*192.0*	*140.0*	*244.0*	***238.8***	***201.6***	***277.8***
HH26	**218.7**	**148.0**	**353.2**	279.0	232.0	332.0	na	na	na	**294.5**	**245.9**	**347.4**
HH27	226.8	142.4	330.4	230.0	210.0	247.0	211.0	169.0	252.0	213.4	180.0	249.1
HH28	**173.1**	**110.6**	**267.9**	152.0	132.0	173.0	147.0	90.0	203.0	**158.8**	**130.2**	**189.3**
HH29	186.9	179.1	208.2	213.0	175.0	271.0	na	na	na	131.7	101.5	164.5
HH30	214.7	182.5	232.6	201.0	182.0	228.0	na	na	na	208.0	169.2	248.0

[1] The participating organisations knew the measured HTC for HH10, HH13, HH19, HH21, HH22, and HH25.

Four participating organisations provided calculated HTC results that were 100% successful in TEST 4, based on the above criteria (Table 26): SMETER B, SMETER E, SMETER G and SMETER H. Only SMETER G was 100% successful in TEST 6 but did not report results for all homes in either TEST 4 or TEST 6. Overall (TEST 4 and TEST 6), SMETERs were successful for between 77% and 100% of the homes.

The average of the reported confidence interval (Table 26) ranged from 6% (SMETER F in TEST 4) to 49% (SMETER E in TEST 4). The maximum declared confidence interval ranged from 10% (SMETER F in TEST 4) to 89% (SMETER E in TEST 6). The confidence interval may be expected to be smaller in TEST 4 where SMETER products were physically installed in (or allocated to) 10 homes, however this pattern was not observed across all participating organisations (Figure 8). It is noteworthy that in TEST 4, SMETER C (8 out of 10 successful), SMETER D (6 out of 10 successful), SMETER F (9 out of 10 successful), and SMETER H (10 out of 10 successful) report confidence intervals that are smaller than those for the measured HTC (Figure 8).

5.2 Analysis of the differences between the SMETER result and the measured HTC

The difference between each calculated HTC and the associated measured HTC was calculated for every case. The distribution of the differences for each SMETER illustrates the range and bias of the variations (Figure 9).

The mean of these differences (mean difference or mean bias error, MBE) was normalised by the average result[38] (normalised MBE, NMBE) (Table 27). The NMBE ranged from 0% (SMETER B in TEST 4) to -19% (SMETER C in TEST 4). The differences were used to calculate the root mean square error (RMSE), and this was divided by the average result (coefficient of variation of RMSE, CVRMSE). The CVRMSE ranged from 12% (SMETER H in TEST 4) to 45% (SMETER A in TEST 4).

The NMBE quantifies the magnitude and direction of the average bias in the calculated HTC. This is a measure of the trueness, or systematic agreement, of the measurement and would ideally be zero. The CVRMSE is a comparative measure of the precision of the calculated HTC. A lower CVRMSE is better. CVRMSE was expected to be lower in TEST 4 (where SMETER products were installed in homes) than in TEST 6, but this was not always the case[39].

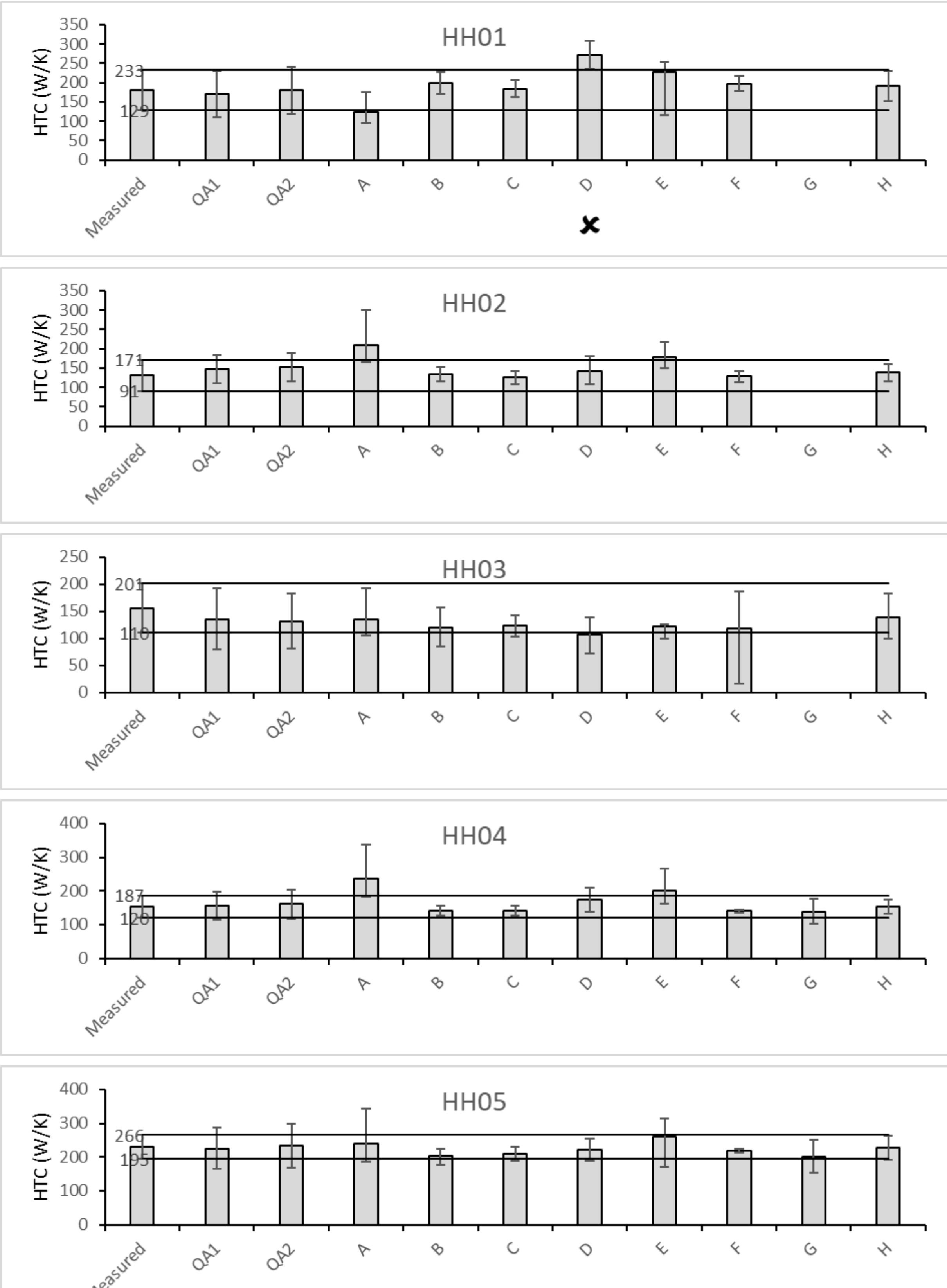
The standard deviation of the difference between each calculated HTC and the associated measured HTC was calculated to estimate upper and lower levels of agreement, after the method described by Bland and Altman (1986) (Table 27). Assuming a normal distribution, the limits of agreement were the mean difference ±1.96 x standard deviation (Bland and Altman, 1986). The smallest range of agreement was +15.6 W/K to -49.3 W/K (SMETER G in TEST 4) i.e., the SMETER-calculated HTC and the measured HTC agreed within that range.

[38] For each case, the average result is calculated from the calculated HTC and the measured HTC, following the method described in the literature (Bland and Altman, 1986).

[39] The TEST team were not able to provide the same data as measured by the SMETER in every case and so TEST 4 was expected to give better results, however this was not conclusive and comparisons across TEST 4 and TEST 6 were deemed valid.

Table 25: Length (in days) of the self-reported period of data (first monitoring period) used by each SMETER for TEST 4 (bold) and TEST 6 (na = not reported)

	SMETER							
Home ID	A	B	C	D	E	F	G	H
HH01	**29**	21	**49**	**21**	157	46	na	23
HH02	30	**21**	51	**21**	168	54	na	22
HH03	30	**21**	46	21	152	30	na	20
HH04	30	21	149	21	**255**	**125**	7	**102**
HH05	30	21	52	21	**254**	**125**	7	**123**
HH06	31	21	86	21	**189**	**59**	na	**56**
HH07	31	21	93	21	**204**	**74**	7	**72**
HH08	**31**	21	**27**	21	255	76	**7**	122
HH09	**31**	21	**64**	21	255	120	**7**	119
HH10	30	21	79	21	**232**	92	7	**99**
HH11	30	21	110	**21**	255	**41**	na	37
HH12	30	21	86	**21**	231	**96**	na	93
HH13	30	**21**	139	21	255	**105**	**7**	116
HH14	30	**79**	10	21	255	104	na	108
HH15	30	21	117	21	255	**117**	**7**	118
HH16	**31**	21	**63**	21	239	73	**7**	81
HH17	**20**	21	**63**	21	239	76	**7**	107
HH18	30	21	131	21	239	**107**	**7**	101
HH19	30	21	129	21	239	**95**	**7**	107
HH20	**20**	21	**55**	**21**	239	80	7	84
HH21	**31**	21	**63**	**21**	231	96	7	98
HH22	**28**	21	**64**	**21**	209	23	na	34
HH23	30	**21**	84	21	**204**	55	na	**69**
HH24	30	**22**	78	21	**189**	41	na	**55**
HH25	30	**21**	89	21	**189**	42	7	**57**
HH26	31	**21**	79	21	**189**	43	na	**49**
HH27	30	**13**	79	**21**	189	46	7	50
HH28	30	**14**	75	21	**182**	41	7	**44**
HH29	**31**	21	**53**	**21**	161	23	na	29
HH30	**28**	21	**53**	**21**	156	56	na	24
Average	29	22	77	21	216	72	7	74



QA1 Multiple linear regression, QA2 Siviour method

Figure 2: HTC results (first monitoring period) for HH01 to HH05 with error bars for 95% confidence interval; the horizontal lines show the 95% confidence interval for the measured HTC and ✖ indicates an unsuccessful result based on overlapping confidence intervals

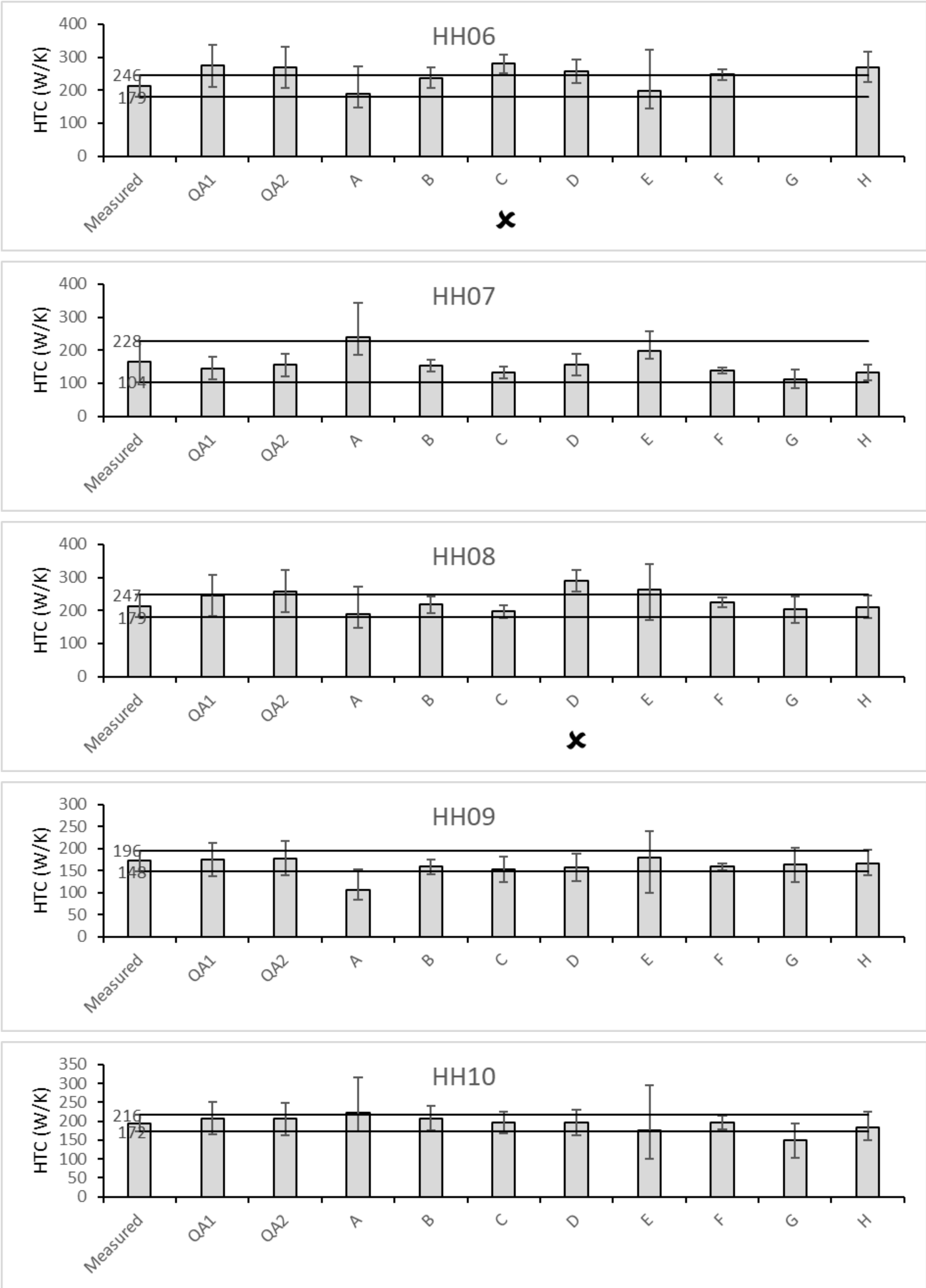


Figure 2: HTC results (first monitoring period) for HH06 to HH10 with error bars for 95% confidence interval; the horizontal lines show the 95% confidence interval for the measured HTC and ✗ indicates an unsuccessful result based on overlapping confidence intervals

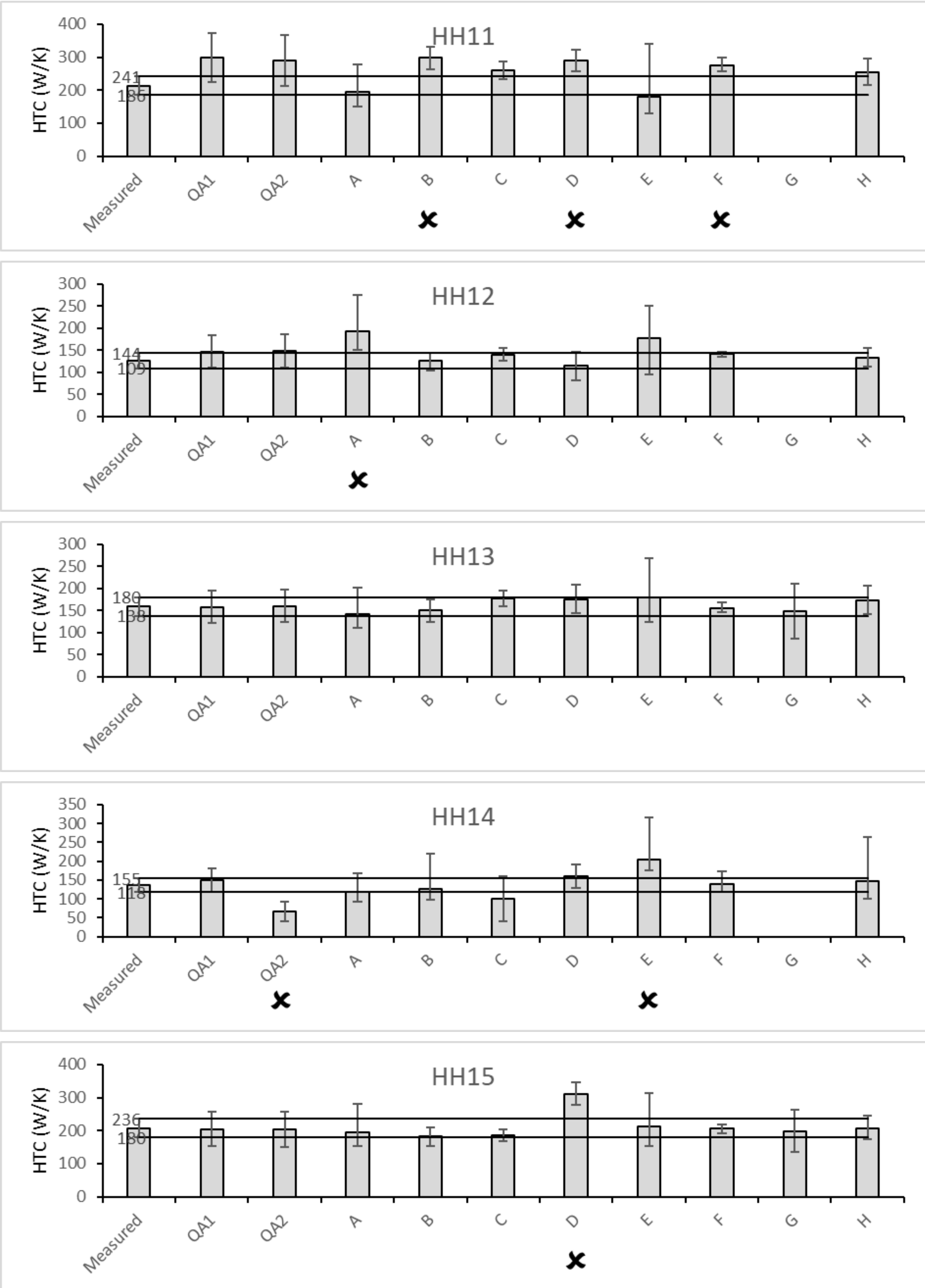


Figure 3: HTC results (first monitoring period) for HH11 to HH15 with error bars for 95% confidence interval; the horizontal lines show the 95% confidence interval for the measured HTC and ✖ indicates an unsuccessful result based on overlapping confidence intervals

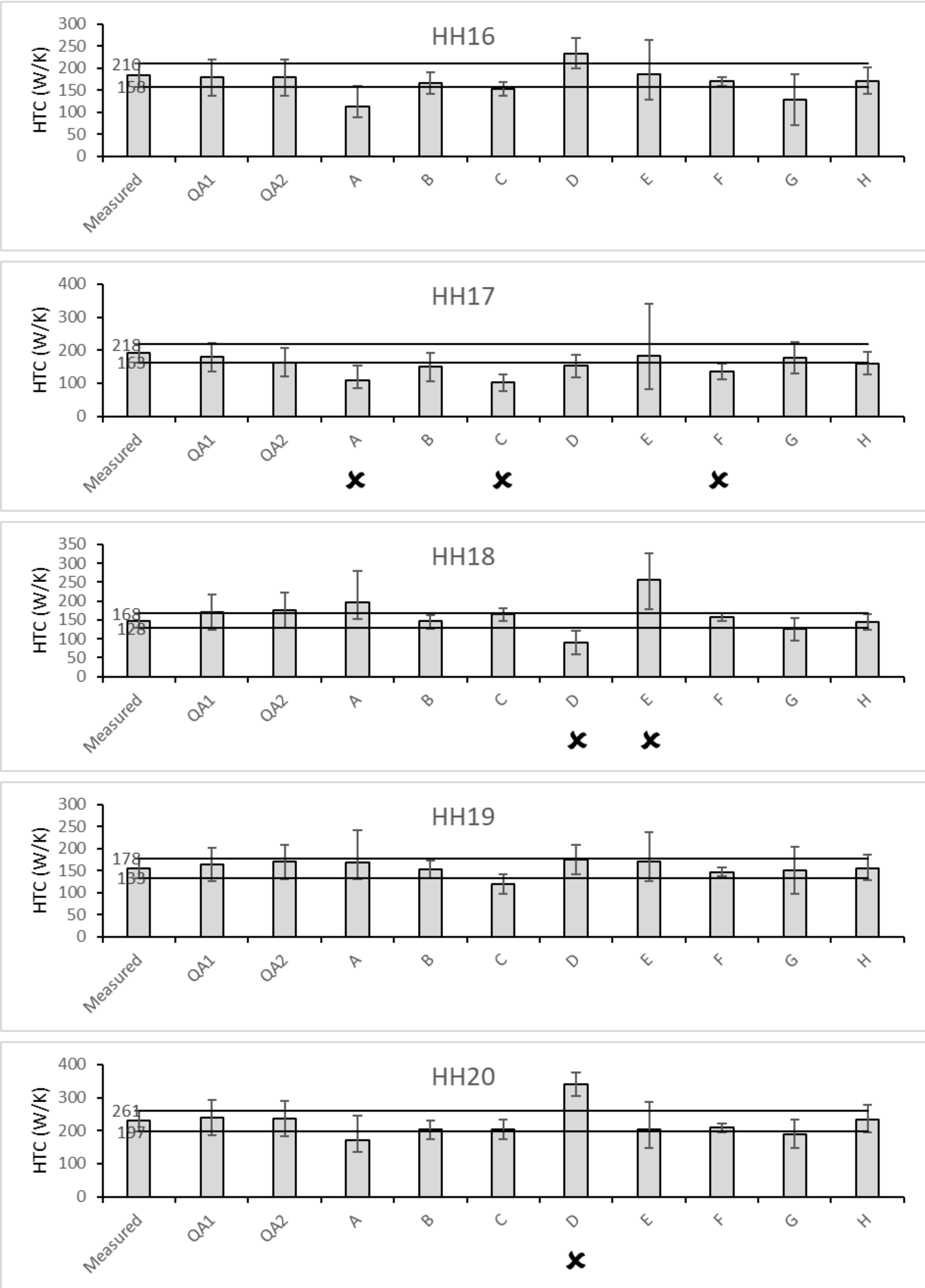


Figure 4: HTC results (first monitoring period) for HH16 to HH20 with error bars for 95% confidence interval; the horizontal lines show the 95% confidence interval for the measured HTC and ✖ indicates an unsuccessful result based on overlapping confidence intervals

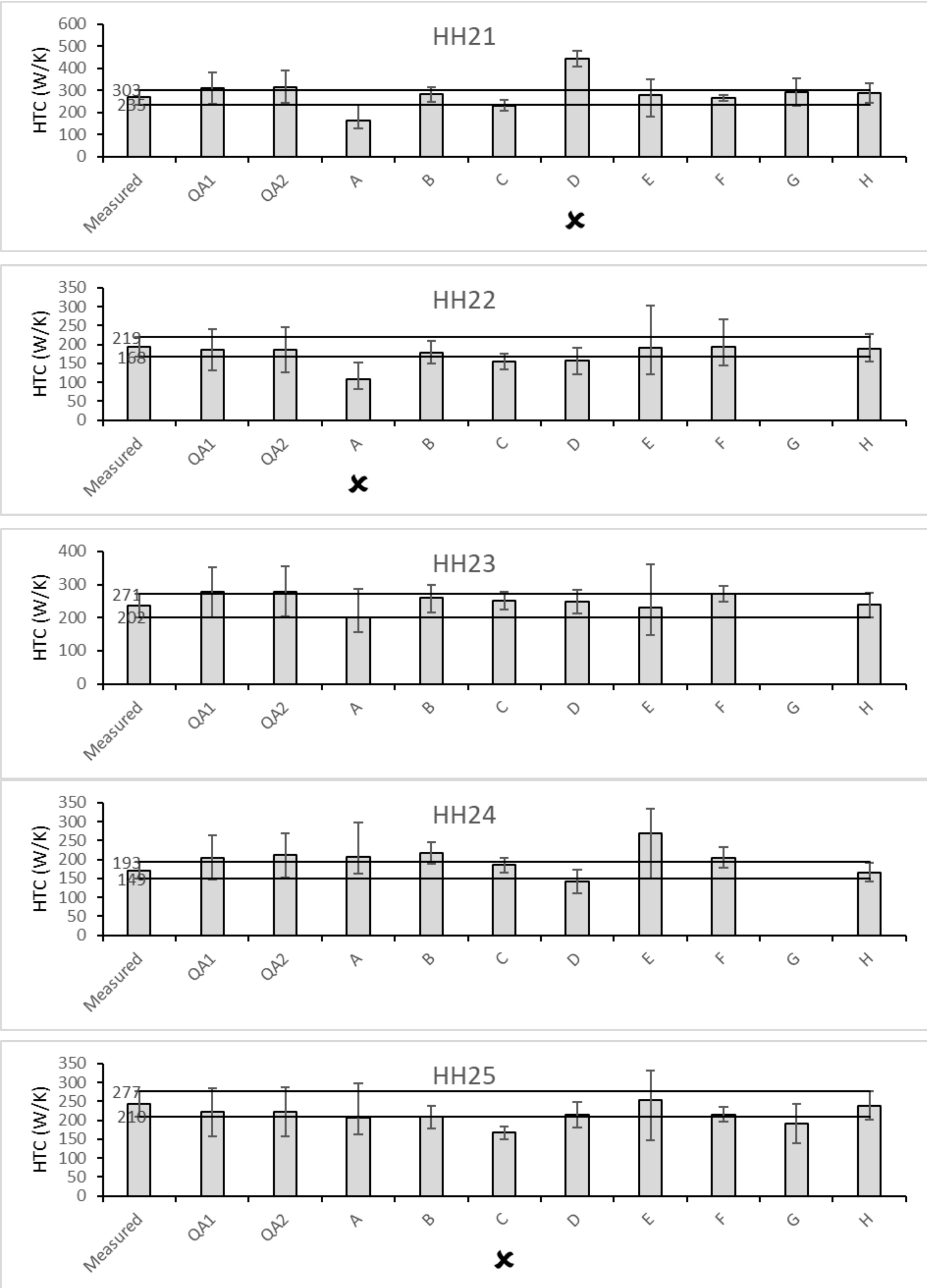


Figure 5: HTC results (first monitoring period) for HH21 to HH25 with error bars for 95% confidence interval; the horizontal lines show the 95% confidence interval for the measured HTC and ✖ indicates an unsuccessful result based on overlapping confidence intervals

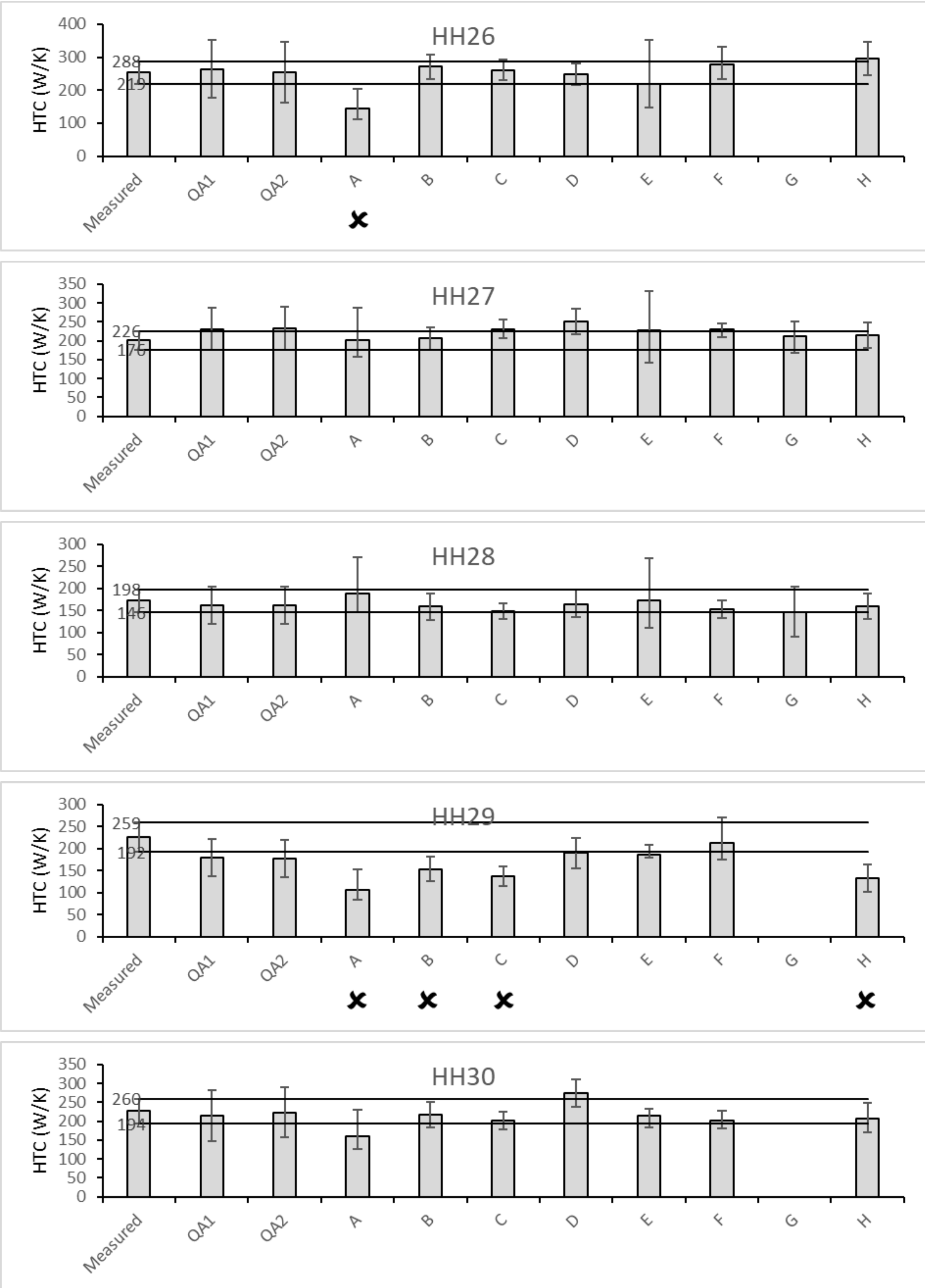


Figure 6: HTC results (first monitoring period) for HH25 to HH30 with error bars for 95% confidence interval; the horizontal lines show the 95% confidence interval for the measured HTC and ✖ indicates an unsuccessful result based on overlapping confidence intervals

Table 26: Summary of results (first monitoring period) - tick indicates successful SMETER result (confidence interval of SMETER and measured HTCs overlap), TEST 6 & TEST 4 (bold)

Home ID	Number of results with overlapping CI	QA1	QA2	SMETER							
				A	B	C	D	E	F	G	H
HH01	8	✓	✓	**✓**	✓	**✓**	**✗**	✓	✓	na	✓
HH02	9	✓	✓	✓	**✓**	✓	**✓**	✓	✓	na	✓
HH03	9	✓	✓	✓	**✓**	✓	✓	✓	✓	na	✓
HH04	10	✓	✓	✓	✓	✓	✓	**✓**	**✓**	✓	**✓**
HH05	10	✓	✓	✓	✓	✓	✓	**✓**	**✓**	✓	**✓**
HH06	8	✓	✓	✓	✓	✗	✓	**✓**	**✓**	na	**✓**
HH07	10	✓	✓	✓	✓	✓	✓	**✓**	**✓**	✓	**✓**
HH08	9	✓	✓	**✓**	✓	**✓**	✗	✓	✓	**✓**	✓
HH09	10	✓	✓	**✓**	✓	**✓**	✓	✓	✓	**✓**	✓
HH10	10	✓	✓	✓	✓	✓	✓	**✓**	✓	✓	**✓**
HH11	6	✓	✓	✓	✗	✓	**✗**	✓	**✗**	na	✓
HH12	8	✓	✓	✗	✓	✓	**✓**	✓	**✓**	na	✓
HH13	10	✓	✓	✓	**✓**	✓	✓	✓	**✓**	**✓**	✓
HH14	7	✓	✗	✓	**✓**	✓	✓	✗	✓	na	✓
HH15	9	✓	✓	✓	✓	✓	✗	✓	✓	**✓**	✓
HH16	10	✓	✓	**✓**	✓	**✓**	✓	✓	**✓**	**✓**	✓
HH17	7	✓	✓	**✗**	✓	**✗**	✓	✓	✗	**✓**	✓
HH18	8	✓	✓	✓	✓	✓	✗	✗	✓	**✓**	✓
HH19	10	✓	✓	✓	✓	✓	✓	✓	**✓**	**✓**	✓
HH20	9	✓	✓	**✓**	✓	**✓**	**✗**	✓	**✓**	✓	✓
HH21	9	✓	✓	**✓**	✓	**✓**	**✗**	✓	✓	✓	✓
HH22	8	✓	✓	**✗**	✓	**✓**	**✓**	✓	✓	na	✓
HH23	9	✓	✓	✓	**✓**	✓	✓	**✓**	✓	na	**✓**
HH24	9	✓	✓	✓	**✓**	✓	✓	**✓**	✓	na	**✓**
HH25	9	✓	✓	✓	**✓**	✗	✓	**✓**	✓	✓	**✓**
HH26	8	✓	✓	✗	**✓**	✓	✓	**✓**	✓	na	**✓**
HH27	10	✓	✓	✓	**✓**	✓	**✓**	✓	✓	✓	✓
HH28	10	✓	✓	✓	**✓**	✓	✓	**✓**	✓	✓	**✓**
HH29	5	✓	✓	**✗**	✗	**✗**	**✓**	✓	✓	na	✗
HH30	9	✓	✓	**✓**	✓	**✓**	**✓**	✓	✓	na	✓
TEST 4	Number attempted			10	10	10	10	10	10	8[1]	10
	Number with overlapping CIs			7	10	8	6	10	9	8	10
	TEST 4 result (as % attempted)			70%	100%	80%	60%	100%	90%	100%	100%
	Maximum CI, % of declared HTC			43%	74%	25%	29%	68%	10%	45%	22%
	Average CI, % of declared HTC			43%	23%	14%	17%	49%	6%	31%	17%
TEST 6	Number attempted			20	20	20	20	20	20	9[1]	20
	Number with overlapping CIs			18	18	18	17	18	19	9	19
	TEST 6 result (as % attempted)			90%	90%	90%	85%	90%	95%	100%	95%
	Maximum CI, % of declared HTC			43%	29%	60%	35%	89%	86%	39%	80%
	Average CI, % of declared HTC			43%	15%	14%	19%	43%	17%	26%	22%
Total	Number attempted	30	30	30	30	30	30	30	30	17[1]	30
	Number with overlapping CIs	30	29	25	28	26	23	28	28	17	29
	Combined result (as % attempted)	100%	97%	83%	93%	87%	77%	93%	93%	100%	97%
	Maximum CI, % of declared HTC	41%	39%	43%	74%	60%	35%	89%	86%	45%	80%
	Average CI as % of declared HTC	26%	27%	43%	18%	14%	18%	45%	13%	29%	21%

Greyed out cells are where TEST 4 and 6 methods are identical (SMETER A and SMETER E), or not applicable (two QA methods).

[1] The participating organisation behind SMETER G commented that this was because of gaps in the data provided by the TEST Project.

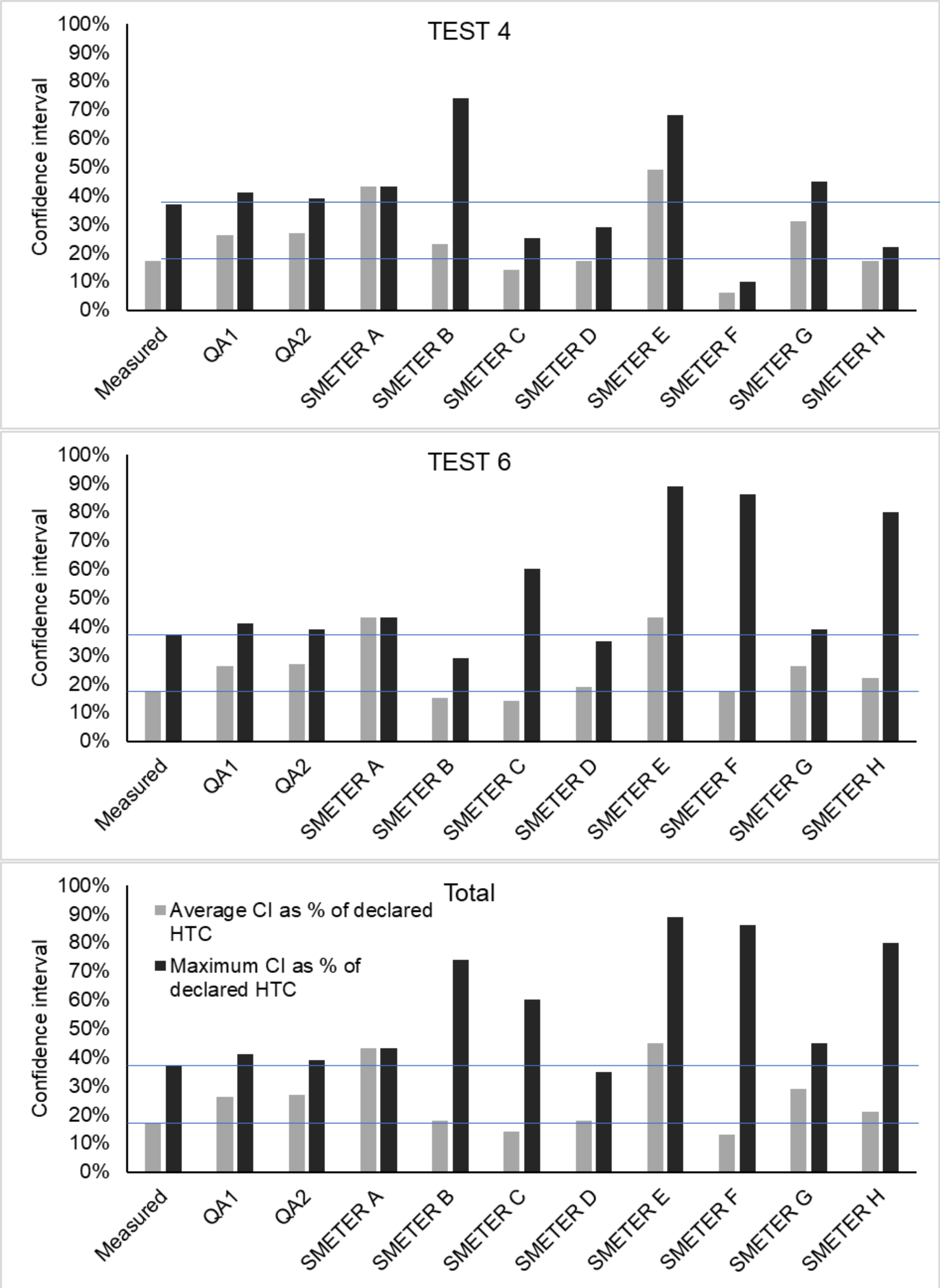
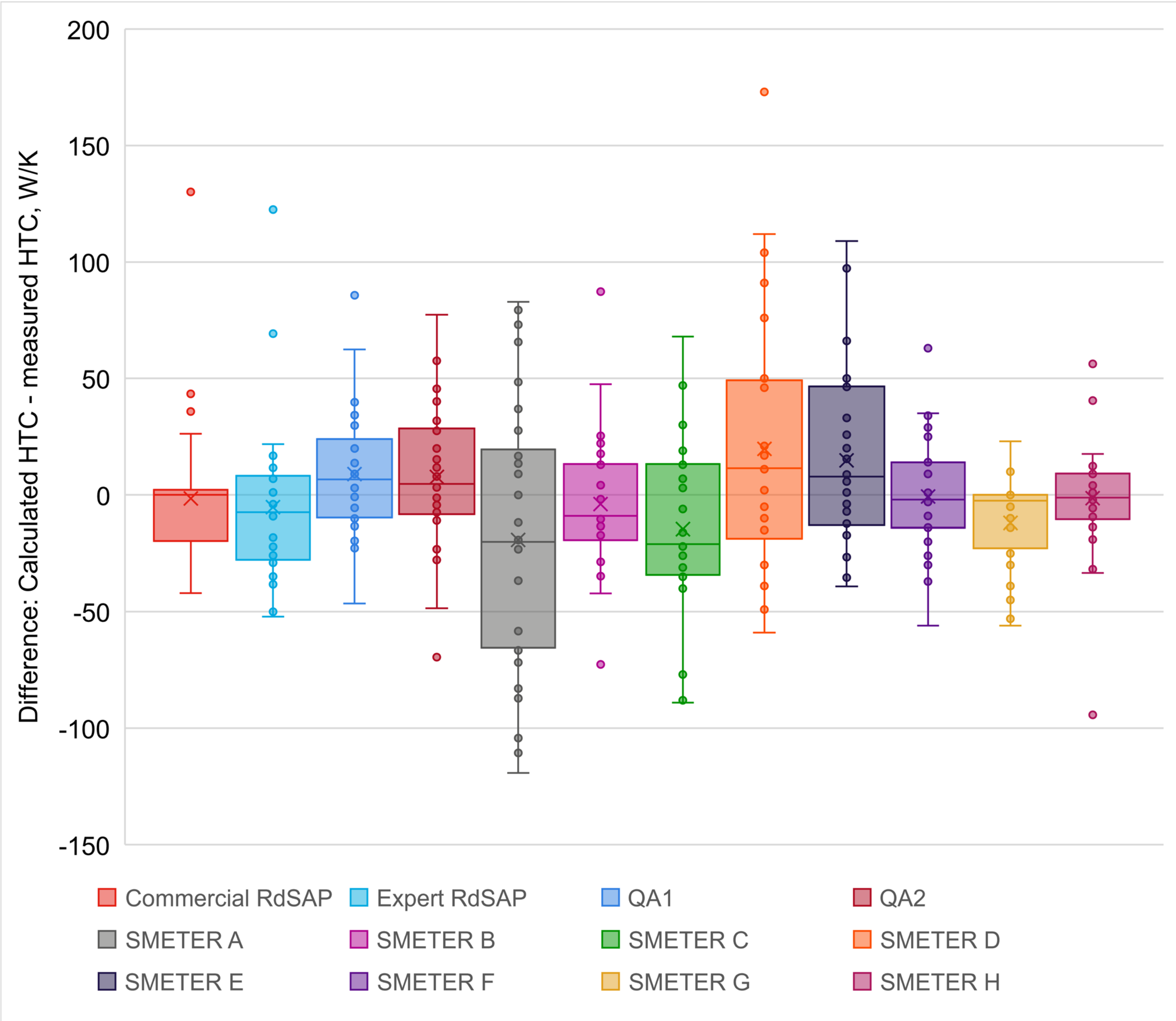


Figure 7: Average and maximum confidence intervals for TEST 4, TEST 6, and the combined total (first monitoring period)

Figure 8: Box-whisker plot of the distribution of the differences from the measured HTC for each home (first monitoring period)



6 Evaluation of SMETER accuracy: second monitoring period

The results were updated following the optional second monitoring period where five participating organisations provided new results for 27 homes. For completeness and ease of comparison, the results are summarised for all eight organisations, removing the three homes that were not part of the second monitoring period (HH07, HH18, and HH30) (Table 28 and Table 29).

New results were provided from SMETER A, SMETER C, SMETER E, SMETER F and SMETER G. Four of these five participating organisations declared that they used data from the second monitoring period only, while the participating organisation behind SMETER E declared using data from both the first and second monitoring periods (Appendix B). Three of the participating organisations used longer periods of data for the new results (Table 30): SMETER A (208 days on average, 179 days more than for the first set of results), SMETER C (101 days on average, 24 days more), and SMETER E (311 days on average, 95 days more). A shorter data period was used by SMETER F (51 days, 21 days fewer than for the first set) and the same length of period was used by SMETER G (7 days).

The same two approaches as described in Section 5 were used for the re-evaluation of the results. Here, as in Section 5, the measured HTC for 6 homes is known to the participating organisations (see Section 2.5).

6.1 Comparison of calculated HTC with measured HTC using confidence intervals

Participating organisations provided HTC calculations for all 27 homes, except for the organisation behind SMETER C which reported results for 26 homes (no result given for HH05) (Table 31). Overall, SMETERs were successful for between 70% and 96% of the homes. The average of the reported confidence interval ranged from 14% (SMETER F) to 33% (SMETER A and SMETER G). The maximum declared confidence interval ranged from 32% (SMETER D) to 80% (SMETER H).

Some of the participating organisations were able to improve their results by using the longer period of data from the second monitoring period (Table 31). For SMETER E, the reported confidence interval was reduced by an average of 19 percentage points, while the percentage of successful results was increased by 3 percentage points. SMETER F also increased the percentage of successful results by 3 percentage points. SMETER G was able to report results for all the homes (compared with 17 out of 30 from the first monitoring period) with only a small increase in the average confidence interval (4 percentage points) and a small decrease in the percentage of correct results (7 percentage points). The results for SMETER C and SMETER

F changed very little. The average confidence interval for SMETER A was reduced by 10 percentage points, but this reduced the percentage of successful results by 13 percentage points. There were small changes to the results for the three SMETERs that did not use data from the second monitoring period due to the change in the analysis from all 30 homes to the subset of 27 – these are reported here for completeness, but the full set of results for 30 homes is used in subsequent comparison (e.g., Section 7, Section 8, and the conclusions in Section 9).

Table 27: HTC results for SMETERS A-D for the first monitoring period, or using data from the second monitoring period (bold)

Home ID	SMETER A			SMETER B			SMETER C			SMETER D		
	HTC	Lower	Upper	HTC	Lower	Upper	HTC	Lower	Upper	HTC	Lower	Upper
HH01	171.1	115.5	226.8	199.6	171.4	227.6	177.1	158.3	195.8	272.0	236.0	307.0
HH02	187.2	126.4	248.0	135.2	115.5	153.3	137.6	125.8	149.3	143.0	107.0	180.0
HH03	123.1	83.1	163.1	120.2	85.3	157.8	125.7	106.1	145.4	106.0	72.0	139.0
HH04	144.8	97.7	191.8	142.0	125.9	156.2	130.2	115.6	144.8	174.0	139.0	209.0
HH05	117.8	79.5	156.1	202.3	176.5	225.1	na	na	na	221.0	188.0	254.0
HH06	93.0	62.8	123.2	237.3	206.0	268.4	189.4	167.4	211.4	258.0	223.0	292.0
HH08	156.1	105.3	206.8	218.7	193.0	241.5	224.4	203.6	245.3	290.0	258.0	322.0
HH09	119.8	80.9	158.7	159.3	142.3	174.2	137.5	116.2	158.7	157.0	126.0	189.0
HH10[1]	139.3	94.0	184.6	208.0	174.8	241.5	144.3	118.2	170.5	196.0	163.0	229.0
HH11	170.3	114.9	225.6	300.2	264.1	332.2	265.8	243.9	287.7	289.0	257.0	322.0
HH12	117.1	79.1	155.2	125.2	104.5	144.4	119.7	104.0	135.4	114.0	81.0	147.0
HH13[1]	127.1	85.8	168.4	150.2	124.7	174.8	126.0	103.7	148.3	176.0	143.0	209.0
HH14	185.9	125.5	246.3	126.7	96.4	220.1	170.6	149.6	191.7	160.0	128.0	191.0
HH15	115.8	78.2	153.4	181.7	152.7	209.2	202.1	175.6	228.6	312.0	278.0	347.0
HH16	106.3	71.8	140.8	166.8	141.6	191.1	154.3	134.4	174.3	233.0	199.0	268.0
HH17	137.6	92.9	182.4	148.9	105.2	192.5	129.6	104.5	154.7	152.0	119.0	185.0
HH19[1]	120.3	81.2	159.4	153.3	131.8	173.9	126.0	106.3	145.8	176.0	142.0	209.0
HH20	158.0	106.6	209.3	203.0	174.7	229.8	180.3	158.0	202.6	341.0	306.0	376.0
HH21[1]	170.4	115.0	225.7	282.0	246.2	313.8	254.8	231.9	277.8	442.0	407.0	478.0
HH22[1]	177.4	119.8	235.1	179.0	149.9	209.0	217.2	194.2	240.1	157.0	122.0	191.0
HH23	174.3	117.7	231.0	259.1	216.5	298.1	189.7	171.5	207.8	248.0	213.0	283.0
HH24	224.3	151.4	297.2	218.5	188.4	245.5	171.2	156.9	185.4	141.0	110.0	172.0
HH25[1]	175.3	118.3	232.3	210.0	178.7	238.9	223.1	204.7	241.5	214.0	180.0	249.0
HH26	152.6	103.0	202.2	271.6	232.2	308.3	231.4	210.5	252.3	249.0	215.0	282.0
HH27	130.3	88.0	172.7	207.0	176.8	235.9	178.3	161.2	195.3	251.0	216.0	286.0
HH28	164.2	110.8	217.5	158.6	127.8	189.3	143.3	127.6	159.0	165.0	134.0	197.0
HH29	116.1	78.4	153.9	153.4	125.5	182.4	119.3	95.8	142.8	190.0	154.0	225.0

[1] The participating organisations knew the measured HTC for HH10, HH13, HH19, HH21, HH22, and HH25.

Table 28: HTC results for SMETERS E-H for the first monitoring period, or using data from the second monitoring period (bold)

Home ID	**SMETER E**			**SMETER F**			**SMETER G**			SMETER H		
	HTC	Lower	Upper	HTC	Lower	Upper	HTC	Lower	Upper	HTC	Lower	Upper
HH01	**204.7**	**157.0**	**255.0**	**182.0**	**157.0**	**206.0**	**180.0**	**135.0**	**225.0**	190.0	152.3	229.9
HH02	**210.2**	**161.0**	**262.3**	**142.0**	**131.0**	**153.0**	**136.0**	**103.0**	**170.0**	138.1	115.1	161.1
HH03	**184.7**	**136.6**	**237.7**	**148.0**	**118.0**	**184.0**	**102.0**	**62.0**	**142.0**	139.0	98.7	183.3
HH04	**202.4**	**157.1**	**249.3**	**137.0**	**127.0**	**147.0**	**112.0**	**74.0**	**150.0**	153.0	132.6	173.6
HH05	**216.0**	**166.8**	**267.4**	**215.0**	**196.0**	**234.0**	**165.0**	**109.0**	**221.0**	228.8	193.2	264.4
HH06	**196.8**	**150.4**	**246.1**	**185.0**	**169.0**	**201.0**	**187.0**	**127.0**	**246.0**	268.3	224.5	316.5
HH08	**223.3**	**172.8**	**275.8**	**233.0**	**219.0**	**250.0**	**231.0**	**188.0**	**273.0**	211.1	176.7	246.2
HH09	**183.9**	**139.9**	**230.8**	**140.0**	**117.0**	**165.0**	**141.0**	**98.0**	**185.0**	166.3	139.4	196.5
HH10[1]	***175.8***	***133.5***	***221.1***	***149.0***	***108.0***	***190.0***	***143.0***	***97.0***	***189.0***	*184.6*	*150.3*	*224.3*
HH11	**184.5**	**139.7**	**232.5**	**261.0**	**241.0**	**281.0**	**232.0**	**186.0**	**277.0**	255.4	215.5	296.8
HH12	**191.0**	**146.3**	**238.3**	**118.0**	**107.0**	**129.0**	**128.0**	**90.0**	**167.0**	133.7	113.7	154.0
HH13[1]	***179.2***	***134.3***	***228.1***	***130.0***	***111.0***	***149.0***	***150.0***	***75.0***	***225.0***	*171.8*	*142.0*	*205.9*
HH14	**187.1**	**143.7**	**232.7**	**163.0**	**130.0**	**196.0**	**207.0**	**160.0**	**254.0**	146.8	99.8	263.6
HH15	**212.9**	**164.7**	**263.0**	**219.0**	**182.0**	**258.0**	**177.0**	**110.0**	**244.0**	207.5	173.6	243.9
HH16	**177.0**	**134.1**	**223.0**	**156.0**	**134.0**	**177.0**	**144.0**	**102.0**	**186.0**	170.3	141.8	201.4
HH17	**163.9**	**123.2**	**207.9**	**160.0**	**106.0**	**212.0**	**115.0**	**62.0**	**168.0**	159.3	127.1	195.4
HH19[1]	***194.1***	***147.9***	***243.1***	***139.0***	***118.0***	***162.0***	***124.0***	***75.0***	***172.0***	*155.4*	*127.4*	*185.9*
HH20	**196.0**	**150.7**	**243.4**	**202.0**	**172.0**	**232.0**	**170.0**	**115.0**	**224.0**	233.1	194.5	276.6
HH21[1]	***260.3***	***195.3***	***330.8***	***261.0***	***239.0***	***281.0***	***292.0***	***202.0***	***382.0***	*286.6*	*244.9*	*331.1*
HH22[1]	***216.0***	***157.1***	***282.2***	***197.0***	***175.0***	***217.0***	***167.0***	***114.0***	***219.0***	*188.7*	*153.7*	*228.3*
HH23	**208.7**	**159.3**	**261.1**	**187.0**	**169.0**	**204.0**	**193.0**	**129.0**	**257.0**	237.9	200.4	275.6
HH24	**203.3**	**153.8**	**256.5**	**166.0**	**153.0**	**179.0**	**175.0**	**139.0**	**212.0**	166.2	140.9	190.9
HH25[1]	***270.5***	***193.2***	***358.9***	***219.0***	***200.0***	***240.0***	***263.0***	***142.0***	***384.0***	*238.8*	*201.6*	*277.8*
HH26	**212.0**	**159.6**	**268.7**	**241.0**	**198.0**	**281.0**	**217.0**	**172.0**	**263.0**	294.5	245.9	347.4
HH27	**206.2**	**157.8**	**257.4**	**187.0**	**172.0**	**202.0**	**175.0**	**116.0**	**234.0**	213.4	180.0	249.1
HH28	**204.2**	**156.7**	**254.3**	**156.0**	**134.0**	**180.0**	**138.0**	**91.0**	**185.0**	158.8	130.2	189.3
HH29	**167.1**	**123.7**	**215.1**	**135.0**	**114.0**	**157.0**	**106.0**	**53.0**	**159.0**	131.7	101.5	164.5

[1] The participating organisations knew the measured HTC for HH10, HH13, HH19, HH21, HH22, and HH25.

Table 29: Length in days of the self-reported period of data used by each SMETER for the first monitoring period, or using data from the second monitoring period (bold)

Home ID	SMETER							
	A	B	**C**	D	**E**	**F**	**G**	**H**
HH01	**208**	21	**110**	21	**58**	**57**	**7**	23
HH02	**208**	21	**100**	21	**58**	**51**	**7**	22
HH03	**208**	21	**112**	21	**58**	**81**	**7**	20
HH04	**208**	21	**89**	21	**421**	**50**	**7**	102
HH05	**208**	21	**na**	21	**421**	**59**	**7**	123
HH06	**208**	21	**77**	21	**364**	**54**	**7**	56
HH08	**208**	21	**105**	21	**58**	**46**	**7**	122
HH09	**208**	21	**110**	21	**58**	**50**	**7**	119
HH10	**208**	21	**99**	21	**407**	**52**	**7**	99
HH11	**208**	21	**91**	21	**421**	**50**	**7**	37
HH12	**208**	21	**109**	21	**406**	**51**	**7**	93
HH13	**208**	21	**108**	21	**421**	**51**	**7**	116
HH14	**208**	79	**88**	21	**421**	**51**	**7**	108
HH15	**208**	21	**104**	21	**421**	**22**	**7**	118
HH16	**208**	21	**101**	21	**413**	**50**	**7**	81
HH17	**208**	21	**90**	21	**413**	**50**	**7**	107
HH19	**208**	21	**71**	21	**413**	**50**	**7**	107
HH20	**208**	21	**74**	21	**413**	**50**	**7**	84
HH21	**208**	21	**83**	21	**406**	**50**	**7**	98
HH22	**208**	21	**86**	21	**384**	**50**	**7**	34
HH23	**208**	21	**90**	21	**379**	**55**	**7**	69
HH24	**208**	22	**140**	21	**364**	**50**	**7**	55
HH25	**208**	21	**103**	21	**364**	**52**	**7**	57
HH26	**208**	21	**144**	21	**364**	**50**	**7**	49
HH27	**208**	13	**146**	21	**364**	**50**	**7**	50
HH28	**208**	14	**110**	21	**58**	**50**	**7**	44
HH29	**208**	21	**97**	21	**58**	**50**	**7**	29
Average	**208**	23	**101**	21	**311**	**51**	**7**	75
Difference[1]	**+179**		**+24**		**+95**	**-21**	**0**	

[1] the difference in the length of the self-reported period of data used, compared to the number given for the first data collection period (Table 25)

Table 30: Summary of results – tick indicates successful SMETER result (confidence interval of SMETER and measured HTCs overlap), for the first monitoring period, or using data from the second monitoring period (bold)

Home ID	SMETER							
	A	B	**C**	D	**E**	**F**	**G**	H
HH01	✓	✓	✓	✗	✓	✓	✓	✓
HH02	✓	✓	✓	✓	✓	✓	✓	✓
HH03	✓	✓	✓	✓	✓	✓	✓	✓
HH04	✓	✓	✓	✓	✓	✓	✓	✓
HH05	✗	✓	**na**	✓	✓	✓	✓	✓
HH06	✗	✓	✓	✓	✓	✓	✓	✓
HH08	✓	✓	✓	✗	✓	✓	✓	✓
HH09	✓	✓	✓	✓	✓	✓	✓	✓
HH10[1]	✓	✓	✗	✓	✓	✓	✓	✓
HH11	✓	✗	✗	✗	✓	✗	✓	✓
HH12	✓	✓	✓	✓	✗	✓	✓	✓
HH13[1]	✓	✓	✓	✓	✓	✓	✓	✓
HH14	✓	✓	✓	✓	✓	✓	✗	✓
HH15	✗	✓	✓	✗	✓	✓	✓	✓
HH16	✗	✓	✓	✓	✓	✓	✓	✓
HH17	✓	✓	✗	✓	✓	✓	✓	✓
HH19[1]	✓	✓	✓	✓	✓	✓	✓	✓
HH20	✓	✓	✓	✗	✓	✓	✓	✓
HH21[1]	✗	✓	✓	✗	✓	✓	✓	✓
HH22[1]	✓	✓	✓	✓	✓	✓	✓	✓
HH23	✓	✓	✓	✓	✓	✓	✓	✓
HH24	✓	✓	✓	✓	✓	✓	✓	✓
HH25[1]	✓	✓	✓	✓	✓	✓	✓	✓
HH26	✗	✓	✓	✓	✓	✓	✓	✓
HH27	✗	✓	✓	✓	✓	✓	✓	✓
HH28	✓	✓	✓	✓	✓	✓	✓	✓
HH29	✗	✗	✗	✓	✓	✗	✗	✗
Number attempted	**27**	27	**26**	27	**27**	**27**	**27**	27
Number with overlapping CIs	**19**	25	**22**	21	**26**	**26**	**25**	26
Combined result (as % attempted)	**70%**	93%	**85%**	78%	**96%**	**96%**	**93%**	96%
Maximum CI, % of declared HTC	**33%**	74%	**20%**	32%	**33%**	**34%**	**50%**	80%
Average CI as % of declared HTC	**33%**	18%	**12%**	18%	**26%**	**14%**	**33%**	21%
Change in result (as % attempted)[2]	**-13%**	-0.7%	**-2.1%**	+1.1%	**+3.0%**	**+3%**	**-7.4%**	-0.4%
Change in average confidence interval	**-10%**	+0.4%	**-2.0%**	-0.6%	**-19%**	**+0.5%**	**+3.9%**	+0.3%

[1] The participating organisations knew the measured HTC for HH10, HH13, HH19, HH21, HH22, and HH25.

[2] The change in the result compared with the results given in the first monitoring period for 30 homes. SMETER B, SMETER D, and SMETER H did not submit new results but the results they gave for 30 homes have been reanalysed for the 27 homes shown in this table.

6.2 Analysis of the differences between the SMETER result and the measured HTC

The difference between each calculated HTC and the associated measured HTC was re-calculated for the 27 homes (Figure 10). The new results submitted after the second monitoring period were used for SMETER A, SMETER C, SMETER E, SMETER F, and SMETER G (Table 32). This did not improve the results in most cases. However, for SMETER E, the NMBE was reduced by 3.3 percentage points and the CVRMSE was reduced by 2.6 percentage points. This improvement could be due to the availability of a longer period of data to analyse (Table 30). The change from analysing 30 homes to analysing 27 made only a small difference to the NMBE and CVRMSE for the three SMETERs that did not use data from the second monitoring period - these are reported here for completeness, but the full set of results for 30 homes is used in subsequent comparison (e.g., Section 7, Section 8, and the conclusions in Section 9).

Figure 9: Box-whisker plot of the distribution of the differences from the measured HTC for each home (* using data from second monitoring period)

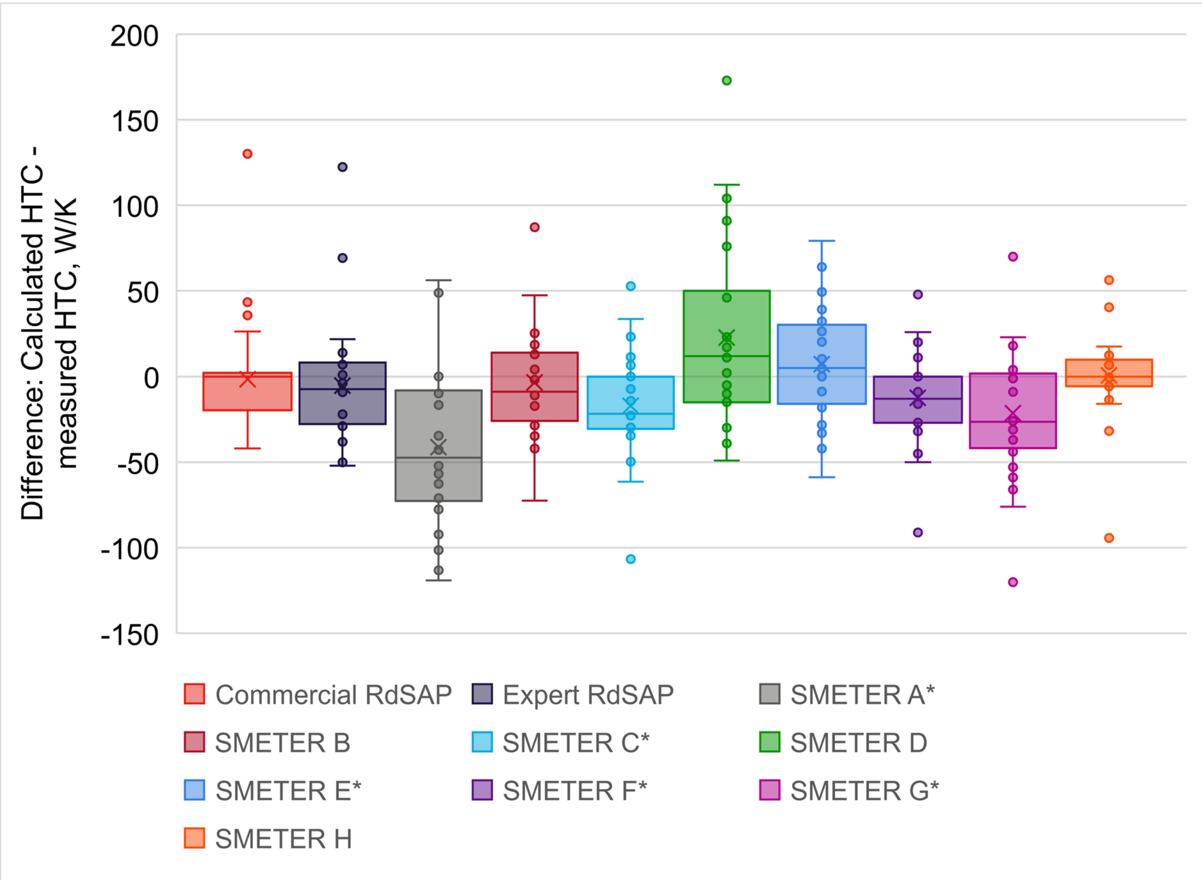


Table 31: Analysis of difference between calculated and measured HTCs, for the first monitoring period, or using data from the second monitoring period (bold)

	SMETER							
	A	B	**C**	D	**E**	**F**	**G**	H
Number attempted	**27**	27	**26**	27	**27**	**27**	**27**	27
Mean difference, MBE	**-45.8**	-3.5	**-19.7**	22.7	**8.1**	**-14.3**	**-23.8**	0.6
Normalised mean difference, NMBE	**-26.9%**	-1.8%	**-10.9%**	11.1%	**4.1%**	**-7.7%**	**-13.1%**	0.3%
Std.Dev. of difference, W/K	**48.6**	30.7	**31.8**	54.5	**34.0**	**26.5**	**37.9**	26.5
Root mean square error, RMSE, W/K	**66.2**	30.3	**36.9**	58.1	**34.3**	**29.7**	**44.2**	26.0
RMSE/average result, CVRMSE	**38.9%**	15.9%	**20.3%**	28.4%	**17.4%**	**15.9%**	**24.4%**	13.4%
Est. upper limit of agreement, W/K	**49.4**	56.7	**42.5**	129.6	**74.8**	**37.7**	**50.5**	52.5
Est. lower limit of agreement, W/K	**-141.1**	-63.7	**-82.0**	-84.1	**-58.6**	**-66.2**	**-98.1**	-51.3
Absolute change in NMBE, pp	**+16**	-0.2	**+2.9**	+1.3	**-3.3**	**+7.3**	**+1.6**	-0.4
Change in CVRMSE, pp	**+6.3**	+0.6	**-0.3**	+0.2	**-2.6**	**+3.0**	**+7.8**	0.0

7 Comparing the results with RdSAP HTCs

The results for the SMETERs were compared with the expert RdSAP HTC (predicted using RdSAPv12, based on the expert survey of all 30 homes) and with the commercial RdSAP HTC (predicted using a commercial domestic energy assessor and Elmhurst Energy RdSAP platform Version 9.94 for 22 homes). This approach sought to answer the question: which SMETERs are more accurate than RdSAP?

Then, the results were re-analysed but ignoring the six homes in which the measured HTC was known by participating organisations (Section 2.5) to show that these six results did not materially change the evaluation of SMETER accuracy. Further analysis identified that it was particularly problematic to calculate the HTC for one of the homes (HH14), but again this did not materially change the evaluation of SMETER accuracy.

7.1 Analysis of results

The differences between the expert and commercial RdSAP HTC and the associated measured HTC were calculated for every case, along with the NMBE, CVRMSE and limits of agreement (Table 27). Interestingly, both the expert and commercial RdSAP HTCs show good agreement with measured HTCs. Previous work (Crawley et al., 2019) has revealed large discrepancies in the SAP ratings produced by different assessors. However, this is the first time that the accuracy of HTCs predicted using RdSAP survey data has been quantified. The commercial assessor was very experienced, and the homes were relatively simple to assess being small in size and without complicated features such as rooms-in-the-roof, extensions with different wall types, or conservatories. It is hypothesised that the average commercial RdSAP HTC would not be as accurate as observed here.

The calculated NMBE and CVRMSE and the limits of agreement for the two RdSAP estimates provide the benchmarks against which the HTCs calculated by each SMETER can be compared. If a SMETER has a NMBE closer to zero and a CVRMSE lower (i.e., better) than the corresponding values produced by RdSAP then the SMETER is more accurate than RdSAP for this sample of 30 homes.

The two best-performing SMETER technologies (the lowest CVRMSE and NMBE closest to zero) were more accurate than the expert RdSAP: SMETER B (NMBE -2.1%, CVRMSE 15.2%) and SMETER H (NMBE -0.7%, CVRMSE 13.4%) (Table 33). However, compared with the commercial RdSAP HTC results for 22 of the homes (NMBE -1.1%, CVRMSE 19.6%) only SMETER H (NMBE -0.7%, CVRMSE 13.4%) was more accurate as the NMBE of SMETER B was further from zero. Two SMETER technologies had a lower CVRMSE than either of the

RdSAP HTC predictions, but their NMBE was not as good: SMETER E (NMBE 4.1%, CVRMSE 17.4%) and SMETER F (NMBE -7.7%, CVRMSE 15.9%).

Table 32: Comparison of calculated HTC results with expert and commercial RdSAP HTCs (results in bold are using the second monitoring period)

Criterion	SMETER							
	A	B	**C**	D	**E**	**F**	**G**	H
Number attempted	**27**	30	**26**	30	**27**	**27**	**27**	30
Success rate (% of attempted)	**70%**	93%	**85%**	77%	**96%**	**96%**	**93%**	97%
Average CI	**33%**	18%	**12%**	18%	**26%**	**14%**	**33%**	21%
NMBE	**-27%**	-2.1%	**-11%**	9.8%	**4.1%**	**-7.7%**	**-13%**	-0.7%
CVRMSE	**38.9%**	15.2%	**20.3%**	28.2%	**17.4%**	**15.9%**	**24.4%**	13.4%
Est. upper limit of agreement, W/K	**49.4**	53.2	**42.5**	126.0	**74.8**	**37.7**	**50.5**	49.7
Est. lower limit of agreement, W/K	**-141.1**	-61.1	**-82.0**	-86.6	**-58.6**	**-66.2**	**-98.1**	-52.4
NMBE better than expert RdSAP (-2.8%)[1]?	✗	✓	✗	✗	✗	✗	✗	✓
CVRMSE better than expert RdSAP (18.2%)[1]?	✗	✓	✗	✗	✓	✓	✗	✓
Limits of agreement better than expert RdSAP (+62.4 W/K to -73.0 W/K)[1]?	✗	✓	✗	✗	✗	✓	✗	✓
NMBE better than commercial RdSAP (-1.1%)[2]?	✗	✗	✗	✗	✗	✗	✗	✓
CVRMSE better than commercial RdSAP (19.6%)[2]?	✗	✓	✗	✗	✓	✓	✗	✓
Limits of agreement better than commercial RdSAP (+74.9 W/K to -67.7 W/K)[2]?	✗	✓	✗	✗	✓	✓	✗	✓

[1] Values for expert RdSAP HTC are based on the results for all 30 homes (Table 27).

[2] Values for the commercial RdSAP HTC are based on the results for 22 homes (Table 27).

7.2 Reanalysis of the results ignoring the six known HTC results

The results for each SMETER technology were re-analysed after removing the six homes (HH10, HH13, HH19, HH21, HH22 and HH25) for which the HTC was known following the interim assessment (Section 2.5). This had only a small effect on the outcome of the comparisons with the expert and commercial RdSAP HTCs (Table 34). However, the CVRMSE of SMETER E increased (from 17.4% to 19%) and was therefore no longer better than the CVRMSE for the expert RdSAP HTC (18.2%), but still better than the commercial RdSAP HTC (19.6%). The limits of agreement for SMETER C and SMETER G were slightly improved resulting in both SMETERs having a smaller range than the commercial RdSAP HTCs, and SMETER G a smaller range than the expert RdSAP HTCs (Table 34).

7.3 Re-analysis of the results ignoring HH14

The calculated HTC results for one home (HH14) had a wider reported confidence interval on average than the other homes after the first monitoring period (Figure 11). The QA method QA2 also failed to successfully calculate the HTC of this home (Figure 4), while both QA methods were successful on all other homes. The reasons for this issue are not understood but could be because the home was not fully occupied until the end of March, reducing the availability of data collected on cold days when the heating system was operating.

The results for each SMETER technology, based on the first monitoring period, were re-analysed after removing HH14. This showed that there was no change to the outcome of the comparison with expert RdSAP HTC (Table 35) and so no further action was taken.

Table 33: Comparison of calculated HTC results with expert RdSAP-predicted HTC – the six known HTCs have been removed (bold where the outcome changed)

Criterion	SMETER A	SMETER B	SMETER C	SMETER D	SMETER E	SMETER F	SMETER G	SMETER H
Number attempted	21	21	20	21	21	21	21	21
Success rate	70%	93%	85%	78%	96%	96%	93%	96%
Average CI	33%	18%	12%	18%	26%	14%	33%	21%
NMBE	-26%	-1.6%	-11%	11.1%	3.4%	-6.9%	-15%	0.13%
CVRMSE	40.7%	17.5%	21.6%	26.1%	19.0%	16.7%	26.9%	15.2%
Est. upper limit of agreement, W/K	99.1	58.4	58.8	115.1	98.2	54.6	15.5	54.2
Est. lower limit of agreement, W/K	-130	-65.5	-81.5	-77.9	-63.8	-53.1	-60.5	-58.5
NMBE better than expert RdSAP (-2.8%)[1]?	✗	✓	✗	✗	✗	✗	✗	✓
CVRMSE better than expert RdSAP (18.2%)[1]?	✗	✓	✗	✗	**✗**	✓	✗	✓
Limits of agreement better than expert RdSAP (+62.4 W/K to -73.0 W/K)[1]?	✗	✓	✗	✗	✗	✓	**✓**	✓
NMBE better than commercial RdSAP (-1.1%)[2]?	✗	✗	✗	✗	✗	✗	✗	✓
CVRMSE better than commercial RdSAP (19.6%)[2]?	✗	✓	✗	✗	✓	✓	✗	✓
Limits of agreement better than commercial RdSAP (+74.9 W/K to -67.7 W/K)[2]?	✗	✓	**✓**	✗	**✗**	✓	**✓**	✓

[1] Values for expert RdSAP are based on the results for all 30 homes (Table 27).

[2] Values for the commercial RdSAP HTC are based on the results for 22 homes (Table 27).

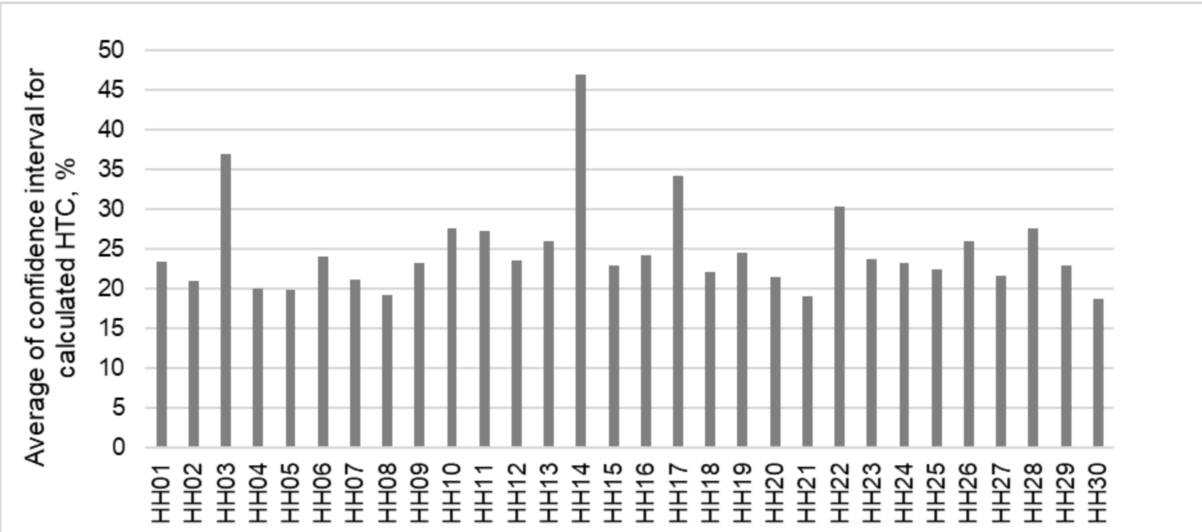


Figure 10: Average confidence interval for each home, as reported by participating organisations (first monitoring period)

Table 34: Comparison of calculated HTC results (first monitoring period) with expert RdSAP-predicted HTC – the results for HH14 has been removed

Criterion	SMETER							
	A	B	C	D	E	F	G	H
Number attempted (total)	29	29	29	29	29	29	17	29
Success rate (total)	83%	93%	86%	76%	97%	93%	100%	97%
Average CI (TEST 4), W/K	43%	17%	14%	17%	49%	6%	31%	17%
NMBE (total), W/K	-10.4%	-1.9%	-7.4%	9.6%	6.5%	-0.4%	-11.6%	-0.9%
NMBE better than expert RdSAP (-2.8%)[2]?	×	✓	×	×	×	✓	×	✓
CVRMSE (total)	31%	15%	20%	30%	20%	13%	16%	13%
CVRMSE better than expert RdSAP (18.2%)[2]?	×	✓	×	×	×	✓	✓	✓
Est. upper limit of agreement, W/K	94.5	54.4	56.5	127.8	85.4	49.3	22.6	50.1
Est. lower limit of agreement, W/K	-133.0	-61.8	-84.3	-88.6	-59.3	-51.0	-65.0	-53.5
Limits of agreement better than expert RdSAP (+62.4 W/K to -73.0 W/K)[2]?	×	✓	×	×	×	✓	✓	✓

[1] SMETER G reported 8 out of 9 in TEST 4 and 9 out of 20 in TEST 6

[2] Values for expert RdSAP are based on the results for all 30 homes.

8 Discussion

The overall attributes of the SMETER technologies and the calculated HTCs are compared and some general observations about their performance are made. The new results, following the second monitoring period, were used for the five participating organisations that submitted them (SMETER A, SMETER C, SMETER E, SMETER F, and SMETER G). The implications of these results for the implementation of SMETER technologies are discussed, and the limitations of the field trials and suggestions for further work are made.

8.1 Comparing the SMETERs

Overall, the SMETER technologies were successful for between 70% and 97% of the homes, with average confidence intervals between 12% and 33% (Table 36). Five participating organisations provided calculated HTC results that were more than 90% successful overall: SMETER B (28 out of 30 homes, with an average confidence interval of +/-18%), SMETER E (26 out of 27 homes, with an average confidence interval of +/-26%), SMETER F (26 out of 27 homes, with an average confidence interval of +/-14%), SMETER G (25 out of 27 homes, with an average confidence interval of +/-33%), and SMETER H (29 out of 30 homes, with an average confidence interval of +/-21%).

The suitability of a SMETER technology for a particular application will depend on various factors including (Table 36): accuracy (success rate, reported confidence interval, NMBE and CVRMSE), duration (average length of data period required for calculation), and cost/convenience (number of sensors in the home, whether a professional installation or survey is required). The survey requirement may be less expensive and less intrusive if the information can be taken from an existing EPC survey. The table (Table 36) has been colour coded (green, amber, red) to ease interpretation: success rate >90%/>80%/<80%; average confidence interval <15%/<20%/>20%; NMBE <5%/<10%/>10%; CVRMSE <20%/<30%/>30%; average length of data period required <14days/<31 days/>31 days; and Type T1/T2&T3/T4. The comparison with the RdSAP row has been colour coded green if the calculated HTC is more accurate than the RdSAP value or red if it is less accurate.

This colour coding is not indicative of fundamental problems with any SMETER technology. For example, SMETERs which use longer periods of data will be well-suited to many applications and SMETERs that are integrated in heating controllers (T4) offer little additional cost/inconvenience to a household choosing that controller. The required accuracy will depend on the application, and it may be possible to calibrate the SMETER technologies to improve the accuracy.

Table 35: Summary comparison of SMETER technologies (results in bold are using the second monitoring period)

Aspect	Criterion	SMETER							
		A BRE	B BTS	**C CAR**	D CSE	**E EDF**	**F HOA**	**G PAS**	H SWI
Accuracy	Number attempted	**27**	30	**26**	30	**27**	**27**	**27**	30
	Success rate	**70%**	93%	**85%**	77%	**96%**	**96%**	**93%**	97%
	Average CI declared	**33%**	18%	**12%**	18%	**26%**	**14%**	**33%**	21%
	NMBE	**-26.9%**	-2.1%	**-10.9%**	9.8%	**4.1%**	**-7.7%**	**-13.1%**	-0.7%
	CVRMSE	**38.9%**	15.2%	**20.3%**	28.2%	**17.4%**	**15.9%**	**24.4%**	13.4%
	NMBE better than expert RdSAP (-2.8%)?	✗	✓	✗	✗	✗	✗	✗	✓
	CVRMSE better than expert RdSAP (18.2%)?	✗	✓	✗	✗	✓ [1]	✓	✗	✓
	NMBE better than commercial RdSAP (-1.1%)?	✗	✗	✗	✗	✗	✗	✗	✓
	CVRMSE better than commercial RdSAP (19.6%)?	✗	✓	✗	✗	✓	✓	✗	✓
Duration	Average length of data period used by participants (self-reported, days)	**208**	22	**101**	21	**311**	**51**	**7**	74
Cost or convenience	Total number of room sensors	**0**	5	**5**	8	**0**	**4**	**2**	0
	Heating controller included			✓					✓
	Professional install?			✓ [2]	✓ [3]				✓ [2]
	Uses data from an EPC survey	✓	✓	✓			✓	✓	✓
	Requires additional home survey	✓	✓				✓	✓	✓
	Type[4] of SMETER product	**T1b**	T3	**T4**	T2	**T1a**	**T3**	**T3**	T4

1 The CVRMSE of SMETER E (EDF) increased (from 17.4% to 19%) when the six homes with a known HTC were removed and was therefore no longer better than the CVRMSE for the expert RdSAP HTC (18.2%), but still better than the commercial RdSAP HTC (19.6%). The CVRMSE of all SMETER technology results changed (some up and some down) because of the smaller sample, but this was the only one that was so close to the RdSAP value.

2 Professional installation required as central heating controller electrically connected to boiler.

3 Professional installed deemed required as an external temperature sensor was mounted above head height on the outside of the home.

4 Type relates to increasing cost/inconvenience: T1a=only smart meter data required, T1b=only smart meter data and survey information required, T2=smart meter data and room sensors required, T3= smart meter data and room sensors and survey data required, T4= smart meter data and heating controller (with sensors) and survey required.

8.2 Generalising the performance

Generalising the performance of SMETER technologies, there was a weak relationship between the reported confidence interval and the CVRMSE of the results (Figure 12, upper graph): those SMETERs with a larger reported confidence interval did not always have a larger CVRMSE.

There was a weak relationship between the average length of the data period used and the CVRMSE of the results (Figure 12, middle graph): there was no clear trend for longer data periods to reduce CVRMSE.

There was a weak relationship between SMETER technology type and the CVRMSE of the results (Figure 12, lower graph). The two SMETER technologies which were most successful, SMETER B and SMETER H were of Type T3 and Type T4 respectively, and so required a home survey and additional installed equipment. More complex SMETER products may be more vulnerable to hardware failure and all Type T3 SMETERs (SMETER B, SMETER F, and SMETER G) experienced some problems with sensors that were not reporting as expected[40], while some households reported that they had problems using the heating controller that was part of the Type T4 SMETER H.

[40] In at least one case, the equipment was unplugged by the household.

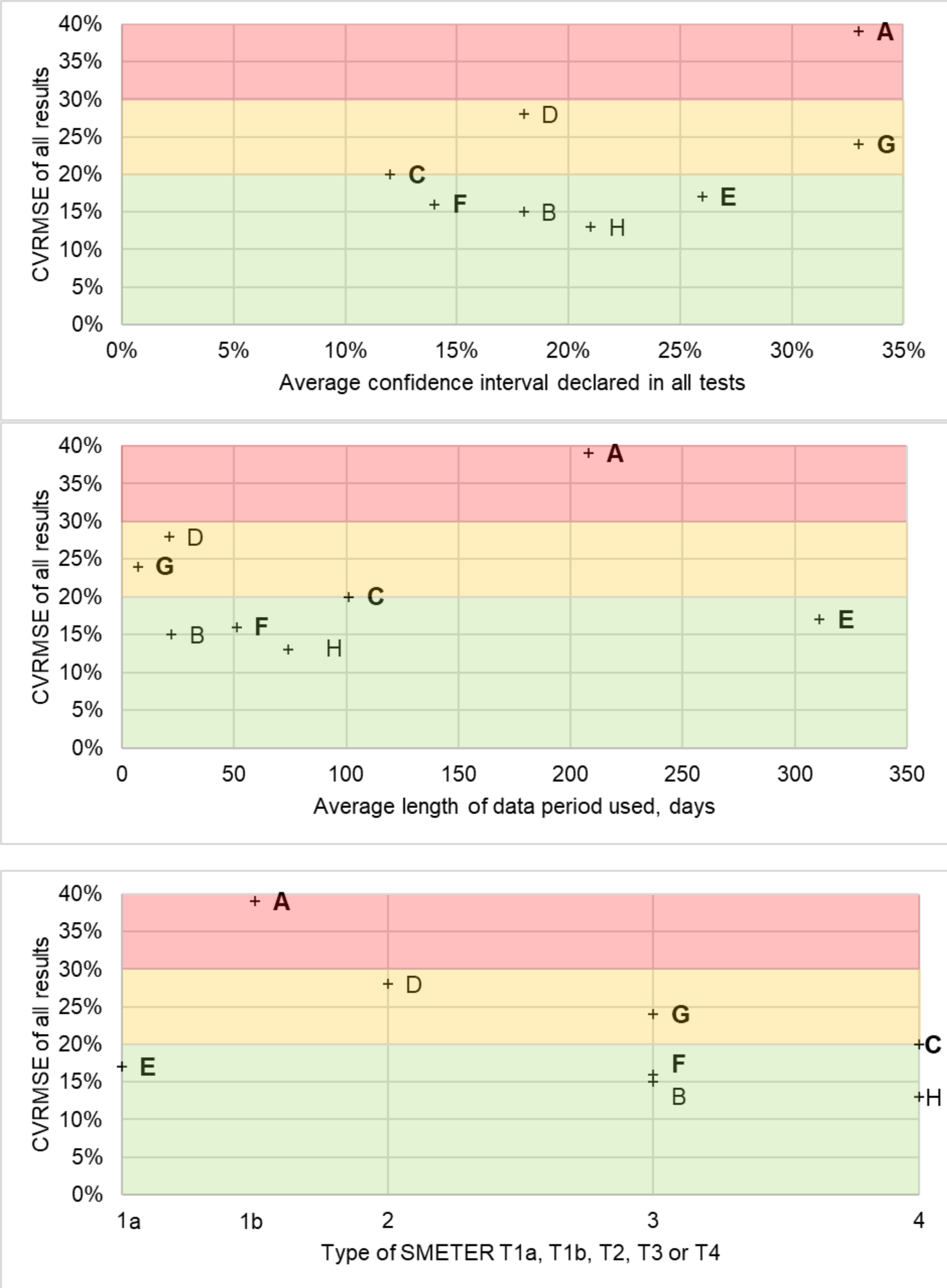


Figure 11: The relationship between SMETER accuracy (CVRMSE) and: the declared confidence interval in TEST 4; the length of data period used; and Type of SMETER (results in bold (A, C, E, F, G) are using data from the second monitoring period)

8.3 Implications for the general implementation of SMETER technologies

This work has shown that the concept of using smart meter data to calculate HTCs clearly has merit. The use of SMETERs might provide a more robust procedure, with more clearly defined error characteristics, than HTCs derived by surveyors and RdSAP. The SMETER approach might also be more discriminating than RdSAP surveys, e.g., between nominally similar homes, where one was constructed with missing sections of cavity wall insulation that cannot be seen, and one that was not. SMETERs could also overcome difficulties associated with the need for presumptions in RdSAP, e.g., about loft insulation where the loft is not accessible. SMETERs could play a role, not only in the energy rating of homes, but also in quantifying the improvement to energy efficiency following refurbishment and identifying under-performance of new homes.

The SMETER technologies with no product in the home (Type T1a and T1b) did not always perform as well as those with sensors (Type T2, T3 and T4). This suggests that the measurement of internal temperatures is likely to lead to more accurate SMETER-calculated HTCs. However, the cost, intrusiveness and reliability of SMETER products must be considered. Integrating the SMETER technology into a new heating controller may offer an unobtrusive solution; but is only possible if the household want a new heating controller as the costs are relatively high. The requirement to collect survey data from the home (Type T1b, T3 and T4) is another potential barrier for some SMETER technologies, but this is eliminated if these data can be obtained from existing EPCs. SMETER technologies that use only smart meter data and no survey data (Type T1a) offer advantages in their ease of mass deployment and low costs.

The view of households is crucial to the success of SMETER deployment. Based on the response of the 30 households in this study, almost all would have no problem having SMETER products installed in their home, and especially if the use of plug sockets is minimised and sensors are unobtrusive.

If used for rating homes and other regulatory purposes, all the SMETER technologies will need protocols that define the homes to which they can be reliably applied and those to which they cannot and give guidance on the uncertainty. Such protocols would also describe how to deal with other diverse matters such as: unmetered heat sources, e.g., wood burners; large energy using appliances that are outside of the heated envelope, e.g., hot tubs and electric vehicles; and homes with an ill-defined thermal envelope, e.g., homes with conservatories, and especially those that are unheated but have internal doors that could be left open by the household. Other situations in which SMETERs cannot be reliably used, or that require guidance on the interpretation of raw SMETER data, are likely to emerge once SMETER use becomes more widespread.

The findings from this work are ground-breaking given the small amount we know about the thermal performance of our housing stock – homes are very rarely measured. The SMETER approach opens up the prospect of a consistent and more reliable national database of domestic home energy efficiency.

There were some limitations to the field trials. The homes, while typical and with a diverse range of occupant types, were not representative of the UK housing stock. They had EPC ratings of C or D and there were no flats, new-build homes, or larger homes (the maximum floor area was 83m2). All homes had the same, or very similar, gas combi, central heating, boilers, and there was little use of secondary heating.

There were some data problems, such as occasional spikes in energy data – these were identified by the QA procedures, but any data cleaning was left to participating organisations. Similar data problems should be expected when SMETER technologies are deployed at scale. Indeed, some participating organisations used indoor air temperature data from the TEST Project due to hardware problems, so their own sensor performance is not tested.

Further validation of SMETER technologies in more homes of varied types should be seen as an important short-term priority. The co-heating test, together with blower door tests to account for ventilation heat loss, should be used as the benchmark value as this remains the most accurate method for measuring the HTC. The further validation should include more highly insulated homes, perhaps homes that comply with the current and future Building Regulations. In well-insulated homes, and potentially flats that are bounded on most sides by other heated spaces, the proportion of all energy use that is for domestic hot water is greater, and internal heat gains from the sun and occupants' activities substantially contribute to space heating. This may lead to systematic errors, as well as greater uncertainty, in the SMETER-calculated HTC. To reduce uncertainty, the un-metered heat gains, such as those from the occupants and the sun, and un-metered heat losses, such as those from hot water going down the drain, should be much less than the metered heat gains from using gas and electricity to heat the home (Li et al., 2019).

The impact of party-wall heat transfer on the accuracy of SMETER-calculated HTCs is still not fully understood and mid-terraced homes, back-to-back terraces and flats should be investigated further. Homes with a wider range of energy technologies[41] should also be investigated, such as those with solar thermal and solar PV systems, heat pump heating systems, and mechanical ventilation systems, including those with heat recovery.

The repeatability of HTC calculations from SMETER technologies should be assessed. Where a home is not physically altered, SMETER technologies should be able to provide consistent calculations of HTC. The ability of SMETER technologies to calculate changes in HTC should also be assessed. The HTC will change when a home undergoes an insulation retrofit, has new windows or doors fitted, or an extension added. It is hypothesised that relatively small changes in the HTC may be identifiable using SMETER technologies as the uncertainty in the change may be smaller than the absolute uncertainty in each calculated HTC. Thus, the

[41] It is relatively straightforward to calculate the heat input to a home from a gas boiler using smart meter data. SMETERs may require additional measurement hardware or new assumptions when some other energy technologies are present.

calculated change in energy demand might be more reliable than the change calculated by an RdSAP calculation.

9 Conclusions

Eight participating organisations (A-H) developing their own SMETER technologies took part in the Phase 2 field trial:

SMETER A	Building Research Establishment
SMETER B	Build Test Solutions
SMETER C	Cambridge Architectural Research
SMETER D	Centre for Sustainable Energy
SMETER E	EDF
SMETER F	Hoare Lea
SMETER G	Passiv UK
SMETER H	Switchee

A ninth SMETER (SMETER I) joined the project too late for their product to be installed in the field trial homes and was evaluated separately (see Appendix I).

The accuracy of each SMETER product was evaluated by comparison with the measured HTC in two ways. Firstly, the SMETER result (calculated HTC, including the 95% confidence interval) was compared directly to the measured HTC. Where the confidence intervals of these two results overlapped, this was deemed to be a successful SMETER result. Secondly, the difference between each SMETER result (central estimate, ignoring confidence intervals) and the corresponding measured HTC was analysed.

All participating organisations were able to report confidence intervals for every calculated HTC. Overall, the SMETER technologies were successful for between 70% and 97% of the homes, with average confidence intervals between 12% and 33%. Five participating organisations provided calculated HTC results that were more than 90% successful overall: SMETER B (28 out of 30 homes, with an average confidence interval of +/-18%), SMETER E (26 out of 27 homes, with an average confidence interval of +/-26%), SMETER F (26 out of 27 homes, with an average confidence interval of +/-14%), SMETER G (25 out of 27 homes, with an average confidence interval of +/-33%), and SMETER H (29 out of 30 homes, with an average confidence interval of +/-21%).

For each SMETER technology, the difference between each calculated HTC and the associated measured HTC was determined. The normalised mean bias error (NMBE) quantifies the magnitude and direction of the average bias in the calculated HTC. This is a measure of the trueness, or systematic agreement, of the measurement and would ideally be zero. The coefficient of variation of root mean square error (CVRMSE) is a comparative measure of the precision of the calculated HTC. A lower CVRMSE is better. The NMBE ranged from -0.7% (best) to -27% (worst) and the CVRMSE from 13.4% to 38.9%.

For all 30 homes, the difference between the HTC predicted from the expert RdSAP was compared with the measured HTCs, yielding an NMBE of -2.8% and CVRMSE of 18.2%. The two best-performing SMETER technologies (the lowest CVRMSE and NMBE closest to zero) were more accurate than the expert RdSAP: SMETER B (NMBE -2.1%, CVRMSE 15.2%) and SMETER H (NMBE -0.7%, CVRMSE 13.4%). However, compared with the commercial RdSAP HTC results for 22 of the homes (NMBE -1.1%, CVRMSE 19.6%) only SMETER H (NMBE -0.7%, CVRMSE 13.4%) was more accurate. Two SMETER technologies had a lower CVRMSE than either of the RdSAP HTC predictions, but their NMBE was not as good: SMETER E (NMBE 4.1%, CVRMSE 17.4%) and SMETER F (NMBE -7.7%, CVRMSE 15.9%).

The suitability of a SMETER technology for a particular application will depend on various factors, including accuracy (success rate, reported confidence interval, NMBE and CVRMSE), duration (average length of data period required for calculation), and cost/convenience (number of sensors in the home, whether a professional installation or survey is required). The survey requirement may be less expensive and less inconvenient if the information can be taken from an existing EPC survey. The SMETER technologies have been compared based on these characteristics but SMETERs which use longer periods of data may be well-suited to many applications and SMETERs that are integrated in heating controllers offer little additional cost/inconvenience to a household choosing that controller. The required accuracy will depend on the application, and it may be possible to calibrate the SMETER technologies to improve the accuracy.

Households were interviewed during the field trial to understand their views on the SMETER products:

- Initially, 97% of households reported that they did not notice the SMETER products in their home.
- 93% of households said they would be happy to have a SMETER product in their home forever and the remaining 7% would be happy for a SMETER product to be installed for 6 months.
- 7% of households found the SMETER product's use of a plug socket to be inconvenient.
- 13% of the household reported that they did not like the flashing light on some of the sensors that were installed for the TEST Project monitoring.

The two SMETER technologies which were most successful, SMETER B and SMETER H, were of Type T3 and Type T4 respectively, and so required a home survey and additional installed equipment. More complex SMETER products such as these may be more vulnerable to hardware failure and all Type 3 SMETERs (SMETER B, SMETER F, and SMETER G) experienced some problems with sensors that were not reporting as expected, while some households reported that they had problems using the heating controller that was part of the Type T4 SMETER H.

There were some limitations to the field trial. The homes, while typical and with a diverse range of occupant types, were not representative of the UK housing stock. They had EPC ratings of C or D and there were no flats, new-build homes, or larger homes. All homes had the same, or very similar, gas combi, central heating, boilers. Therefore, it is not yet possible to comment on the reliability of SMETERs for more energy efficient homes (Rated A and B) or with more complex energy technologies (e.g., heat pumps). It is expected that the range of application, and the accuracy of SMETER algorithms will improve with experience and the collection of new data. Sharing the measured HTCs, dwelling characteristics, and ancillary measurement for more homes from this project, will stimulate further innovation.

While further field work is required to extend our understanding of SMETER technologies, this does not preclude their immediate use for homes of the type monitored here. In fact, some participating organisations are already offering this service and potentially we stand to learn much about the thermal performance of our housing stock this way

10 References

Bland, J.M., Altman, D.G., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 327, 307–310. https://doi.org/10.1016/S0140-6736(86)90837-8

BRE, 2020. Building Energy Performance Assessment - Support Website [WWW Document]. URL https://www.ncm-pcdb.org.uk/sap/ (accessed 9.10.20).

Crawley, J., Biddulph, P., Northrop, P.J., Wingfield, J., Oreszczyn, T., Elwell, C., 2019. Quantifying the Measurement Error on England and Wales EPC Ratings. Energies 12, 3523. https://doi.org/https://doi.org/10.3390/en12183523

Li, M., Allinson, D., Lomas, K., 2019. Estimation of building heat transfer coefficients from in-use data: Impacts of unmonitored energy flows. Int. J. Build. Pathol. Adapt. 38. https://doi.org/10.1108/IJBPA-02-2019-0022

ONS, 2021. Families and households in the UK: 2020 [WWW Document]. Off. Natl. Stat. URL https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/families/bulletins/familiesandhouseholds/2020

Appendix A: Monitored temperature data fulfilment

TEST Project monitored temperature data in the living room and kitchen (first monitoring period)

Home ID	Living room				Kitchen				Notes
	Start date	End date	Total days	Fulfilment (%)	Start date	End date	Total days	Fulfilment (%)	
HH01	03/03/20	06/08/20	156	92.1	03/03/20	06/08/20	156	91.3	
HH02	21/02/20	06/08/20	167	95.3	21/02/20	06/08/20	167	95.7	
HH03	08/03/20	06/08/20	151	94.9	08/03/20	06/08/20	151	90.3	
HH04	26/11/19	06/08/20	254	96.0	26/11/19	14/06/20	201	93.3	
HH05	26/11/19	17/02/20	83	71.7	26/11/19	05/02/20	71	88.6	
HH06	26/11/19	31/03/20	126	85.6	26/11/19	13/04/20	139	81.3	
HH07	26/11/19	06/07/20	223	92.7	26/11/19	06/08/20	254	96.1	
HH08	26/11/19	06/08/20	254	95.2	26/11/19	29/02/20	95	72.5	
HH09	26/11/19	06/08/20	254	91.1	26/11/19	06/08/20	254	96.3	
HH10	26/11/19	06/08/20	254	96.2	26/11/19	06/08/20	254	92.8	
HH11	26/11/19	06/08/20	254	84.3	26/11/19	06/08/20	254	96.4	
HH12	26/11/19	22/03/20	117	95.4	26/11/19	06/08/20	254	92.4	
HH13	26/11/19	06/08/20	254	92.4	26/11/19	06/08/20	254	95.9	
HH14	26/11/19	08/03/20	103	96.8	26/11/19	13/05/20	169	94.7	
HH15	26/11/19	28/03/20	123	93.5	26/11/19	30/03/20	125	93.0	
HH16	12/12/19	06/08/20	238	96.2	12/12/19	06/08/20	238	90.5	
HH17	12/12/19	06/08/20	238	96.4	12/12/19	06/03/20	85	92.6	
HH18	12/12/19	06/08/20	238	95.9	12/12/19	06/08/20	238	91.7	
HH19	12/12/19	06/08/20	238	76.5	12/12/19	04/08/20	236	89.9	
HH20	12/12/19	06/08/20	238	91.3	12/12/19	06/08/20	238	96.1	
HH21	18/12/19	06/08/20	232	95.3	18/12/19	29/04/20	133	75.9	
HH22	12/12/19	22/07/20	223	96.2	12/12/19	06/08/20	238	96.2	
HH23	16/01/20	06/08/20	203	91.0	16/01/20	06/08/20	203	92.1	
HH24	31/01/20	06/08/20	188	91.4	31/01/20	06/08/20	188	91.1	
HH25	31/01/20	06/08/20	188	95.0	31/01/20	06/08/20	188	82.3	
HH26	02/03/20	06/08/20	157	95.8	31/01/20	06/08/20	188	91.1	
HH27	31/01/20	06/08/20	188	95.9	02/03/20	06/08/20	157	95.1	
HH28	31/01/20	06/08/20	188	85.8	31/01/20	06/08/20	188	95.1	
HH29	28/02/20	06/08/20	160	95.3	28/02/20	06/08/20	160	94.3	
HH30	04/03/20	06/08/20	155	93.3	04/03/20	06/08/20	155	88.7	

Note: “Start date” is the first day of data that was provided to participating organisations. Data were released from dates after the co-heating test was completed and the home was assumed occupied.

TEST Project monitored data in the hall and bathroom (first monitoring period)

Home ID	Hall				Bathroom				Notes
	Start date	End date	Total days	Fulfilment (%)	Start date	End date	Total days	Fulfilment (%)	
HH01	03/03/20	06/08/20	156	91.3	NA	NA	0	N/A	
HH02	21/02/20	23/05/20	92	93.0	21/02/20	06/08/20	167	94.7	
HH03	08/03/20	06/08/20	151	80.2	08/03/20	06/08/20	151	96.1	
HH04	26/11/19	06/08/20	254	62.4	28/11/19	16/03/20	109	23.2	
HH05	26/11/19	22/12/19	26	71.5	26/11/19	28/01/20	63	42.6	
HH06	26/11/19	06/08/20	254	96.0	26/11/19	06/08/20	254	90.8	
HH07	26/11/19	06/08/20	254	96.1	26/11/19	06/08/20	254	92.1	
HH08	26/11/19	25/06/20	212	87.2	26/11/19	10/07/20	227	94.9	
HH09	NA	NA	0	NA	26/11/19	06/08/20	254	92.7	
HH10	26/11/19	06/08/20	254	91.2	26/11/19	06/08/20	254	96.5	
HH11	02/03/20	06/08/20	157	95.9	26/11/19	06/08/20	254	92.8	
HH12	26/11/19	07/12/19	11	93.7	26/11/19	06/08/20	254	92.1	
HH13	NA	NA	0	NA	26/11/19	06/08/20	254	92.6	
HH14	26/11/19	10/05/20	166	90.9	26/11/19	06/08/20	254	92.1	
HH15	26/11/19	06/08/20	254	63.0	26/11/19	06/08/20	254	93.3	
HH16	12/12/19	06/08/20	238	69.0	12/12/19	02/03/20	81	92.7	
HH17	12/12/19	06/08/20	238	95.3	12/12/19	06/08/20	238	92.4	
HH18	12/12/19	06/08/20	238	96.1	12/12/19	06/08/20	238	92.5	
HH19	12/12/19	06/08/20	238	97.2	12/12/19	06/08/20	238	64.3	
HH20	12/12/19	06/08/20	238	88.3	12/12/19	06/08/20	238	96.6	
HH21	NA	NA	0	NA	18/12/19	16/07/20	211	93.6	
HH22	12/12/19	06/08/20	238	96.0	12/12/19	06/08/20	238	91.9	
HH23	16/01/20	06/08/20	203	91.5	16/01/20	06/08/20	203	95.7	
HH24	31/01/20	06/08/20	188	94.7	31/01/20	06/08/20	188	94.7	
HH25	31/01/20	06/08/20	188	91.0	31/01/20	06/08/20	188	83.1	
HH26	31/01/20	06/08/20	188	95.6	31/01/20	06/08/20	188	92.4	
HH27	31/01/20	06/08/20	188	92.0	31/01/20	06/08/20	188	94.2	
HH28	31/01/20	06/08/20	188	95.3	31/01/20	06/08/20	188	94.9	
HH29	09/03/20	06/08/20	150	96.0	28/02/20	06/08/20	160	95.0	
HH30	04/03/20	14/05/20	71	75.1	04/03/20	06/08/20	155	96.0	

Note: “Start date” is the first day of data that was provided to participating organisations. Data were released from dates after the co-heating test was completed and the home was assumed occupied.

TEST Project monitored temperature data in bedroom 1 and bedroom 2 (first monitoring period)

Home ID	Bedroom 1				Bedroom 2				Notes
	Start date	End date	Total days	Fulfilment (%)	Start date	End date	Total days	Fulfilment (%)	
HH01	03/03/20	06/08/20	156	91.6	03/03/20	06/08/20	156	91.8	
HH02	21/02/20	06/08/20	167	92.0	NA	NA	NA	NA	
HH03	08/03/20	06/08/20	151	95.5	08/03/20	06/08/20	151	96.4	
HH04	26/11/19	06/08/20	254	96.2	NA	NA	NA	NA	
HH05	26/11/19	17/02/20	83	78.2	26/11/19	11/12/20	381	94.2	
HH06	26/11/19	23/05/20	179	94.4	26/11/19	11/06/20	198	90.2	
HH07	26/11/19	06/08/20	254	95.9	NA	NA	NA	NA	
HH08	26/11/19	27/06/20	214	92.1	26/11/19	06/08/20	254	93.6	
HH09	26/11/19	06/08/20	254	96.7	NA	NA	NA	NA	
HH10	26/11/19	06/08/20	254	95.6	26/11/19	06/08/20	254	46.8	
HH11	26/11/19	06/08/20	254	96.0	26/11/19	29/06/20	216	91.8	
HH12	26/11/19	06/08/20	254	92.2	NA	NA	NA	NA	
HH13	26/11/19	06/08/20	254	91.8	NA	NA	NA	NA	
HH14	26/11/19	28/04/20	154	95.8	26/11/19	30/03/20	125	90.6	
HH15	26/11/19	06/08/20	254	91.2	26/11/19	06/08/20	254	95.7	
HH16	12/12/19	06/08/20	238	93.8	12/12/19	06/08/20	238	93.2	
HH17	12/12/19	06/08/20	238	92.7	12/12/19	06/08/20	238	95.2	
HH18	12/12/19	06/08/20	238	92.3	NA	NA	NA	NA	
HH19	12/12/19	06/08/20	238	79.4	12/12/19	23/07/20	224	94.4	
HH20	12/12/19	27/01/20	46	99.2	12/12/19	06/08/20	238	90.8	
HH21	18/12/19	08/07/20	203	95.7	18/12/19	06/08/20	232	95.9	
HH22	12/12/19	06/08/20	238	96.0	12/12/19	06/08/20	238	89.8	
HH23	16/01/20	06/08/20	203	95.9	16/01/20	06/08/20	203	95.7	
HH24	31/01/20	06/08/20	188	92.2	NA	NA	NA	NA	
HH25	31/01/20	06/08/20	188	95.5	31/01/20	06/08/20	188	84.2	
HH26	31/01/20	06/08/20	188	90.1	02/03/20	06/08/20	157	95.9	
HH27	31/01/20	06/08/20	188	91.6	31/01/20	06/08/20	188	89.6	
HH28	31/01/20	06/08/20	188	92.2	31/01/20	06/08/20	188	94.8	
HH29	28/02/20	06/08/20	160	95.4	28/02/20	06/08/20	160	94.1	
HH30	04/03/20	06/08/20	155	93.4	04/03/20	15/07/20	133	93.4	

Note: “Start date” is the first day of data that was provided to participating organisations. Data were released from dates after the co-heating test was completed and the home was assumed occupied.

TEST Project monitored temperature data in bedroom 3 and other rooms (first monitoring period)

Home ID	Bedroom 3				Other room (see notes)				Notes
	Start date	**End date**	**Total days**	**Fulfilment (%)**	**Start date**	**End date**	**Total days**	**Fulfilment (%)**	
HH01	NA	NA	NA	NA	NA	NA	NA	NA	
HH02	NA	NA	NA	NA	NA	NA	NA	NA	
HH03	NA	NA	NA	NA	NA	NA	NA	NA	
HH04	NA	NA	NA	NA	NA	NA	NA	NA	
HH05	26/11/19	17/02/20	83	96.5	NA	NA	NA	NA	
HH06	26/11/19	06/08/20	254	95.7	26/11/19	06/08/20	254	95.0	Other room = “smallliv”
HH07	NA	NA	NA	NA	NA	NA	NA	NA	
HH08	26/11/19	21/07/20	238	91.9	26/11/19	21/05/20	177	85.1	Other room = “util”
HH09	NA	NA	NA	NA	NA	NA	NA	NA	
HH10	NA	NA	NA	NA	NA	NA	NA	NA	
HH11	26/11/19	06/08/20	254	91.8	NA	NA	NA	NA	
HH12	NA	NA	NA	NA	NA	NA	NA	NA	
HH13	NA	NA	NA	NA	NA	NA	NA	NA	
HH14	NA	NA	NA	NA	NA	NA	NA	NA	
HH15	26/11/19	06/08/20	254	95.2	NA	NA	NA	NA	
HH16	NA	NA	NA	NA	NA	NA	NA	NA	
HH17	NA	NA	NA	NA	NA	NA	NA	NA	
HH18	NA	NA	NA	NA	NA	NA	NA	NA	
HH19	NA	NA	NA	NA	NA	NA	NA	NA	
HH20	12/12/19	06/08/20	238	97.2	NA	NA	NA	NA	
HH21	NA	NA	NA	NA	18/12/19	27/03/20	100	80.1	Other room = “backroom”
HH22	NA	NA	NA	NA	NA	NA	NA	NA	
HH23	16/01/20	10/07/20	176	94.6	16/01/20	06/08/20	203	91.6	Other room = “toilet”
HH24	NA	NA	NA	NA	NA	NA	NA	NA	
HH25	NA	NA	NA	NA	NA	NA	NA	NA	
HH26	02/03/20	06/08/20	157	96.0	NA	NA	NA	NA	
HH27	31/01/20	06/08/20	188	91.8	NA	NA	NA	NA	
HH28	NA	NA	NA	NA	NA	NA	NA	NA	
HH29	28/02/20	06/08/20	160	91.6	NA	NA	NA	NA	
HH30	04/03/20	17/05/20	74	64.9	04/03/20	22/05/20	79	69.8	Other room = “dining”

Note: “Start date” is the first day of data that was provided to participating organisations. Data were released from dates after the co-heating test was completed and the home was assumed occupied.

TEST Project monitored temperature data in the living room and kitchen (second monitoring period)

Home ID	Living room				Kitchen				Notes
	Start date	End date	Total days	Fulfilment (%)	Start date	End date	Total days	Fulfilment (%)	
HH01	01/08/20	01/03/21	212	82.6	01/08/20	01/03/21	212	81.6	
HH02	01/08/20	01/03/21	212	96.9	01/08/20	01/03/21	212	80.2	
HH03	01/08/20	01/03/21	212	73.4	01/08/20	01/03/21	212	77.8	
HH04	01/08/20	01/03/21	212	90.4	05/10/20	01/03/21	147	95.1	
HH05	15/10/20	18/02/21	126	100	15/10/20	18/02/21	126	100	
HH06	05/10/20	01/03/21	147	93.6	05/10/20	01/03/21	147	88.5	
HH08	01/08/20	01/03/21	212	70.7	09/11/20	01/03/21	69	94.9	
HH09	01/08/20	01/03/21	212	96.2	01/08/20	01/03/21	212	83.8	
HH10	01/08/20	01/03/21	212	96.7	01/08/20	01/03/21	212	92.7	
HH11	01/08/20	01/03/21	212	86.3	01/08/20	01/03/21	212	85.2	
HH12	06/10/20	01/03/21	146	80.0	01/08/20	01/03/21	212	82.0	
HH13	01/08/20	01/03/21	212	82.6	01/08/20	01/03/21	212	96.2	
HH14	02/11/20	01/03/21	119	90.4	02/11/20	01/03/21	119	90.7	
HH15	06/10/20	17/01/21	103	88.6	06/10/20	08/01/21	94	58.6	
HH16	01/08/20	01/03/21	212	81.1	01/08/20	01/03/21	212	67.7	
HH17	01/08/20	12/01/21	164	71.5	11/10/20	06/01/21	87	75.5	
HH19	01/08/20	29/01/21	181	87.7	01/08/20	01/03/21	212	53.3	
HH20	01/08/20	01/03/21	121	74.9	01/08/20	01/03/21	212	82.0	
HH21	01/08/20	18/02/21	201	86.3	11/11/20	18/02/21	99	91.8	
HH22	05/10/20	01/03/21	147	89.2	01/08/20	01/03/21	212	75.2	
HH23	01/08/20	01/03/21	212	92.7	01/08/20	01/03/21	212	81.7	
HH24	01/08/20	01/03/21	212	74.0	01/08/20	01/03/21	212	92.7	
HH25	01/08/20	01/03/21	212	76.7	01/08/20	01/03/21	212	89.9	
HH26	01/08/20	01/03/21	212	96.8	01/08/20	01/03/21	212	83.6	
HH27	01/08/20	01/03/21	212	96.3	01/08/20	01/03/21	212	85.2	
HH28	01/08/20	01/03/21	212	84.2	01/08/20	01/03/21	212	84.0	
HH29	01/08/20	01/03/21	212	93.7	01/08/20	01/03/21	212	82.4	

Note: “Start date” is the first day of data that was provided to participating organisations.

TEST Project monitored temperature data in the hall and bathroom (first monitoring period)

Home ID	Hall				Bathroom				Notes
	Start date	End date	Total days	Fulfilment (%)	Start date	End date	Total days	Fulfilment (%)	
HH01	01/08/20	01/03/21	212	83.9	04/11/20	01/03/21	117	95.2	
HH02	03/11/20	14/02/21	103	78.2	01/08/20	01/03/21	212	90.9	
HH03	01/08/20	01/03/21	212	79.5	01/08/20	01/03/21	212	70.1	
HH04	01/08/20	01/03/21	212	82.4	07/11/20	01/03/21	114	93.8	
HH05	15/10/20	18/02/21	126	100	15/10/20	18/02/21	126	100	
HH06	01/08/20	01/03/21	212	76.4	01/08/20	01/03/21	212	95.9	
HH08	09/10/20	01/03/21	143	94.5	09/10/20	01/03/21	143	95.9	
HH09	02/11/20	01/03/21	119	95.4	01/08/20	01/03/21	212	97.0	
HH10	01/08/20	01/03/21	212	81.0	01/08/20	01/03/21	212	83.3	
HH11	01/08/20	01/03/21	212	88.7	01/08/20	01/03/21	212	92.3	
HH12	04/11/20	01/03/21	117	95.3	01/08/20	01/03/21	212	81.9	
HH13	01/08/20	01/03/21	212	74.2	01/08/20	01/03/21	212	96.9	
HH14	02/11/20	01/03/21	119	73.6	01/08/20	16/12/20	137	79.9	
HH15	01/08/20	07/01/21	159	62.2	01/08/20	01/03/21	212	71.3	
HH16	01/08/20	01/03/21	212	62.3	09/11/20	01/03/21	143	94.6	
HH17	01/08/20	14/01/21	166	94.9	01/08/20	14/01/21	166	73.0	
HH19	01/08/20	01/03/21	212	96.8	01/08/20	01/03/21	212	96.8	
HH20	01/08/20	01/03/21	212	77.8	01/08/20	01/03/21	212	75.3	
HH21	NA	NA	NA	NA	09/11/20	18/02/21	101	91.7	
HH22	01/08/20	01/03/21	212	96.9	01/08/20	01/03/21	212	92.5	
HH23	01/08/20	01/03/21	212	74.4	01/08/20	01/03/21	212	77.1	
HH24	01/08/20	01/03/21	212	95.5	01/08/20	01/03/21	212	96.4	
HH25	01/08/20	01/03/21	212	79.6	01/08/20	01/03/21	212	85.1	
HH26	01/08/20	01/03/21	212	85.6	01/08/20	01/03/21	212	85.6	
HH27	01/08/20	01/03/21	212	92.5	01/08/20	01/03/21	212	81.5	
HH28	01/08/20	01/03/21	212	84.5	01/08/20	01/03/21	212	84.4	
HH29	01/08/20	01/03/21	212	88.7	01/08/20	01/03/21	212	93.1	

Note: “Start date” is the first day of data that was provided to participating organisations.

TEST Project monitored temperature data in bedroom 1 and bedroom 2 (second monitoring period)

Home ID	Bedroom 1				Bedroom 2				Notes
	Start date	End date	Total days	Fulfilment (%)	Start date	End date	Total days	Fulfilment (%)	
HH01	01/08/20	01/03/21	212	82.2	01/08/20	01/03/21	212	75.4	
HH02	01/08/20	14/02/21	197	72.1	NA	NA	NA	NA	
HH03	01/08/20	01/03/21	212	79.9	01/08/20	01/03/21	212	73.8	
HH04	01/08/20	01/03/21	212	64.6	NA	NA	NA	NA	
HH05	15/10/20	18/02/21	126	100	15/10/20	10/12/20	56	100	
HH06	05/10/20	01/03/21	147	82.5	01/08/20	01/03/21	212	95.1	
HH08	09/10/20	01/03/21	133	94.9	09/10/20	01/03/21	143	76.0	
HH09	01/08/20	01/03/21	212	96.9	NA	NA	NA	NA	
HH10	01/08/20	01/03/21	212	56.4	01/08/20	18/01/21	170	96.6	
HH11	01/08/20	01/03/21	212	96.9	22/10/20	01/03/21	130	90.6	
HH12	01/08/20	01/03/21	212	86.9	NA	NA	NA	NA	
HH13	01/08/20	01/03/21	212	75.0	NA	NA	NA	NA	
HH14	02/11/20	01/03/21	119	90.9	02/11/20	01/03/21	119	94.6	
HH15	01/08/20	18/01/21	170	92.8	01/08/20	30/01/21	182	93.1	
HH16	01/08/20	01/03/21	212	80.8	01/08/20	21/02/21	204	72.3	
HH17	01/08/20	07/01/21	159	93.1	01/08/20	05/02/21	188	59.4	
HH19	01/08/20	01/03/21	212	69.9	15/10/20	01/03/21	137	90.6	
HH20	09/10/20	01/03/21	143	91.4	01/08/20	08/02/21	191	79.2	
HH21	11/11/20	18/02/21	99	91.7	01/08/20	18/02/21	201	86.7	
HH22	01/08/20	01/03/21	212	96.8	01/08/20	01/03/21	212	91.4	
HH23	01/08/20	01/03/21	212	88.6	02/11/20	01/03/21	119	94.7	
HH24	01/08/20	01/03/21	212	83.4	NA	NA	NA	NA	
HH25	01/08/20	01/03/21	212	97.1	01/08/20	01/03/21	212	81.0	
HH26	01/08/20	01/03/21	212	83.8	01/08/20	01/03/21	212	96.9	
HH27	01/08/20	01/03/21	212	91.8	01/08/20	01/03/21	212	93.1	
HH28	01/08/20	01/03/21	212	72.8	01/08/20	01/03/21	212	96.7	
HH29	01/08/20	01/03/21	212	88.7	01/08/20	01/03/21	212	86.0	

Note: “Start date” is the first day of data that was provided to participating organisations.

TEST Project monitored temperature data in bedroom 3 and other rooms (second monitoring period)

Home ID	Bedroom 3				Other room (see notes)				Notes
	Start date	End date	Total days	Fulfilment (%)	Start date	End date	Total days	Fulfilment (%)	
HH01	NA	NA	NA	NA	NA	NA	NA	NA	
HH02	NA	NA	NA	NA	NA	NA	NA	NA	
HH03	NA	NA	NA	NA	NA	NA	NA	NA	
HH04	NA	NA	NA	NA	NA	NA	NA	NA	
HH05	15/10/20	18/02/21	126	100	NA	NA	NA	NA	
HH06	05/10/20	01/03/21	147	95.0	01/08/20	01/03/21	212	77.1	Other room = “smallliv”
HH08	09/10/20	01/03/21	143	94.9	09/10/20	01/03/21	143	94.2	Other room = “util”
HH09	NA	NA	NA	NA	NA	NA	NA	NA	
HH10	NA	NA	NA	NA	NA	NA	NA	NA	
HH11	01/08/20	01/03/21	212	62.9	NA	NA	NA	NA	
HH12	NA	NA	NA	NA	NA	NA	NA	NA	
HH13	NA	NA	NA	NA	NA	NA	NA	NA	
HH14	NA	NA	NA	NA	NA	NA	NA	NA	
HH15	01/08/20	07/01/21	159	74.5	NA	NA	NA	NA	
HH16	NA	NA	NA	NA	NA	NA	NA	NA	
HH17	NA	NA	NA	NA	NA	NA	NA	NA	
HH18	NA	NA	NA	NA	NA	NA	NA	NA	
HH19	NA	NA	NA	NA	NA	NA	NA	NA	
HH20	01/08/20	01/03/21	212	62.8	NA	NA	NA	NA	
HH21	25/11/20	18/02/21	85	100	09/11/20	18/02/21	101	97.7	Other room = “backroom”
HH22	NA	NA	NA	NA	NA	NA	NA	NA	
HH23	NA	NA	NA	NA	01/08/20	01/03/21	212	85.0	Other room = “toilet”
HH24	NA	NA	NA	NA	NA	NA	NA	NA	
HH25	NA	NA	NA	NA	NA	NA	NA	NA	
HH26	01/08/20	01/03/21	212	96.9	NA	NA	NA	NA	
HH27	01/08/20	01/03/21	212	81.3	NA	NA	NA	NA	
HH28	NA	NA	NA	NA	NA	NA	NA	NA	
HH29	01/08/20	01/03/21	212	89.8	06/10/20	05/02/21	122	81.9	Other room = “dining”

Note: “Start date” is the first day of data that was provided to participating organisations.

Appendix B: Data periods declared by participating organisations in TEST 4 and TEST 6 (first monitoring period) and then after the second monitoring period

SMETER (A-D) self-reported number of days and period of data used to calculate the HTC (first monitoring period).

	SMETER A				SMETER B				SMETER C				SMETER D			
Home ID	Days	from	to	period	Days	from	to	period	Days	from	to	period	Days	from	to	period
HH01	29	3/3/20	31/3/20	29	21	6/3/20	26/3/20	21	39	9/3/20	26/4/20	49	21	16/3/20	5/4/20	21
HH02	30	1/4/20	30/4/20	30	21	6/3/20	26/3/20	21	36	6/3/20	25/4/20	51	21	16/3/20	5/4/20	21
HH03	30	1/4/20	30/4/20	30	21	8/3/20	28/3/20	21	31	14/3/20	28/4/20	46	21	10/3/20	30/3/20	21
HH04	30	1/4/20	30/4/20	30	21	26/11/19	16/12/19	21	116	2/12/19	28/4/20	149	21	10/3/20	30/3/20	21
HH05	30	1/4/20	30/4/20	30	21	26/11/19	16/12/19	21	29	13/12/19	2/2/20	52	21	14/1/20	3/2/20	21
HH06	31	1/5/20	31/5/20	31	21	2/2/20	22/2/20	21	53	1/2/20	26/4/20	86	21	10/2/20	1/3/20	21
HH07	31	1/5/20	31/5/20	31	21	2/2/20	22/2/20	21	65	17/1/20	18/4/20	93	21	10/2/20	1/3/20	21
HH08	31	1/1/20	31/1/20	31	21	26/11/19	16/12/19	21	27	26/2/20	23/3/20	27	21	10/2/20	1/3/20	21
HH09	31	1/1/20	31/1/20	31	21	27/11/19	17/12/19	21	32	25/2/20	28/4/20	64	21	10/2/20	1/3/20	21
HH10	30	1/4/20	30/4/20	30	21	20/12/19	9/1/20	21	53	26/12/19	13/3/20	79	21	10/2/20	1/3/20	21
HH11	30	1/4/20	30/4/20	30	21	10/2/20	1/3/20	21	55	8/1/20	26/4/20	110	21	2/3/20	22/3/20	21
HH12	30	1/4/20	30/4/20	30	21	20/12/19	9/1/20	21	73	27/12/19	21/3/20	86	21	10/2/20	1/3/20	21
HH13	30	1/4/20	30/4/20	30	21	27/12/19	16/1/20	21	109	2/12/19	18/4/20	139	21	10/2/20	1/3/20	21
HH14	30	1/4/20	30/4/20	30	12	29/11/19	15/2/20	79	10	17/12/19	26/12/19	10	21	10/2/20	1/3/20	21
HH15	30	1/4/20	30/4/20	30	21	1/12/19	21/12/19	21	97	2/12/19	27/3/20	117	21	10/2/20	1/3/20	21
HH16	30	1/1/20	31/1/20	31	21	7/1/20	27/1/20	21	47	26/2/20	28/4/20	63	21	10/2/20	1/3/20	21
HH17	20	12/12/19	31/12/19	20	21	7/1/20	27/1/20	21	41	26/2/20	28/4/20	63	21	10/2/20	1/3/20	21
HH18	30	1/4/20	30/4/20	30	21	22/12/19	11/1/20	21	98	18/12/19	26/4/20	131	21	10/2/20	1/3/20	21
HH19	30	1/4/20	30/4/20	30	21	22/12/19	11/1/20	21	65	21/12/19	27/4/20	129	21	10/2/20	1/3/20	21
HH20	20	12/12/19	31/12/19	20	21	10/1/20	30/1/20	21	32	26/2/20	20/4/20	55	21	1/3/20	21/3/20	21
HH21	31	1/1/20	31/1/20	31	21	23/12/19	12/1/20	21	50	26/2/20	28/4/20	63	21	1/3/20	21/3/20	21
HH22	27	1/2/20	28/2/20	28	21	3/3/20	23/3/20	21	47	25/2/20	28/4/20	64	21	16/3/20	5/4/20	21
HH23	30	1/4/20	30/4/20	30	21	28/1/20	17/2/20	21	56	3/2/20	26/4/20	84	21	10/2/20	1/3/20	21
HH24	30	1/4/20	30/4/20	30	21	30/1/20	20/2/20	22	48	11/2/20	28/4/20	78	21	10/2/20	1/3/20	21
HH25	30	1/4/20	30/4/20	30	21	31/1/20	20/2/20	21	61	31/1/20	28/4/20	89	21	10/2/20	1/3/20	21
HH26	31	1/5/20	31/5/20	31	21	31/1/20	20/2/20	21	46	10/2/20	28/4/20	79	21	10/2/20	1/3/20	21
HH27	30	1/4/20	30/4/20	30	13	21/2/20	4/3/20	13	47	8/2/20	26/4/20	79	21	2/3/20	22/3/20	21
HH28	30	1/4/20	30/4/20	30	14	21/2/20	5/3/20	14	41	14/2/20	28/4/20	75	21	10/2/20	1/3/20	21
HH29	31	1/3/20	31/3/20	31	21	28/2/20	19/3/20	21	40	7/3/20	28/4/20	53	21	1/3/20	21/3/20	21
HH30	28	4/3/20	31/3/20	28	21	5/3/20	25/3/20	21	40	7/3/20	28/4/20	53	21	1/4/20	21/4/20	21
Average	**29**			**29**	**20**			**22**	**53**			**77**	**21**			**21**

SMETER (E-H) self-reported number of days and period of data used to calculate the HTC (first monitoring period).

	SMETER E				SMETER F				SMETER G				SMETER H			
Home ID	Days	from	to	period	Days	from	to	period	Days	from	to	period	Days	from	to	period
HH01	5	3/3/20	6/8/20	157	42	8/3/20	22/4/20	46					20	5/3/20	27/3/20	23
HH02	7	21/2/20	6/8/20	168	49	8/3/20	30/4/20	54					18	6/3/20	27/3/20	22
HH03	5	8/3/20	6/8/20	152	16	3/3/20	1/4/20	30					16	8/3/20	27/3/20	20
HH04	59	26/11/19	6/8/20	255	116	27/11/19	30/3/20	125	7	5/1/20	12/1/20	7	96	26/11/19	6/3/20	102
HH05	61	27/11/19	6/8/20	254	117	27/11/19	30/3/20	125	7	9/2/20	16/2/20	7	122	26/11/19	27/3/20	123
HH06	13	31/1/20	6/8/20	189	47	1/2/20	30/3/20	59					53	1/2/20	27/3/20	56
HH07	21	16/1/20	6/8/20	204	64	17/1/20	30/3/20	74	7	19/1/20	26/1/20	7	68	16/1/20	27/3/20	72
HH08	63	26/11/19	6/8/20	255	66	27/11/19	10/2/20	76	7	29/12/19	5/1/20	7	113	27/11/19	27/3/20	122
HH09	62	26/11/19	6/8/20	255	111	28/11/19	26/3/20	120	7	12/1/20	19/1/20	7	109	27/11/19	24/3/20	119
HH10	43	19/12/19	6/8/20	232	97	20/12/19	20/3/20	92	7	23/2/20	1/3/20	7	87	20/12/19	27/3/20	99
HH11	59	26/11/19	6/8/20	255	39	19/2/20	30/3/20	41					46	10/2/20	27/3/20	46
HH12	46	20/12/19	6/8/20	231	86	25/12/19	29/3/20	96					88	21/12/19	22/3/20	93
HH13	63	26/11/19	6/8/20	255	93	13/12/19	26/3/20	105	7	19/1/20	26/1/20	7	87	3/12/19	27/3/20	116
HH14	64	26/11/19	6/8/20	255	59	30/11/19	12/3/20	104					22	29/11/19	15/3/20	108
HH15	64	26/11/19	6/8/20	255	90	5/12/19	30/3/20	117	7	29/12/19	5/1/20	7	113	1/12/19	27/3/20	118
HH16	54	12/12/19	6/8/20	239	69	14/1/20	26/3/20	73	7	19/1/20	26/1/20	7	76	7/1/20	27/3/20	81
HH17	54	12/12/19	6/8/20	239	65	11/1/20	26/3/20	76	7	5/1/20	12/1/20	7	88	12/12/19	27/3/20	107
HH18	53	12/12/19	6/8/20	239	93	15/12/19	30/3/20	107	7	29/12/19	5/1/20	7	98	18/12/19	27/3/20	101
HH19	55	12/12/19	6/8/20	239	45	23/12/19	26/3/20	95	7	12/1/20	19/1/20	7	102	12/12/19	27/3/20	107
HH20	55	12/12/19	6/8/20	239	73	7/1/20	26/3/20	80	7	5/1/20	12/1/20	7	79	4/1/20	27/3/20	84
HH21	46	20/12/19	6/8/20	231	89	22/12/19	26/3/20	96	7	5/1/20	12/1/20	7	94	21/12/19	27/3/20	98
HH22	23	11/1/20	6/8/20	209	21	4/3/20	26/3/20	23					27	23/2/20	27/3/20	34
HH23	18	16/1/20	6/8/20	204	47	1/2/20	26/3/20	55					61	19/1/20	27/3/20	69
HH24	15	31/1/20	6/8/20	189	35	15/2/20	26/3/20	41					53	2/2/20	27/3/20	55
HH25	15	31/1/20	6/8/20	189	36	14/2/20	26/3/20	42	7	9/2/20	16/2/20	7	56	31/1/20	27/3/20	57
HH26	14	31/1/20	6/8/20	189	34	1/2/20	14/3/20	43					47	31/1/20	19/3/20	49
HH27	15	31/1/20	6/8/20	189	43	10/2/20	26/3/20	46	7	9/2/20	16/2/20	7	49	7/2/20	27/3/20	50
HH28	13	7/2/20	6/8/20	182	38	15/2/20	26/3/20	41	7	23/2/20	1/3/20	7	43	13/2/20	27/3/20	44
HH29	6	28/2/20	6/8/20	161	16	28/2/20	21/3/20	23					28	28/2/20	27/3/20	29
HH30	5	4/3/20	6/8/20	156	54	6/3/20	30/4/20	56					23	4/3/20	27/3/20	24
Average	**36**			**216**	**62**			**72**	**7**			**7**	**66**			**74**

SMETER A, SMETER C, SMETER E, SMETER F, and SMETER G: self-reported number of days and period of data used to calculate the HTC after the second monitoring period.

Home ID	SMETER A				SMETER C				SMETER E				SMETER F				SMETER G			
	Days	from	to	period	Days	from	to	period	Days	from	to	period	Days	from	to	period	Days	from	to	period
HH01	30	1/8/20	25/2/21	208	82	6/11/20	24/2/21	110	20	2/12/20	29/1/21	58	40	1/11/20	28/12/20	57	7	18/2/21	25/2/21	7
HH02	30	1/8/20	25/2/21	208	69	5/11/20	13/2/21	100	20	2/12/20	29/1/21	58	40	1/11/20	22/12/20	51	7	29/10/20	5/11/20	7
HH03	30	1/8/20	25/2/21	208	81	4/11/20	24/2/21	112	22	2/12/20	29/1/21	58	40	1/11/20	21/1/21	81	7	26/11/20	3/12/20	7
HH04	30	1/8/20	25/2/21	208	54	25/11/20	22/2/21	89	51	5/12/19	29/1/21	421	40	1/11/20	21/12/20	50	7	26/11/20	3/12/20	7
HH05	30	1/8/20	25/2/21	208	na	na	na	na	41	5/12/19	29/1/21	421	40	1/11/20	30/12/20	59	7	29/10/20	5/11/20	7
HH06	30	1/8/20	25/2/21	208	48	22/11/20	7/2/21	77	23	31/1/20	29/1/21	364	40	1/11/20	25/12/20	54	7	24/12/20	31/12/20	7
HH08	30	1/8/20	25/2/21	208	74	11/11/20	24/2/21	105	42	2/12/20	29/1/21	58	40	1/11/20	17/12/21	46	7	11/2/21	18/2/21	7
HH09	30	1/8/20	25/2/21	208	66	4/11/20	22/2/21	110	54	2/12/20	29/1/21	58	40	1/11/20	21/12/20	50	7	24/12/20	31/12/20	7
HH10	30	1/8/20	25/2/21	208	63	7/11/20	14/2/21	99	49	19/12/19	29/1/21	407	40	1/11/20	23/12/20	52	7	31/12/20	7/1/21	7
HH11	30	1/8/20	25/2/21	208	66	25/11/20	24/2/21	91	39	5/12/19	29/1/21	421	40	1/11/20	21/12/20	50	7	19/11/20	26/11/20	7
HH12	30	1/8/20	25/2/21	208	81	7/11/20	24/2/21	109	45	20/12/19	29/1/21	406	40	1/11/20	22/12/20	51	7	24/12/20	31/12/20	7
HH13	30	1/8/20	25/2/21	208	67	6/11/20	22/2/21	108	58	5/12/19	29/1/21	421	40	1/11/20	22/12/20	51	7	24/12/20	31/12/20	7
HH14	30	1/8/20	25/2/21	208	56	6/11/20	2/2/21	88	44	5/12/19	29/1/21	421	40	3/11/20	24/12/20	51	7	29/10/20	5/11/20	7
HH15	30	1/8/20	25/2/21	208	41	1/10/20	13/1/21	104	46	5/12/19	29/1/21	421	40	1/11/20	23/11/20	22	7	19/11/20	26/11/20	7
HH16	30	1/8/20	25/2/21	208	67	11/11/20	20/2/21	101	49	13/12/19	29/1/21	413	40	1/11/20	21/12/20	50	7	31/12/20	7/1/21	7
HH17	30	1/8/20	25/2/21	208	51	8/10/20	6/1/21	90	50	13/12/19	29/1/21	413	40	1/11/20	21/12/20	50	7	31/12/20	7/1/21	7
HH19	30	1/8/20	25/2/21	208	48	18/11/20	28/1/21	71	52	13/12/19	29/1/21	413	40	1/11/20	21/12/20	50	7	26/11/20	3/12/20	7
HH20	30	1/8/20	25/2/21	208	55	25/11/20	7/2/21	74	50	13/12/19	29/1/21	413	40	1/12/20	20/1/21	50	7	26/11/20	3/12/20	7
HH21	30	1/8/20	25/2/21	208	74	26/11/20	17/2/21	83	49	20/12/19	29/1/21	406	40	1/12/20	20/1/21	50	7	4/2/21	11/2/21	7
HH22	30	1/8/20	25/2/21	208	50	26/11/20	20/2/21	86	28	11/1/20	29/1/21	384	40	1/11/20	21/12/20	50	7	29/10/20	5/11/20	7
HH23	30	1/8/20	25/2/21	208	62	25/11/20	23/2/21	90	23	16/1/20	29/1/21	379	40	1/12/20	25/1/21	55	7	12/11/20	19/11/20	7
HH24	30	1/8/20	25/2/21	208	74	7/10/20	24/2/21	140	12	31/1/20	29/1/21	364	40	1/11/20	21/12/20	50	7	29/10/20	5/11/20	7
HH25	30	1/8/20	25/2/21	208	76	13/11/20	24/2/21	103	22	31/1/20	29/1/21	364	40	1/11/20	23/12/20	52	7	29/10/20	5/11/20	7
HH26	30	1/8/20	25/2/21	208	78	3/10/20	24/2/21	144	15	31/1/20	29/1/21	364	40	5/11/20	25/12/20	50	7	5/11/20	12/11/20	7
HH27	30	1/8/20	25/2/21	208	82	1/10/20	24/2/21	146	21	31/1/20	29/1/21	364	40	1/11/20	21/12/20	50	7	19/11/20	26/11/20	7
HH28	30	1/8/20	25/2/21	208	67	6/11/20	24/2/21	110	24	2/12/20	29/1/21	58	40	1/11/20	21/12/20	50	7	12/11/20	19/11/20	7
HH29	30	1/8/20	25/2/21	208	52	8/10/20	13/1/21	97	22	2/12/20	29/1/21	58	40	1/11/20	21/12/20	50	7	31/12/20	7/1/21	7
Average	**30**			**208**	**65**			**101**	**36**			**311**	**40**			**51**	**7**			**7**

Appendix C: Detailed description of the measurement, analysis and uncertainty of the measured HTC

The co-heating test was first used in late 1970s, where Sonderegger and Modera (1979) and Sonderegger et al. (1980) used electric heaters and the buildings own heating system to determine the on-site efficiency of duct heating and cooling systems under actual boundary conditions. It follows that the term co-heating originates from the multiple use of heating provisions within the test. Since inception, the principle established has been used to estimate thermal characteristics of the building envelope, notwithstanding the fact that the methods to heat the building are now often limited to a single source of electric heaters. More recently, the basic principle of establishing a HTC through co-heating test has been applied to better understand the performance of the whole building and building elements (Bell & Lowe, 1997; Masy & Lebrun, 2004; Francisco et al., 2006; Lowe et al., 2007; Bell et al., 2010; Palmer et al., 2011; Stamp, 2011; Bauwens, Standaert, et al., 2012; Deconinck and Leunis, 2012; Bauwens and Roels, 2014; Johnston et al. 2014: 2015; Gorse et al., 2017; Jack et al., 2018, , Johnston et al. 2020; Glew et al., 2020).

To better understand the accuracy of the method, the NHBC (Butler & Dengel, 2013) undertook a series of co-heating tests, performed by different groups on the same building. The weather conditions during the tests represented a confounder, with solar radiation having a major impact. From the tests undertaken, a maximum uncertainty in the HTC of 17% was reported. The testing periods for the different groups ranged from 12-18 days. The spread of results was principally due to the different methods of analysis, as opposed to significant variations in the way that the different participants set up and conducted the test (Butler & Dengel, 2013). Jack (2015) subsequently reviewed the sensitivity of analysis and reproducibility of the method using data from the same study and found that when the precise data collection and methods are followed the HTC can be reported with an uncertainty of ±10%. Thus, within the work reported here, further consideration is given to the analysis and uncertainty of the results.

The LBU's co-heating methodology has evolved from the early studies Sonderegger et al. (1979) with the first method described by Wingfield et al. (2010). From this, a more elaborate description of the experimental set up has evolved, with factors to be considered including methods to limit the impact of solar and wind, the latent heat from materials that may be drying out, the heat exchanges as a result of party elements, and the thermal capacity of ground floors etc. (Johnston et al. 2013). While variations in the method do exist, the method provided by LBU was considered sufficiently robust for this study. However, adaptations were considered necessary for the practical delivery of the programme, and to ensure uncertainty in measurements and analysis were given due consideration.

The adapted in-use reference measured HTC as reported here is considerate of the variation likely to be experienced in measurement due to operational changes to the fabric when in-use. The inclusion of these variations results in each HTC being assigned a confidence interval, the upper and lower bounds of which provide the range of HTCs considered reliable for this this study.

The co-heating test method (henceforth, referred to as the co-heating test) was the primary method used to measure the reference HTCs, as it has been shown to be a reliable method of determining the HTC of a building (Jack et al., 2018). In the UK, the co-heating test method has become established as a recognised in situ test method to obtain an estimate of the overall heat loss coefficient (HLC) of an unoccupied dwelling and formed a requirement of the recent Technology Strategy Board's Building Performance Evaluation Programme, whose aim was to understand the key factors that influence the in-use performance of buildings (Technology Strategy Board, 2010).

Experience from previous co-heating tests suggests that reliable test results can be obtained within a 2-week test window if environmental conditions are satisfactory and no operational difficulties are encountered. Consequently, properties were made available for a minimum of 3 weeks duration to allow for a co-heating test and subsequent QUB tests to be carried out. However, for the purpose of calculating the HTC, not all data collected through the test period was used. The initial heat-up period, often in the region of 3 days, was removed from the analysis, as were interruptions due to people entering the building and days which suffered from equipment failure.

Simulation work by Alexander and Jenkins (2015) and Stamp et al. (2017), as well as LBU's past experience of undertaking numerous co-heating tests, has demonstrated that such periods of quasi-steady conditions can be sufficient to obtain a confident HTC estimate. As 3 weeks was the maximum testing window, where weather patterns were unseasonable the whole sample of data collected were not always appropriate – in two of the tests, a reliable HTC from co-heating alone was not obtained.

Measured HTC: Analysis

Electric co-heating is a quasi-steady state test which involves artificially heating the internal environment of a building to an elevated, homogenous, and constant temperature with electric resistance heaters. The electric power input to the house, as well as the internal and external environmental conditions, are measured throughout the test period.

To derive a value for the HTC, the total power input to the property over each 24 hour period is calculated, along with the average temperature difference between the internal and external environments (ΔT) over each 24 hour period. Performing linear regression with daily average ΔT values as the independent variable and the daily power input as the dependent variable will then result in a straight line, whose gradient is a rough estimate of the HTC. This linear fit is forced through zero on the basis that there should be no heat loss if there is no temperature difference. However, adopting such an approach is too simplistic in practice, as it does not take into several complicating factors which occur in real co-heating tests (for a full overview of the co-heating test and data analysis, refer to Bauwens and Roels (2014).

The first complicating factor in co-heating tests is solar heating, as solar radiation from the sun will heat the test dwelling in addition to the space heating system. To account for solar gains, the solar radiation each day is recorded using a nearby weather station. Multiple linear regression is then performed on the daily power input as a function of both daily solar radiation and ΔT. The coefficient which is derived for the solar radiation term can then be multiplied by the daily measured radiation to estimate the solar power received by the property over the 24-hour period. The input power to the property can then be corrected by taking the known electrical power input and adding the estimated solar power input. An uncertainty on the estimate of solar power is output from the linear fitting procedure, and this uncertainty can be fed through to the estimate for the total power input using standard error propagation. There will also be an uncertainty on the total power input caused by the uncertainty in the measurement devices. However, the accuracy of the measurement devices means that their contribution to overall uncertainty is very small (<1% of the overall uncertainty) and was therefore not included in this analysis.

The second complicating factor are the dynamics that occur when heat energy is added to a property. The co-heating test aims to achieve a steady temperature to mitigate these dynamic effects. The analysis also averages data over a 24-hour aggregation period so that short-term dynamic effects are not significant. However, at the start of a co-heating test, the property will be in a heat-up phase and, although internal air temperatures stabilise earlier, a steady thermal gradient through the building fabric will not be achieved. We therefore manually analyse the early days of the co-heating and remove days where the data suggests the building fabric is not yet heat soaked. Typically, this means removing the first 3 days of data. Even when the property is well-heated however, the solar gains mentioned previously can introduce dynamics during the daylight hours. We mitigate against this by dividing out data into 24-hour aggregation periods starting at 06:00 am as, by this time, solar heating from the previous day will have largely been lost and any dynamics introduced will therefore be contained within the same 24-hour aggregation period.

The final complicating factor is uncertainty in the ΔT value. The co-heating analysis uses the mean temperature difference between the internal and external environment over a 24-hour period, but this value will be associated with some uncertainty. The mean internal temperature will, by design, be very stable and be associated with a small degree of uncertainty. The mean external temperature on the other hand will vary naturally over the day and will therefore be more uncertain. If the external temperature follows a gaussian distribution, the standard error of the mean (SEM) external temperature can be calculated from equation 1

Equation 1

$$SEM = \frac{\sigma}{\sqrt{n}}$$

Where, σ is the standard deviation of the external temperature and n the number of data points. However, the external temperatures were observed to not follow a gaussian distribution and we therefore use a bootstrap procedure to estimate the uncertainty in the mean, as this bootstrap procedure does not assume a particular underlying probability distribution for the data. In practice, we found that the SEM produced by the bootstrap procedure was often very close to the SEM calculated from the equation above.

After the above corrections have been made, the data can be fit with a linear model. Ordinary least squares regression is not suitable for data which has uncertainty in the x-values. Instead, we use Deming regression which can include uncertainty in both the x and y variables (see Figure C1). All fits were forced through zero and residuals of the fitting process were tested for heteroskedasticity.

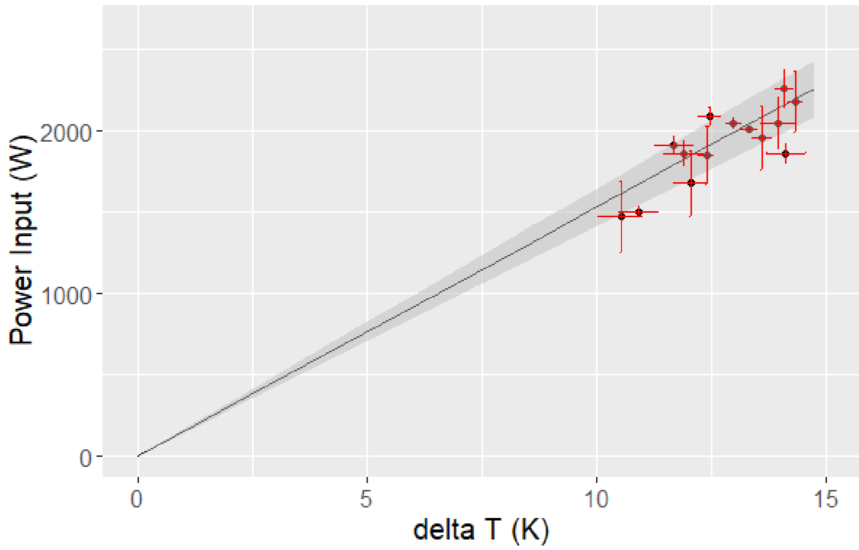


Figure C1: Example of Deming regression in a co-heating analysis. The gradient of the line gives an estimate for the HTC.

Measured HTC: Accounting for seasonal Uncertainty

In addition to the uncertainty output from the regression analysis, it was noted from modelling work that the time of year at which the co-heating test takes place can also affect the result. This appears to be due to solar heating introducing dynamics which cannot be properly accounted for the in the co-heating analysis. To quantify this seasonal uncertainty, 59 homes were simulated as undergoing a consistent co-heating test between September and February. Multiple co-heating analyses were then performed on these data using a 21-day rolling window and the results collated to determine the average co-heating result throughout the period. The uncertainty required to ensure that, 95% of the time, a single co-heating result agrees with the mean was then determined. The uncertainty required to meet this criterion was found to be ±10.7%.

As the seasonal uncertainty depends on solar radiation, it will increase the further from the winter period a test is conducted. Three of the co-heating tests performed during this research were conducted in April-May and, as such, the seasonal uncertainty will be larger. Repeating the above procedure for the extended period of September to May resulted in an uncertainty of ±26.0% being required to account for seasonal variations.

Measured HTC: Accounting for in-use operation of purpose-built ventilation

The LBU co-heating test method stipulates that all purpose-provided ventilation openings should be sealed when undertaking a test. The HTC resulting from a LBU co-heating test (Johnston et al. 2013) is thus a metric which quantifies the aggregate fabric and unintended air infiltration heat loss rates. Consequently, the LBU co-heating test method does not include the heat loss rate through purpose provided ventilation, such as designed ventilation openings within the building fabric. However, both SMETER technologies and RdSAP assume that purpose provided ventilation will largely be left open. The LBU co-heating test protocol therefore needs to be modified to include the heat loss rate through designed ventilation openings within the building fabric that are left open (e.g., intermittent extract fans, air bricks, flues, fireplaces, trickle vents, passive vents, etc.).

Using modelling to correct a standard co-heating test result for purpose provided ventilation is challenging. This is because assumptions based upon the ventilation rate through designed ventilation openings can differ widely (e.g., an open chimney in SAP 2012 is assumed to have an airflow rate of 40 m^3/h (BRE, 2014), though a study by BRE suggested a typical airflow rate of 75 to 80 (± 20) m^3/h (BRE, 2015)). In addition, the condition of these openings is difficult to establish through a cursory survey, such as that undertaken for RdSAP. For instance, chimney flues may be blocked or airflow through airbricks could have been compromised by retro-filled cavity wall insulation or changes in external ground level, which introduces further uncertainty.

Instead, adjustment of the co-heating test results was performed experimentally, to allow for ventilation that might be encountered under habitation. Initially, the co-heating test protocol itself was modified to incorporate designed infiltration by leaving purpose-built ventilation open. However, after collecting data on 6 houses in this manner, it became apparent that wind was introducing too large an uncertainty into the co-heating measurements for the less airtight dwellings. Therefore, the co-heating test was deployed in the standard manner (with all vents closed), and multiple air pressurisation tests were instead used to calculate differences in heat loss due to ventilation. Air pressurisation tests were first carried out in accordance with ATTMA TS1 (ATTMA, 2016) with all purpose-provided ventilation sealed. The uncertainty in the infiltration was calculated using the procedure in ISO9972:2015. The results of the air pressurisation tests were then used to disaggregate the HTC obtained from the co-heating test into its fabric and air infiltration heat loss components, via Equation 2:

Equation 2

$$HTC = \Sigma U.A + 0.33N.V$$

$$\Sigma U.A = HTC - 0.33N.V$$

In Equation 2, $\Sigma U.A$ is the fabric heat loss term in which U is the U-value for each building element and A is the area. $0.33N.V$ is the background ventilation heat loss term in which N is the air leakage in air changes per hour during the co-heating test, V is the volume of the building and 0.33 approximates the density of air multiplied by its specific heat capacity at 25°C. As the pressurisation test gives a value of N at 50 Pa pressure difference, the $N_{50}/20$ *'rule of thumb'*, originally devised by Kronvall and Persily (cited by Jacobson et al. 1984 and Sherman, 1987), was used to convert this to more typical pressure differences of 4Pa.

In addition to the air pressurisation test with all ventilation sealed, a pressurisation test was carried out with all purpose-provided ventilation open. This new pressurisation test gave a value of N which could be used to approximate an average background ventilation heat loss term under more typical, lived-in conditions. This was then added back on to the fabric heat loss term to give an estimate of the HTC of the lived-in property (see Figure C2).

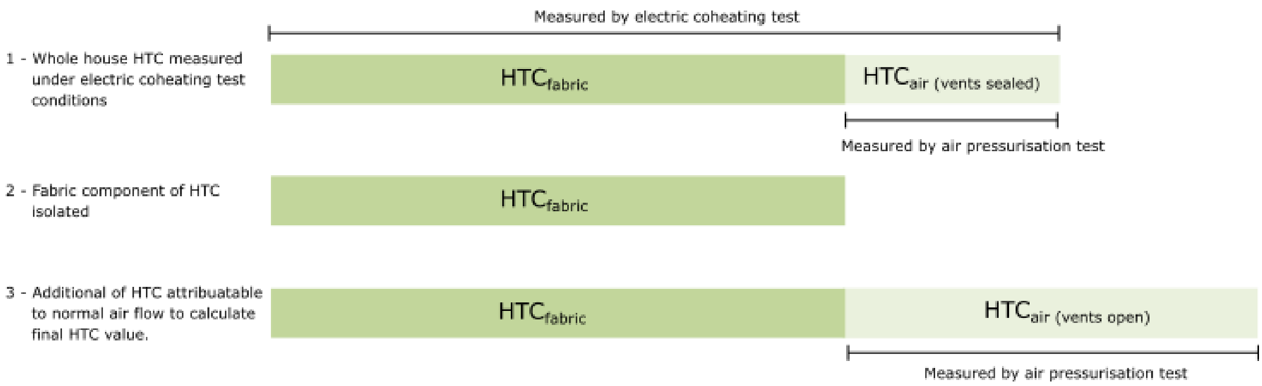


Figure C2: Schematic of HTC's obtained from co-heating and air pressurisation tests used to provide HTCs for fabric, fabric and infiltration with vents closed and fabric, infiltration and ventilation through purpose provided ventilation (vents open)

An additional complication arises from the fact that some purpose provided ventilation (i.e., trickle vents) can be changed from open to closed by the occupant. As the status of these vents, when buildings are occupied, cannot be known an uncertainty is introduced into the value of HTC $_{\text{(air vents open)}}$. To account for this, a final air pressurisation test was conducted with all ventilation open except for trickle vents. A second HTC was calculated for each house assuming both that all vents were open, except for these trickle vents. The average of this HTC, and the HTC assuming all vents were open was taken as the final value of the HTC for a lived-in property. The uncertainty on this final value has to reflect the fact that the trickle vents may all be open, may all be closed, or may be a mixture of the two. The confidence interval (CI) on this final HTC value is therefore taken as the lower bound of the 95% CI assuming that trickle vents were closed, and the upper bound of the CI assuming that trickle vents were open. After including the additional seasonal uncertainty described above, this value of HTC and its associated confidence interval will represent the range of continuous ventilation systems that could be operating during in-use monitoring. The use of intermittent mechanical ventilation (extract fans) was not accounted for in this analysis.

In two properties, the initial surveys found suspended timber floors with no covering over the floorboards. When occupied, these floorboards would be carpeted thus increasing the air tightness. For these 2 properties, the air tightness in the occupied dwelling was estimated by first laying craft paper to mimic carpeting before the air pressurisation tests described above.

Heat flux density measurements were undertaken in the majority of tested houses, Measurements were taken using Hukseflux HFP01 heat flux plates, connected to either an Omni Instruments DT80 datataker or an Eltek Squirrel 450/850 datalogger. Heat flux plates were fitted following guidance in BS ISO 9869:2014, ensuring each was in good thermal contact with the surface to which it was affixed. To minimise the risk of surface damage, Low tac double sided adhesive tape was used to secure the heat flux plates in place. Where heat flux was measured through party walls, it was not possible to measure any temperatures in the adjacent home.

Data limitations

There are two main limitations to the data obtained as part of the SMETER trials. The first of these pertains to the external temperatures recorded. For the majority of properties, a single weather station was used to record the solar and external temperatures. Although the weather station was located a maximum of 5.8 km from the test houses, it is possible that the temperature recorded was not representative of the local conditions near a property. However, the uncertainty applied to the mean ΔT value should help to mitigate this limitation.

The second potential limitation arises due to the neighbours. The majority (29 out of 30) of houses in which co-heating was performed were semi-detached or end terraces, and therefore had a party wall (PW) which was shared with a neighbour. During the co-heating test, the elevated temperature means that heat will be monodirectional through the PW element. However, when the houses are occupied, the heating PW heat flow will not necessarily be monodirectional, meaning that the co-heating test and SMETER products may be experiencing different environments. However, the PW only represents a small proportion of the total heat loss area of the tested dwellings and the heat lost or gained through the PW is therefore expected to be small. Indeed, an analysis of the 29 tested houses suggested that the average decrease in HTC after accounting for PW heat loss was only (6.4±1.40) W/K. The PW corrected HTC values are very often within the uncertainty of the original HTC values, highlighting that this limitation of the data is seldom significant.

Equipment used for co-heating tests

Device	LBU		LU	
	Make/model	**Uncertainty**	**Make/model**	**Uncertainty**
Indoor dry-bulb temperature	Eltek GS52 Transmitter with Pt100 RTD Sensor Eltek GC10 T/RH Sensor/Transmitter Eltek SRV250 wireless data loggers	± 0.3°C ± 0.4°C	Grant U-type thermistor wired to DT85 DataTaker	± 0.3°C
Temperature controller	InstCube PID digital temperature controller	± 0.5°C	InstCube PID digital temperature controller	± 0.5°C
Thermostat for temperature controller	Pt100 RTD	± 0.1°C	PT100 resistance thermometer	± 0.1°C
Heater	Stanley ST-02 & Sealey 2kW industrial fan heaters	N/A	Dimplex DXFF30TSN 3000W/1500W	N/A
Circulation fan	14” & 18” floor mounted fans	N/A	45 cm floor mounted fan	N/A
Electricity consumption measurement	Elster A100 kWh meter	< ± 0.1%	Plogg metering plug	< ± 0.5%
Heat flux	Hukseflux HFP01 Wired to DataTaker DT80	± 0.5%	N/A	N/A

Details of the heat flux measurements

Home ID	Start date	End date	Wall monitored	Floor monitored
HH01	15/03/19	05/04/19	Living room, kitchen, front and rear bedroom	Living room, kitchen
HH02	15/04/19	03/05/19	Living room, bedroom	Bedroom
HH03	24/04/19	17/05/19	Living room, front and rear bedroom	Living room
HH04	13/09/19	03/10/19	Bedroom, bathroom	None
HH05	13/09/19	30/09/19	Front and rear bedroom	None
HH06	30/09/19	14/10/19	Front bedroom, rear bedroom (north)	None
HH07	23/09/19	04/10/19	None	None
HH08	30/09/19	21/10/19	Rear bedroom (west), Bathroom	Hall
HH09	01/10/19	18/10/19	None	None
HH10	07/10/19	24/10/19	Bedroom (north), airing cupboard	None
HH11	16/10/19	05/11/19	None	None
HH12	21/10/19	08/11/19	Living room, kitchen, bedroom	Living room, kitchen
HH13	04/11/19	15/11/19	Living room, kitchen	Living room
HH14	06/11/19	22/11/19	Kitchen, hall, landing	Living room
HH15	05/11/19	18/11/19	Living/dining room	Living room
HH16	05/11/19	25/11/19	None	None
HH17	18/11/19	28/11/19	None	None
HH18	12/11/19	03/12/19	Living room, kitchen	Living room
HH19	12/11/19	29/11/19	Living room, kitchen	Living room
HH20	20/11/19	06/12/19	Living room, kitchen	Living room
HH21	20/11/19	10/12/19	Living room	Living room
HH22	25/11/19	10/12/19	Living room	Living room
HH23	03/12/19	16/12/19	Bedroom (north-east)	Hall
HH24	26/11/19	11/12/19	None	None
HH25	06/12/19	19/12/19	Living room	Living room
HH26	02/12/19	18/12/19	None	None
HH27	18/12/19	07/01/20	Kitchen	Living room, kitchen
HH28	18/12/19	06/01/20	None	None
HH29	04/02/20	18/02/20	None	None
HH30	03/02/20	17/02/20	Living/dining room, front bedroom	Living/dining room

References

Alexander, D.K. and Jenkins, H.G (2015) The validity and reliability of coheating tests made on highly insulated dwellings, Energy Procedia, v79, 6th International Building Physics Conference, IBPC 2015 Nov. pp 1732-1737 https://doi.org/10.1016/j.egypro.2015.11.282

Alzetto, F., Pandraud, G., Fitton, R., Heusler, I. and Sinnesbichler, H. (2018) QUB: A fast dynamic method for in-situ measurement of the whole building heat loss. Energy and Buildings 174:124-133; DOI: 10.1016/j.enbuild.2018.06.002

ATTMA (2016) Technical Standard L1. Measuring Air Permeability of Building Envelopes (Dwellings) [Internet]. Airtightness Testing and Measurement Association, Northampton. September 2016 Issue. Available from: https://www.bcta.group/shop/attma-technical-standard-l12016/ [Accessed: 29th April 2020].

Bauwens, G. and Roels, S. (2014) Co-heating test: A state-of-the-art. Energy and Buildings, Volume 82, October 2014, p 163-172.

Bauwens, G. and Roels, S. (2017) On site thermal performance characterisation of building envelopes: How important are heat exchanges with neighbouring zones. Energy Procedia 132 (2017) 339-334; doi.org/10.1016/j.egypro.2017.09.744 Everett, R. (1985). Rapid Thermal Calibration of Houses, Technical Report, Open University Energy Research Group, Milton Keynes, UK, 1985, ERG 055.

Bauwens, G., Standaert, P., Decluve, F., Roels, S., (2012) Reliability of Co-Heating Measurements, in: Wright, J., Cook, M. (Eds.), Proceedings of the 2012 Building Simulation and Optimization Conference; 10-11 September 2012, Loughborough, Leicestershire, UK. Loughborough University, Loughborough, pp. 348–355.

Bell, M. and Lowe, R.J. (1997) The York Energy Demonstration Project. Final Report on a DOE Greenhouse Project, CEBE Research Report No 3, Leeds Metropolitan University.

Bell, M., Wingfield, J., Miles-Shenton, D., (2010) Low carbon housing Lessons from Elm Tree Mews. Joseph Rowntree Foundation, York.

BRE (2014), The Government's Standard Assessment Procedure for Energy Rating of Dwellings: 2012 edition (updated Feb. 2014). Building Research Establishment, Watford: BRE on behalf of the Department for Energy and Climate Change.

Butler, D and Dengel A. (2013) A review of coheating test methodologies, NHBC Foundation, Milton Keynes, ISBN 978-1-84806-351-8

Cen TC89 (2020) CEN/TC 89 THERMAL PERFORMANCE OF BUILDINGS AND BUILDING COMPONENTS, European Committee for Standardization, accessed 27 4 20 https://standards.cen.eu/dyn/www/f?p=204:7:0::::FSP_ORG_ID:6072&cs=1838C2B207A6B7D88562A05376928F458

Cooper, C., Zheng, X., Wood, C., Gillot, M., Tetlow, D., Riffat, S. and De Simon, L. (2016): Field trialling of a pulse airtightness tester in a range of UK homes, International Journal of Ventilation, DOI: 10.1080/14733315.2016.1252155

Deconinck, A.H. and K. Leunis (2012). Co-heating: a critical evaluation of data analysis methods. Masters Thesis. Katholieke Universiteit Leuven.

Everett, R. (1985) Rapid Thermal Calibration of Houses. Technical Report ERG 055 for the Science and Engineering Research Council, Milton Keynes, UK.

Francisco, P., Siegel, J., Palmiter, L., Davis, B., (2006) Measuring residential duct efficiency with the short-term coheat test methodology. Energy and Buildings 38(9), 1076–1083. doi:10.1016/j.enbuild.2005.12.008.

Ghiaus, C. & Alzetto, F. (2019) Design of experiments for Quick U-building method for building energy performance measurement, Journal of Building Performance Simulation, 12:4, 465-479, DOI: 10.1080/19401493.2018.1561753

Glew, D, et al. (2020) Thin Internal Wall Insulation (TIWI), Measuring Energy Performance Improvements in Dwellings Using Thin Internal Wall Insulation. Department for Business Energy and & Industrial Strategy – in print.

Gorse, C. Glew, D. Johnston, D. Fylan, F. Miles-Shenton, D. Smith, M., Brook-Peat, M. Framer., D. Stafford, A., Parker, J., Fletcher, M. and Felix, T. (2017), Core Cities Green Deal Monitoring Project Leeds. BEIS Buildings Energy Efficiency Technical Research, 9 November 2017.https://www.gov.uk/government/collections/buildings-energy-efficiency-technical-research

Jack, R. (2015) Building diagnostics: practical measurement of the fabric thermal performance of houses, PhD, University of Loughborough, Loughborough, UK.

Jack, R., Loveday, D., Allinson, D. & Lomas. K. (2018) First evidence for the reliability of building co-heating tests, Building Research & Information, 46:4, 383-401, DOI: 10.1080/09613218.2017.1299523

Jacobson, D. I., Dutt, D. S. and Socolow, R. H (1984) Pressurlzatlon Testing, Infiltration Reduction, and Energy Savings, in Measured Air Leakage of Buildings, ASTM STP904.

Johnston, D. Farmer, D. Brooke-Peat, M. and Miles-Shenton, D. (2014) Bridging the Domestic Building Fabric Performance Gap. Building Research & Information. Volume 44, Issue 2, pp.147-159. DOI:10.1080/09613218.2014.979093.

Johnston, D. Miles-Shenton, D. and Farmer, D. (2015) Quantifying the Domestic Building Fabric ‘Performance Gap’. Building Services Engineering Research & Technology (BSER&T). Volume 36, No.5, September 2015, pp.614–627. DOI:10.1177/014362441557034.

Johnston, D., Miles-Shenton, D., Farmer, D. & Wingfield, J. (2013) Whole House Heat Loss Test Method (Coheating), Leeds Metropolitan University, 2013, June 2013.Available from: https://www.leedsbeckett.ac.uk/as/cebe/projects/cebe_coheating_test_method_june2013.pdf

Johnston, D. Siddall, M. Ottinger, O. Peper, S. and Feist W. (2020) Are the energy savings of the Passive House standard reliable? A review of the as-built thermal and space heating performance of Passive House dwellings from 1990 to 2018. Energy Efficiency Journal. DOI:10.1007/s12053-020-09855-7.

Jones, B. Persily, A. and Sherman, M. (2016) The origin and application of leakage-infiltration ratios. In Proceedings of the 37th AIVC-ASHRAE-IAQ joint Conference 2016: Defining Indoor Air Quality: Policy, Standards and Best Practices. 12-14th September 2016, Alexandria, Virginia, USA.

Kronvall, J. (1978). Testing of houses for air leakage using a pressure method. In: ASHRAE Transactions 84, pp. 72–79.

Lowe, R.J., Wingfiled, J., Bell, M., and Bell, J.M. (2007). Evidence for heat losses via party wall cavities in masonry construction". Building Services Engineering Research and Technology 28,2, pp. 161–181.

Masy, G. and Lebrun, J. (2004) Passive Solar Houses: Availability of weather suitable for calibration in the UK. Proceedings of the Fourth International Conference for Enhanced Building Operations, Paris, France, October 18th - 19th, 2004.

Meulemans J. (2019) An Assessment of the QUB/e Method for Fast In Situ Measurements of the Thermal Performance of Building Fabrics in Cold Climates. In: Johansson D., Bagge H., Wahlström Å. (eds) Cold Climate HVAC 2018. CCC 2018. Springer Proceedings in Energy. Springer

Palmer, J., Pane, G., Wingfield, J., Bell, M. (2011) Comparing Primary and Secondary Terms Analysis and Re-Normalisation (PStar) Test and Co-Heating Test Results. Final report: BD2702. Department for Communities and Local Government, London

Persily, A. and Linteris, G. (1983) A comparison of measured and predicted infiltration rates. ASHRAE Transactions 89, pp. 183–200.

Sharpe, T. McGill, G. Gupta, R. Gregg, M. and Mawditt, I.(2016) Characteristics and performance of MVHR systems - A meta study of MVHR systems used in the Innovate UK Building Performance Evaluation Programme. Innovate UK.

Sherman M. H. (1987) Estimation of Infiltration for Leakage and Climate Indicators. Energy and Buildings, 10, pp. 81-86.

Siviour, J. (1985) Experimental Thermal Calibration of Houses. In: Everett R Rapid Thermal Calibration of Houses. Technical Report ERG 055 for the Science and Engineering Research Council, Milton Keynes, UK.

Sonderegger, R. C. Condon, P. E. and Modera, M. P. (1979) In-situ Measurements of Residential Energy Performance Using Electric Co-heating. ASHRAE Transactions, 1979; Vol. 86 (I).

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

Sonderegger, R. C., Condron, P. E., and Modera, M. P. (1980) In-situ measurements of residential energy performance using electric co-heating. In: American Society of Heating, Refrigerating and Air Conditioning Engineers Transactions 86, p. 394.United States: N.P. Web https://www.osti.gov/servlets/purl/5492967

Sougkakis, V., Meulemans, J., Alzetto, F., Wood, C., Gillott, M. & Cox, T. (2017) An assessment of the QUB method for predicting the whole building thermal performance under actual operating conditions. International

SEEDS Conference 2017: Sustainable Ecological Engineering Design for Society, Leeds Beckett University, Sep 2017, Leeds, United Kingdom.

Stamp, S. (2011). The Error and Uncertainty in Coheating Tests. Masters Thesis. University College London.

Stamp, S. F., (2015) Assessing uncertainty in co-heating tests: Calibrating a whole building steady state heat loss measurement method, PhD, University College London.

Stamp, S., Almirano-Medina, H. and Lowe, R. (2017) Assessing the Relationship between Measurement Length and Accuracy within Steady State Co-Heating Tests. Buildings 2017, 7(4), 98; doi:10.3390/buildings704009

The Technology Strategy Board (2010) Building Performance Evaluation, Domestic Buildings - Guidance for Project Execution. Swindon, The Technology Strategy Board.

Wingfield, J. (2010) In-situ Measurement of Whole House Heat Loss using Electric Coheating. Technical Report. Leeds, UK: Leeds Beckett University.

Zheng, X.F., Cooper, E. Zu, Y., Gillott, M., Tetlow, D., Riffat, S., Wood, C. (2019) Experimental Studies of a Pulse Pressurisation Technique for Measuring Building Airtightness. Future Cities and Environment, 5(1): 10, 1–17. DOI: https://doi.org/10.5334/fce.66

Appendix D: Description of the QA methods - Multiple Linear Regression (MLR) and Siviour

Multiple Linear Regression (MLR) and Siviour are based on the steady state heat balance equation, given below:

$$\mathrm{P_{heat}} = HTC(T_{in} - T_{ex}) - A_{sol}I_{sol} - \mu_B P_B - \Phi_0 + P_B$$

This particular formulation of the heat balance equation is adapted from Equation 10 from (Chambers & Oreszczyn, 2019). The nomenclature is as follows:

P_{meas} = measured power demand (Watts of fuel)
μ_{HS} = efficiency of heating system
P_{heat} = heat input to dwelling (Watts of heat)
P_f = total power (Watts of fuel)
T_{in} = internal temperature (C)
T_{ex} = external temperature (C)
A_{sol} = effective solar aperture (m^2)
I_{sol} = solar irradiance (W/m^2)
P_B = baseload power (W)
Φ_0 = metabolic gains (W)

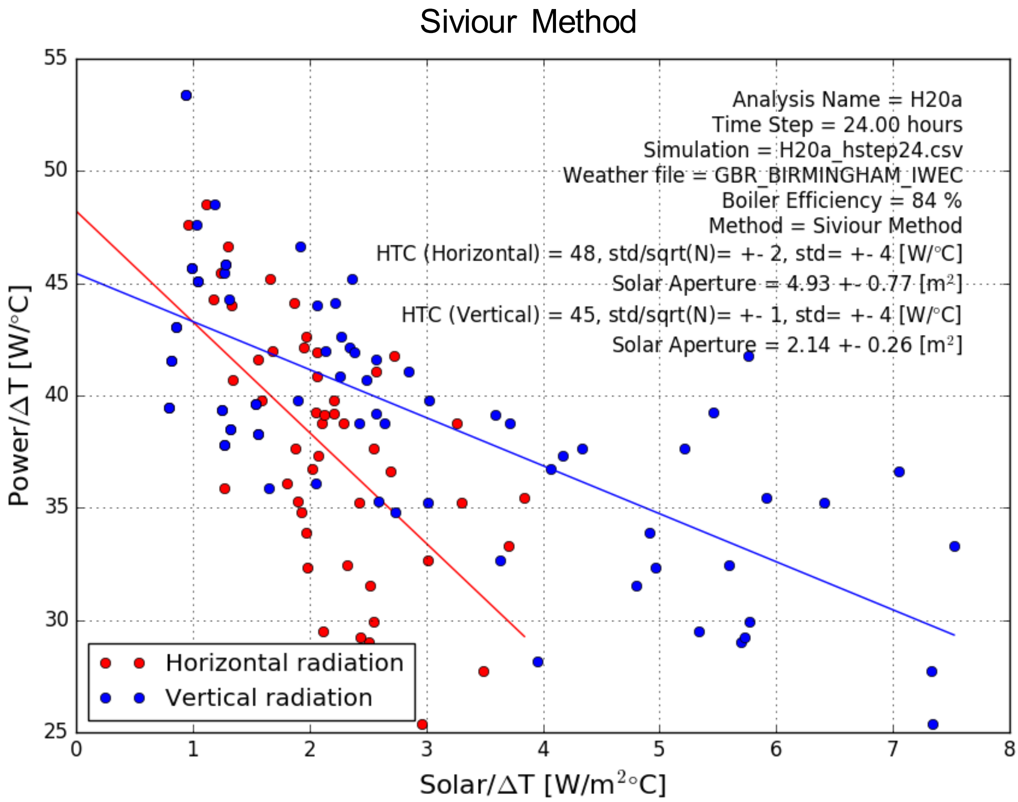
MLR and the Siviour method are regression methods, in which data are fitted to different rearrangements of the above steady state heat balance equation. The data are usually aggregated to at least daily level to retain the steady state assumptions of the equation. Data consist of energy, internal and external temperature and solar gain.

The two methods use very similar rearrangements of the equation and therefore give similar answers, although the different rearrangements of the heat balance equation lead to different treatment of errors (Stamp, 2015).

The Siviour rearrangement of the heat balance equation is given below (adapted from Equation 2.7 (Stamp, 2015)). The HTC is obtained from the intercept of the plot on the y-axis.

$$\frac{P_{heat}}{T_{in} - T_{ex}} = -A_{sol}\frac{I_{sol}}{T_{in} - T_{ex}} + HTC$$

Different values for HTC are obtained depending on how solar irradiation on the dwelling is calculated. The below plot illustrates the difference made by using vertical (south facing) and horizontal direct solar radiation on one of the SMETER archetypes.

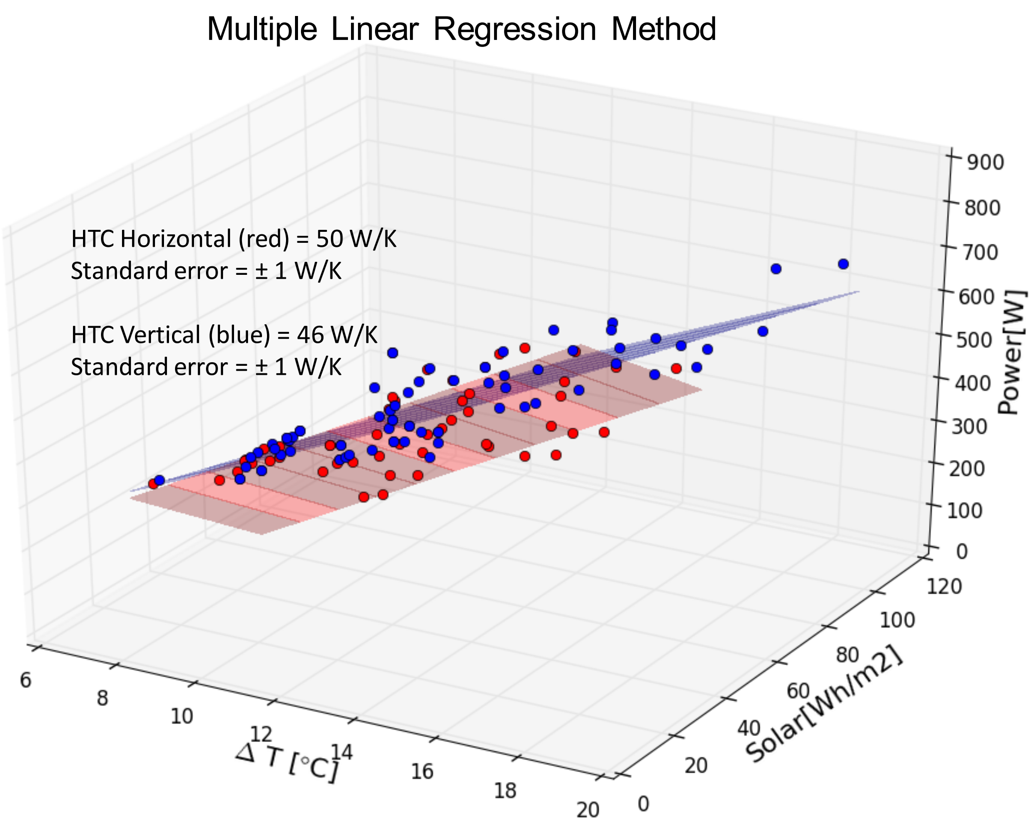


The MLR equation is obtained by multiplying the Siviour equation by $T_{in} - T_{ex}$:

$$\mathrm{P_{heat}} = -A_{sol}\mathrm{I_{sol}} + \mathrm{HTC}(T_{in} - T_{ex})$$

The MLR equation separates out the two effects of solar heat gain and heat loss through the building fabric by treating them as orthogonal variables. This is illustrated below. The HTC is now formulated as one of the regression coefficients, i.e. the partial derivative of power with respect to temperature.

For the data suitability checks carried out on the simulated data using MLR and the Siviour Method, the period 1st January – 28 February was used, and data were aggregated daily.



Parameter uncertainty and confidence intervals

In each of the three example plots above, the HTC is presented with an estimate of fit between the model and the data. This latter quantity is a fitting error, and in all three above examples is calculated as the standard error on the parameter: its standard deviation divided by the square root of the number of data points.

References

Chambers, J. D. & Oreszczyn, T. (2019). Deconstruct: a scalable method of as-built heat power loss coeffient interference for UK dwellings using smart meter data. *Energy and Buildings*, Volume 183, pp. 443-453.

Stamp, S., 2015. Assessing uncertainty in co-heating tests: Calibrating a whole building steady state heat loss measurement method, UCL: PhD Thesis.

Appendix E: TEST 5 installer survey for each SMETER Participant

Respondent name	Gavin Roberts	Date	10 /3 /2020
SMETER Participant	A – no physical product required to be installed		
Is this normally installed by a professional? What professional skills did they require? (e.g. plumbing/electrical/other qualifications, please specify below): No:		Yes ☐ No ☐	

Average duration (time) to install the **entire product** (including router)			00h 00m
Average duration (time) to install the **router only**			00h 00m
Number of people required to install the entire product			0
Hardware item installed	Size (h x w x d) (mm)	Room installed	Sockets used
Total number of hardware items	0	Total number of sockets	0
Identify any hazards e.g. trailing wires			
Wireless communications (yes/no)?			
Wireless infrastructure			
Ease of wireless comms set-up (1-easy, 2-OK, 3-acceptable, 4-difficult, 5-problematic)			
Ease of wireless comms set-up (please write why you chose the above rating)			
Gas had to be turned off (yes/no)?	n	Duration of outage?	00h 00m
Electricity had to be turned off (yes/no?)	n	Duration of outage?	00h 00m
What was the most time-consuming aspect of the installation?			
Identify any problems encountered with the installation			

Identify any potential problems with the installation that you could envisage occurring in an occupied home
List any deviations from the method statement provided by the SMETER participant n/a
Did everything go as planned? Why/Why not?
Once you became experienced in the installation process did it take less time? How much more or less?
List any house types or households where you envisage installation problems?
Please attach example photographs of all equipment in-situ (via email) N/A – no physical device
Any other comments (e.g. anything not covered in the form)

Respondent name	Gavin Roberts	Date	10 /3 /2020
SMETER Participant	B		
Is this normally installed by a professional? What professional skills did they require? (e.g. plumbing/electrical/other qualifications, please specify below): No: A grey box to be connected to a plug socket and placed any where in the house and 4 sensors to be located in specific locations attached via adhesive strips (not supplied)		Yes ☐ No ☐	

Average duration (time) to install the **entire product** (including router)			00h 50m
Average duration (time) to install the **router only**			00h 15m
Number of people required to install the entire product			1
Hardware item installed	Size (h x w x d) (mm)	Room installed	Sockets used
Small sensor	50x30x20	Lounge	0

Small sensor	50x30x20	Bedroom	0
Small sensor	50x30x20	Thermostat	0
Small sensor	50x30x20	Kitchen	0
Comms box- containing router, Modem Hub and 2g extension lead	200x400x250	Suitable location	1
Total number of hardware items	7	Total number of sockets	1

Identify any hazards e.g. trailing wires Location of the comms box was awkward as it’s use of space and sockets

Wireless communications (yes/no)?	yes
Wireless infrastructure	Bluetooth but this changed to RF with the new sensors (I think)
Ease of wireless comms set-up (1-easy, 2-OK, 3-acceptable, 4-difficult, 5-problematic)	2

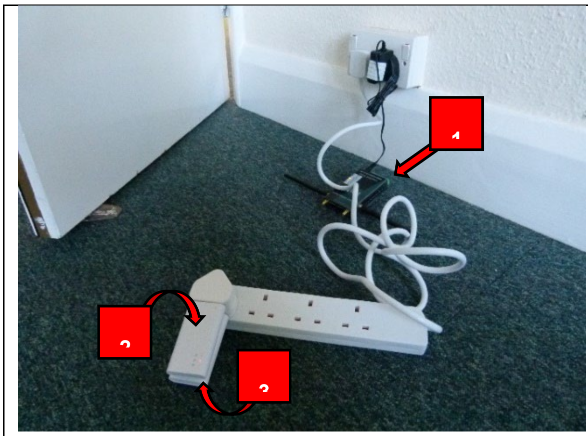
Ease of wireless comms set-up (please write why you chose the above rating) pre-configured

Gas had to be turned off (yes/no)?	n	Duration of outage?	00h 00m
Electricity had to be turned off (yes/no?)	n	Duration of outage?	00h 00m

What was the most time-consuming aspect of the installation? Setting up the grey box at the start but after speaking to the company they started sending the boxes pre-assembled making my task much easier. The property also required a survey and a form returning with orientation, windows area and floor area
Identify any problems encountered with the installation Having to replace the sensors and refusing to install 2 variants of sensors at the start Some sensors were lost by customers due to the non-mechanical fixing method.
Identify any potential problems with the installation that you could envisage occurring in an occupied home Finding somewhere for the comms box.
List any deviations from the method statement provided by the SMETER participant n/a

Did everything go as planned? Why/Why not? As previously stated all the sensors need changing.
Once you became experienced in the installation process did it take less time? How much more or less? Slightly as I got used to the form – about 5 mins less
List any house types or households where you envisage installation problems? Would be suitable for all
Please attach example photographs of all equipment in-situ (via email)

Any other comments (e.g. anything not covered in the form) Initially tried to get us to install double sensors in every location and then requested we changed all the sensors in the trial for a different model due to comms issues.

	1. GSM modem/router 2. USB temperature sensor hub 3. CAD (these were all boxed for installation in homes as shown in the upper photograph of the survey form above)
	Five temperature sensors ready for placement on walls. Black/red wires are sensor aerials.
	USB temperature sensor hub.

Photographs of SMETER B hardware

Respondent name	Gav Roberts	Date	10/ 3 /2020
SMETER Participant	C		
Is this normally installed by a professional? What professional skills did they require? (e.g. plumbing/electrical/other qualifications, please specify below): Yes, requires an electrician and possibly a plumber dependent upon existing radiator valves.		Yes □ No □	

Average duration (time) to install the **entire product** (including router)	2h 00m
Average duration (time) to install the **router only**	20h 00m
Number of people required to install the entire product	

Hardware item installed	Size (h x w x d) (mm)	Room installed	Sockets used
Coms box (router and separate gateway and extension cable)	200x250x400	Suitable location	1
Honeywell Hub	40x60x40	Living room	1
Radiator valve	30x40x60	All main living spaces on radiators	0
Sensor	30x30x30	All main living spaces on radiators	0

Total number of hardware items	10	Total number of sockets	3

Identify any hazards e.g. trailing wires Only the one wire to the comms box

Wireless communications (yes/no)?	y
Wireless infrastructure	Wifi and zigbee
Ease of wireless comms set-up (1-easy, 2-OK, 3-acceptable, 4-difficult, 5-problematic)	2

Ease of wireless comms set-up (please write why you chose the above rating) 2 systems arrived not configured so required setup, this took a lot of trouble shooting as there was a change of staff. If there is a problem then I would suggest a 5

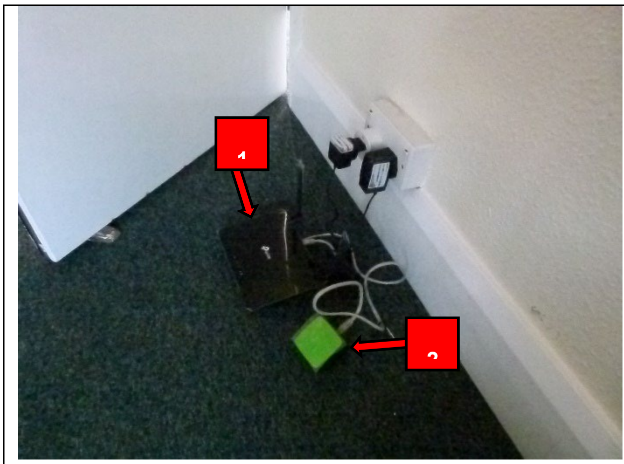
Gas had to be turned off (yes/no)?	n	Duration of outage?	00h 00m
Electricity had to be turned off (yes/no?)	y	Duration of outage?	1h 00m

What was the most time-consuming aspect of the installation? Un boxing of the kit it took ages and then every item required batteries to be installed. There was also a lot of rubbish to dispose of.
Identify any problems encountered with the installation Very fiddly and some of the radiators were not compatible with the Honeywell TRVs meaning that they could not be installed.

Identify any potential problems with the installation that you could envisage occurring in an occupied home Finding somewhere for the comms box- accessing all the radiators. Finding suitable places for the sensors as the brackets are awkward and replacing an existing heating control system
List any deviations from the method statement provided by the SMETER participant none
Did everything go as planned? Why/Why not? No some of the devices lost Wi-Fi coms and had to be reprogrammed, some of the devices arrived mot registered so it meant a lot of trouble shooting.
Once you became experienced in the installation process did it take less time? How much more or less? No difference
List any house types or households where you envisage installation problems? Properties that had no gas central heating
Please attach example photographs of all equipment in-situ (via email) 

Any other comments (e.g. anything not covered in the form)

A floorplan was required before the parts could be sent out and this added delays- it meant that in 2 cases we had to re allocate properties with different SMETERs

	1. GSM modem/router. 2. Temperature sensor hub.
	Temperature sensor (placed on a picture rail in this instance). Mounting bracket not attached in this image.
	Heating and TRV programmer.
	Smart TRV.
	Boiler relay – would usually be wired into the boiler (no socket needed).

Photographs of SMETER C hardware

Respondent name	Gavin Roberts	Date	10 /3/2020
SMETER Participant	D		
Is this normally installed by a professional? What professional skills did they require? (e.g. plumbing/electrical/other qualifications, please specify below): Yes. Required drilling external walls to mount the external sensor		Yes ☐ No ☐	

Average duration (time) to install the **entire product** (including router)			00h 3000m
Average duration (time) to install the **router only**			00h 00m
Number of people required to install the entire product			1
Hardware item installed	Size (h x w x d) (mm)	Room installed	Sockets used
USB sensor	100x20x20	Every room in the house	0
External USB sensor	200x 50x50	external	0
Total number of hardware items	1 more than the number of rooms	Total number of sockets	0

Identify any hazards e.g. trailing wires none

Wireless communications (yes/no)?	no
Wireless infrastructure	
Ease of wireless comms set-up (1-easy, 2-OK, 3-acceptable, 4-difficult, 5-problematic)	1

Ease of wireless comms set-up (please write why you chose the above rating) n/a

Gas had to be turned off (yes/no)?	n	Duration of outage?	00h 00m
Electricity had to be turned off (yes/no?)	n	Duration of outage?	00h 00m

What was the most time-consuming aspect of the installation? Drilling the external wall, but it was a very easy install
Identify any problems encountered with the installation

The items that arrived were not labelled so during the de-install I shall have to label them for return. They do have a constant flashing light at 1 min intervals
Identify any potential problems with the installation that you could envisage occurring in an occupied home. None they would work well on furniture and had a mechanical fixing option
List any deviations from the method statement provided by the SMETER participant non
Did everything go as planned? Why/Why not? Yes although I did screw into place rather than use the command strips provided
Once you became experienced in the installation process did it take less time? How much more or less? No difference
List any house types or households where you envisage installation problems? None
Please attach example photographs of all equipment in-situ (via email) 


Any other comments (e.g. anything not covered in the form) Data is only gathered at the end of the project, once they have been shipped back to the participant.

	Temperature sensor.

Photograph of SMETER D hardware

Respondent name	Gavin Roberts	Date	10 /3 /2020
SMETER Participant	E – no physical product required to be installed		
Is this normally installed by a professional? What professional skills did they require? (e.g. plumbing/electrical/other qualifications, please specify below): No:		Yes ☐ No ☐	

Average duration (time) to install the **entire product** (including router)	00h 00m
Average duration (time) to install the **router only**	00h 00m
Number of people required to install the entire product	0

Hardware item installed	Size (h x w x d) (mm)	Room installed	Sockets used
Total number of hardware items	0	Total number of sockets	0

Identify any hazards e.g. trailing wires	
Wireless communications (yes/no)?	
Wireless infrastructure	
Ease of wireless comms set-up (1-easy, 2-OK, 3-acceptable, 4-difficult, 5-problematic)	
Ease of wireless comms set-up (please write why you chose the above rating)	

Gas had to be turned off (yes/no)?	n	Duration of outage?	00h 00m
Electricity had to be turned off (yes/no?)	n	Duration of outage?	00h 00m

What was the most time-consuming aspect of the installation?
Identify any problems encountered with the installation
Identify any potential problems with the installation that you could envisage occurring in an occupied home
List any deviations from the method statement provided by the SMETER participant n/a

Did everything go as planned? Why/Why not?
Once you became experienced in the installation process did it take less time? How much more or less?
List any house types or households where you envisage installation problems?
Please attach example photographs of all equipment in-situ (via email) N/A – no physical device
Any other comments (e.g. anything not covered in the form)

Respondent name	Gavin Roberts	Date	10/3/2020
SMETER Participant	F		
Is this normally installed by a professional? What professional skills did they require? (e.g. plumbing/electrical/other qualifications, please specify below): No, simple install	Yes ☐ No ☐		

Average duration (time) to install the **entire product** (including router)			00h 45m
Average duration (time) to install the **router only**			00h 00m
Number of people required to install the entire product			1
Hardware item installed	Size (h x w x d) (mm)	Room installed	Sockets used
Sensor	60x20x6	Lounge	
Sensor	60x20x6	Kitchen	
Sensor	60x20x6	Bedroom	
Sensor	60x20x6	landing	
Coms Box (including router , gateway and 2gang power adapter)	200x250x400	Suitable location	1
Total number of hardware items	5	Total number of sockets	1

Identify any hazards e.g. trailing wires

Wireless communications (yes/no)?	y
Wireless infrastructure	Not sure
Ease of wireless comms set-up (1-easy, 2-OK, 3-acceptable, 4-difficult, 5-problematic)	1

Ease of wireless comms set-up (please write why you chose the above rating) Plug and play

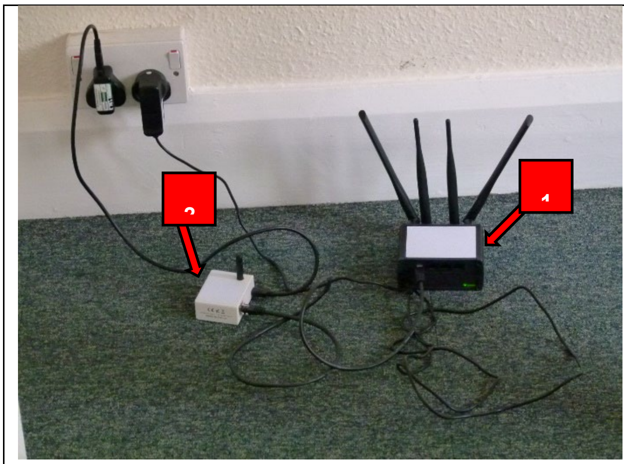
Gas had to be turned off (yes/no)?	n	Duration of outage?	00h 00m
Electricity had to be turned off (yes/no?)	n	Duration of outage?	00h 00m

What was the most time-consuming aspect of the installation? Logging into the web portal uploading a floor plan and then putting the sensors in each of the locations. Although this was straight forward.

Identify any problems encountered with the installation None, the sensors were hung on a single screw
Identify any potential problems with the installation that you could envisage occurring in an occupied home Finding a suitable location for the comms box
List any deviations from the method statement provided by the SMETER participant None
Did everything go as planned? Why/Why not? Yes, clear instructions and a helpful team to guide me through the initial installs
Once you became experienced in the installation process did it take less time? How much more or less? n/a
List any house types or households where you envisage installation problems? non
Please attach example photographs of all equipment in-situ (via email)

Any other comments (e.g. anything not covered in the form)

no

	1. GSM modem/router. 2. Sensor hub.
	Temperature sensors (x3).
	Temperature sensor on wall (stuck in this example, but note the hole for screw).

Photographs of SMETER F hardware

Respondent name	Gavin Roberts	Date	10/3/2020
SMETER Participant	G		
Is this normally installed by a professional? What professional skills did they require? (e.g. plumbing/electrical/other qualifications, please specify below):		Yes ☐ No ☐	

No simple, they are all wall mounted using 2 screws and a raw plug, these could easily be put on a shelf or sideboard	

Average duration (time) to install the **entire product** (including router)			00h 10m
Average duration (time) to install the **router only**			00h 00m
Number of people required to install the entire product			
Hardware item installed	Size (h x w x d) (mm)	Room installed	Sockets used
Sensor	100x100x30	Landing	0
Sensor	100x100x30	Living room	0
Total number of hardware items	2	Total number of sockets	0
Identify any hazards e.g. trailing wires			

Wireless communications (yes/no)?			y
Wireless infrastructure	Sigfox – must have Sigfox signal		
Ease of wireless comms set-up (1-easy, 2-OK, 3-acceptable, 4-difficult, 5-problematic)			1
Ease of wireless comms set-up (please write why you chose the above rating) N/A just a signal test with the company			
Gas had to be turned off (yes/no)?	n	Duration of outage?	00h 00m
Electricity had to be turned off (yes/no?)	n	Duration of outage?	00h 00m
What was the most time-consuming aspect of the installation? Speaking to a representative to check the signal (short amount of time)			
Identify any problems encountered with the installation Although some said they had signal it fluctuates with the time of day, so we are going to have to share our sensor data back with them			
Identify any potential problems with the installation that you could envisage occurring in an occupied home			

None, possibly a slight amount of dust if wall mounted.
List any deviations from the method statement provided by the SMETER participant None
Did everything go as planned? Why/Why not? Yes, it is a simple 2 sensors and a basic floor plan with locations on that needs to be sent to Passiv UK
Once you became experienced in the installation process did it take less time? How much more or less? It took approximately 10 mins to fit the sensors and 10 mins to complete the associated paperwork. The install time did not vary with experience.
List any house types or households where you envisage installation problems? n/a
Please attach example photographs of all equipment in-situ (via email)

Any other comments (e.g. anything not covered in the form)

A document was sent to Passiv UK with the install locations and the boiler type in the property

	Temperature/humidity sensor mounted on wall.

Photograph of SMETER G hardware

Respondent name	Gavin Roberts	Date	10/3/2020
SMETER Participant	H		
Is this normally installed by a professional? What professional skills did they require? (e.g. plumbing/electrical/other qualifications, please specify below): Professional, requires an electrician	Yes ◻ No ◻		

Average duration (time) to install the **entire product** (including router)			1h 30m
Average duration (time) to install the **router only**			00h 00m
Number of people required to install the entire product			2
Hardware item installed	Size (h x w x d) (mm)	Room installed	Sockets used
thermostat	160x70x30	Central location	1
Total number of hardware items	1	Total number of sockets	1
Identify any hazards e.g. trailing wires			
Wireless communications (yes/no)?			y
Wireless infrastructure	GSM		
Ease of wireless comms set-up (1-easy, 2-OK, 3-acceptable, 4-difficult, 5-problematic)			1

Ease of wireless comms set-up (please write why you chose the above rating) plug and play			
Gas had to be turned off (yes/no)?	n	Duration of outage?	00h 00m
Electricity had to be turned off (yes/no?)	y	Duration of outage?	1h 00m
What was the most time-consuming aspect of the installation? Rewiring the boiler to accommodate the thermostat			
Identify any problems encountered with the installation Finding a suitable location for thermostat with a plug socket.			
Identify any potential problems with the installation that you could envisage occurring in an occupied home Finding a location to install it and what to do about decommissioning the existing boiler controls			
List any deviations from the method statement provided by the SMETER participant No the only thing to note was that each device had to be registered, firmware updated then programmed by Switchee before it would work- it took sometimes up to an hour and required a revisit to site to check the device			
Did everything go as planned? Why/Why not? Mostly- a lot of customers seem to have struggled with understanding the device and how it worked. A lot of sites required more than one visit and calls to Switchee directly. I think a lot of our customers are not used to thermostatic control systems.			
Once you became experienced in the installation process did it take less time? How much more or less? Same time as it was completed by an experienced electrician			
List any house types or households where you envisage installation problems? Houses that have no gas central heating system			
Please attach example photographs of all equipment in-situ (via email)			

Any other comments (e.g. anything not covered in the form)

The 3 hour reset to program feature built in to the device has upset a few customers and it nearly had to be removed from one property as the customer just did not understand the device.

	Smart thermostat.

Photograph of SMETER H hardware

Appendix F: TEST 5 Interviews

TEST 5 Householder interview 1

Householder interview 1

House ID number	HH00			Date	DD	/MM	/2020	
SMETER referred to:	A	B	C	D	E	F	G	H

Introduction

"Thank you for participating in this important research project about SMETERs. SMETERs are devices which use smart meter data to measure how energy efficient your house is. We can use the data to help you understand how to reduce your energy bills".

"We will now ask you a few questions about how you have found living with the SMETER in your home".

Demographics

How many people live in your home?	00 Male adults , 00 Female adults and 00 Children (under 18)

	Employed full time	Employed part time	Unemployed	Retired	Education	Other (specify)
Employment status (write numbers of people)						

Household ages (write number of people between each age range below relevant range)	0-4	5-9	10-14	15-19	20-29	30-39	40-49	50-59	60-69	70-79	80+

Does anyone in your household have a long-standing illness, disability or infirmity? (Anything that has troubled you or them over a period of time or that is likely to affect you or them over a period of time)	Yes	No	Don't know	Refused to answer
	Pay on receipt of bill	Direct debit	Pre-pay	Other (please specify)
How do you pay for your electricity?				

How do you pay for your gas?				

Basic householder information

What date did you *permanently* move in?	DD /MM /YYYY
Do you have a working internet connection at home?	Yes, broadband router (e.g. WiFi)
	Yes, Mobile internet (e.g. 3G/4G via mobile phone only)
	Yes, other (please specify)
	No

Do you have a smart meter?	Yes	No	Don't know	Other

Attitudes to technology and energy use

	Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Don't know
I like new technology						
I know how to save energy at home						
I try to save energy at home						
I would dismiss a home if I knew it wasn't energy efficient						
I would dismiss a home if I knew the energy bills were going to be high						
It would help me to budget if I knew in advance what my energy bills were going to be						

Householder behaviour/confounding factors

"We are asking these important questions about the way you live as we believe they will affect how well the SMETERs perform".

Heating and hot water

Do you set your central heating to heat to a particular temperature?	Yes	No	Don't know	Other
If yes to previous question, what temperature is your central heating set to?	°C			
If you have a set central heating temperature, do you ever change the heating temperature?	Yes	No	Don't know	Other

How often do you change the central heating temperature?	Daily	Weekly	Monthly	Season-ally	Yearly	Never	Other

What is your usual weekday central heating schedule?	Please list on/off times

Is your central heating schedule different at weekends?	Yes	No	Don't know	Other

If yes, what is your usual weekend central heating schedule?	Please list on/off times						
Do you periodically or ever change your heating schedule?	Yes	No	Don't know	Other			
How often do you change the central heating schedule?	Daily	Weekly	Monthl y	Season -ally	Yearly	Never	Other
Do you ever change the TRVs temperature? *Photo of TRV in MS Forms version.*	Yes	No	Don't know	Other			
How often do you change your TRVs temperature?	Daily	Weekly	Monthly	Season-ally	Yearly	Never	Don't know
Do you ever turn off TRVs completely?	Yes	No	Don't know	Other			
If yes, in which rooms do you turn off TRVs completely?							
Do you use any **fixed** secondary heating such as a gas fire or electric heaters?	Yes	No	Don't know	Other			
If so, what rooms and what type/duration of heating, estimated Wattage?							
Do you use any **portable** secondary heating such as a gas fire or electric heaters? If so, what rooms and what type/duration of heating?	Yes	No	Don't know	Other			
If so, what rooms and what type/duration of heating, estimated Wattage?							
Laundry							
Do you have a washing machine?	Yes	No	Don't know	Other			
If so, how many times per week do you use it?							
How do you usually dry your clothes **in winter**?	Tumble dryer	Airer/rack/ clothes horse	Directly on radiator	Outside	Other (please specify)		
Do you have a tumble dryer? If so, how many times per week do you use it?							
Do you use the tumble dryer the whole year or just in the winter?	Whole year	Seasonal ly, winter only	Seasonal ly,	Seasonal ly, winter/au	Other (please specify)		

			winter/autumn	tumn/spring	
Dishwashing					
Do you have a dishwasher or wash by hand?	Wash by hand	Dishwasher	Both dishwasher and washing by hand	Other	
If dishwasher, how many times per week do you use it?					
Personal hygiene					
How many showers per day in this household?	Per week				
How many baths per day in this household?	Per week				
Cooking					
How many meals do you **cook** at home per week (e.g. something you would use to oven or hob for)?	*1 household meal = a meal for the entire household, not per person.*				
Gas or electric hob?	Gas	Electric	Other (please specify)		
Gas or electric oven/grill?	Gas	Electric	Other (please specify)		
Other appliances and electricity use					
How many televisions/computers in the house?					
How many hours per day is the television/computer(s) on?	Enter a per DAY value as a whole number only (e.g. 1) for EACH television/computer, (e.g. "6" for 2 televisions x 3 hours on per day).				
Any uses of electricity outside of the main house?	1. In the garage/workshop 2. Outdoor lighting (how much?) 3. Charging car batteries (e.g. flat starter battery) 4. Electric vehicle 5. Other (please specify)				
Windows and ventilation					
Do you open the windows in the autumn/winter?	Yes	No	Other		
Is this a regular window opening or random?	Regular	Random	Other		
Why do you open your windows?	Please write why. E.g. to get cooking smells out / or randomly when hot				

In which rooms do you usually open windows in the autumn/winter?	Living room	Kitchen	Bathroo m	Adult bedroo m	Child bedroo m	Other room	
On average, for how long per day in the autumn/winter do you usually open the windows in each room?	Please enter a value per DAY in hours as a number without units, e.g. "Living room 0.5" for half an hour in the living room, or "Kitchen 1; Bathroom 0.5" for 2 hours in the kitchen and 30 mins in the bathroom.						
Do you have trickle vents on your windows?	Yes	No	Don't know	Other			
Do you adjust (open and close) the trickle vents?	Yes	No	Don't know	Other			
How often do you adjust your trickle vents?	Daily	Weekly	Monthly	Season al-ally	Yearly	Never	Other
In which rooms do you adjust trickle vents in?	List all rooms trickle vents adjusted in						
Are trickle vents usually open or usually closed – which rooms?	Please write the specific rooms and whether they are usually open or closed.						
Do you *have* an extractor fan in the kitchen?	Yes	No	Don't know	Other			
Do you *use* an extractor fan in the kitchen?	Yes	No	Don't know	Other			
Do you use extractors fans in the kitchen (e.g. only when cooking, all the time, never)?	Continuously on	Only when cooking	Never	Other (please specify)			
Do you *have* an extractor fan in the bathroom?	Yes	No	Don't know	Other			
Do you *use* an extractor fan in the bathroom?	Yes	No	Don't know	Other			
Do you use extractors fans in the bathroom (e.g. only when washing, all the time, never)?	Continuously on	Only when washing	Never	Other (please specify)			
Lifestyle							

<table>
<tr><td>Did you have any long periods away from home since you moved into the property? (Or any planned in the future?)</td><td colspan="7">Please specify how many people (inc. gender/age), when, how long. Could be holiday/hospital stay/away with work etc. E.g. more than a few days away.</td></tr>
<tr><td>Has anyone recently moved out or expected to move in?</td><td colspan="4">Yes (how many/gender/age?)</td><td colspan="3">No</td></tr>
<tr><td>Do you have any pets? (please specify)</td><td colspan="7"></td></tr>
<tr><td colspan="8">Installation (hypothetical based on experience of other installs)</td></tr>
<tr><td>What are the important issues for you when someone is in your home installing something?</td><td colspan="7">(Prompts if needed: e.g. time spent in the house, rooms entered, mess/dust/drilling, anything else? etc.)</td></tr>
<tr><td colspan="8">Perceptions of SMETER</td></tr>
<tr><td>What do you think about the appearance and positioning of the SMETER device? Specifically what do you LIKE and DISLIKE about the SMETER?</td><td colspan="7">Try and determine which SMETER they are talking about. Prompts if needed: e.g. it looks discreet, too big, flashing lights, I like/don't like the colour, well/poorly positioned, it makes a noise. Ask "anything else?"</td></tr>
<tr><td>How long would you be willing to have a SMETER device in your home?</td><td>1 day</td><td>1 week</td><td>1 month</td><td>6 months</td><td>1 year</td><td>Forever</td><td>Other</td></tr>
<tr><td colspan="8">Interactions with the SMETER</td></tr>
<tr><td>Did it alter/affect your daily life in any way?</td><td colspan="7"></td></tr>
<tr><td>Did the SMETER affect your use of plug sockets?</td><td colspan="7"></td></tr>
<tr><td>Did you received advice on how to use the SMETER?</td><td colspan="7"></td></tr>
<tr><td>Did the SMETER make you change your behaviour in any way? E.g. turn the heating down, switch off lights.</td><td colspan="7"></td></tr>
<tr><td>Did you receive advice/instructions on how to use your SMETER?</td><td colspan="7"></td></tr>
</table>

TEST 5 Householder interview 2

Householder interview 2								
House ID number	HH00		Date	DD	/MM	/2020		
SMETER referred to:	A	B	C	D	E	F	G	H

Demographics		
Has anyone moved in/out since the last interview, or are they expected to do so soon?	Yes	No
Has anyone moved in/out since the last interview, or are they expected to do so soon?	00 Male adults , 00 Female adults and 00 Children (under 18) and date moved out: DD/MM/YY	

Thinking about the COVID-19 restrictions, have you changed the amount of time you are at home compared to the same time last year?	At home much less	At home slightly less	No change from last year	At home slightly more	At home much more
During the full lockdown period (23 March to 13 May)					
Since 13 May, as lockdown restrictions ease					

Have you turned your boiler off at any time, for any reason?	Yes	No
When did you turn your boiler off?	Specify date from when the boiler was turned off, to the date turned back on again	
Have you made any changes to the structure/insulation/windows/doors/ etc. of the property since the last interview?	Yes	No
What changes have you made and when did you make them?		

Are you interested and able to improve the thermal performance of your home?	Yes, I'm interested and able to improve my home	Yes, I'm interested, but not able to improve my home	Yes, I'm able to improve my home, but I'm not interested	No

Now that you have lived with a SMETER for a few months, what do you like/dislike about the SMETER appearance/size/position/functionality?			
What could be done to improve the SMETER with respect to appearance/size/position/functionality?			
How much did you notice the sensors on the wall? Do they bother you?	Yes I noticed them, but I'm not bothered by them	Yes I noticed them, and I am bothered by them	No, I didn't notice them
Which sensors bothered you and why?			
Have you switched the SMETER device off at all?	Yes	No	Don't know
If so, why?			

Appendix G: Evaluation of alternative building performance evaluation methods QUB and Pulse

This appendix describes the two additional tests that were carried out in each of the TEST Project field trial homes for comparison with the standard methods: QUB, a shorter duration alternative to the co-heating test; and Pulse, an alternative to the blower door test which is carried out at lower pressure differences using compressed air.

G1 Secondary HTC measurement – QUB test

Introduction

QUB tests were carried out by a TEST team researcher from LBU using three different prototype QUB kits supplied by Saint Gobain (Table G1).

Table G1: Details of QUB tests

Home ID	QUB 1 date	Vents	QUB 2 date	Vents	QUB 3 date	Vents	QUB 4 date	Vents	QUB 5 date	Vents	QUB 6 date	Vents	QUB 7 date	Vents
HH01	05/04/19	O	06/04/19	O	07/04/19	O								
HH02	03/05/19	O	04/05/19	O	05/05/19	O	06/05/19	O						
HH03	17/05/19	O	18/05/19	O	19/05/19	O								
HH04	17/09/19	C	20/09/19	O	21/09/19	O	22/09/20	O						
HH05	30/09/19	C	01/10/19	C	02/10/19	C	03/10/19	O	04/10/19	O	05/10/19	O	06/10/19	O
HH06	23/09/19	O	24/09/19	O	26/09/19	C	27/09/19	C	28/09/19	C	29/09/20	C		
HH07	09/10/19	C	10/10/19	C	11/10/19	C	12/10/19	C	13/10/19	C				
HH08	07/10/19	C	08/10/19	C										
HH09	18/10/19	O	19/10/19	O	20/10/19	O	21/10/19	C	22/10/19	C	23/10/19	C		
HH10	28/10/19	C	29/10/19	C	30/10/19	C	31/10/19	C						
HH11	05/11/19	C	06/11/19	C	07/11/19	C								
HH12	08/11/19	O	09/11/19	O	10/11/19	O								
HH13	18/11/19	C	19/11/19	C	20/11/19	C	21/11/19	C						
HH14	22/11/19	C	23/11/19	C	24/11/19	C								
HH15	18/11/19	C	19/11/19	C	20/11/19	C	21/11/19	C						
HH16	26/11/19	C	27/11/19	C	28/11/19	C								
HH17	29/11/19	C	30/11/19	C	01/12/19	C								
HH18	03/12/19	C	04/12/19	C	05/12/19	C								
HH19	29/11/19	C	30/11/19	C	01/12/19	C								
HH20	06/12/19	C	07/12/19	C	08/12/19	C	09/12/19	C						
HH21	10/12/19	C	11/12/19	C	12/12/19	C								
HH22	10/12/19	C	11/12/19	C										
HH23	16/12/19	C	17/12/19	C	18/12/19	C								
HH24	16/12/19	C	17/12/19	C										
HH25	19/12/19	C	20/12/19	C	21/12/19	C	22/12/19	C	23/12/19	C				
HH26	19/12/19	C	20/12/19	C	21/12/19	C	22/12/19	C	23/12/19	C				
HH27	07/01/20	F	08/01/20	F	09/01/20	F								
HH28	06/01/20	F	07/01/20	F	08/01/20	F	09/01/20	F						
HH29	18/02/20	C	19/02/20	C	20/02/20	C								
HH30	18/02/20	C	19/02/20	C	20/02/20	C								

C = Closed (vents)

O = Open (vents)

F = Failed test, HH27 and HH28 failed all tests and no QUB result is available for these homes.

The QUB test method, developed by Saint-Gobain (Alzetto et al., 2018a), is a test method capable of measuring the HTC of a dwelling in one night. It achieves this by heating the property under dynamic conditions and treating the property as a simple Resistor-Capacitor (RC) system (see Figure G1). The internal temperature (T_{in}) and external temperature (T_{out}) are measured, along with the heating power input into the property (P(t)). By conducting the test overnight, little power will be input from the sun and the input heating power can therefore be closely controlled. Established equations which govern electrical circuits can then be used to determine the values for the thermal resistance and thermal capacitance of the property.

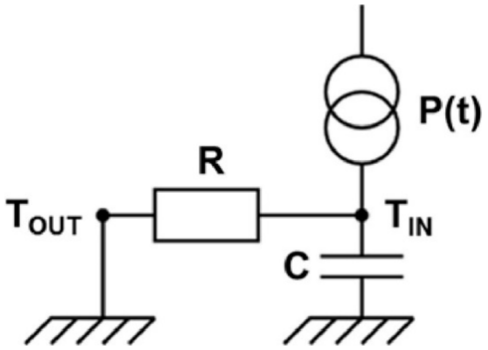


Figure G1: RC system used in the QUB test. T_{in} and T_{out} are the measured internal and external temperatures. P(t) is the measured heating power. R is the thermal resistance of the property which is used to calculate the HTC. C is the thermal capacitance of the property.

To ensure that this RC approximation is as valid as possible, the QUB kit aims to create a homogenous internal temperature throughout the property by appropriate location of multiple small, low-inertia electrical heaters. The number of heaters used varies depending on the property size, and these heaters are placed in such a way as to uniformly heat the space. Although the external temperature cannot be controlled, it is measured in 2 locations (ground level and top-floor level), to ensure an accurate representation of the external environment is obtained.

When the experiment is set up, a controller unit is programmed to heat the property for half of the night, and let it cool for the remaining half (see Figure G2). All temperatures and powers are recorded by the controller unit, which also performs an automated analysis to generate the value for HTC. The QUB algorithm further calculates an “alpha” parameter, which can be used to assess the accuracy of the QUB test. The alpha parameter is calculated from the equation.

$$\alpha = 1 - \frac{HLC_{ref}\Delta T_0}{P_1}$$

Where HLC_{ref} is a reference HLC for the property (often calculated theoretically), ΔT_0 is the initial temperature difference between internal and external environments, and P_1 is the power input into the property during the heating phase. It has been shown that the QUB test is most successful for alpha parameters in the range 0.4 to 0.7 (Pandraud et al., 2014; Meulemans et al., 2016).

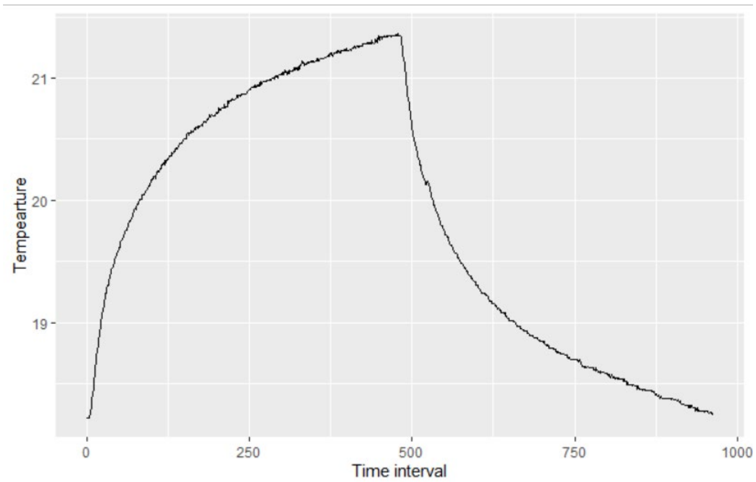


Figure G2: Example an internal air temperature profile during a QUB test. The property is heated for half the test period, and left to cool for the remaining half.

Previous work has shown that the HLC from the QUB test is in reasonable agreement with that from the co-heating test when both are conducted in controlled conditions (Alzetto et al. 2018a, Meulemans et al. 2017). However, little work has been done to compare co-heating and QUB on real dwellings (Alzetto et al., 2018b, Sougkakis et al., 2018). Furthermore, previous work suggests that poorer performing properties are more likely to cause disagreement between QUB and co-heating (Alzetto et al., 2018b). This disagreement is due, at least in part, to the higher convective losses which occur during the co-heating test (Meulemans et al. 2017). The work being conducted as part of SMETER presented an opportunity to conduct a comparison between co-heating and QUB at scale, as well as deploy QUB in properties with poorer thermal performance. QUB kits where therefore deployed in all SMETER properties, so that QUB HTCs could be compared to the HTCs from the co-heating test.

Method

QUB kits were deployed in all 30 of the SMETER properties. Typically, the test was set up immediately after the co-heating test and using the same building conditions - If the building was coheated with vents open, the QUB kit was deployed with the vents open. The number of heaters used was determined with the assistance of the QUB kit controller unit. This controller unit can take as input the floor area and external temperature forecast, and use these to determine the predicted number of required heaters.

Heaters were placed in similar locations to those during the co-heating test. Experimenters exercised their judgement on where the heaters should be placed to ensure a homogenous temperature. The internal temperature sensors were placed on tripods in the middle of the rooms. The external temperature sensors were placed first in weatherproof boxes. One external temperature sensor was then placed on the ground near the property. The second external temperature sensor was affixed to the external wall on the highest level of the property.

The QUB tests were typically run for 3 consecutive nights to test consistency between individual QUB results. Upon completion of the 3-night period, it became apparent that several of the QUB tests failed to complete. Some of these failures were due to a software error, and these tests were therefore discarded. Other failures occurred due to sensors dropping out. When a sensor failure occurred, it was sometimes possible to use other sensors to compensate. If this was not possible, the tests in which the sensors failed were also discarded. In total, 26 of the 30 properties in SMETER had QUB tests which completed and reported a value for the HTC.

Results

When all tests were complete, Saint-Gobain took the raw data from the QUB kits for analysis. The results for each completed QUB test are listed in table 1 alongside the result of the co-heating test. If the errors bars of a co-heating test HTC and QUB test HTC overlap, we cannot detect a statistical difference between them. We colour these values in table 1 as green. It was noted by the experts at Saint-Gobain that some data was of poor quality and is likely resulting in incorrect HTCs. For the purposes of this report, we nonetheless include these tests to study the reasons behind the possible bad data.

Table G2: Results from the completed QUB tests. Note that H5 and H7 did not have a co-heating test.

Home ID	Co-heating	QUB1	QUB2	QUB3	QUB4	QUB5	QUB6
HH01	183.9 ± 11.2	169.1 ± 31.8	150.4 ± 46.8	180.1 ± 52.9			
HH02	133 ± 15.4	90.4 ± 2.2	88.1 ± 1.6	87.9 ± 1.8	89.2 ± 1.6		
HH03	159.2 ± 5.1	118.7 ± 7.1	114.1 ± 12.2	119 ± 6.5			
HH04	149.7 ± 26.7	157.6 ± 22.7	151.9 ± 28	165.8 ± 41.4			
HH05	*	186 ± 16.3	208.7 ± 30.0	198.2 ± 9.8	230.4 ± 20.8		
HH06	204.3 ± 17.8	212 ± 11.6	193.4 ± 4.9	195.6 ± 8.6	192.7 ± 6.3	234.8 ± 11.4	233.2 ± 13.7
HH07	*	140.2 ± 12.3	182.2 ± 19.8	153.2 ± 13.3	139.5 ± 14.1	137.9 ± 12.0	
HH08	197.9 ± 13.2	228.2 ± 9.8	214.3 ± 17.8				
HH09	150.1 ± 6.2	144.4 ± 3.5	123.3 ± 3.1	125.2 ± 3.7	129.8 ± 3.3	140.2 ± 3.9	123.8 ± 4.7
HH10	194.5 ± 3.8	165.9 ± 9.8	184.3 ± 6.3	173.7 ± 9	169.6 ± 8.2		
HH11	214 ± 13.9	168.9 ± 4.5	159.4 ± 1.8	176.3 ± 6.3			
HH12	122.7 ± 6.9	113.9 ± 11.2	113.7 ± 6.1	128.3 ± 14.3			
HH13	157.5 ± 4.7	105.3 ± 7.6	105.3 ± 6.1	100.8 ± 2.9	98.9 ± 2.7		
HH14	133.5 ± 6.9	124.4 ± 4.9	125.2 ± 7.8	128.2 ± 8.2			
HH16	187.6 ± 8.9	137.4 ± 10	133.2 ± 9.4	129.1 ± 17.1			
HH17	186.1 ± 14.8	147.5 ± 9.6	137.8 ± 5.5	134.9 ± 9.8			
HH18	145.4 ± 7.7	112.5 ± 2.7	109.6 ± 2	138.6 ± 9			
HH19	148.7 ± 8.7	125.4 ± 4.7	135 ± 3.3	120.8 ± 4.1			
HH20	222.1 ± 13.8	210.2 ± 25.7	224.5 ± 13.9	194.9 ± 15.5	176.4 ± 13.5		
HH21	260 ± 7.8	186.3 ± 9.2	156.7 ± 9	168.9 ± 10.6			
HH22	189.8 ± 10.6	137.8 ± 2.4	149.6 ± 4.7				
HH23	231.4 ± 19	175 ± 2.9	148 ± 2.4	158.4 ± 6.7			
HH24	166.3 ± 7.8	117.7 ± 11.2	128.2 ± 7.8				
HH25	239.1 ± 15.3	188.9 ± 3.7	181.1 ± 2.9	193.1 ± 6.7	192.3 ± 4.1	173.9 ± 4.5	
HH26	242.5 ± 11.8	181 ± 14.7	172.1 ± 15.7	171.3 ± 18.4	187.6 ± 9.4	170.8 ± 12.3	
HH30	221.9 ± 17.4	163.8 ± 5.9	160.9 ± 8.8	146.1 ± 6.5			

Comparison of QUB and Co-heating.

One noticeable feature of the data in Table G2, is that the HTC from the QUB test is almost always lower than that derived from the co-heating test. To illustrate this, Figure G3 shows the QUB HTCs plotted as a function of the co-heating HTCs. The solid line denotes y=x and the majority of the points lie below this line or have error bars which extend below the line. Meulemans et al. (2017) also found this pattern, and posited that it is due to increased convective heat loss during the co-heating test. The co-heating test employs large heaters and fans to circulate warm air, which may also force air from the building. The QUB test meanwhile uses small, low-inertia heaters which should cause less air exchange. This problem will be worse in properties with lower air-tightness values.

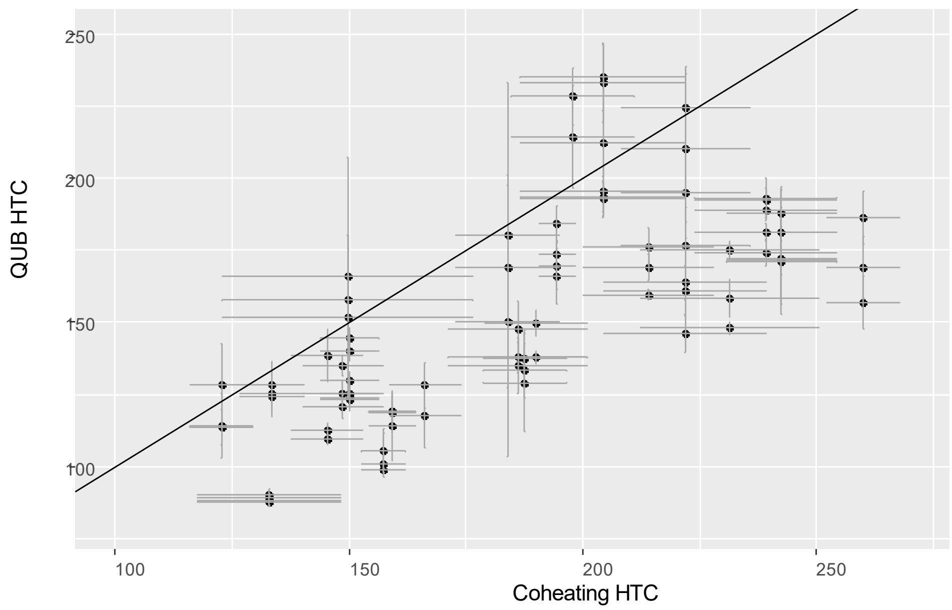


Figure G3: Comparison of co-heating and QUB HTCs

Another notable feature of the data in Table G2, is that no QUB tests from H21 onwards had agreement with the co-heating test. These tests correspond to those completed in December 2019 onwards. In this month, multiple QUB tests were being completed in a short space of time. As the set-up procedure of the QUB test is dependent on the field-testers judgement, it is possible that time pressure caused this set-up quality to decrease. A standard operating procedure (SOP) is being prepared by Saint-Gobain to ensure future consistency of QUB set-up and avoid this issue above.

Predicting HTCs for H5 and H7

Houses 5 and 7 had no co-heating test. In isolation, it is difficult to determine if these QUB tests are reliable or not. However, the GBT algorithm has the potential to identify reliable tests. The predictions from the algorithm are shown in Table G3. The algorithm predicts that 1 of the tests for house 5 is reliable, but none of the tests for houses 7 were correct. However, it is probable that at least one of the tests for house 7 are acceptable, but currently the recall for the algorithm is not high enough to identify them.

Table G3: Assessment of houses 5 and 7 from the gradient boosted tree

Home ID	Coheating	QUB1	QUB2	QUB3	QUB4	QUB5	QUB6
HH05		186 ± 16.3	208.7 ± 30.0	198.2 ± 9.8	230.4 ± 20.8		
HH07		140.2 ± 12.3	182.2 ± 19.8	153.2 ± 13.3	139.5 ± 14.1	137.9 ± 12.0	

Conclusion

It has previously been shown that the QUB test is effective in buildings with good thermal performance. In this work, the QUB method was deployed on homes with poorer thermal performance, yet many of the QUB tests were found to be in agreement with the co-heating test. Of the QUB tests which disagree with co-heating, analysis of the data suggests that this is partially due to inhomogeneous heating of the property. This inhomogeneous heating of the property may be due to a poor experimental setup, in which case greater training and guidance could alleviate this problem. This inhomogeneous heating could also be due to the poorer thermal performance of the building, where heat loss from some areas of the property could well exceed others. This is a harder problem to solve, but inclusion of heat flux data into the QUB analysis may help account for the areas where heat loss is greater.

Encouragingly, it seems possible to tell if a QUB test has been successful purely from the QUB data gathered. This may allow QUB tests to be used in place of co-heating tests in certain circumstances. However, 30% of the tests which the algorithm judges as "incorrect" are in fact correct. There is therefore considerable room for improvement in this algorithm in parallel to the improvements in the QUB test set-up.

References

Pandraud, G., Gossard, D. and Alzetto, F. (2014) Experimental optimization of the QUB method. In: IEA–EBC Annex 58, 6th Expert meeting, 14–16 April 2014, Ghent (Belgium)

J. Meulemans, F. Alzetto, D. Farmer and C. Gorse (2017). QUB/e – A novel transient experimental method for in situ measurements of the thermal performance of building fabrics. In: Gorse, C. and Dastbaz, M. (eds), *Building Information Modelling, Building Performance, Design and Smart Construction*, Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-319-50346-2_9

V. Sougkakis, J. Meulemans, F. Alzetto, C. Wood, M. Gillott and T. Cox (2017). An assessment of the QUB method for predicting the whole building thermal performance under actual operating conditions. In: Gorse, C. (Ed.) *International Sustainable Ecological Engineering Design for Society (SEEDS) Conference 2017 –* Conference Proceedings, 13 – 14 September 2017, Leeds (UK). Leeds Beckett University, Leeds, UK.

F. Alzetto, D. Farmer, R. Fitton, T. Hughes and W. Swan (2018). Comparison of whole house heat loss test methods under controlled conditions in six distinct retrofit scenarios. *Energy and Buildings*, vol. 168, pp 35 – 41. https://doi.org/10.1016/j.enbuild.2018.03.024

F. Alzetto, G. Pandraud, R. Fitton, I. Heusler and H. Sinnebichler (2018). QUB: a fast dynamic method for in-situ measurement of the whole building heat loss. *Energy and Buildings*, vol. 174, pp 124 – 133. https://doi.org/10.1016/j.enbuild.2018.06.002

V. Sougkakis, C. Wood, J. Meulemans, M. Gillott, and T. Cox. QUB/e: Validation of a transient method for determining whole building thermal performance and building element U-values in situ under actual operating conditions. In S. Riffat, Y. Su, D. Liu, and Y. Zhang, editors, *The 17th International Conference on Sustainable Energy Technologies – SET 2018*, volume 3, pages 174–84, Wuhan, China, 21–23 August 2018. University of Nottingham. https://nottingham-repository.worktribe.com/output/1607959

F. Alzetto, J. Meulemans, G. Pandraud and D. Roux (2018). A perturbation method to estimate building thermal performance. *Comptes Rendus Chimie*, vol. 21 (10), pp 938 – 942. https://doi.org/10.1016/j.crci.2018.09.003

J. Meulemans (2019). An assessment of the QUB/e method for fast in situ measurements of the thermal performance of building fabrics in cold climates. In: Johansson D., Bagge H., Wahlström Å. (eds) *Cold Climate HVAC 2018*, Springer Proceedings in Energy. Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-030-00662-4_27

G2 Pulse test

Introduction

Pulse tests were carried out by a TEST team researcher from LBU (Table G4) using the BTS Model 1. The Pulse test (Cooper, et al., 2019) is a recently developed novel airtightness test method which measures the air leakage of a house at a pressure differential which can be considered more typical of a house under natural conditions (4 Pa), rather than that experienced during an air pressurisation test (50 Pa). Thus, n4 measured by a Pulse test can be considered a more appropriate value of the air change rate than the air pressurisation test derived n50/20 value using the Kronvall Persily rule (Jones, et al., 2016).

The Pulse method has been proven to provide consistent results for new build UK housing (Cooper et al., 2016; Zheng et al., 2019) with purpose provided ventilation either closed or temporarily sealed. However, in this project the TEST team were investigating airtightness characteristics of existing dwellings which are inherently less airtight than new build, and under a range of ventilation strategies – both with and without some or all the purpose provided ventilation open. Under these less stable conditions the Pulse method appeared to be somewhat less reliable, suffering more from adverse environmental conditions and peculiarities of some more unpredictable building elements.

Acoustic effects have always been suspected to be an issue with pulse testing. The pulse theory assumes that at each time interval measured, the pressure is consistent throughout the entire test property. Early research concentrated on single room testing to avoid such issues (Nishioka, 2000), Carey and Etheridge (2001) suggested using simultaneous release systems to counter for this effect in larger buildings. Whilst airflow around dwellings during a pulse test is not assumed to be an issue in new-build or existing housing, there are concerns about turbulent airflow into buffer zones between the habitable space and the external environment (knee wall voids, floor and sub-floor voids, cellars, etc.) causing acoustic interference and resonance. Most modern buildings will have a designed air barrier on the inner surface of the envelope, this is often not the case with older housing where pressurisation and depressurisation of these buffer zones may result in non-uniform pressures being experienced within the test dwelling during the measurement period.

Table G4: Details of Pulse tests

Home ID	Test date	Number of successful tests		
		All vents open	All vents closed	Trickle vents open only
HH01	08/04/19	3	1	1
HH02	07/05/19	1	1	1
HH03	17/05/19	1	1	1
HH04	01~03/10/19	1	1	1
HH05	03/10/19	1	1	1
HH06	18/10/19	1	1	1
HH07	18/10/19	1	1	1
HH08	21/10/19	1	1	2
HH09	18/10/19	1	1	1
HH10	24~28/10/19	1	1	0
HH11	08/11/19	1	1	0
HH12	11/11/19	3	3	0
HH13	15/11/19	1	1	1
HH14	22/11/19	1	1	0
HH15	18/11/19	2	1	1
HH16	26/11/19	1	1	1
HH17	29/11/19	1*	1	1
HH18	06/12/19	1	1	1
HH19	29/11/19	1	1	1
HH20	13/12/19	1*	1	0
HH21	13/12/19	0*	1	0
HH22	13/12/19	1	2	1
HH23	19/12/19	1	1	0
HH24	16/12/19	1	1	1
HH25	13/01/20	1	1	0
HH26	19/12/19	1	1	0
HH27	10/01/20	1	1	0
HH28	10/01/20	1	1	0
HH29	18/02/20	1	1	0
HH30	18/02/20	1	1	0
Total		34	33	18

* Indicates that these results may not be reliable as the air leakage with vents open appeared to be too great for the Pulse system, although the Pulse system listed them as successful tests in 2 of the 3 cases.

In this review, comparisons are made between the blower tests at 50 Pa and Pulse methodology at 4 Pa, and while the relationships reported are not erroneous, the Pulse and Blower door operating at these different pressures distort the building fabric in different ways resulting in different fabric behaviour. At higher pressures, typical of those experience with blower doors, windows, doors, hatches, and seals may be pushed open (under pressurisation) and closed (when depressurised) whereas at lower pressures, such as those experienced with Pulse, the changes in the fabric will be considerably less. The lower pressures experienced with Pulse, do not tend to open windows, doors, seals or lift moveable fabrics (such as floorcoverings). Thus, while conversion factors do suggest a non-linear relationship, the Pulse and Blower door test measure the air permeability of the building in different physical states.

The Pulse test operates at pressure differentials a house would experience under normal conditions, while the blower door tests induces pressures far in excess of those that would be experienced even on very windy days. As the Blower door test exerts considerably greater air pressures, it is also more stable when the air permeability of the building is high (the building is leaky) and the velocity of air flowing into or out of the building is high.

Method

The Pulse tests were conducted using a 3-step Pulse methodology, where 3 pulses of air are emitted in sequence and their pressurisation-relaxation rates analysed. Under favourable conditions, and using the most appropriate tank size, the Pulse method provided 3 "good" steps and satisfactory results. Where less favourable conditions were encountered the results often had to be determined from just 2 steps, where only 1 step was deemed "good" the result was not recorded.

Observations during fieldwork

Using a 3-step Pulse methodology, the Pulse test regularly failed to reproduce all 3 "good" steps in less airtight dwellings and those with purpose provided ventilation open. Strong or gusty winds appeared to create issues with Pulse test reliability and repeatability, most noticeably in some of the less airtight scenarios. Leaky and less rigid suspended timber ground floors also appeared to proffer consistency issues, where pressurisation of the void/cellar below appeared to produce a "bounce-back" acoustic effect, an effect only observed in uncarpeted properties and diminished by mounting the Pulse unit in a different area of the dwelling.

Whilst performing multiple tests in higher humidity dwellings, the air release valve at the top of the Pulse unit was susceptible to icing up for the 2nd and/or 3rd pulse step, particularly when the test house was unheated.

Results

Both Pulse and Blower Door methodologies were observed to provide reliable and repeatable results in the more airtight dwellings tested under favourable environmental conditions. The observation that strong or gusty winds appeared to create issues with reliability and repeatability of the Pulse test in some of the less airtight scenarios was not an issue for the Blower Door method - where higher pressure differentials could be used in strong winds, and on gusty days readings were only taken between gusts[42].

A distinct advantage of the Blower Door method was the ability of the testers to get a feel for where the majority of air leakage was occurring from, even when no formal leakage detection was undertaken. Whilst performing the blower door tests the TEST team were aware of significant airflows through observations such as carpets lifting or whistling noises, these were not apparent under the pulse method. Also, thermal imaging under depressurisation (when a

[42] Measurement was carried out after a gust had subsided, based on observation of surroundings e.g. movement of trees. If a gust interrupted the measurement, then the measurement was repeated.

sufficient temperature differential existed) proved to be a useful diagnostic tool under blower door test conditions, allowing the TEST team to distinguish whether thermal anomalies identified previously were due purely to construction or moisture issues or were due to air movement within the structure.

Performing multiple tests in higher humidity dwellings also proved more difficult using Pulse than with Blower Door. The air release at the top of the Pulse unit was susceptible to icing up, particularly when the test house was unheated. Having to use a fan heater to de-ice the release valve before repeating tests cancelled out any time advantages that the Pulse method may have had over performing a blower door test. By comparison, the Blower Door test often had a beneficial effect of purging some of the long-vacant dwellings of stale moist air.

Actual results from tests performed under a variety of conditions on 46 individual existing properties (from this project, the current BEIS DEEP project and a study on Park Homes) are shown in Table G5. In each of these dwellings, pulse and blower door tests were undertaken with the dwelling in identical conditions, although there may have been variations in both internal and external environments between tests. Although Table G5 contains all “positive” test results, in a number of cases the researchers were not happy with the results obtained from both test methodologies and repeat tests were undertaken. These repeat tests were performed for various reasons; in HH01 pulse tests were repeated with the unit in 3 different locations due to “bouncy” floors and concerns over acoustic effects, in other dwellings pulse tests were repeated as the testers were unsure whether the severe frosting around the release valve had affected the pulse result, in HH17 & HH20 the Pulse unit appeared to be unable to successfully test larger and less airtight dwellings with all vents open even though the unit itself registered these as successful tests, blower door tests were also often repeated at the start and end of co-heating and where initial results appeared questionable (HH29).

Using the results from Table G5, these data suggest that relationship between the two test methodologies is not a simple geometric one. Figure G5 shows an improved correlation using a 7√(4Pa) conversion factor to transform Pulse results into an equivalent permeability @50Pa value (for the buildings in this project). It is noted that an AP50 = AP4^0.9241 x 5.2540 conversion has been introduced by the BRE as a means of comparison for new-build UK dwellings that fall within the Building Regulation’s requirements for new build of <10 $m^3/(h.m^2)$@50Pa and the results shown in Figure G5 confirm that this is not an unreasonable approximation. However, the majority of existing buildings tested in the SMETER and DEEP projects had air permeability values in excess of current new build compliance levels, and with SMETER tests also being undertaken under a variety of ventilation strategies (vents closed and open) this issue was exacerbated.

Since carrying out the testing the manufacturer has developed a new Pulse 2.0 device. Build Test Solutions commented that this seeks to mitigate the issues experienced with icing up of the main air release valve, includes software improvements, and has better guidance on the number of air receivers required to deliver sufficient flow rates in more leaky properties.

Discussion

The main advantage of Pulse is that it directly measures ventilation rate at a pressure much more similar to those experienced under real world conditions than the blower door method with its quite extreme pressure differentials. The other advantage is that the overall time taken to perform the test is slightly shorter. The Pulse test takes around 12~15 minutes for the compressor to charge the cylinder, time which could also be used to prepare (and measure) the house. Once charged the test itself only takes a few minutes, and dismantling the equipment also only takes a couple of minutes. With a blower door test it takes 10 minutes for assembly, 15 minutes preparation (and additional measuring if required), 10~20 minutes running the test (10 if just pressurisation or depressurisation, 20 mins if both pressurisation/depressurisation), and a further 10 minutes dismantling. The Pulse test takes under ½ hour with 1 person, the Blower Door test takes ¾~1 hour with 1 person, ½~¾ hour with 2 people.

Table G5: Results from Pulse and Blower Door tests performed on 46 individual dwellings.

Test Dwelling and Ventilation Strategy		Pulse Test		Blower Door Test	
		$m^3/(h.m^2)$ @ 4Pa	h^{-1} @ 4Pa	$m^3/(h.m^2)$ @ 50Pa	h^{-1} @ 50Pa
HH01	All Vents Open	4.590	4.956	16.99	18.35
		3.593	3.880		
		5.559	6.002		
	Trickle Vents Only Closed	4.150	4.480	15.91	17.18
	All Vents Closed	3.105	3.352	15.52	16.76
HH02	All Vents Open	1.830	2.740	9.17	13.76
	Trickle Vents Only Closed	1.770	2.660	7.33	11.00
	All Vents Closed	1.530	2.290	6.54	9.81
HH03	All Vents Open	2.180	2.388	10.50	11.50
	Trickle Vents Only Closed	1.172	1.284	7.98	8.74
	All Vents Closed	1.243	1.362	6.86	7.51
HH04	All Vents Open	1.589	2.252	8.90	12.62
	All Vents Closed	1.174	1.664	6.53	9.26
HH05	All Vents Open	6.875	7.435	20.26	21.91
	All Vents Closed	3.536	3.824	14.82	16.02
HH06	All Vents Open	3.676	3.711	13.91	14.04
	All Vents Closed	2.210	2.231	10.47	10.57
HH07	All Vents Open	2.174	3.122	8.56	12.27
	Trickle Vents Only Closed	1.322	1.898	7.96	11.40
	All Vents Closed	1.434	2.060	5.99	8.58
HH08	Trickle Vents Only Closed	3.319	3.432	9.84	10.18
		2.645	2.735		
HH09	All Vents Open	5.539	8.099	14.93	22.03
		6.460	9.446		
	Trickle Vents Only Closed	3.446	5.039	14.49	21.39
	All Vents Closed	2.933	4.289	11.92	17.59
HH10	All Vents Open	2.348	2.595	12.02	10.45
	All Vents Closed	1.321	1.460	10.43	9.21
HH11	All Vents Open	2.945	3.486	14.64	14.36
	All Vents Closed	2.762	3.269	12.96	12.72
				13.83	13.57
HH12	All Vents Open	1.127	1.665	8.42	12.44

		1.322	1.954		
		1.111	1.642		
	All Vents Closed	0.588	0.868	5.76	8.51
		0.224	0.331		
		0.734	1.085		
HH13	All Vents Open	4.632	6.771	14.01	20.60
	Trickle Vents Only Closed	3.159	4.615	12.17	17.89
		2.777	4.059		
	All Vents Closed	2.111	3.086	9.52	14.00
HH14	All Vents Open	1.990	2.107	9.06	9.59
	All Vents Closed	1.277	1.352	7.49	7.93
HH15	All Vents Open	4.816	5.075	16.05	16.91
		5.265	5.548		
	All Vents Closed	3.511	3.700	12.73	13.41

HH16	All Vents Open	2.396	2.544	13.56	13.11
				13.12	12.68
	All Vents Closed	1.740	1.847	10.26	9.91
				9.89	9.46
HH17	All Vents Open	13.869	15.396	20.10	19.97
				19.38	19.25
	All Vents Closed	3.540	3.930	17.25	17.13
				16.70	16.18
HH18	All Vents Open	2.301	3.362	11.37	16.71
	All Vents Closed	1.814	2.651	9.58	14.08
HH19	All Vents Open	4.993	5.590	14.54	16.29
HH20	All Vents Open	11.451	11.994	13.46	14.65
	All Vents Closed	1.444	1.512	9.79	10.66
HH21	All Vents Closed	2.041	2.311	9.84	11.14
HH22	All Vents Open	1.469	1.577	11.44	12.28
		1.222	1.312		
	All Vents Closed	1.202	1.290	9.57	10.28
		1.023	1.099		
HH23	All Vents Open	2.313	2.383	10.99	11.32
	All Vents Closed	1.686	1.737	8.66	8.92
HH24	All Vents Open	2.574	3.780	9.97	12.20
				9.98	12.21
	All Vents Closed	1.687	2.478	7.33	8.96
				7.03	8.60
HH25	All Vents Open	2.565	2.731	13.44	14.31
	All Vents Closed	1.832	1.951	11.01	11.72
HH26	All Vents Open	5.638	6.156	20.71	19.23
				20.48	19.01
	All Vents Closed	2.867	3.134	14.91	13.84
				14.77	13.71
HH27	All Vents Open	2.095	2.470	10.94	11.64
	All Vents Closed	1.184	1.396	9.20	9.79
HH28	All Vents Open	2.553	2.718	11.60	10.67
				11.32	10.41

	All Vents Closed	1.529	1.628	7.63	7.02
				7.57	6.97
HH29	All Vents Open	3.306	3.494	38.95	42.64
				20.79	18.99
	All Vents Closed	3.011	3.182	15.67	14.31
				15.55	14.21
HH30	All Vents Open	3.553	3.737	15.60	16.40
	All Vents Closed	2.822	2.968	13.04	13.72
HX01	All Vents Open	1.658	1.776	9.79	10.49
	All Vents Closed	1.271	1.361	7.17	7.68
D01	All Vents Closed	2.853	3.081	16.78	18.13
D02	All Vents Closed	1.513	1.642	9.61	10.43
D04	All Vents Closed	2.836	3.156	12.58	14.00
D06	All Vents Closed	3.921	6.021	14.74	22.63
D07	All Vents Closed	2.895	3.794	12.36	16.20
D10	All Vents Closed	0.089	0.089	0.93	0.89
D11	All Vents Closed	2.910	3.233	13.98	15.53
D12	All Vents Closed	1.907	1.844	9.29	9.59
D14	All Vents Closed	1.507	1.312	9.33	9.63
A01	All Vents Closed	1.690	2.687	5.96	9.48
A02	All Vents Closed	0.804	1.298	5.07	8.19
A03	All Vents Closed	0.572	0.993	4.91	8.70
A04	All Vents Closed	1.878	2.918	8.47	13.17
A05	All Vents Closed	1.304	2.262	6.66	11.56
A06	All Vents Closed	1.311	2.086	5.93	9.44

HH01 – HH30 and HX01 are from the TEST Project
D01 to D14 are standard masonry buildings from the BEIS-funded DEEP Project (D10 was an EnerPHit retrofit)
A01 to A06 are from mobile/park homes (Johnston and Miles-Shenton, 2017)

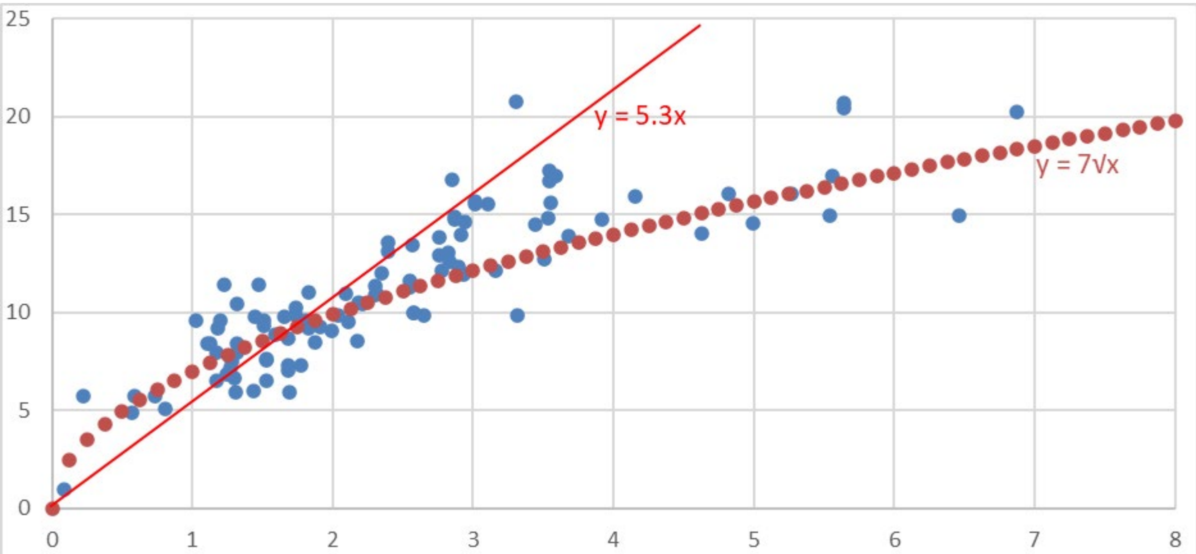


Figure G5: Plotting blower door permeability results ($m^3/(h.m^2)$@50Pa – *y-axis*) plotted against Pulse test results ($m^3/(h.m^2)$@4Pa – *x-axis*) suggests the relationship between them is not a simple geometric one.

The results from the Pulse test appear to be fairly repeatable for dwellings which have all vents closed (as per ATTMA TS1 test conditions), particularly those recently constructed under new-build Building Regs airtightness requirements (air permeability of less than 10 $m^3/h.m^2$ @50Pa). For leakier existing buildings and those with vents open it appears less reliable, though the spread of results is not enormous. The simple conversion factor to air permeability @50Pa of 5.3(4Pa) does not fit with existing dwellings above about 10~12 $m^3/(h.m^2)$@50Pa; our data show that 7√(4Pa) is a more suitable conversion; however, some care should be given to this result as the dataset is small and from a limited sample. For many of the leakier buildings tested floor coverings were not in place, and it is not known if the relationship may be closer for leaky buildings with sealed floors. However, it is expected that this would have a positive impact on the correlation between the two methodologies.

The pulse test is useful in measuring ventilation under normal living conditions at that moment in time. There remains an issue when converting the 4Pa ventilation rate figure to a 50Pa permeability value, which is not a simple linear conversion. Needless to say the Blower door test has similar inadequacies when using n/20 to convert a 50Pa permeability figure to an average ventilation rate figure under normal living conditions.

References

ATTMA (2016) Technical Standard L1. Measuring Air Permeability of Building Envelopes (Dwellings) [Internet]. Airtightness Testing and Measurement Association, Northampton. September 2016 Issue. Available from: https://www.bcta.group/shop/attma-technical-standard-l12016/ [Accessed: 29th April 2020].

Carey, P. and Etheridge, D. (2001) Leakage measurements using unsteady techniques with particular reference to large buildings, Building Services Engineering Research & Technology, Vol. 22, No. 2, (2001) pp. 69-82.

Cooper, C., Zheng, X., Wood, C., Gillot, M., Tetlow, D., Riffat, S. and De Simon, L. (2016): Field trialling of a pulse airtightness tester in a range of UK homes, International Journal of Ventilation, DOI: 10.1080/14733315.2016.1252155

Johnston, D., Miles-Shenton, D. (2018): The airtightness and air leakage characteristics of new UK holiday homes, Building Services Engineering Research and Technology, Vol. 39, No.3, pp343-361, DOI: 10.1177/0143624417748238

Nishioka, T. (2000) A new technique for measuring airtightness of the building envelope using pulse pressurization, 2000 Proceedings of Ventilation 2000, Volume 1, proceedings of the 6th International Symposium on Ventilation for Contaminant Control, held Helsinki, Finland, 4-7 June 2000. https://www.aivc.org/sites/default/files/airbase_13551.pdf

Zheng, X.F., Cooper, E. Zu, Y., Gillott, M., Tetlow, D., Riffat, S., Wood, C. (2019) Experimental Studies of a Pulse Pressurisation Technique for Measuring Building Airtightness. Future Cities and Environment, 5(1): 10, 1–17. DOI: https://doi.org/10.5334/fce.66

Appendix H: Definition of the Heat Transfer Coefficient

The widely recognised metric for building heat loss is the Heat Transfer Coefficient (HTC), which expresses the time averaged rate at which heat is lost per degree Kelvin temperature difference between the inside and outside of a building in units of W/K. It includes the heat loss through the fabric and by ventilation and infiltration. A lower HTC demonstrates a lower average rate of heat loss and therefore better thermal performance.

The Heat Transfer Coefficient (HTC) used in this project is defined in Equation 1, which is based on the *building heat transfer coefficient* described in BS EN ISO 13789:2017 *Thermal performance of buildings - Transmission and ventilation heat transfer coefficients - Calculation method*. The transmission and ventilation heat transfer coefficients are as defined in Equations 2 and 3 (BS EN ISO 13789:2017). It should be noted that: the direct transmission heat transfer coefficient represents heat losses through the building fabric, including thermal bridges; and the ventilation heat transfer coefficient includes infiltration and ventilation.

$$\text{HTC} = H_{tr} + H_{ve} \quad \text{Equation 1}$$

Where:

HTC Building Heat Transfer Coefficient (W/K)
H_{tr} Transmission heat transfer coefficient (W/K)
H_{ve} Ventilation heat transfer coefficient (W/K)

$$H_{tr} = H_d + H_g + H_u + H_a \quad \text{Equation 2}$$

Where:

H_{tr} Transmission heat transfer coefficient (W/K)
H_d Direct transmission heat transfer coefficient (W/K)
H_g Transmission heat transfer coefficient through the ground
H_u Transmission heat transfer coefficient through unconditioned spaces
H_a Transmission heat transfer coefficient to adjacent buildings

$$H_{ve} = \rho_{air} \times c_p \times q_v \quad \text{Equation 3}$$

Where:

H_{ve} Ventilation heat transfer coefficient (W/K)
ρ_{air} Density of air (kg/m^3)
c_p Heat capacity of air (J/kgK)
q_v Air flow rate through the building (m^3/s)

BS EN ISO 13789:2017 assumes that the heat transfer through the ground will vary on a monthly basis and therefore that the value of HTC changes slightly each month giving 12 values. For the purposes of this project the monthly variation will be included in the estimation of the 95% confidence interval and therefore only one value of measured HTC will be defined for each home.

The HTC predicted using the SAP and RdSAP methods is also based on the BS EN ISO 13789:2017 definition but varies in two subtle ways:

1. The ventilation heat transfer coefficient includes a specified value for intermittent use of mechanical extract fans in the kitchen and bathroom.
2. The ventilation heat transfer coefficient varies on a monthly basis due to changes in the assumed windspeed (as modified by the shelter factor for the building) – giving 12 different values.

It will still be reasonable to compare the HTC values measured in this project with those from SAP and the differences in definition are not expected to be significant when compared to the error in measurement.

To ensure consistency with the definition of HTC in this project, the ‘traditional’ co-heating test methodology has been modified to account for the designed ventilation openings (e.g., extract fan openings, air bricks, flues, fireplaces, trickle vents, passive vents). This introduces additional uncertainty. Further uncertainty has also been added to account for monthly variation in the measured HTC.

There are potentially two reasons why the HTC of a building in-use may vary from the value defined here:

1. Increased ventilation heat losses, for example opening and closing of windows and the regular use of intermittent extract fans, may lead to a higher HTC than when the house was not occupied.
2. Occupants own furnishings, floor coverings, curtains and blinds may lead to reduced transmission heat loss and a lower HTC than measured while the house was empty.

Appendix I: Evaluation of SMETER I technology from a ninth participating organisation

A ninth participating organisation (SMETER I, Knauf Energy Solutions) joined the project part way through the Phase 2 field trial and too late for their product to be installed in field trial homes. Therefore, the participating organisation supplied three additional homes (named HX01, HX02, and HX03). HX01 was located in Manchester (UK) while HX02 and HX03 were located in Genk and Munsterbilzen (Belgium) respectively. The TEST team at LBU carried out co-heating and blower door tests for HX01 to determine the measured HTC. A team at KU Leuven (Belgium) had previously carried out co-heating tests on HX02 (November 2017) and HX03 (February 2018) as part of another project and the data were procured from them for LBU to determine the measured HTC. The ninth participating organisation was not told any of the measured HTC results.

It was not possible for the SMETER I product to be evaluated in TEST 5 as the product was not installed by the TEST team and the TEST team had no contact with the households in any of the three homes. However, it was understood that the SMETER product included a heat meter that was installed in the distribution pipework of the hydronic central heating system as well as temperature sensors installed in the home. Therefore, it was deemed to be Type T4 given the need for professional installation. The participating organisation did not take part in TEST 6.

The measured HTC for HX03 was complicated to calculate as the property had an airtightness retrofit after the co-heating test was carried out due to a construction fault. Blower door tests carried out before and after the airtightness retrofit were used to adjust the co-heating test result. However, there was a large discrepancy between this measured HTC (87 W/K) and that calculated by the participating organisation (201 W/K). The measured HTC was deemed to be unreliable, and it was not possible to re-measure the HTC of HX03. Therefore, HX03 was removed from further analysis.

The participating organisation used a 74-day data period of data on average in their calculations of the HTC (Table I1). The confidence intervals reported were smaller than those for the measured HTC (Table I2), with an average confidence interval of ±6% and a maximum confidence interval of ±8%. The confidence intervals overlapped for HX01 and HX02 (Figure I1). Direct comparison with the results from other SMETER technologies is difficult as there were only two cases.

Table I1: The self-reported period of data used by SMETER I

Home ID	Days	from	to	period
HX01	52	26/02/2020	01/05/2020	66
HX02	71	11/02/2020	01/05/2020	81
Average	62			74

Table I2: Measured and calculated HTC results with confidence intervals

Home ID	**Measured HTC**							**SMETER I calculated HTC**				Confidence interval overlap?
	Central estimate (W/K)	Experimental uncertainty			Experimental uncertainty and seasonal variation			Central estimate (W/K)	Reported 95% confidence interval			
		lower bound (W/K)	upper bound (W/K)	expressed as +/-	lower bound (W/K)	upper bound (W/K)	expressed as +/-		lower bound (W/K)	upper bound (W/K)	expressed as +/-	
HX01	169	142	196	16%	136	202	20%	169	155.8	182.2	8%	✓
HX02	235	196	275	17%	188	282	20%	224	216.6	231.4	3%	✓

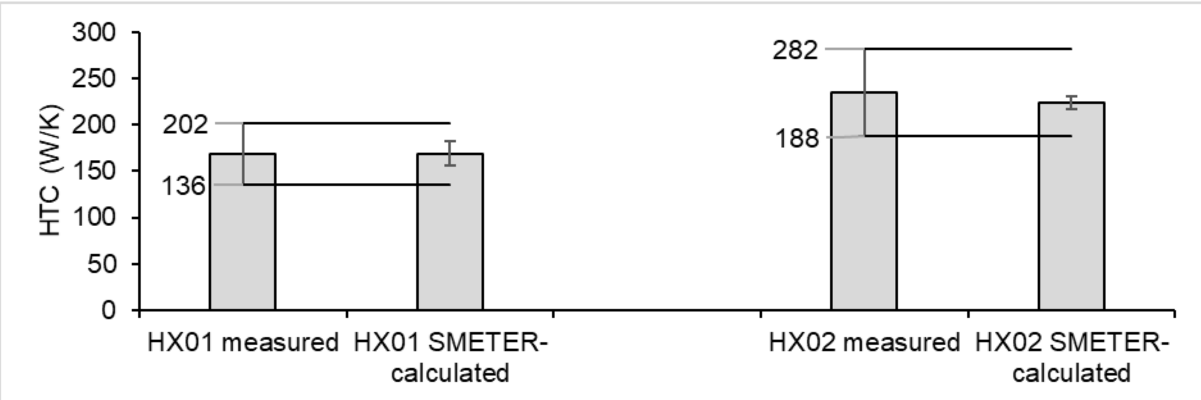


Figure I1: SMETER I HTC results for HX01and HX02 with error bars for 95% confidence interval; the horizontal lines show the 95% confidence interval for the measured HTC

This publication is available from: www.gov.uk/government/publications/smart-meter-enabled-thermal-efficiency-ratings-smeter-technologies-project-technical-evaluation

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