Department for
Business, Energy
& Industrial Strategy

Technical Evaluation of SMETER Technologies (TEST) Project

Executive Summary

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

January 2022

Contents

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.

[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.

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 (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:

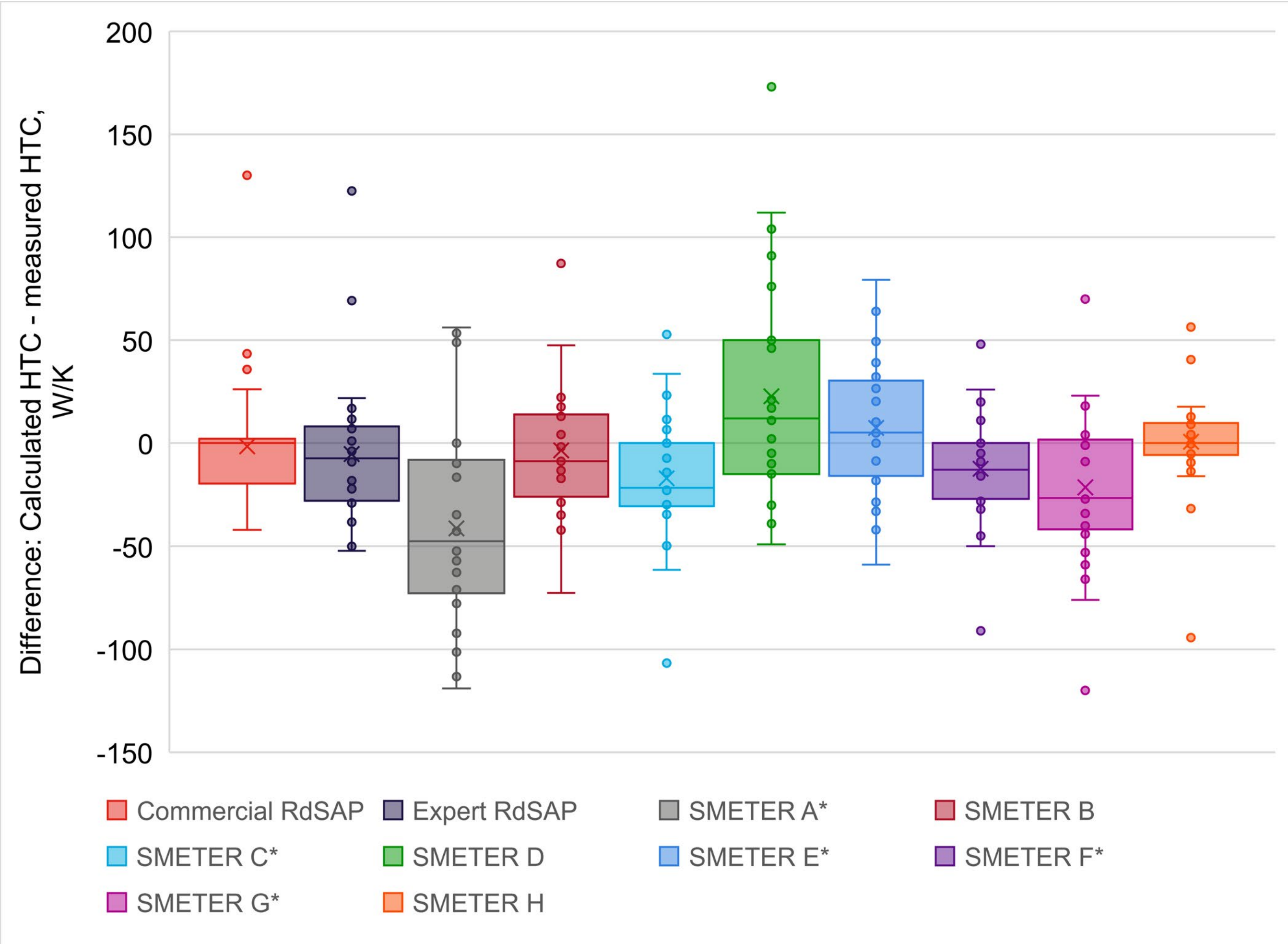
- SMETER B (BTS, 28 out of 30 homes, with an average confidence interval of +/-18%),
- SMETER E (EDF, 26 out of 27 homes, with an average confidence interval of +/-26%),
- SMETER F (HOA, 26 out of 27 homes, with an average confidence interval of +/-14%),
- SMETER G (PAS, 25 out of 27 homes, with an average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

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

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%.

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%).

[5] The NMBE of SMETER B (BTS) being further from zero.

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 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.

[6] Quantifying the Measurement Error on England and Wales EPC Ratings: https://doi.org/10.3390/en12183523
[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	SMETER B BTS	SMETER C CAR	SMETER D CSE	SMETER E EDF	SMETER F HOA	SMETER G PAS	SMETER 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 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.

[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.

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