In-use thermal performance metrics workshop: 12th April 2021
Executive summary

The workshop was convened to share findings of the Technical Evaluation of SMETER Technologies (TEST) project and consider how to design and deliver an effective in use thermal performance metrics system in the UK.

The TEST project ran alongside the SMETER[1] Innovation Competition, which aimed to accelerate the delivery of SMETER products to the market and provide BEIS with confidence that these products can meet accuracy, effectiveness and acceptability user requirements. Eight products were tested in a total of 30 occupied homes; participating organisations were asked to measure the Heat Transfer coefficient (HTC) in a blind trial against HTC as measured using a gold standard physical method (co-heating test).

Results showed that three out of the eight SMETER products showed very little bias and relatively high precision (better than an RdSAP assessment carried out by an expert). A late joining SMETER method also successfully predicted the HTC of two separate homes.

On the **accuracy requirements for thermal performance metrics**, participants argued that SMETER accuracy should be as good as or better than alternative methods, and accuracy requirements should be determined in relation to the application. Higher accuracy will be needed for regulatory purposes, but less precise methods would have other uses (e.g., detecting badly performing outliers in the stock, evaluating impacts of measures over multiple homes). Consistency, robustness and reliability are also important, especially for regulatory purposes, as is the ability to take account of wider developments such as heat pumps.

Consumer engagement and support options should be considered, in order to make full use of accurate metrics to deliver consumer benefits; **metrics should be easily understandable** by consumers, and the **need for positive consumer acceptance** of devices and data sharing mean that the benefits should be clear to them.

On **market development, data access and communications requirements**, participants' comments suggested that **integrated approaches to data collection** (using the same communications routes for multiple data requirements to produce in use metrics) and **open /easy access to data** (including "open data" principles to avoid the need to duplicate existing data streams) are desirable.

Policy levers were seen as important to market development: including **aligning policy with the use of SMETER-enabled products and services** to reward accurate metrics (e.g. within EPCs) and encouraging adoption through pathfinder policies, publicly funded retrofit schemes and trials; and considering mandating **changes to the smart metering rollout** to support SMETERs.

Some form of **validation or approval mechanism** was generally felt to be required, to ensure accuracy and provide confidence to consumers and other stakeholders. This might, for instance, involve establishing a **central certification body, to oversee arrangements** for accuracy testing and validation. Differing suggestions were made about the technical approach to validation. As well as accuracy, assessment of ease-of-use, usefulness, robustness, representativeness and repeatability may need to be taken into account.

[1] Smart Meter Enabled Thermal Efficiency Ratings: https://www.gov.uk/guidance/smart-meter-enabled-thermal-efficiency-ratings-smeter-innovation-programme

In-use thermal performance metrics workshop, summary report

The workshop was convened by BEIS and Loughborough, UCL and Leeds Beckett Universities to share findings of the Technical Evaluation of SMETER Technologies (TEST) project and consider how to design and deliver an effective in use thermal performance metrics system in the UK.

The following sections summarise the TEST project findings, followed by summaries of the workshop discussion sessions including:

1) Introducing requirements for a new system of metrics
2) Accuracy
3) Market development, data access and communications requirements
4) Validation, standards and auditing/ quality assurance (QA)
5) Future research, innovation and other priorities.

About this report: this report provides a summary of the workshop and individual inputs by participants: it should not be seen as representing consensus among participants or endorsement by the organisers.

The context for the workshop was introduced by Professor Paul Monks (BEIS Chief Scientific Adviser) in terms of Government commitments to Net Zero and the 10 Point Plan. Feedback on progress such as through SMETERs will be critical to achieve these commitments as part of a 'system of systems' approach.

TEST project findings

The Technical Evaluation of SMETER Technologies (TEST) project sought to evaluate the performance of the technologies in measuring the Heat Transfer Coefficient (HTC) of domestic buildings in-use and occupied. A modified version of the Co-Heating Test method was used as the gold standard comparison that provided the baseline 'measured HTC' against which the technologies performance could be evaluated. The project included a field trial with eight SMETER technologies in occupied homes (six requiring extra in-home monitoring, and two smart metering data alone). 30 detached, semi-detached or end-terrace properties, constructed between 1927 and 1990, were surveyed and tested, and had monitoring equipment installed along with SMETER technologies. Participating organisations were asked to calculate the HTC, in a blind trial against the measured HTC from the modified version of the Co-Heating Test method.

The modified version of the co-heating test method added in a calculation of purpose-provided ventilation and included additional uncertainty to account for seasonal variation in the HTC. This provided a better comparator for the in-use HTC calculated by participants and is commensurate with The Standard Assessment Procedure (SAP). The HTC was also calculated using the Reduced (RdSAP) methodology in SAP using data collected by an expert member of the team.

Analysis of differences between the HTC calculated by SMETERs and the measured HTC revealed that three out of eight SMETER methods showed lower bias and better relative precision than expert RdSAP for the homes included in this study, which is a very promising result. Self-declared uncertainty of individual predictions ranged from +/- 6% to +/- 49%. A late-joining SMETER method successfully calculated the HTC of two other homes with individual uncertainties of 3% and 8%.

Further work is now underway, including comparisons with HTC measurements from fresh commercial EPC assessments, the testing of other SMETER methods developed in the Annex 71 project using data from some of the homes, and a Green Homes Grant (GHG) SMETER trial to trial the use of SMETER methods to evaluate energy efficiency measures and collect further data.

An independent project evaluation based on questionnaire surveys and follow-up discussions with project participants brought forward ideas on: policy and market development; further testing; improved data access; incorporation into SAP calculations; testing new capabilities; development of QA methods; and, crucially, making the resulting metrics relevant to consumers and impactful in terms of behaviour change.

Introducing requirements for a new system of metrics

Accurate in-use thermal performance metrics would create new possibilities for supporting the delivery of Net Zero: in particular, metrics can serve a number of functions, including diagnosis, public information and “pay for performance” [see associated slide pack for further details].

Measured heat loss metrics could be incorporated into the existing models (in particular SAP and RdSAP) and metrics which underpin current policies, to provide greater accuracy and validate modelled predictions (which will still be necessary for design and other purposes) through feedback on actual outcomes, supporting existing and new functions.

The design criteria for any new system to measure in use thermal performance metrics should be steered by their purpose and functions: **criteria for effective metrics as identified by workshop participants** most frequently mentioned consumer use/understandability as the key criterion, followed by accuracy, repeatability and consistency. Participants also identified a range of interested stakeholders, and wider developments such as heat pumps, demand management and future changes to the regulatory system, which would be relevant to any new system of in use performance metrics.

Accuracy

The introductory presentation explored the requirements of accuracy for different purposes and the dimensions of accuracy (true accuracy, precision, repeatability and reproducibility). Reasons for variation in HTC estimates were reviewed in the context of ISO 13789:2017, followed by some of the drivers of uncertainty, including confounding heat gains and losses, data-related issues and storage. Additional presentations provided information on the repeatability of HTC measurements over time, sensor accuracy and the assumed relationship between measurement duration, number of data inputs and estimation accuracy. [see slide packs for details]

Participants were asked: **what should be the accuracy requirements for thermal performance metrics, and how could these vary for different purposes?** The following is a summary of the key points made on this question:

- SMETER accuracy should be **as good as/better than alternative methods** (e.g., better than RdSAP).
- Accuracy requirements should be primarily **determined in relation to the application**:
  - E.g., higher accuracy is needed where used for regulatory purposes.
  - Lower accuracy (significantly greater than +/- 10%) may be acceptable for diagnosis pre-retrofit, e.g., in the context of observed underperformances against predictions of 50% plus.
  - Less accurate measurements may also be useful to detect outliers in the stock.
  - Accuracy should be sufficient to distinguish successful from unsuccessful retrofits. This will be challenging for single measures, given savings (e.g., 7% for cavity wall insulation, estimated through NEED); however aggregated analysis for multiple homes could be used for evaluation and reporting where individual uncertainty is high. Using SMETERs to investigate a change in thermal performance may deliver lower uncertainty in that change than in the absolute HTC estimates.

- **Consistency, robustness and reliability** may be more important than accuracy (and need to be assessed alongside it). If used for regulatory purposes, methods would have to be robust enough to reproduce a similar HTC value under different occupancy scenarios.
- Accuracy should be **correctly declared**, so that methods perform in accordance with their advertised capabilities.
- Accuracy may be more difficult to achieve for more efficient homes, which may need different accuracy criteria. E.g., **absolute accuracy (W/K) may be more relevant** than a percentage of the HTC measurement, for such homes.
- There is a **cost/complication/accuracy trade-off**. I.e. there is no point chasing the best accuracy if it becomes a barrier to practical action.
- There may be a **difference between point in time accuracy** (i.e., of a one-off assessment or review), and ongoing (and potentially improving) estimates.
- There would be value in being able to **accurately predict savings to occupiers from retrofit** under existing occupancy conditions; and also, measurement of heating system and hot water efficiency measures, to complement measurement of the HTC.

Participants were also asked two further questions, and the bullets below each summarise key points made in response:

1) **What is the likely variation in HTCs estimated by SMETERs over time, considering physical effects and occupants or neighbours' behaviours? How can we mitigate this?**

- Robustness to **changes in occupancy, and that of neighbouring properties**, are important, these are likely to have bigger effects in better insulated properties, and those with more party wall elements. Such issues should be addressed in the SMETER or the CI and further research may help reduce their impacts.
- It was suggested that we may be able to partially mitigate occupant behaviour change through the use of simple but structured user questionnaires.
- Conversely, it was suggested that occupant behaviour may be hard to even categorise and that use of the wrong assumptions may significantly impede SMETER accuracy. Calculations over a long period of time may take out variation due to occupants. Also, the suggestion was made that we should accept that the HTC varies, and consideration of the average and trend will be useful.
- There can be strong **variations in spring and autumn seasons** due to differing behavioural responses to quick changes in weather; weather conditions for optimal measurements should be defined. I**nternal door use** may also change with seasons and occupants, leading to significantly different internal airflows.
- Seasonal **changes will vary for different forms/types of construction**: e.g., dwellings with suspended ground floors, partial fill external walls or party walls with thermal bypass will be more affected by seasonal variations in wind speed and direction.
- **Thermal storage** can have a large effect with short monitoring periods (less than 1-2 weeks).
- **Other causes of increased uncertainty** that were identified included: short-term changes to behaviours, such as over Christmas; irregular working hours, leading to difficulty establishing patterns of behaviour; unusual appliance use (e.g., tumble dryers) or hobbies (e.g., keeping reptiles); homes with very intermittent heating (such that thermal mass effects are always significant) – e.g., homes with smart thermostats with very intermittent occupancy.
- **Sub-metering** may enable SMETER development to reduce variation in HTC estimation, by accounting for the behaviours of occupants.
- An overarching comment: physical and occupant-related variation is an engineering challenge that SMETER providers have to demonstrate they can resolve – which requires a validation approach, which is an essential next step.

- [An additional, broader comment on this session] Isolating and removing the benefit of solar gains may underscore a well-designed home: this is an argument for a broader approach to in-use performance measurement, going beyond fixed fabric heat loss to include built form and the benefit of solar gain.

2) **what do you expect would be the other key sources of bias, and what needs to be done to address them?**

- **Number and location of sensors**; including the effect of being in direct sunlight (although it was also suggested that this could be diluted by having 4-5 sensors in different parts of the house); also: position of sensors near to the ceiling, proximity to heat emitters, or being coupled to the building structure and its thermal storage.
- The **accuracy of sensors** in measuring indoor temperature (a bias of 2° C will lead to a difference in HTC of around 10%).
- Links to unheated areas such as a connected garage, or where large amounts of energy is used outside of the main envelope (e.g., home offices in garden sheds, workshops, hot tubs etc.).
- Existence of unheated spaces within the main envelope, or the relation of where temperatures are recorded to where heating is used.
- Heat metering is essential for heat pumps.
- The assumptions that underpin a SMETER method are critical e.g., different assumptions about party wall heat loss and hot water use can cause a large difference in HTC estimates. Cross-validation of output calculations in a wide range of homes, with a range of occupants, levels of fabric heat saturation, leakiness, weather etc. should be built into SMETER approval
- Defining whether issues are a matter of variation in HTC or bias was not clearly aligned across the participants; however, the importance of addressing them was agreed.

Market development, data access and communications requirements

The introductory presentations set the smart metering context in terms of system design and rollout progress; short presentations then proposed [see slide packs for details]:

- Development of a real-time HTC feed into the EPC register, and making actual performance levels available as open data[2].
- The development of a zero contact, zero setup solution for remote sensing of energy and internal environments, with sensors pre-paired to smart meters and posted out, with data backhauled through the DCC architecture.
- Incorporating temperature sensing into cloud-connected IHDs, as a means of supporting rollout to smart metered homes, linking to the IHD mandate, ECO obligations, EPC compilation and smart control ratings.
- Using smart thermostats, which are being rolled out in social housing, to collect data.

Participants were asked **what specific aspects of implementing in use thermal performance metrics could build on the smart metering rollout**. The following is a summary of the suggestions on this question:

- **Data collection:**
  - Collection of (some combination of) gas, electricity, internal temperature, humidity, thermostat set point data, domestic hot water data, potentially all through the same comms route (e.g. DCC or a connected IHD).

[2] It was not precisely defined within the workshop what would constitute open data, but it was discussed that making data available in an open and transparent form would be helpful as a general principle.

- Collection of data from new boilers, heat pumps (including return water temperature) and heat meters on heat networks, integrated in the above data collection process.
- Metering local energy generation and submetering for EV's.
- Mandating systems to be open, if they don't actually run on the DCC.

- **Consumer engagement and support:**
  - Use of IHDs to communicate with/engage occupants.
  - User-centred energy and carbon savings information/analytics (e.g. projected energy costs for upcoming week).
  - Targeting of the most vulnerable in society with appropriate energy measures.
  - Also, assessment of mould risk, compliance with Buildings Regulations Part F and the new Part X overheating standard.

A summary of key points made by participants on the question of **how generally market development should be encouraged**:

- **Ease of data access/openness:**
  - Making it easy for third parties to connect to the HAN and use the DCC data network, e.g. via an API, with customer consent.
  - Being proactive in opening up data, to avoid having to install yet more hardware to duplicate data streams that already exist.
  - Ensuring households can access the data themselves.
  - Making SAP-predicted HTCs available through an open access API.

- **Policy levers:**
  - **Aligning policy with the use of SMETER-enabled products and services** to reward accurate metrics (e.g. augment EPCs to include) and encourage their adoption through pathfinder policies, e.g. building into an ECO commitment and adopting SMETER on publicly funded retrofit schemes (which require PAS2035 compliance) and trials, as part of funding terms and conditions.
  - Mandating **change to the smart metering rollout** to incorporate SMETERs.
  - Generally, using regulation in order to make in use performance measurement mainstream.

- **Certification:**
  - Making a **standardised "test" available** for new SMETER methods, e.g., using test houses and known reference models.
  - Developing a **Quality Mark** to accompany in use HTC measurements
  - **Introduce device standards** (otherwise measurements from add-on devices will be insufficiently accurate)

- **Market incentives/offerings:**
  - Opportunities for energy suppliers and others to develop "guaranteed" energy performance products and retrofit installations – with the guarantee being demonstrated by savings in energy bills combined with data from SMETERs.
  - Promote the use of accurate performance measurement within the property market.

A summary of key points made by participants on the questions of **consumers could be involved in implementing in use metrics, to achieve the greatest uptake and impact, and how other stakeholders (e.g. landlords) could be best involved?:**

- Important that metrics are **understandable by consumers** (e.g. MPG for vehicles analogy). Different metrics may be needed for different types of user (research requirement). HTC itself may not be directly relevant to consumers.
- **Consumer interest** is essential, including willingness to permit data access and to share information about actual energy use habits in order to get the best out of a reporting system, e.g., through receiving actionable insights which makes sense.
- **Perceived benefit to consumers** is important to enable uptake/something consumers actively seek. There may be scope to add extra value services, using the data collected.
- May be possible to **link implementation to retrofit measures or the availability of grants**: to project savings from measures, encourage uptake, help consumers understand how the measures are working/delivering benefits, including monitoring performance over time, and for policy evaluation.
- In rented housing, landlords will benefit from in use metrics to enable **targeting of retrofit, prioritising budgets, providing insights, and use of data to address well-being issues**. Implementation could be linked to **financial incentives for landlords** to upgrade properties.

A summary of key points made by participants on the question "**what risks or unintended consequences should Government consider** in developing the market in this area?":

- The **privacy of personal data** will require careful presentation of the implications of collecting sensor data, alongside explanation of benefits. If not addressed, privacy worries could lead to consumers becoming more negative about smart metering.
- Development of a **market of non-standardised products** could create confusion, while a monopoly is undesirable. There should therefore be a single open platform for developers, and market surveillance of solutions to avoid differing outcomes.
- Preferably **market development** should be driven by the demand side, and the overall objective of reducing emissions. There will be scope for linking new services and opportunities (e.g. metered energy savings, links to health) which will change over time, so the market needs to be able to evolve and grow.
- There is a **risk of gaming/fake calculations** where there are financial incentives. The system needs to include **quality control** to ensure that deviations from accurate assessments do not erode trust. Transparency/approval by an independent third party would build confidence.

Validation, standards and auditing/QA

The introductory presentations outlined the purpose of validation, in supporting metrics' business functions and ensuring reliability and trust. Validation approaches would need to take account of the inherent variability of HTC. Potential approaches include comparisons of in-use HTC with physical measurement including assessment against houses, modules and configurations of known quantities, assessment of repeatability; cross validation between different methods; or and the use of self-validation. It was noted that the testing of in-use methods against physical measurements of HTCs for a wide range of housing types would require a larger and more representative test data set of houses. In-use metrics could contribute to the new Building Performance Evaluation British Standard which is under development [see slide packs for details].

The ability to reliably disaggregate energy data into occupancy, space heating and fabric performance is essential if HTCs are to be considered valid and accepted. Beyond validation of HTC, it becomes increasingly important to have reliable energy data on building use and operation with net zero carbon in mind. Whole building performance metrics and valid diagnostics for services and fabric are essential for building owners to make informed maintenance and improvement decisions. Once measurements of fabric (HTC) and services can be relied on, the ability to influence ethical and responsible changes in occupant behaviour are possible and can also be measured. Central to the above is the valid assessment of the building fabric's energy efficiency based on reliable HTC data. Studies have found erroneous data in EPCs that limit reliability. The use of a validated smart data driven system offers potential to improve on the current position; however, trust in those systems used to measure the HTC must exist.

A summary is set out below of key points made by participants on the questions "what requirements would a system of validation need to meet?"; "What specific approaches do you think should be included?"; "How could a validation system be delivered in practice?":

- **Role of a central certification body:** there should be oversight and approval by an independent, non-commercial third-party body, with clear governance, and possibly a supporting expert panel. This body would enable competing methodologies to be validated, with an "open door" certification approach, with auditing regimes that report on and incentivise accuracy. This could approve applications for different purposes, depending on their accuracy. Such a body could have the right to examine the internal working of different methodologies.
- **Coverage/constraints**: a validation system should cover in-use measurements for all/a wide variety of housing types/ages; also, wide variation in occupancy, weather conditions, orientation and heating system types. Any exclusions should be clear (e.g. by providing a method for identifying where an algorithm is unlikely to work).
- **Approach to accuracy testing/validation**: a number of different requirements and approaches were suggested, including:
    - A requirement to demonstrate strong external validity against physically measured HTCs (focusing testing on the relation to the true value).
    - Cross validation as well as external validation are essential to build confidence.
    - Establish different test cells and configurations to represent the range of house types, occupancy and orientations, used to validate accuracy of measurements within various parameters.
    - The use of blind testing, and blind test datasets, including for more difficult to test properties (mid terraces and flats).
    - Use of some existing co-heating test data (some under standard co-heating test conditions and others with modified in-use co-heating analysis) , use of inter-SMETER comparison and evaluation, making use of empty homes or test cells with well-characterised synthetic occupancy.
    - A mix of validation against unoccupied co-heating tests and in-use energy consumption data.
    - A combined validation system, with a central data repository of data collected from validated hardware that has been tagged with HTC values and household characteristics, and blind tests which technologies were required to pass.
    - Controlled testing in artificial environments (test houses) – could be a property with variable fabric, orientation/shading and window size.
    - Testing to include a simple but standardised user input on the use of the building
    - The ability to detect deliberate gaming of the system.
    - Results should be compared with the system's own listed capabilities rather than an arbitrary standard.
    - Social landlords could be involved in validation, e.g. through a sector-wide innovation network.

- A request for proportionality, and easy to use validation approaches (not locked into a single source) with a focus on fitness for purpose.

- **The value of setting an objective to deliver increasing accuracy over time**: this was not generally seen as a suitable objective for a central validation system, as fitness for purpose doesn't necessarily require ever-increasing accuracy. Accuracy needs to be considered alongside ease-of-use, usefulness, robustness, representativeness and repeatability. Enabling innovation and promoting continuous improvement would be more suitable objectives.
- **Enhancing EPCs vs. assessing thermal upgrades:** SMETER TEST demonstrates the potential of smart technology to inform policy and support reliable EPCs assessment. Further work is required, as technologies are used to assess building performance improvements following changes and retrofits. Such work where subtle and interrelated fabric changes are made encompasses different performance requirements and may require further environmental parameters to be considered. A valid and reliable assessment of performance resulting from individual and or combinations of element and component changes requires further consideration on the precision and accuracy of tests. The BEIS Demonstration of Energy Efficiency Potential and the BEIS/UCL SMETER Green Homes Grant research projects should offer further insight into the issues to be considered.

Research, innovation and other priorities

Participants were asked about future research and innovation priorities, potential risks or unintended consequences and what key overall messages they would like to give following the session. Future research, innovation and other priorities should include:

- Exploring SMETERs with more diverse dwellings including a wider range of the housing stock and diversity of heating systems (including electrical, low/zero carbon systems and unmeasured gains/losses, e.g. oil heating, EVs) to improve confidence. This should be undertaken over longer time periods and enable comparisons to co-heating tests and existing systems. An open data approach would help facilitate more research as well as scrutiny of results.
- Exploring use of HTC measurements in practice for different applications, including EPC/SAP, retrofit design and targeting (at individual building and local area levels), identifying and resolving performance issues (e.g. the performance gap for new homes), including the possibility of a new home commissioning product which is halfway between co-heating and in use SMETER.
- Within the previous point, there was a focus on producing actionable insights for different actors (e.g. householders, building professionals) and exploring differences in accuracy/validation needs for different applications.
- Exploring the role of SMETER within future innovative applications including demand side response (DSR), decentralised generation/ storage and smart technologies (e.g. smart thermostats).
- Recognition that the EPC is well known and understood, so needs building on (as opposed to replacing) to ensure this is not lost. Furthermore, ways to 'game' the system need to be considered during further development. Future development will require convergence between the range of interested Government departments (BEIS, MHCLG, Treasury).
- Striking a balance between accuracy/validation and ensuring a system is implementable is important, as well as ensuring developments have a clear business model/ value proposition
- Designing systems around the occupant and considering the development of the supply chain are key issues to consider within future developments.
- Finally, there is considerable market interest in future SMETER developments and its applications.

Appendix: Organisation list

Active Building Centre Research Programme
Association for Decentralised Energy (ADE)
Association for Environmentally Conscious Building
Bath University
British Board of Agrement (BBA)
Blue Yonder
Building Research Establishment
British Gypsum
Build Test Solutions (BTS)
Cambridge Architectural Research (CAR)
Centre for Sustainable Energy (CSE)
Chameleon
Chartered Institution of Building Services Engineers (CIBSE)
Citizens Advice
City Science
Committee on Climate Change (CCC)
Data Communications Company (DCC)
Department for Business, Energy and Industrial Strategy (BEIS)
EDF
Edinburgh University
Elmhurst Energy Management
Energy Systems Catapult
Energy UK
EPSRC
Etude
Evergreen Energy
Four Walls
Future Climate
Good Homes Alliance (GHA)
Green Energy Options (GEO)
Halton Housing
Hildebrand Technology
Hoare Lea LLP
Igloo
Imperial College
Innovate UK / UKRI
Interfacing
KIWA
Knauf
Knauf Energy Services
KU Leuven
Leeds Beckett University
Loughborough University
Ministry of Housing, Communities and Local Government's (MHCLG)
Mott Macdonald
Newcastle University
Northern Ireland Assembly Nottingham University
Octopus Net Zero
Ofgem
Oxford Brookes University
Parity Projects
PassivSystems
Purmetrix
Quidos
Saint Gobain
Salford University
Scottish Government
Sero Homes
SOAP Retrofit
Stroma
Sustainability First
Sustainable Energy Association
Swansea University
Switchee
Tado
Trilliant
Twinn Sustainability Innovation
UCL
UKRI
University of Geneva
Useable Buildings Trust
Welsh Government