EBC

Energy in Buildings and Communities Programme

IEA EBC Annex 94:

Validation and verification of in-situ building energy performance measurement techniques

Report

Preparation Meeting

November 18-19, 2024 - Leuven, Belgium

Contents

1. Participants

Nine countries were represented at this Kick-off Meeting, yielding 44 participants from 31 organizations.

First Name	Last Name	Organization	Country
David	Allinson	Loughborough University	United Kingdom
Beñat	Arregi	Tecnalia	Spain
Peder	Bacher	DTU	Denmark
Hans	Bloem	DYNASTEE	Netherlands
Aimee	Byrne	TU Dublin	Ireland
Lior	Carno	Kestrix	United Kingdom
Mark	Collett	Leeds Beckett	United Kingdom
Joshua	Cooper	Hildebrand Technology Limited	United Kingdom
Jamie	Corson	UCL	United Kingdom
Maarten	De Strycker	Seco Belgium	Belgium
Cliff	Elwell	UCL Energy Institute	United Kingdom
David	Farmer	University of Salford	United Kingdom
Richard	Fitton	UoS	United Kingdom
Catalina	Giraldo-Soto	University of the Basque Country - UPV/EHU	Spain
David	Glew	Leeds Beckett University	United Kingdom
Vincent	Gonneau	Saint-Gobain Research	France
Matt	Goodridge	Kestrix	United Kingdom
Adam	Hardy	Leeds Beckett University	United Kingdom
Grant	Henshaw	University of Salford	United Kingdom
Frances	Hollick	UCL	United Kingdom
Myriam	HUMBERT	Cerema	France
Richard	Jack	Build Test Solutions	United Kingdom
María José	Jiménez	CIEMAT	Spain
Carlos	Jimenez-Bescos	University of Westminster	United Kingdom
David	Johnston	Leeds Beckett University	United Kingdom
Sarah	Juricic	Centre Scientifique et Technique du Bâtiment	France
Matthew	Li	Loughborough University	United Kingdom
Iñigo	Lopez	Tecnalia Research & Innovation	Spain
Johann	Meulemans	Saint-Gobain Research Paris	France
Johannes	Pernpeintner	Deutsches Zentrum für Luft- und Raumfahrt e.V. DLR	Germany
Martin	Prignon	Buildwise	Belgium
Katia	Ritosa	KU Leuven	Belgium
Staf	Roels	KU Leuven	Belgium
Twan	Rovers	Saxion University of Applied Sciences	Netherlands
Juile	Runser	Cerema	France
Mojgan	Sami	Homelink	United Kingdom
Patrick	Schalbart	Mines Paris, CEEP	France

First Name	Last Name	Organization	Country
Stan	Schellekens	Stichting W/E adviseurs	Netherlands
Justinas	Smertinas	Technical University of Denmark	Sweden
Sam	Stamp	UCL Institute for Environmental Design and Engineering	United Kingdom
Christian	Struck	Saxion University of Applied Sciences	Netherlands
Simon	Thebault	CSTB	France
Liesje	Van Gelder	BCCA	Belgium
Matthious	Van Hove	DTU	Denmark

2. Agenda

Monday November 18, 2024

09.00 Arrival and coffee
09.30 Welcome and aim of the workshop
09.45 Introduction from our sponsor - Knauf Energy Solutions
09.55 KU Leuven – ConstructThor
10.15 Presentations of work relevant to the annex (ST2-4) – part 1
11.15 Coffee break
11.30 Presentations of work relevant to the annex (ST2-4) – part 2
12.30 Networking lunch
13.30 Description of sub-tasks, outputs, and instructions for breakout sessions
14.00 Breakout sessions for forming working groups on ST2, ST3, and ST4
15.50 DYNASTEE
16.00 End of day 1 / Optional city tour

Tuesday November 19, 2024

09.15 Instructions for completing breakout sessions
09.30 Breakout sessions for agreeing tasks for ST2, ST3, and ST4
10.30 Coffee break and group photograph
11.10 Reports and actions from sub-task leaders for ST2, ST3, and ST4
11.30 ST1 briefing
12.30 Networking lunch
13.30 ST5 briefing
14.05 Presentations of work relevant to the annex
15.00 Discussion and next steps
15.50 Closing remarks
16.00 Finish

3. Day 1 – November 18, 2024

Welcome and aim of the workshop

Richard Fitton, Cliff Elwell, and David Allinson welcomed everyone. Thanks were given to everyone for attending, KU Leuven for hosting, Knauf Energy Solutions for sponsoring the lunches, teas, coffees and snacks. Grant Henshaw and Katia Ritoša were acknowledged for all their help with organizing this meeting.

The objectives of the annex were summarised:

- Objective 1 (Subtask 1) To develop new knowledge and understanding of the breadth of real-world applications for in-situ building energy performance measurement techniques and the technical requirements of those applications across different sectors. Stakeholders include the building industry, government policy and regulation, innovators from around the world, and applications in heating and

cooling climates. The requirements of these applications include the requisite accuracy and how different techniques can be validated and verified for use.

- Objective 2 (ST2) To extend the current HTC estimation methods to new building typologies and climates while improving their accuracy, repeatability and robustness. This work will include high performing homes (e.g. new-build and Passivhaus, where methods are limited in accuracy due to the low heat losses) and apartments in larger buildings (with many party elements). Methods that can be used outside of the main heating season, and in cooling dominated climates, will be investigated.
- Objective 3 (ST3) To co-create a new framework for the verification and validation of in-situ building energy performance measurement techniques. This will reconcile the current disparate approaches and provide the playbook of methods required for estimating uncertainty and evaluating accuracy and repeatability in the field. Recommendations will be made for auditing, accreditation schemes, and standards that may be required to assure delivery in practice.
- Objective 4 (ST4) To develop a new research area on building performance diagnostics, that identify the reason for the HTC performance gap. Novel in-situ methods are needed that disaggregate HTC estimates to identify the root-causes of underperformance and negate the need for expensive forensic examinations.
- Objective 5 (ST5) To collect and curate data sets to support the work of this EBC annex and to create a legacy resource that is free to access and that accelerates innovation and adoption of in-situ building energy performance measurement techniques. This will include new and existing simulated data, data from field trials in occupied homes and data from test houses. Best practice methods for the collection, reporting and curation of these data will be developed.

The aims of this initial meeting were explained:

1. Help developing the details of the work to be done.
2. Agree what the outputs should be (at least in the short term).
3. Give commitments to help deliver this work, including letters of participation.

Introduction from our sponsor - Knauf Energy Solutions

Barry Lynham and Steven Heath outlined the related work of Knauf Energy Solutions. Energy efficiency is essential to the energy transition but there are errors in EPCs and consumers are sceptical about energy efficiency. Doing retrofit work properly vs badly has a big impact on real performance. Scalable measurement methods are required, such as the method developed by Knauf. To be trusted, these methods all require validation, including re-testing the same house over different seasons. Annex 94 can help by demonstrating that these methods work.

KU Leuven – ConstructThor

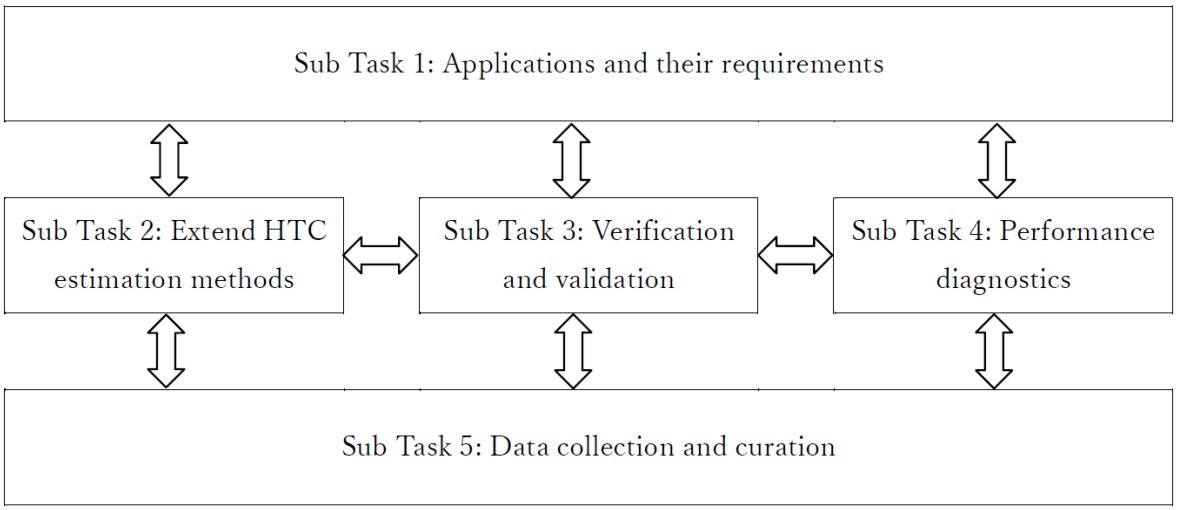
Staf Roels welcomed everyone to Leuven and introduced the town and University. He explained their impressive new test facility called ConstrucThor, located in Ghenk on the same site as the THOREAQ twin house. They are rebuilding some old buildings as well as new buildings working with recuperated steel, wood, bricks etc. Energy systems include bore holes and heat pumps. There are 12 test slots for simultaneously testing different building facades. Construction ranges from robotic construction of façade panels to rebuilding older homes to the drawings and standards of the day. Artificial occupants are used that both sense and mimic people. The site should be up and running by January 2026.

Breakout sessions for forming working groups on ST2, ST3, and ST4

An overview of the subtasks was given and the attendees were split into three working groups with sub-task

leaders:

- ST1 Applications and their requirements - Liesje Van Gelder, Mark Collett.
- ST2 Extend HTC methods -Katia Ritoša and Maria Jose Jimenez
- ST3 Verification - Sarah Juricic, Frances Hollick, and Simon Thebault.
- ST4 Performance diagnostics - Grant Henshaw and Richard Jack.
- ST5 Data Collection and Curation - Matt Li and Josh Cooper



DYNASTEE

Hans Bloem presented the DYNASTEE Centre of Excellence. DYNASTEE is the Network for DYNamic Analysis Simulation and Testing of Energy and Environmental performance of buildings. Formed in 2005, DYNASTEE's roots go back to 1985 and PASSYS 1. DYNASTEE organize an annual summer school that is very relevant to Annex 94. DYANASTEE offered also to organize a symposium in 2025 or 2026 to showcase some of the initial findings of this annex.

4. Day 2 November 19, 2024

Reports and actions from all sub-task leaders.

Sub Task 1 Applications and their requirements - Liesje Van Gelder, Mark Collett.

Main applications:

- Verification
- Retrofit
- Heating system sizing
- EPC

Research areas:

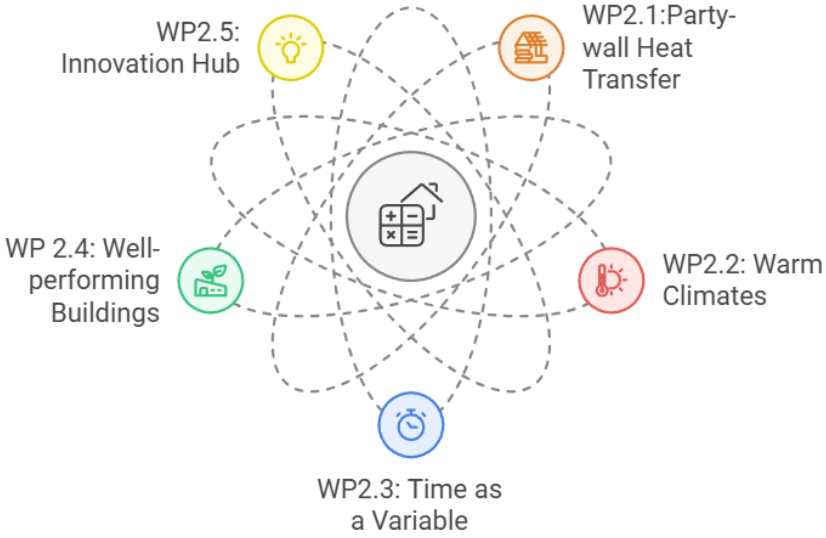
- Required accuracy for application
- Accessibility of data
- Link with finance - mortgage / pay for performance / zero bills / EU taxonomy
- Roll out - skills / labour / scalability
- Diagnostics - link with ST4 & other measurement techniques
- HTC definition - in use / for specific application
- Policy
- Understanding of results including uncertainty

Activity	Deliverable	Year	Volunteers
Mapping related projects and have interaction (Annex 88, IEA ES Task 43, …)	One pager	Y1	STL
Vocabulary	White paper	Y1	
Stakeholder survey (interaction from other subtasks needed)	Dataset	Y1	All
In-dept interviews/meetings	Paper	Y2	
Form advisory panel	Report	Y2	
Pathways to impact	Report/paper	Y3	

ST2 Extend HTC estimation methods -Katia Ritoša and María José Jiménez

Main objective

Enhance and further develop HTC assessment methods by extending their applicability to diverse building typologies, climates, and operational scenarios, while accounting for dynamic effects and the importance of monitoring time. In the process include innovative ideas, approaches, statistical tools and data sources.



Working Packages (WPs)

WP2.1: Party-wall heat transfer

The aim is to adapt HTC calculation methodologies to address challenges posed by differently heated neighbouring zones, including flats, multifamily, and terraced dwellings. The problem is especially pronounced in dwellings where poorly insulated party-wall heat losses are on the same scale as outdoor environment losses. This issue concerns also the definition of the heat loss area/building envelope and can be important in different applications such as energy labelling or evaluation of retrofits.

Research Questions:

- How important is party-wall heat loss or gain to HTC evaluation for specific, studied, archetypes? What generalised, qualitative, guidance be developed to address the importance of this issue in the broader stock?
- How can party walls and neighbouring zones in multifamily or terraced buildings be integrated into statistical models? What measurements are necessary to be included in the assessment? Is the HTC of individual such properties an appropriate measure and if not, what may be more appropriate?

Tasks:

- Conduct studies (simulations and dedicated experiments) to quantify the magnitude of the party-wall heat losses compared to other physical effects.
- Develop models to include neighbouring effects, shared boundaries, and party walls. Explore which input data is necessary and what information can be extracted from existing or limited data (e.g. collective heating data, inter-zonal and internal temperatures).

WP2.2: Warm climates

Warm climates are characterized by short heating seasons and usually exhibit low temperature differences (ΔT), high levels of solar radiation, or/and relatively high influence of non-linear phenomena typically linearizable or negligible, limiting the applicability of HTC assessment by traditional methodologies. Such climates require the extension of suitable periods to conduct informative measurements and pairing with statistical approaches which are more sensitive to such phenomena. Moreover, dominant cooling demand opens the door for the development of innovative characterization approaches.

Research Questions:

- Does the HTC assessment depend on climatic conditions? Is the variability inherent to the modelling approach?
- How to correctly include the building physics into models to properly describe the phenomena?

Tasks:

- Conduct studies (simulations and dedicated experiments) to account for and quantify the impact of small temperature differences, and other relevant phenomena typical of these climates (e.g. air leakage and its variability with wind, long wave effects) in statistical approaches and how to correctly incorporate them.
- Develop methodologies suitable for warm climates with dominant cooling demands that lower the uncertainty associated with HTC estimation in these conditions.

WP2.3: Time as a variable in HTC assessment

Time (reflected in data/measurement frequency, averaging, length, starting point and period of measurement) plays a crucial role in all types of statistical models applied in HTC assessment and needs to be thoroughly examined. Pairing of adequate measurements to statistical models affects their performance, reliability and repeatability. However, pairing the appropriate time aspects and the examined building is equally important because of the characteristics specific to the building, such as time constants, accumulation, delay in dynamic effects, and seasonal variability of HTC.

Research Questions:

- How do the frequency, averaging time, length, starting point and period of measurements affect the HTC estimation for different methods?
- How can time-dependent parameters (dynamic effects) be incorporated into different statistical models to improve HTC estimation and reveal supplementary insights (e.g. response times to heating and cooling, thermal mass)?

Tasks:

- Systematically explore the impact of frequency, averaging time, length, starting point and period of measurements on a variety of case studies reflecting different typologies and climates (connection with WP2.1 and WP2.2) and using different statistical models.
- Develop insights and criteria for the time aspects of measurement combined with the thermal dynamics of the building and the applied model to support the evaluation of HTC in buildings with different dynamic characteristics (e.g. thermal mass).

WP2.4: Well-performing buildings

Well-performing and passive-houses pose similar constraints to statistical assessment methodologies as for WP2.2 and 2.3. In such houses, high levels of solar radiation, or/and the influence of usually unmetered heat sources play an important role in the overall heat balance. Due to small heat losses, they also show less correlation between the indoor and outdoor environments. Moreover, such well-insulated houses often exhibit a very different performance over time, usually presenting slower dynamics.

Research Questions:

- How do the dynamic and physical phenomena (investigated in WP2.2 and WP2.3) affect statistical methodologies for HTC assessment in the special case of well-performing houses?

Tasks:

- Investigate the effects of slower thermal dynamics and reduced indoor-outdoor correlation on HTC estimation on test cases reflecting well-performing houses.
- Combining it with the approaches investigated in WP2.2 and WP2.3 proposes adjustments to current assessment methodologies to improve accuracy.

WP2.5: Innovation hub

WP2.1-2.4 focuses on extending existing approaches; this subtask also proposes to encourage and enable the

development of innovative ideas, tools and applications that are not extensions to these existing methods. Novel approaches can target the HTC estimation directly in the form of new statistical tools (e.g. machine learning, AI, dynamic HTC re-estimation, …) or as supporting tools to provide new data sources (e.g. smart meters, thermal scans, weather platforms, …) and principles to evaluate importance and feasibility in the HTC assessment (e.g. correlation and causation, importance of different inputs, …).

Research Questions:

- Which unconventional methods, approaches and principles can contribute to HTC estimation?
- Do different data sources provide consistent and comparable HTC estimates?

Tasks:

- Explore innovative approaches which can advance HTC estimation and evaluate their usefulness and applicability using common datasets.
- Compare with established methodologies and seek insight for application or plug-in.

Deliverables

WP	Item	Type	Timeline
WP2 / common deliverable from different STs	State-of-the-art review and catalogue of methods	Journal paper	Year P (in progress)
WP2.1	Implementation and importance of party-wall heat transfer in statistical models	White papers, journal papers	Year 1-3
WP2.2	Extending the applicability of the methods to different climates	White papers, journal papers	Year 1-3
WP2.3	Time as a variable in HTC assessment	White papers, journal papers	Year 1-3
WP2.5	Innovative approaches facilitating HTC assessment	White papers, journal papers	Year 1-3
WP2.5	Comparison of the impact of different weather sources	Internal review, white papers	Year 1-2

Collaborative work across STs

ST5

Define for each WP: sources of data (simulation and real life), types of measurements, level of detail, availability, requirements for different estimation approaches, and dealing with missing/unmeasured data or new data such as smart meters.

ST1 & ST3

Comparison of estimates from different methods, causes and range of variability, seasonal/climate variability, optimal combination of data + approach + applicability + achievable accuracy, scaling of the methods. This applies across the full range of property types, including high-performance homes, and in accounting for time aspects of buildings, measurement and methods.

ST4

Impact of different aspects in the HTC disaggregation: party-walls, infiltration, occupancy, and other effects dominant in warm climates.

ST3 Verification and validation - Sarah Juricic, Frances Hollick, and Simon Thebault.

As in-situ HTC estimates are more widely used and new methods for their estimation developed, it is crucial to have a mutual understanding of what the uncertainty estimates represent, and trust in both the HTC estimation method and its uncertainty. Frameworks for uncertainty calculation and validation of HTC estimation methods will enable widespread verification and comparison of results and build confidence in the use of these values for widespread applications.

Based on a common understanding of notions and nomenclature (WG3.1), subtask 3 aims at creating a consensual framework for HTC measurement validation (WG3.3). To do so, a deeper understanding of the sources of uncertainty and their quantification (WG3.2) is needed, acting as a verification protocol for all stakeholders. To consolidate the findings from these, subtask 3 aims to develop the frameworks into an accreditation scheme (WG3.4).

WG 3.1 Use and share a common language on uncertainty, verification and validation	**Deliverable**	**When?**
(with ST2) catalogue of in-situ building energy performance measurement techniques within a unified description	Internal report	1st semester 2025 + update 2027
Review and assemble vocabulary and nomenclature Expression of uncertainty in measurement (use GM definitions), verification, validation, auditing, etc.	"Vademecum" white paper + online training webinar?	1st semester of 2025
(with ST2) HTC definition, HPLC, GPLC, data dependent definitions...	White paper and/or research paper?	Y1/Y2

WG 3.2 Robust uncertainty calculation framework for all methods	**Deliverable**	**When?**
Review of good practice: uncertainty budgeting, GUM methodology	White paper or published review paper	Y1
(with ST2) Open call: each method owner share boundary conditions, data and metadata needs	Report	Y1
General survey on "± quantification" with associated knowledge (publication, recent work, expert rule...)	Report	Y1
Experimental work to complete knowledge on "± quantification"	Report	Y2
Framework for global uncertainty propagation and SA for any method (data sources, regression methods, bias…)	Report	Y2
Catalogue of methods with detailed application of framework, peer review	Report	Y2
Implementation of detailed application for each method (decide whether open, private, API, ...)	softwares	Y3

WG 3.3 Method validation	Deliverable	When?
(with WG 3.1) What is validation?	White paper	Y1
(with ST1) Needs and requirements for validation: define the minimal requirements needed for the validation dataset. What requirements on the inputs? relating to the requirements on the results expected by the stakeholders	Report (paper?)	Y1
What has been done? And how? Modelling, synthetical data, controlled experiments and field data	Literature review	Y1
How to define a reference value? E.g. blind tests, different households, different years, different measurands (HTC with/without ventilation, infiltration…)	Paper	Y2
(With ST5) List available test facilities and their possible variations	Internal report	Y2
Method-agnostic framework: validating models, validating methods	White paper	Y3
(with ST5) Validation dataset/tool: proposal, implementation (virtual & in situ), requirements for input data uncertainty	Report	V0 Y1 Y3
(with ST1) Comparison of results: when is good enough good enough? What is satisfactory?	Report	Y3

WG 3.4 Compliance with HTC methods and results users (with ST1)	Deliverable	When?
(ST1) Lit review & if need be survey of acceptability (uncertainty, validation) for different applications and stakeholders	Report	Y1-2
Accreditation: List all risks of a "bad implementation" of methods (take example from EPC & airtightness)	Report	Y2
Scheme for accreditation of operators and softwares	Report	Y3
Catalogue of applications with corresponding list of all HTC methods with validity domain (archetypes, durations, weather)	Report	Y3
Preparation of standard: • normative references (WG3.1: definition of terms) • apparatus: sensors required for using the methods in the field • calibration procedure • measurements: installation of the sensors, sensor location criteria, data acquisition, … • analysis of the data: description of the principles used in the methods • accuracy (WG3.2) • test report	Report	Y3

ST4 Performance diagnostics - Grant Henshaw and Richard Jack.

Proposed 6 key activities to be completed during the annex:

1. **State-of-the-art review** to understand what metrics would be good to measure, what tools are available to do so and where this sort of data is currently stored and used
2. **Consumer review** to understand who the stakeholders involved are
3. **Stakeholder survey** to understand the demand for building diagnostic, we want to be problem led and understand what the most valuable solutions are
4. Design a **standard diagnostic process** that others can follow
5. Organize **testing events** where organizations can demonstrate their technology at a test facility
6. **Repeat state-of-the-art review** at the completion of the annex

Activity	Deliverable	When	Who
State of the art review	Journal paper Industry white paper	1st draft Spring '25	ALL
Data storage and formatting review	Journal paper Industry white paper	1st draft Spring '25	Richard Jack Grant Henshaw ?
Consumer review	Internal information paper for survey development	Spring '25 for review at 2nd meeting	Matthew Goodridge Josh Cooper
Stakeholder survey	Survey ready for distribution	Draft survey for review at 2nd meeting	Matthew Goodridge Josh Cooper
Stakeholder survey results	Journal paper Industry white paper Presentation which can be given at industry events	Autumn '25	Matthew Goodridge Josh Cooper
Diagnostic process draft	Draft document	1st draft Autumn '25 for review at meeting 3	Twan Rovers Sam Stamp Peder Bacher Vincent Gonneau
Diagnostic process document	Standard document?	Review meeting 5, Autumn '26	
Specify testing event structure and locations	Test plan	Autumn '25	
Carry out testing events	Report of events	Autumn '26	
Final state of the art review	Journal paper Industry white paper	Autumn '27 (end of annex)	

ST5 Data Collection and Curation - Matt Li and Josh Cooper

Activity	Output(s) & outcome(s)	Timeline
Review of existing datasets relevant to the activities of the Annex (publicly available and/or held by Annex members)	Output: Internal table/database of datasets Outcomes: Understanding of currently available data.	Y1 (table to be produced ahead of next meeting)
Review of data requirements of Annex participants (including survey of Annex members and review of requirements of existing HTC characterisation methods)	Output: Internal report Outcomes: Understanding of the data requirements for Annex activities; contributions to action plans for data gathering & generation exercises.	Y1 (to report at next meeting)
Review of methods for data collection and management, including options for data sharing	Output: internal report Outcome: establishment of best-practice protocols & guidelines for data gathering, curation and sharing within the Annex.	Y1
In collaboration with Subtasks 2–4, establish action plans for gathering new data — simulated, test house experiments, field trials.	Output: action plans for data gathering/generation to meet needs of each Subtask. Outcome: Identification of personnel, means and methods for meeting data gathering/generation needs.	Y1
Production of new simulated data sets	Output: open source datasets. Outcome: enabling delivery of planned activities of Subtasks 2–4.	Y2
Data gathering from new test house experiments		Y2-3
Data gathering from new field trials		Y2-3
Curation of annex-relevant data sets	Output: a shared (and open-source, where possible) repository of data sets. Outcomes: Readily accessible, consistently curated datasets to enable Annex activities; legacy resource for future research.	Y1-3

5. Presentations of work relevant to the annex

These can be found here:
https://annex94.iea-ebc.org/event-link?LinkID=383

Presenter	Description
Matt Li presentation	Measuring the In-Use Thermal Performance of Dwellings: Improving the Precision by Accounting for Variability.
Simon Thebault & Sarah Juricic	• Quantifying Uncertainty in HTC Measurement • Benchmarking Methods from Numerical Virtual Experiments
Mark Collett	Metrics for Evaluating Accuracy and Precision of Measurement Methods.
Johannes Pernpeintner	Leakages in Building Envelopes - Localization by Acoustic and Thermographic Methods.
Peder Bacher	Statistical Level Testing.
Benat Arregi	Characterisation of Building Energy Performance (BEP).
Matt Goodridge	Thermal imaging and AI to scale building retrofit.
Justinas Smertinas	Taking Advantage of Data: EXRAY solution.
Katia Ritoša	Characterising the actual performance of buildings: A review.
Liesje Van Gelder	Opportunities and risks for heat loss coefficient measurements associated with EPBD recast.
Mojgan Sami	Time To Lose 1°C: Evaluating Energy Efficiency.
María José Jiménez	Building Energy Performance Assessment Methodologies and Experiences under Sunny Weather.
Joshua Cooper	Overview of Heat Transfer Coefficient (HTC) Applications.

6. Discussion and next steps

Richard Fitton thanked everyone for attending, KU Leuven for hosting, Knauf Energy Solutions for sponsoring the lunches, teas, coffees and snacks.

Important next steps:

- Letters of participation are needed from each participant country though their local rep.
- Full action plans are to be sent from each sub group leader within 14 days of the meeting
- An offer was made from CIEMAT to host the next meeting, and further offers are expected in the coming months
- It was agreed that meetings will be held using a hybrid approach with IT systems allowing for those who could not travel to attend virtually.
- In future meetings it was agreed that “mini conference papers” would be generated in addition to slides, the organizing group will seek a special edition of a suitable journal as an outlet for these works