Plan Overview

A Data Management Plan created using DMPonline.be

Title: IMPACT: Integrated Methodology for Performance Assessment and Characterisation of Thermal efficiency of buildings

Creator: Katia Ritoša

Principal Investigator: Katia Ritoša

Data Manager: Katia Ritoša

Project Administrator: Katia Ritoša

Affiliation: KU Leuven (KUL)

Funder: KU Leuven (KUL)

Template: KU Leuven BOF-IOF

Principal Investigator: Katia Ritoša

Data Manager: Katia Ritoša

Project abstract:

At the current state, the majority of residential buildings are underperforming in terms of energy. To assess the current state of the building sector and bridge the gap between the designed and operative energy demand, quantifying the as-built building behaviour is essential. Thus, the main aim of this project is developing a methodology to assess the energy performance of existing in-use buildings, by combining statistical approaches with optimised on-site measured data. In contrast to
the state of the art where the focus was typically on one KPI or the total energy use without investigating the underlying physical phenomena, this study aims to provide insight into the different components that are essential to characterize the energy efficiency of dwellings: 1) system
operation, 2) user behaviour and 3) building fabric. Starting from the well-researched Heat Loss Coefficient as a building fabric performance indicator, the proposed approach extends to include indicators covering all aspects of the overall building performance. To substantiate the development of advanced statistical methods, both detailed artificially simulated datasets and real-life collected measurements will be used. With the main intent to bridge the building energy performance gap, the
outcome of the project will allow quantification and guidelines to identify the sources that cause specific buildings to underperform, therefore contributing to effective building design and construction practices.

ID: 211233

Start date: 01-10-2024

End date: 30-09-2025

Last modified: 25-11-2024

IMPACT: Integrated Methodology for Performance Assessment and Characterisation of Thermal efficiency of buildings

Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

Dataset name / ID	Description	New or reuse	Digital or Physical data	Data Type	File format	Data volume	Physical volume
		Indicate: ***N****(ew data) or* ***E****(xisting data)*	Indicate: **D**(igital) or **P**(hysical)	Indicate: **A**udiovisual **I**mages **S**ound **N**umerical **T**extual **M**odel **SO**ftware **O**ther (specify)		Indicate: <1GB <100GB <1TB <5TB >5TB NA	
Districts	Simulated energy demand for two residential neighbourhoods	E	D	N	.csv	<100GB	
Terraced house	Simulation of a middle terraced house from the 1930s with different renovation scenarios	N	D	M	.mos, .csv	<100GB	
SMETER	Measured collected as part of the SMETER study in the UK	E	D	N	.csv	NA	

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

The "Districts" dataset was generated and used during the PhD study and uploaded on a public repository during the postdoctoral study: https://doi.org/10.5281/zenodo.14191440
The "SMETER" dataset was generated during a government research project in the UK. Part of the data is publicly available: https://repository.lboro.ac.uk/articles/report/Technical_evaluation_of_SMETER_technologies_TEST_project/19169027?file=34059095
In case the research stay at the UCL Energy Institute is approved, additional data with restricted access might be used at the host institution.

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.

- Yes, human subject data (Provide SMEC or EC approval number below)

The data from the SMETER study is collected from in-use houses. The data contains indoor environment measurements, such as indoor temperature, humidity and energy use. The part of the data available online is public, and the data which will be potentially used in collaboration with UCL Energy Institute will undergo the approval process after the research stay will be confirmed.

Will you process personal data? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).

- No

All data is related to houses or households, not individuals.

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- No

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material or Data transfer agreements, Research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- Yes

A research collaboration agreement will be needed in case of the research stay at UCL Energy Institute which will allow access to real-life monitored inhabited houses.

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- Yes

In case of collaboration with UCL Energy Institute, the used data is property of the host institution.

Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g. in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, codebook.tsv etc. where this information is recorded).

All datasets include or will include detailed metadata, recorded in a README.txt file, describing the purpose of the data, collection methodology, data structure, variable definitions, and units of measurement. In case the data was or will be part of a publication, or published as separate dataset information will be written in form of a journal/conference paper.
Observational and simulated data will be recorded systematically, with clear documentation of tools, sensors, and platforms used (e.g., smart meters, simulation software like Modelica). All code (e.g., statistical models, data processing scripts) will be shared alongside the data with embedded comments and a supplementary documentation file explaining dependencies, execution instructions, and workflow.

Will a metadata standard be used to make it easier to find and reuse the data?
If so, please specify which metadata standard will be used.

If not, please specify which metadata will be created to make the data easier to find and reuse.

- No

Metadata will adhere to the FAIR principles (Findable, Accessible, Interoperable, Reusable). Depending on the source of data, simulated or real-life, original or from third parties different approaches ensuring information describing data is structured, consistent, and interoperable will be adopted.

Data Storage & Back-up during the Research Project

Where will the data be stored?

- Personal network drive (I-drive)
- OneDrive (KU Leuven)
- Other (specify below)

Data and associated documentation will be stored on secure, backed-up platforms (e.g., work laptop and institutional data repositories).
In case of publicly shared data persistent identifiers (e.g., DOIs) will be assigned to datasets to ensure long-term accessibility.

How will the data be backed up?

- Standard back-up provided by KU Leuven ICTS for my storage solution

Is there currently sufficient storage & backup capacity during the project?

If no or insufficient storage or backup capacities are available, explain how this will be taken care of.

- Yes

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

During the research the data will available only when logging in with personal credentials or by our technical staff member Patricia Elsen.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

With the repository system at the Building Physics and Sustainable Design section we can guarantee
that data is well documented and stored over longer periods. The foreseen volume of data does not provide additional cost to the existing repository.

Data Preservation after the end of the Research Project

Which data will be retained for 10 years (or longer, in agreement with other retention policies that are applicable) after the end of the project?

In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

- All data will be preserved for 10 years according to KU Leuven RDM policy

Where will these data be archived (stored and curated for the long-term)?

- Shared network drive (J-drive)
- Other (specify below)

The "Districts" dataset is available on the Zenodo public repository.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

The foreseen volume of data does not provide additional cost to the existing internal repository in the research unit.
The data might be deposited on the KU Leuven Research Data Repository or in case of publishing on other public repositories a free option will be selected.

Data Sharing and Reuse

Will the data (or part of the data) be made available for reuse after/during the project?
Please explain per dataset or data type which data will be made available.

- Yes, as restricted data (upon approval, or institutional access only)

The original data will be available in case of internal collaboration within the research unit under supervision.

If access is restricted, please specify who will be able to access the data and under what conditions.

During the research the data will available only when logging in whit personal credentials or by our technical staff member Patricia Elsen.
In case of internal collaboration within the research unit, only the data produced inside the research unit will be shared with colleagues under supervision.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

Please explain per dataset or data type where appropriate.

- Yes, privacy aspects
- Yes, intellectual property rights
- Yes, ethical aspects

In case of collaboration with UCL Energy Institute, the used data is property of the host institution.

Where will the data be made available?

If already known, please provide a repository per dataset or data type.

- Other data repository (specify below)

Internal repository of the research unit.
In case the original data will be considered useful for other studies there is a possibility the data will be published on the KU Leuven Research Data Repository or other public repositories.

When will the data be made available?

- Upon publication of research results

Which data usage licenses are you going to provide?

If none, please explain why.

- Other (specify below)

Data usage license will be provided in case the original generated data will be considered useful for other studies.

Do you intend to add a persistent identifier (PID) to your dataset(s), e.g. a DOI or accession number? If already available, please provide it here.

- Yes, my dataset already has a PID

The "Districts" dataset was generated and used during the PhD study and uploaded on a public repository during the postdoctoral study: https://doi.org/10.5281/zenodo.14191440
Other originally generated datasets might not be shared publicly at the end of the research.

What are the expected costs for data sharing? How will these costs be covered?

In case of publishing on public repositories a free option will be selected.

Responsibilities

Who will manage data documentation and metadata during the research project?

The author (Katia Ritosa).

Who will manage data storage and backup during the research project?

The author (Katia Ritosa) with the support of the technical staff member Patricia Elsen.

Who will manage data preservation and sharing?

The author (Katia Ritosa) with the support of the technical staff member Patricia Elsen.

Who will update and implement this DMP?

The author (Katia Ritosa).