Presentation of new geospatial datasets for renewable thermal energy systems modelling in Switzerland

Jonathan Chambers¹, Mercedes Rittman-Frank², Martin Patel¹

¹ Chair for Energy Efficiency, Institute for Environmental Sciences and Department F.-A. Forel for Environmental and Aquatic Sciences, University of Geneva, Switzerland
² Institute for Solar Technology, Eastern Switzerland University of Applied Sciences, Oberseestrasse10, 8640 Rapperswil

Abstract. Decarbonising heating and cooling energy buildings means going beyond individual buildings to geospatial analysis of regions and the country. This creates a need for higher resolution geospatial datasets to perform energy systems modelling. In this work we present open heating and cooling demand geospatial raster dataset produced as part of the FEEB&D research project. We discuss challenges in the production and sharing of such datasets and discuss future work towards more comprehensive databases for thermal energy modelling.

1. Introduction
Decarbonising heating and cooling energy use in buildings to meet the Swiss Energy Strategy 2050 goals requires an approach that moves beyond individual buildings to consider districts and regions using geospatial analysis methods. This is because thermal energy cannot be transported over long distances, requiring a regional matching between thermal resources and demands using district heating to transport energy. This is a presents a research challenge with respects to modelling the pathways towards a decarbonised thermal energy system, especially considering the changing spatiotemporal distributions of demand through renovation and new construction.

Modelling low carbon thermal energy system therefore requires detailed geospatial datasets as inputs. The SCCER Future Energy in Buildings and Districts (FEEB&D) Work Package 3 (WP3) included extensive work on geospatially resolved energy demand for heating and cooling as well as sources of renewable heat energy, which produced a range of geospatial datasets. The datasets generated include mapping of building envelope efficiency, space heating and hot water, and energy saving potentials. The WP also developed models to map the potential for thermal networks. FEEB&D WP3 focused on geospatial modelling for demands at the ‘meso-scale’, with the goal of taking advantage of advances in computational tools to allow higher resolution analysis to be conducted over much large geographic areas using raster (pixel) based geospatial analysis.

In this paper, selected datasets from FEEB&D WP3 have been adapted from their original formats through re-mapping and aggregation onto a raster grids, with per-pixel values disaggregated by building type. These results have been checked, extensive metadata has been added, and the data uploaded to the a long term data archive. We summarize challenges that have arisen throughout the management of this
project to archiving these kinds of data. We furthermore reflect on future work on this topic under the new SWiss Energy research for the Energy Transition (SWEET) program Decarbonisation of Heating and Cooling in Switzerland (DeCarbCH) project.

2. Datasets on heating and cooling demand

Here we present selected datasets from the FEEB&D project pertaining to heat supply and demand in the building stock. These were archived in the “frictionless data package” format [11], consisting of a folder containing a ‘datapackage.json’ metadata file in JSON format. The metadata includes title, descriptions, authors, modification dates for the package of data overall as well as for each individual file in the dataset. The individual dataset files include file names and descriptions for each column (e.g. column name and data type). The tables containing the actual data are stored in the same folder in CSV format. Geospatial raster data is stored as (x, y) pixel coordinates columns in the Swiss EPSG:2056 coordinate reference system (CRS). Only pixels containing buildings are stored. The data are available at the UNIGE Yareta archive service under CC BY-NC 4.0 license. Building stock characteristics and heating demand models: https://doi.org/10.26037/yareta:sluzlbgmjfxji7qmymptxs7zoa.

Cooling demand in the service sector: https://doi.org/10.26037/yareta:5pdsn2n5yzesj6xjda4w7keby

2.1. Raster aggregate summary of building stock characteristics building typology

This dataset summarises the characteristics of the swiss building stock on a raster basis at a resolution of 200x200m. Buildings from the Swiss building registry are categorised into 63 archetypes according to their age, use, and location following the work in [1–4]. The location aspect from the original work was converted into pixel coordinates. Table 1 summarises the identifier fields in the output data that are used in Sections 2.1 and 2.2. Only pixels containing at least one building were stored producing resulting in 189’315 pixels and 695’488 points with unique pixel, age, and building type.

| name         | description                      | units |
|--------------|----------------------------------|-------|
| x            | x coordinate of pixel in CRS EPSG:2056 | m     |
| y            | y coordinate of pixel in CRS EPSG:2056 | m     |
| age          | Building construction year       | -     |
| buildingtype | Type or use of building           | -     |

Parameters were aggregated for each archetype per pixel were calculated. The data was collated from multiple sources, starting with the Swiss building registry [5]. Floor areas are particularly important [6], and were calculated using residence areas and building footprints and numbers of floors. Missing values were filled using the Swiss tlm3D building geometry dataset. Energy Reference Areas (ERA) were calculated, being the building area that is thermally conditioned and account for features such as stairwells, storage areas, etc. For residential buildings, statistical work was used to assess the relation between total area and heated area [6]. For non-residential buildings (especially service sector buildings), the ERA is the total floor area. The dataset has 7 fields for 63 archetypes although most pixels include only a subset of archetypes, resulting in 4’868’416 data points in total (Table 2).

| name                  | description                                    | units |
|-----------------------|-----------------------------------------------|-------|
| footprint_area        | building footprint area                        | m²    |
| total_area            | building total estimate area                   | m²    |
| residential_area      | total residential area                          | m²    |
| non_residential_area  | total non-residential area                     | m²    |
| era                   | Energy Reference Area                          | m²    |
| era_main_residence    | Energy Reference Area of main residences       | m²    |
| n_buildings           | Number of buildings of given type              | -     |
2.2. Raster aggregate summary of heat demand models by building typology

Two estimates for building energy demand for heating and domestic hot water (DHW) were generated and are included in this dataset. The first was based on a regression model derived from measured heating demand for a set of buildings and estimated heat demand totals and load curves as a function of building characteristics [7]. Power loads are summarized as the 25th, 50th, 75th, 95th, 99th percentiles and the maximum for useful and final energy. 23 data fields were produced (Table 3) (15'996'224 points).

Table 3 Data fields for regression-based building heat energy and power demand.

| name                  | description                                                                 | units |
|-----------------------|-----------------------------------------------------------------------------|-------|
| e_h_dhw_final         | Final energy (fuel) consumption for heating and domestic hot water (dhw).    | kWh   |
| q_h_dhw_useful        | Useful energy (heat thermal energy) consumption for heating and DHW          | kWh   |
| q_h_useful            | Useful energy (heat thermal energy) consumption for heating.                 | kWh   |
| q_dhw_useful          | Useful energy (heat thermal energy) consumption for DHW                      | kWh   |
| co2_eq_total          | Greenhouse gas emissions in CO2 equivalent for the heating                   | kg    |
| p<N>_h_useful         | <N>th percentile or max of useful power for heating                         | kW    |
| p<N>_h_dhw_useful     | <N>th percentile or max useful power for heating and DHW                    | kW    |
| p<N>_h_dhw_final      | <N>th percentile or max final power for heating and DHW                     | kW    |

A second heat demand estimate using the ‘SwissRes’ method (an archetype model derived from the building energy efficiency labelling database) [1–4]. This model estimated potential energy savings if the building archetypes were renovated to high energy efficiency standards (Minergie label equivalent). Data from [8] was mapped onto the pixel grid (6 data fields, 4'172928 data points) (Table 4).

Table 4 Data fields for SwissRes heat demand and retrofit savings potential model.

| name                    | description                                                                 | units |
|-------------------------|-----------------------------------------------------------------------------|-------|
| e_h_dhw_primary         | Primary energy consumption for heating and domestic hot water (dhw) before retrofit. | kWh   |
| e_h_dhw_final           | Final energy consumption for heating and DHW before retrofit.                | kWh   |
| q_h_dhw_useful          | Useful energy (heat thermal energy) consumption for heating and DHW before retrofit. | kWh   |
| e_h_dhw_final_savings   | Final energy savings for heating and DHW from retrofit.                     | kWh   |
| q_h_dhw_useful_savings  | Useful energy (heat thermal energy) savings for heating and DHW from retrofit | kWh   |
| cost_retrofit_investment| Total investment costs to perform retrofit for archetype                    | CHF   |
2.3. Raster estimate of future cooling demand in service sector buildings

Cooling demand in service sector buildings - offices, hotels, retail, and health related buildings - was estimated using a Monte Carlo simulation for the present day and under different future climate change scenarios for Switzerland [9,10]. Only service sector buildings were modelled as these are the only ones for which reliable consumption and cooling system adoption data was available. In the residential sector, while there are sources to estimate cooling demand intensity based on the experience of other countries, the adoption of such systems (which is a major determinant of final cooling demand) is highly uncertain.

Temperature projections for 2050 under representative concentration pathways (RCP) RCP2.6, RCP4.5, RCP8.5 were used in addition to a reference scenario for the year 2015. By using Monte Carlo simulation, means and standard deviations for cooling demand in the different scenarios were calculated. Because of the lower number of buildings in the service sector, results were aggregated at a lower 400x400m resolution for each service building category giving 6778 pixels, 33’890 points with unique pixel and sector for each climate scenario (Table 5).

Table 5 Summary of fields for cooling demand simulation results for 2015 and 2050.

| name          | description                                                                 | unit |
|---------------|-----------------------------------------------------------------------------|------|
| x             | x coordinate of pixel in CRS EPSG:2056                                      | m    |
| y             | y coordinate of pixel in CRS EPSG:2056                                      | m    |
| sector        | Sector of building use                                                      | -    |
| rcp           | Climate scenario for the simulation result (only for 2050)                  | -    |
| q_c_useful    | Estimate of useful cooling need (thermal energy to be removed from conditioned space). | kWh  |
| q_c_useful_std| Standard deviation of cooling need estimate.                                | kWh  |

3. Challenges

3.1. Limitations on source data

Data needed for modelling is frequently not available in the form required for the modelling task. A common cause of this is that data relevant for energy modelling in the built environment is not originally collected for this purpose. For example, energy consumptions are usually collected for billing purposes – this may result in a mismatch between the desired modelling unit (e.g. a single contiguous building) and the boundary of the measured consumption (e.g. a residential unit within a building). Many errors may be introduced if the billing practices are not perfectly understood (e.g. if a single bill is given for multiple buildings connected to a single meter). Unfortunately, tracking all these issues is not possible in practice with the limited human and technical resources available. Statistical methods can be used to mitigate their impact, for instance by determining suitable aggregation levels that limit the influence of individual input data errors, as well as filtering for outliers. Experience has shown that a particularly powerful approach to mitigate this problem is to clearly and specifically define the scope of the research questions, the sets of inputs that are considered relevant (e.g. building types), and the scope of the validity of the outputs and conclusions. For example, in a study of building energy certificates [12], we removed statistical outlier buildings with extremely large energy consumption. These readings are not necessarily errors, as there were several large luxury residences within the dataset. However, these buildings are not representative of the bulk of the building stock – by clearly defining the research scope to explicitly exclude such dwellings we are able to better define how to clean the input data and what the scope of validity of the results is.

3.2. Archiving data in dynamic modelling context

Deliverables for the FEEB&D project, as well as increasingly common open data requirements from project funding more broadly, stipulate the delivery of output datasets. The intention behind these deliverables is twofold, firstly to facilitate reproducibility of project research using the same underlying data, and secondly to allow future research to build upon these outputs.
However, modelling for energy in buildings is not a straightforward case of receiving an input dataset and building a model directly on it. Instead, data from various sources may be integrated and transformed to produce an intermediate dataset. This can apply to several of the inputs needed for a model. Model outputs in turn can become inputs for further analysis. The processing, formatting, and storage methods (file types) may be deeply tied to the model itself. Statistical modelling methods, such as Monte Carlo simulations, generate very large intermediate datasets with randomized parameters.

This raises the question of what stage of the data processing should be archived at the outcome of a project. A simplistic approach to archive only the raw, unprocessed data is of limited value, since generally this data is available from the original sources. Therefore, a decision must be made for what level of partially processed data captures the added value from the analysis (i.e. that can capture the useful work and expertise that went into the data processing), while retaining utility for future work. It is a question of whether the main purpose of data archival is only to allow the original research to be reproducible, or instead to support future work.

3.3. Archival formats for large datasets
Data archival aims to use highly compatible file formats that do not depend on specific versions of certain software in order to be read. Plain text formats such as CSV are usually preferred for this, as they can be readily parsed using any present or future software. However, data volumes in energy modelling are growing rapidly. The complexity of the data is also increasing, causing issues even with supposedly simple text-based data through problems such as the use of different character sets (accented characters, non-Latin characters) as well as file format ambiguities (e.g. multiline fields, quoting/escape problems). This drives towards the use of more complex (binary) file formats, such as HDF and netCDF. However, in energy research there are no currently established standards and conventions for using these formats, unlike for example the CF conventions commonly used in climate research.

3.4. Data licensing and ownership
A further challenge to the production of open data archives is issues around data ownership and licensing. For example, in WP3 work was also performed to develop electricity demand maps and load curves. However, as this was developed with an industry partner, the results can only be shared through the partner with corresponding commercial limitations.

4. Future work
Much more data is needed in order to fully describe and model heating and cooling for Switzerland. The new Decarbonisation of Heating and Cooling in Switzerland (DeCarbCH) project includes a task to build a database of common datasets for the technical, cost, and emissions, and spatial (where relevant) characteristics for energy resources, conversion, storage, and demands. This will be achieved through a combination of collecting existing datasets and new research. Datasets will be managed and documented in a shared database in accordance with the Data Management Plan. The database aims to cover: i) Renewable energy resources: solar (heat and electricity), water bodies, waste incineration, wind, hydropower, building envelope efficiency, industrial excess heat, bioenergy, wastewater, and geothermal; ii) Energy conversion technologies: space heating and cooling heat pumps, industrial heat pumps, CHP/tri-generation systems; iii) Thermal energy storage: water tanks and pits (short term & seasonal), aquifer thermal energy storage, latent heat, novel, and mobile storage; iv) Energy transport: thermal networks, energy hubs; v) Energy demand: space heating and hot water, space cooling (see below), heat/cooling for industrial processes.

It is the ambition of this database to provide a common set of inputs for different energy models as well as a common repository for model outputs (which may in turn by used as inputs for other research). It aims to support the DeCarbCH project as well as the wider research community by facilitating the storage, discovery, and retrieval of relevant datasets. In the long term, it is hoped that management of such data archives can be handled by a relevant institutional stakeholder.
5. Conclusion

This work highlights some of the tangible output dataset generated in FEEB&D and aims to encourage others to use this information as a starting point for further research. This inscribes itself in the context of comparable parallel work in the SCCER Joint Activity for Scenario Modelling (JASM) project, which already highlighted the need and value for such kind of data sharing. Publishing, documenting, and promoting the existence of such datasets can help researchers in the future rapidly identify data relevant to their needs, reduce duplication of effort and enhance collaboration and sharing opportunities. It can furthermore contribute to increased consistency and rigour in energy system modelling by allowing different work to use a common basis of data.

Nevertheless, as highlighted in this work, there remain considerable challenges to producing datasets for thermal energy modelling. These occur at every step of the process – from the preparation and interpretation of source data, to processing and storage of datasets though models and analysis code, to releasing data in a way that is technically accessible and with suitable licensing permissions. Lessons learned in the process of archiving and publishing these datasets are important for future projects. There is increasingly a requirement from funding bodies for open data to be published as part of long-term data management plans. In this direction, future work under the aegis of the DeCarbCH project aims to develop comprehensive databases for thermal energy modelling.

References

[1] Streicher K N, Berger M, Chambers J, Schneider S and Patel M K 2019 Combined geospatial and techno-economic analysis of deep building envelope retrofit Journal of Physics: Conference Series vol 1343 p 12028

[2] Streicher K N, Padey P, Parra D, Bürger M C and Patel M K 2018 Assessment of the current thermal performance level of the Swiss residential building stock: Statistical analysis of energy performance certificates Energy Build. 178 360–78

[3] Streicher K N, Padey P, Parra D, Bürger M C, Schneider S and Patel M K 2019 Analysis of space heating demand in the Swiss residential building stock: Element-based bottom-up model of archetype buildings Energy Build. 184 300–22

[4] Streicher K N, Mennel S, Chambers J, Parra D and Patel M K 2020 Cost-effectiveness of large-scale deep energy retrofit packages for residential buildings under different economic assessment approaches Energy Build. 215 109870

[5] Swiss Federation RegBL | Registre fédéral des bâtiments et des logements

[6] Schluck T, Streicher K N and Mennel S 2019 Statistical modelling of the energy reference area based on the Swiss building stock J. Phys. Conf. Ser. 86 35–42

[7] Schneider S, Hollmuller P, Le Strat P, Khoury J, Patel M and Lachal B 2017 Spatial–Temporal Analysis of the Heat and Electricity Demand of the Swiss Building Stock Front. Built Environ. 3 53

[8] Anon JASM - Home Page

[9] Li X, Chambers J, Yilmaz S and Patel M K 2021 A Monte Carlo building stock model of space cooling demand in the Swiss service sector under climate change Energy Build. 233 110662

[10] Chambers J D, Azar E, Hollmuller P, Patel M K, Schueler A, Bouvard O, Schueler A, Scartezzini J-L, Azar E and Patel M K 2019 Evaluating the electricity saving potential of electrochromic glazing for cooling and lighting at the scale of the Swiss non-residential national building stock using a Monte Carlo model Energy 185 136–47

[11] Anon Data Packages | Frictionless Data

[12] Cozza S, Chambers J and Patel M K 2020 Measuring the thermal energy performance gap of labelled residential buildings in Switzerland Energy Policy 137 111085