Automated Generation of Building Stock Databases and High-Resolution Heat Load Profiles for Districts and Municipalities in Germany

Luis Blanco[1,2,3], Philip Groesdonk[1,2,3], Jacob Estevam Schmiedt[1,3] and Bernhard Hoffschmidt[1,2]

[1] *German Aerospace Center (DLR), Institute of Solar Research, Jülich/Cologne, Germany*

[2] *RWTH Aachen University, Chair of Solar Components, Aachen, Germany*

[3] *heatbrAIn GmbH, Jülich, Germany*

Abstract

Reliable data about the building stock and its heating demand are essential for deciding about the investments in the transformation of the heating sector. Following the practical requirements of planning and engineering offices, we designed a method to automatically generate heat demand and building energy models for entire municipalities or districts in Germany.

In this contribution, we present a compilation of data, tools, and methods that enable the generation of a building stock database with building-level spatial resolution and energy demand models that can be simulated in up to hourly resolution. Sources are data about LOD2 geometry, cadastre, building archetypes, and typical weather conditions that are fully open data throughout Germany, and a Random Forest method to inform construction periods for each building. An exemplary application to a small town gives insights about heat demand outcomes as well as about adaptations that were necessary to get consistent results.

Introduction

One of the biggest challenges to move away from fossil fuels is the transformation of the heat supply. Transporting heat energy over long distances is expensive, which is why the heat transition is being planned and implemented locally. In Germany, heating and hot water systems accounted for 200 million t_{CO2} in 2020, which was about 85% of the emissions of the residential sector and amounted to almost a third of the country's CO_2 emissions (Umweltbundesamt 2023).

Following international agreements on emission reduction, the German government has put in place several measures to decarbonize the heating sector. Municipal heat plans (*kommunale Wärmepläne* – KWP) are to be developed on a mandatory basis and municipalities are to submit their plans for the conversion of the heating infrastructure to achieve CO_2 neutrality by 2045. District heating utilities are obliged to make existing networks run on renewable energy from 2045 on. The transformation of existing networks as well as the design and the construction of new ones are subsidized under the federal funding scheme for efficient heat networks (*Bundesförderung effiziente Wärmenetze* – BEW).

Reliable data about the building stock, the heating demand as well as generation potential for specific locations are essential for all steps in the decision process of heat sector transformation in general, including the aforementioned use cases resulting from German legislation. To meet the demand of planning and engineering offices for high-quality data and models, we designed a method to automatically generate heat demand and building energy models for entire municipalities or districts, achieving geolocalized building-level spatial resolution and hourly heat demand resolution. The data resulting from our model can directly be connected to common tools used in practise, such as QGIS, with the goal to save time and costs and thereby to accelerate the heating transition. In this contribution, we present the tools and methodologies we combined for the purpose as well as the insights from an exemplary application.

Tools and Methodology

The state of the art includes various bottom-up methods for modeling building energy use on district or urban scale. Some are based on energy consumption data, some on theoretical demand values, others on a combination (Lim & Zhai 2017, Malhotra et al. 2022). This contribution focuses on demand models for two main reasons. First, they are not influenced by current user behavior that might change in upcoming years; therefore, they can provide comprehensive methods to generate building-specific extrapolations for future energy use. Second, data privacy considerations limit accessibility and usability of data generated from measured consumption.

In order to obtain realistic building energy models for existing buildings, information is required on:

- Geometry of the buildings (e.g. volume, external surfaces, roof type)
- Physical parameters (e.g. U-values, airtightness)
- Building function (e.g. office building, residential building)

The general process is outlined in Figure 1, where the entire workflow for the model is summarized in a total of seven steps going from the selection of the study area to the calculation of heat load profiles according to three different refurbishment state scenarios. The following subsections explain the method in more detail.

Combination of Datasets

Availability of information is regionally different. For application in Germany, we use a combination of official and open data sets, with the aim of automating the process so that the effort involved in the initial data collection for heat transition projects is substantially reduced. This includes:

- LOD2 3D building models
- Building and property polygons from the official cadastre (*ALKIS*)
- Years of construction from the census
- TABULA building archetypes published by Loga et al. (2016)
- OpenStreetMaps (OSM)
- Test reference years DWD & BBSR (2017)

As all these datasets are openly available, geometry (3D building models), use (cadastre and OSM) and building physics (TABULA) parameters can be derived for any given area within Germany. Similar approaches are state of the art and the basis for the model presented in this study (Fuchs 2018, Weiler 2022, Blanco et al. 2024, Gorzalka et al. 2021). The data sources have different spatial resolutions and data structures. Therefore, the compilation of a single database is not straightforward. The model presented in this study aims for the automatic integration of all the different data sources and information

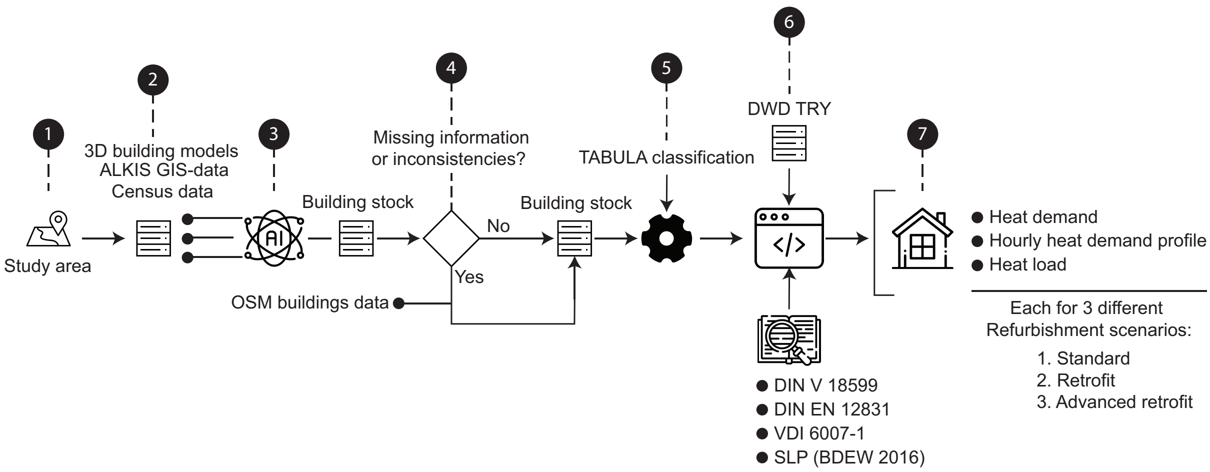


Figure 1: Schematic representation of the model's workflow. **1:** Selection of the study area. **2:** Import of all available open data about the study area. **3:** Restructuring the data and AI-supported dis-aggregation of information. **4:** Cross-check to ensure the database is complete; supplement with OSM data in case of missing data or inconsistencies. **5:** Classification of the building stock into the TABULA building types and assignment of three refurbishment scenarios per building. **6:** Import of DWD test reference years and other standard boundary conditions e.g. for building use. **7:** Results for 3 refurbishment scenarios including calculations of energy requirements and load profiles according to DIN/VDI standards.

about the building stock for a specific area and the direct calculation of the heat demand of individual buildings under different refurbishment scenarios.

Building Geometry

Building 3D Models are digital representations of a built facility including the geometry of building components at various levels of detail (LOD) as seen in Figure 2 (Biljecki et al. 2016). Given the importance of roof geometry and relationships to adjacent buildings, LOD2 provides a better base for urban building energy modeling than LOD1 models. LOD3 models with window and detailed structure information are barely available and much more computationally expensive to handle. In Germany, LOD2 data in CityGML format are generated based on cadastral footprints and digital surface models (Gruber et al. 2014). However, due to state-specific differences in cadastral data, there are variations in data structures.

Figure 2: The five different LoDs for 3D Building Models in CityGML (Biljecki et al. 2016). Licensed under CC BY-NC-ND 4.0 DEED.

Building Age

The census data is given in grid cells of 100 m×100 m with just one total aggregated value for residential buildings (Fig. 3). As it is normally the only source of information about the year of construction of a building, the assignment of building age classes to individual buildings is challenging. To overcome this limitation, we employ a previously developed method (Blanco et al. 2023) which involves training a Random Forest model on buildings with known ages and applying it to others. Open datasets containing information about years of construction and census hectare tiles with homogeneous ages prove to be valuable for training this model.

Building Physics

Another significant data gap lies in the absence of a source for small-scale allocation of refurbishments to buildings. To address this, we propose modeling all buildings based on the three refurbishment statuses outlined in TABULA (Loga et al. 2016). The

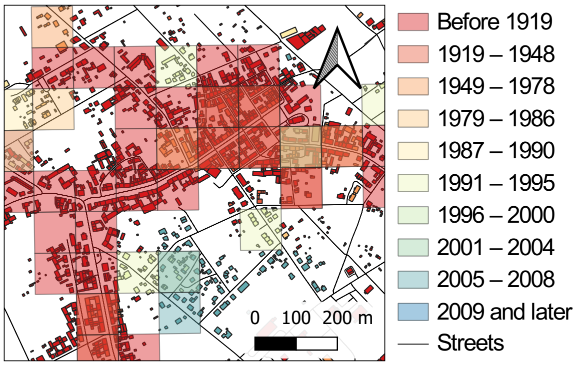


Figure 3: GIS-based data overlay of streets infrastructure (OSM), census data of building age in grid format with resolution of 100 m×100 m (Zensus2011, DL-DE/BY-2-0), and 3D building model footprints (LVermGeo-ST, DL-DE/BY-2-0) colored by building age after classification.

TABULA building typology classifies the residential building stock of Germany and other 20 European countries into different building types based on construction age, building type, location and refurbishment state. This approach for the generation of a building stock database enables the calculation of scenarios with varying current and future refurbishment states, providing essential insights for the planning process.

Weather Conditions

Local weather factors like ambient temperature and solar radiation faced by the buildings play a significant role in determining the total heat demand. We integrate DWD test reference years for average years, years with extreme winters, and such with extreme summers for 2015 and 2045. These datasets include hourly values for temperature, solar radiation, wind speed, and other parameters.

Building Stock Database

The extracted data from various sources is stored in a comprehensive database combining GeoPandas' GeoDataFrame and the open-source *TEASER* Python package (Remmen et al. 2018), with individual buildings serving as a common reference point and geometric neighborhood relationships between buildings to model the absence of heat transfer between adjacent building zones with similar use conditions. Missing building data and inconsistencies are cross-checked and/or supplemented with OSM data.

The database setup enables not only standardized determination of heat demand for useful heat and of heat load values according to the standards DIN V 18599-2 and DIN EN 12831-1 respectively, but also hourly resolved dynamic simulations of energy requirements and load profiles following the guideline VDI 6007-1. The dynamic simulations are run by using the open-source *AixLib* Modelica libraries mainly developed at RWTH Aachen (Maier et al. 2024). Domestic hot water profiles are generated from DIN V 18599-2 annual demand values and standard load profiles, while other possible outputs include map representations and sectioning proposals (BDEW et al. 2016, Blanco et al. 2024).

Practical Applications

The practical relevance of this methodology is highlighted by the fact that numerous municipalities in Germany are still in the initial phase of developing heating plans for their communities. Due to a lack of qualified personnel and standardized procedures, the presented methodology, along with its results, can expedite the first phase of these heating plans, which involves collecting data about the building stock and calculating heat demands. As presented in the introduction, further use cases of the methodology appear within the BEW scheme.

In addition to its use in research studies, our model has been applied in various projects across multiple states in Germany. For reasons of confidentiality, we have selected a random municipality in Germany to demonstrate the results. This case study illustrates the practical application and effectiveness of the presented methodology in real-world scenarios to provide insights into the energy demand and building stock of urban areas. The study area selected is Calvörde, a small municipality in the state of Saxony-Anhalt, with a total population of 3392, and 2084 residential buildings according to the ALKIS database.

Results and Discussion

Our application to Calvörde provides insights into input data quality, heat demand results, and comparisons between load profile and standard demand calculation results. There is a clear difference in data structure among different German states, posing challenges for the respective data integration. Unlike other states, Saxony-Anhalt does not provide specific function labels for non-heated building parts like garages. Consequently, most buildings in Calvörde's stock are labeled as residential. To address this, we applied a filter considering all parts of residential buildings with a height of less than 3.0 meters and/or without an address label as unheated (False), as shown in Figure 5. The heat demand map shown in Figure 4 (resulting from the application of the methodology of 1) repre-

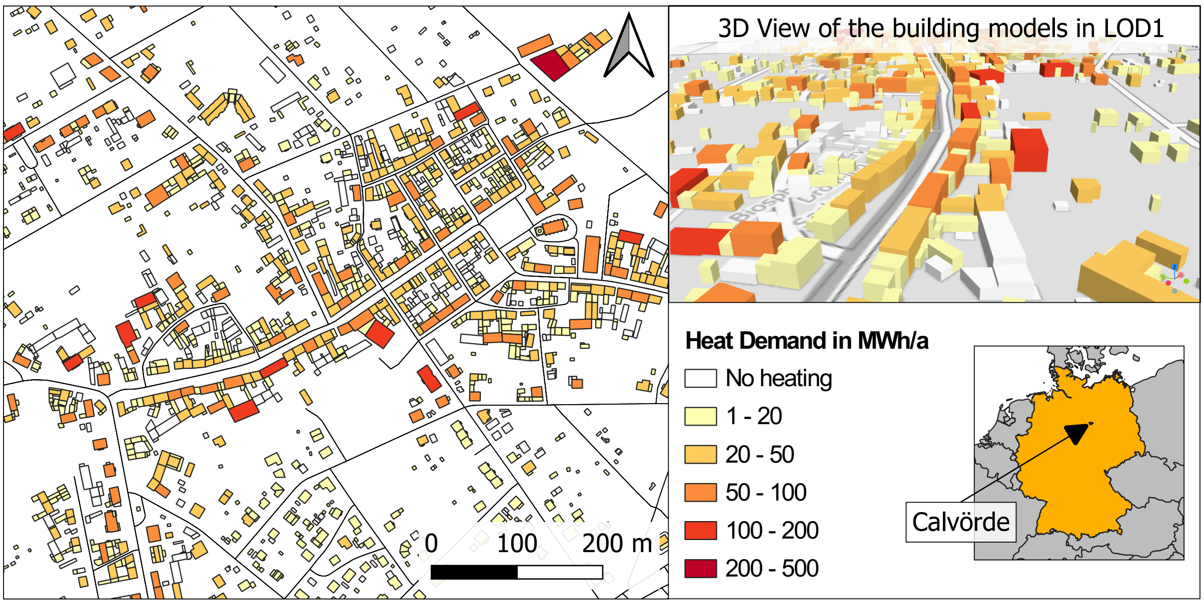


Figure 4: Study area of Calvörde in the state of Saxony-Anhalt showcasing the heat demand for individual buildings under the scenario of no refurbishment and a visualization of the 3D building models in a subsection of the city. Basemap.de and LVermGeo-ST 3D buildings licensed under DL-DE/BY-2-0.

sents the standard scenario of TABULA, i.e. without any retrofit. For the central town of Calvörde, useful heat demand for space heating and hot water in residential buildings sum up to 38.6 GWh/a in the **standard** scenario, 10.2 GWh/a in the **retrofit** scenario, and 6.0 GWh/a in the **advanced retrofit** scenario, considering the local 2015 test reference year. These numbers most likely overestimate the actual demand due to many unheated buildings that were not filtered out. In the practice, the presented model requires local adjustments to the cadastral function entries.

The simulated load profile for space heating in the three retrofit scenarios and for domestic hot water is shown for an exemplary building in Figure 5. For consistency in the practical application of these profiles, it should be made sure that their yearly integration matches the annual demand values calculated using the monthly energy balance method defined in DIN V 18599-2.

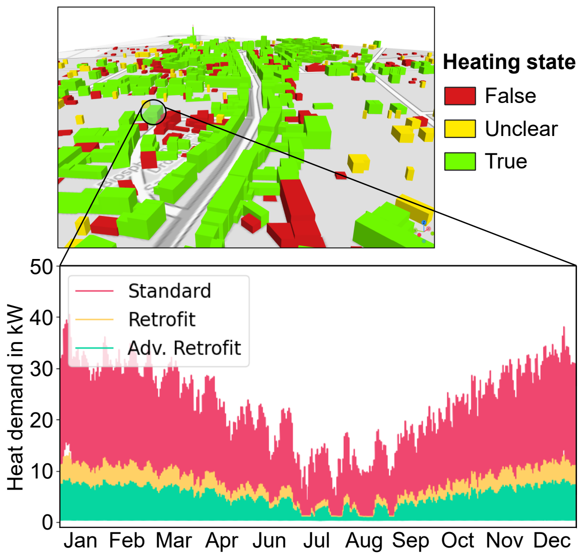
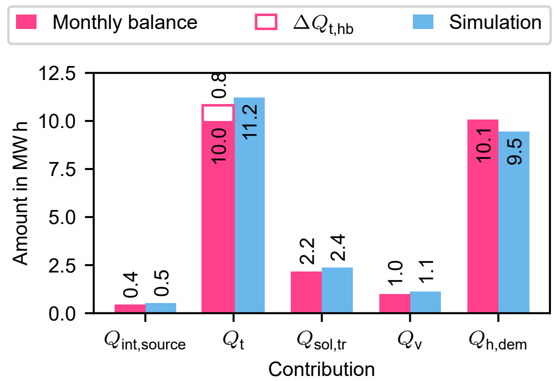
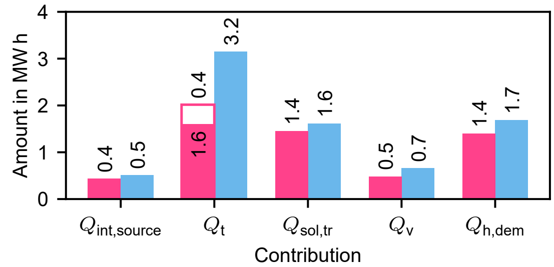


Figure 5: 3D view of heating state classification in Calvörde's old town, with hourly demand profiles for an example building for each refurbishment state.

Figure 6 presents a comparison of the contributions to the annual heat demand $Q_{\mathrm{h,b}}$ for both the *standard* and *advanced retrofit* cases. The figure contrasts the heat sources, including internal gains Q_{I} and solar gains through windows $Q_{\mathrm{S,tr}}$, against the heat losses, which occur through transmission Q_{t} and ventilation Q_{V}. Heat bridges are not covered by the AixLib reduced-order model. The transmission losses due to them $\Delta Q_{\mathrm{t,hb}}$ are therefore pictured separately. It must be noted that the different contributions do not simply sum up to the annual demand. For example, solar gains in summer contribute to the annual sum of heat gains through the windows, but do not reduce the heat demand significantly. To a lesser extend, deviations in the transmission and ventilation losses between simulation and monthly balance in months with a very low heat demand, do influence the contribution bars in the figures, but not the annual sum. From the comparison, it is concluded that the demand calculated by annual simulation matches the monthly balance values satisfyingly well. This was reached by harmonizing the consideration of longwave radiation, air infiltration, and heat transfer through the basement. Challenges remain regarding a slightly overestimated heat demand in the advanced retrofit scenario.



(a) Without retrofit (standard) scenario.



(b) Advanced retrofit scenario.

Figure 6: Contributions to annual heat demand calculated by monthly energy balance (DIN V 18599-2) and simulation.

Summary and Conclusion

This paper presents a comprehensive approach for generating a building stock database with building-level spatial resolution and energy demand models that can be simulated in hourly resolution. The method utilizes open data sources, including LOD2 geometry, cadastre, years of construction, building archetypes, and typical weather data, covering all of Germany. Additionally, we made adaptations to the cadastral data and calculation methods to ensure consistent results. The final method was applied to the town of Calvörde as an example.

In its current state of development, the presented methodology can be applied fully automatically to the modeling of residential buildings in Germany. Following the EU Open Data Directive, building data have been openly available throughout the country since June 2024. Manual adjustments to the basic data are possible and also necessary for non-residential buildings. Further development of the model includes the automation of this step as well as keeping it up-to-date with newly published input parameters. The integration of consumption data can contribute to a calibration of the model to real building conditions, in particular regarding retrofit states, provided that challenges in data quality and granularity can be solved.

Acknowledgments

The authors acknowledge the funding provided by Helmholtz Association under the Helmholtz Enterprise framework (project number HE-2022-15).

References

BDEW, VKU & GEODE (2016), ‘BDEW/VKU/GEODE Leitfaden – Abwicklung von Standardlastprofilen Gas’.

Biljecki, F., Ledoux, H. & Stoter, J. (2016), ‘An improved LOD specification for 3D building models’, *Computers, Environment and Urban Systems* **59**, 25–37.

Blanco, L., Aditya, M., Schiricke, B. & Hoffschmidt, B. (2023), Classification of building properties from the German census data for energy analyses purposes, *in* ‘Proceedings of Building Simulation 2023: 18th Conference of IBPSA’, pp. 817–824.

Blanco, L., Alhamwi, A., Schiricke, B. & Hoffschmidt, B. (2024), ‘Data-driven classification of Urban Energy Units for district-level heating and electricity demand analysis’, *Sustainable Cities and Society* **101**, 105075.

DWD & BBSR (2017), ‘Ortsgenaue Testreferenzjahre von Deutschland für mittlere, extreme und zukünftige Witterungsverhältnisse: Handbuch’.

Fuchs, D. (2018), Berechnung des Energiebedarfs von Wohngebäuden auf Basis von semantischen 3D-Stadtmodellen am Beispiel der Gemeinde Poing, Bachelor’s thesis, Technische Universität München.

Gorzalka, P., Garbasevschi, O. M., Estevam Schmiedt, J., Droin, A., Linkiewicz, M. M., Wurm, M. & Hoffschmidt, B. (2021), Collecting data for urban building energy modelling by remote sensing and machine learning, *in* ‘Proceedings of Building Simulation 2021: 17th Conference of IBPSA’, pp. 1139–1146.

Gruber, U., Riecken, J. & Seifert, M. (2014), ‘Germany on the Way to 3D-Cadastre’, *zfv – Zeitschrift für Geodäsie, Geoinformation und Landmanagement* **4**, 223–228.

Lim, H. & Zhai, Z. J. (2017), ‘Review on stochastic modeling methods for building stock energy prediction’, *Building Simulation* **10**(5), 607–624.

Loga, T., Stein, B. & Diefenbach, N. (2016), ‘TABULA building typologies in 20 European countries—Making energy-related features of residential building stocks comparable’, *Energy and Buildings* **132**, 4–12.

Maier, L., Jansen, D., Wüllhorst, F., Kremer, M., Kümpel, A., Blacha, T. & Müller, D. (2024), ‘AixLib: an open-source Modelica library for compound building energy systems from component to district level with automated quality management’, *Journal of Building Performance Simulation* **17**(2), 1–24.

Malhotra, A., Bischof, J., Nichersu, A., Häfele, K.-H., Exenberger, J., Sood, D., Allan, J., Frisch, J., van Treeck, C., O’Donnell, J. & Schweiger, G. (2022), ‘Information modelling for urban building energy simulation—A taxonomic review’, *Building and Environment* **208**, 108552.

Remmen, P., Lauster, M., Mans, M., Fuchs, M., Osterhage, T. & Müller, D. (2018), ‘TEASER: An open tool for urban energy modelling of building stocks’, *Journal of Building Performance Simulation* **11**(1), 84–98.

Umweltbundesamt (2023), ‘Kohlendioxid-Emissionen im Bedarfsfeld „Wohnen“’.

Weiler, V. (2022), Automatische Simulation von Wärmebedarf und -versorgung auf Quartiersebene, PhD thesis, Karlsruher Institut für Technologie (KIT).