Urban Building Energy Modelling with Combinatorial Optimisation and Microsimulation

Application and Policy Analysis for Hamburg, Germany

A dissertation submitted to HafenCity University Hamburg
in fulfilment of the requirements for the degree of

Doctor-Engineer

Dissertation vorgelegt im Promotionsausschuss der HafenCity Universität Hamburg
zur Erlangung des akademischen Grades

Doktor-Ingenieur

by/von

Ivan Dochev MSc

from/aus

Sofia, Bulgarien

defended/verteidigt

Feb. 2023

Doctoral Dissertation
Published on: 06.08.2023
Version: June 2023
HafenCity University Hamburg
Infrastrukturplanung und Stadttechnik
Author: Ivan Dochev MSc
First Advisor: Prof. Irene Peters PhD
Second Advisor: Prof. Ursula Eicker PhD
DOI: 10.34712/142.42

To my wife, Jenny, who put up with me. To my mom, who first taught me statistics. To my advisor Irene, who introduced me to microsimulation and from whom I learned so much professionally and personally. To my second advisor Ursula, a pioneer in the field of urban building energy modelling. To Lubow, Arne and Roland from the Hamburg BUE , with whom we had such a great cooperation on the *Wärmekataster.* To all the wonderful people I met at the HCU and HAW. To Donald and Ale. To my cousin Velian, who first taught me Python. To my colleague Hannes, who left too soon.

To my dad who never made it to Hamburg.

Summary

Urban Building Energy Models (UBEMs) are representations of building stocks. They can be used by analysts and decision makers in government, private companies, NGOs, and, if made publicly accessible, by the public at large. Their purpose is the energy analysis of the building sector – as it is, as it could and should be in the future, and paths to get there. Insofar as UBEMs are the basis for policy measures that affect many actors in the building sector they play an important role in those actors' communications and negotiations.

In the last two decades, it has become obvious that UBEMs need to explicitly account for space, just as they need to represent the micro-level, i. e. individual buildings. The analysis of centralised heat supply (district heating) calls for spatial analysis. Renewable energies are often used most effectively in decentralised fashion. The interaction of heat demand and supply unfolds at the level of the urban neighbourhood. Indeed, today's UBEMs are most often defined at the level of the individual building (though in publications, buildings are often clustered for data protection).

UBEMs come in different forms. Until recently, they were often created ad hoc for a specific research project or for an individual consulting service to local, regional or national government. The last few years have seen such models in more formalised fashion and made publicly accessible, as "heat cadastres" ("Wärmekataster", as in the city-state of Hamburg), or "Heat Atlas" ("Wärmeatlas", as in the federal Land Bavaria). These cadastres and atlases contain a variety of energy-related information, like buildings, infrastructure, heat and power generating stations, waste heat sources and more. In this thesis, I focus on the residential building stock and its heat demand.

The major challenge in creating a UBEM is the availability of data relating to the building stock. In Germany, there is a variety of data sources with different detail on building characteristics and at different levels of aggregation. Accordingly, UBEMs tend to be created with a Top-Down or a Bottom-Up approach. In simple terms, the former distributes total energy consumption at the city level onto all buildings according to their size. The latter makes use of building typologies and represents individual building characteristics and their specific heat demand by assigning "archetypes" to individual buildings. The Top-Down approach has the drawback of little flexibility in analysing measures that address individual building characteristics. UBEMs of the Bottom-Up type often overestimate total consumption as a consequence of how building types are defined and assigned to individual buildings.

I propose a hybrid approach, based on combinatorial optimisation, inspired and heavily borrowing techniques from the field of spatial microsimulation. My case study is Hamburg. I use three main data sources: the digital building cadastre (ALKIS), 100m Census raster cells, and a sample of building energy audits. When aggregated, my UBEM comes within a couple of percent of consumption data for all of Hamburg. As a micro-model with detail at the building level, it offers the flexibility to analyse various cases and scenarios. I illustrate a number of different applications of my UBEM that show how its main features – the micro-level, the georeferencing (each building with its geographic coordinates), and the consumption-adjusted heat demand – facilitate the analysis of a spectrum of energy policy measures.

Zusammenfassung

Urban Building Energy Models (UBEMs) sind Datenmodelle, die Gebäudebestände abbilden und von Analysten und Entscheidungsträgern in Behörden, Wirtschaft, NGOs und, falls öffentlich zugänglich gemacht, von der breiten Öffentlichkeit verwendet werden können. Sie dienen der energetischen Analyse des Gebäudesektors, sowohl was seinen Ist-Zustand, als auch, was seine Ziel-Zustände und Wege dorthin betrifft. Als Grundlage für Politikmaßnahmen der öffentlichen Hand, die viele Akteure berühren, fällt UBEMs eine wichtige Rolle in deren Verständigung zu.

In den letzten zwei Jahrzehnten ist die Notwendigkeit für die explizite Darstellung des Raums in solchen Modellen deutlich geworden, ebenso wie die explizite Berücksichtigung der „Mikro"-Ebene. Die Analyse der netzgebundenen Wärmeversorgung erfordert die räumliche Betrachtung. Erneuerbare Energien werden oft günstig dezentral genutzt. Das Zusammenspiel zwischen Angebot und Nachfrage erfolgt auf der Ebene des Quartiers. UBEMs werden heute meist auf der Ebene des einzelnen Gebäudes definiert (wobei in Veröffentlichungen die Gebäude aus Gründen des Datenschutzes in Clustern oder Gruppen dargestellt werden).

Solche UBEMs gibt es in unterschiedlichen Formen. Bis vor kurzem wurden sie vielfach ad hoc für ein bestimmtes Forschungsprojekt oder als Dienstleistung für lokale, regionale oder nationale Verwaltungen erstellt. In den letzten Jahren sind solche Modelle mehr formalisiert und öffentlich zugänglich gemacht worden, als „Wärmekataster" (z.B. in Hamburg) oder „Wärmeatlas" (z. B. in Bayern). Diese Kataster und Atlanten können eine Vielzahl von energiebezogenen Informationen abbilden – Gebäude, Infrastruktur, Energieerzeuger, Abwärmequellen und mehr. Darin können sie über ein UBEM hinausgehen. In meiner Arbeit konzentriere ich mich auf den Wohngebäudebestand und dessen Wärmebedarf.

Die große Herausforderung bei der Erstellung eines UBEMs ist die Verfügbarkeit von Daten zum Gebäudebestand. In Deutschland gibt es eine Vielzahl von Quellen in unterschiedlicher Detailschärfe und auf unterschiedlichen Aggregationsebenen. Dementsprechend werden sowohl der „Top-Down"- als auch der „Bottom-Up"-Ansatz bei der Erstellung von UBEMs verfolgt. Ersterer verteilt (vereinfacht gesprochen) den Gesamtenergieverbrauch auf Stadtebene auf alle Gebäude entsprechend ihrer Größe; letzterer nutzt Gebäudetypologien, um energetische Merkmale einzelner Gebäude abzubilden und ihnen typische spezifische Wärmebedarfswerte zuzuordnen. Das Problem des „Top-Down" Ansatzes ist seine mangelnde Flexibilität, wenn es um die Analyse von Maßnahmen geht, die detaillierte Gebäudeeigenschaften adressieren. UBEMs mit „Bottom-Up" Ansatz dagegen überschätzen oft den Gesamtverbrauch aufgrund der Art und Weise, wie Gebäudetypen definiert und einzelnen Gebäuden zugeordnet werden.

Ich schlage einen hybriden Ansatz vor, der auf kombinatorischer Optimierung und Methoden der räumlichen Mikrosimulation basiert. Mein Anwendungsfall ist Hamburg. Ich verwende hauptsächlich drei Datenquellen: Das digitale Gebäudekataster, 100-Meter-Zensus-Rasterzellen sowie eine Stichprobe von Gebäude-Energieausweisen. Aggregiert kommt mein UBEM sehr nah an die Verbrauchsdaten auf Stadtebene heran. Als Mikromodell mit Detail auf Gebäudeebene bietet es die Flexibilität für die Analyse unterschiedlicher Fragestellungen und Szenarien. Ich illustriere verschiedene Anwendungen des UBEM, die zeigen, wie seine Hauptmerkmale – die „Mikro"-Ebene, Georeferenzierung (jedes Gebäude mit seinen geografischen Koordinaten) und verbrauchskorrigierter Wärmebedarf – eine Bandbreite energiepolitischer Analysen ermöglichen.

Prologue

We live in fast times. Conveying information is ever easier, while doing it efficiently seems to have become more difficult. A wise person taught me the quote: "I am in a hurry friend, so my letter is long". With that in mind, here I am listing, in short, some misconceptions and "gotcha" moments related to urban energy modelling that my research confronted me with. All of these I am addressing in the pages that follow. Perhaps these will stir you, my reader, to read on. Or at least leave you with a little more food for thought on the topic.

- "*Urban building energy modelling concerns buildings*". Yes, however, counting buildings per se is tricky. There are multiple definitions about where a building ends and a new one starts. The German Census counts "entrances", the cadastre counts "objects". No, they do not correspond or add up. Best is to use the area, as in X m^2 of floor area. But be careful if it is residential floor area or total floor arear or net (not counting walls), etc.

- "*Heat Demand is well defined and when people discuss it, everyone is on the same page about what it is, like Pi.*" No, heat demand may refer to a dozen different metrics – useful heat (what comes out of the radiators), delivered heat (what comes out of the boiler), final energy (what goes in the boiler). Further, if it is specific (per m^2), then specific to which area (see above)? Sometimes it is used interchangeably with heat consumption, when the two are very different.

- "*3D data is better than 2D data*". If a 2D cadastre has information on the height of buildings or number of storeys it could potentially be more precise than an actual 3D object in some known 3D format, like CityGML. The fact that something has the 3rd dimension in its geometry, does not necessarily mean it is *correct*. The fact that something is represented as a 2D polygon does not necessarily mean it could not have good information on the 3rd dimension.

- "*The IWU residential building typology is not good, because I have tried to apply it to a building stock and it overshot the total consumption from other sources*". To that I would ask – How did you apply it? Did you use all renovation levels when applying the typology or did you use only the "baseline" state? Further note that the IWU Typology presents typical buildings, not average buildings. The difference is the same as between a statistical mode and a statistical mean.
  That being said, the underpinnings of the approach to characterise a building stock with a typology are, at this point in 2022, very shaky. This method was devised in 1990s, when energetic building renovations were just starting. At that time the majority of the buildings were, energetically, much more similar to when they were built. Nowadays, a majority of buildings have had some form of improvement. They no longer fit nicely into types.

Table of Contents

List of Tables

List of Figures

List of Formulas and Equations

List of Abbreviations

ADE	Application Domain Extension
AGEB	*Arbeitsgemeinschaft Energiebilanzen e.V.*
ALKIS	*Amtliches Liegenschaftskatasterinformationssystem* (German cadastral system)
APD	Absolute Percent Deviation
ARGE	*Arbeitsgemeinschaft für zeitgemäßes Bauen e.V.*
aSHC	Average Specific Heat Consumption
AW	*Außenwände* (outer walls)
BIM	Building Information Modelling
BUE	*Behörde für Umwelt und Energie* (Hamburg Ministry of Environment and Energy)
CO	Combinatorial Optimisation
DA/OG	*Dach/Obergeschoß* (roof, attic floor)
DHW	Domestic Hot Water
EFH	*Einfamilienhaus* (single-family house)
FE	*Fenster* (windows, glazing)
EnEV	*Energieeinsparverordnung* (a German energy efficiency ordinance)
GIS	Geographical Information Systems
GMH	*Großes Mehrfamilienhaus* (Large multifamily building)
HAW	*Hochschule für Angewandte Wissenschaften* (university of applied sciences)
HCU	HafenCity University
HH	*Hochhaus* (Highrise building (7+ storeys))
IFBHH	*Hamburgische Investitions- und Förderbank* (Hamburg state investment bank)
IPF	Iterative Proportional Fitting
IWU	*Institut Wohnen und Umwelt* (a German scientific institute)
KE/FB	*Keller/Fußboden* (cellar or base floor)
LiDAR	Light Detection And Ranging
MAPD	Mean Absolute Percent Deviation
MFH	*Mehrfamilienhaus* (multi-family building)
NNI	Nearest Neighbour Index
NP	Nondeterministic Polynomial-time
PV	Photovoltaic
RH	*Reihenhaus* (row house or terraced house)
SHC	Specific Heat Consumption
UBEM	Urban Building Energy Model
WSVO	*Wärmeschutzverordnung* (another German energy efficiency ordinance)
WDVS	*Wärmedämmverbundsystem* (Exterior insulation and finish system (EIFS))

List of Related Publications

As part of my research, I have published a number of research articles. None of these feature the algorithm presented in this thesis, but all relate to the overall topic and complement my work. I will list the most relevant of those with a short explanation of how the content relates to this thesis:

- Dochev, I., Gorzalka, P., Weiler, V., Schmiedt, J. E., Linkiewicz, M., Eicker, U., Hoffschmidt, B., Peters, I., & Schröter, B. (2020). *Calculating urban heat demands: An analysis of two modelling approaches and remote sensing for input data and validation.* Energy and Buildings, 110378. https://doi.org/10.1016/j.enbuild.2020.110378

  In this publication, together with colleagues from the HfT Stuttgart and the DLR Berlin we looked at different ways to estimate heat demand for a neighbourhood in Berlin. It is as part of this research that I looked into the Berlin ALKIS and its peculiarities (see Section 6.4)

- Dochev, I., Seller, H., & Peters, I. (2020). *Assigning Energetic Archetypes to a Digital Cadastre and Estimating Building Heat Demand. An Example from Hamburg,* Germany. Environmental and Climate Technologies, 24(1), 233–253. https://doi.org/10.2478/rtuect-2020-0014

  A large part of my knowledge on the TABULA heat demand estimation method and the IWU Typology was acquired as part of the research presented in this paper (see Section 6.2).

- Dochev, I., Seller, H. & Peters, I. (2019). *Spatial aggregation and visualisation of urban heat demand using graph theory*, International Journal of Sustainable Energy Planning and Management, no. 24. https://doi.org/10.5278/ijsepm.3346

  This paper concerns the use of UBEMs and more concretely the need for aggregating the buildings into groups to satisfy data protection requirements when publishing the model as a "heat demand cadastre".

- Dochev, I., Peters, I., (2019). *Potential for utilising near surface geothermal heat via heat pumps. A case study from Hamburg,* in Wittmann, J. et al.: Simulation in den Umwelt- und Geowissenschaften - Workshop Kassel 2019, Shaker, Aachen

  This paper exemplifies the benefits of having a georeferenced UBEM when analysing the potential for renewable energy (near surface geothermal in this case).

- Dochev, I., Peters, I., Seller, H., & Schuchardt, G. K. (2018). *Analysing district heating potential with linear heat density. A case study from Hamburg.* Energy Procedia, 149, 410–419. doi:10.1016/j.egypro.2018.08.205

  Similarly to above, this one relates to analysing neighbourhoods suitable for district heating.

1 Introduction & Literature

1.1 Motivation: The Role of Urban Building Energy Models (UBEMs)

Around the world, the building sector makes up around 20% of final energy consumption (U.S. EIA, 2016, p. 101); in Germany, that share is 40% (BMWi, 2017, p. 16). With such high shares, policies that target building energy are an important part of climate protection.

For decades, models for energy policy analysis have depicted building energy at the national level (Martinsen, Krey, et al., 2007; Martinsen, Linssen, et al., 2007). More recently, modelling of building energy has taken the city into focus. Municipalities play an important role in building stock development (an objective of urban planning) and heat supply (a local affair, due to the nature of heat, and often provided by municipally controlled companies). With the advent of Geographic Information Systems (GIS), Urban Building Energy Models (UBEMs) nowadays often represent buildings together with their spatial location and thus facilitate the analysis of neighbourhoods, allowing heat planning and urban development measures to be tailored to specific spatial constellations.

1.2 Urban Building Energy Models (UBEMs)

1.2.1 A note on terminology

Firstly, a clarification of some terms is in order. What I mean by an Urban Building Energy Model is a *dataset* representing a building stock as opposed to f. ex. a statistical model[1] or a machine learning model (which usually represent relations between variables). A UBEM consists of data structured in a specific way to represent real-world objects (in this case, usually buildings) and their characteristics specifically related to energy[2]. Statistical models may play a role in the field of UBEMs with regard to energy demand estimation. However, I do not regard them as part of a UBEM itself.

The *urban* part in UBEM refers to the geographic scope of these models – usually an entire city[3]. The term *building* conveys, that although a city or country is modelled, the basic unit of analysis is the building, there are just many of them represented in the model[4]. A good overview of UBEMs is given by Reinhart and Cerezo (2016), Lim and Zhai (2017) and Li et al. (2017).

---

[1] For example, of the sort $Y_i = f(X_i, \beta) + e_i$ (where e is a probability distributed variable)

[2] A closely related, but more general term is *Urban Energy System Model* (Keirstead et al., 2012). It refers to modelling urban energy systems, of which the buildings are a part, but also transport, district heating and electricity grids and other energy systems.

[3] A UBEM can technically encompass any spatial unit, a city, a neighbourhood or even a region or a whole country. Whether then "urban" is appropriate or should be substituted with "regional" is a matter of semantics. The key attribute is the building as the base unit.

[4] Note that representing a city with f. ex. three building types, each with a given weight, representing the number of such buildings in a city can be a "grey area" in the context of terminology. Such a model is definitely *urban*. Whether the aggregated representation of the buildings is enough to say the unit of analysis is the building is a valid question. I do consider it enough. However, I do not believe the semantics of it are critical enough to explore further.

A *simulation* is the imitation of a process incurring with or within the elements of the model. For example, deciding on how to represent a building (as an object with attributes or as an object consisting of other objects with attributes etc.) and filling the data structure with values would be part of *urban building energy modelling.* Using a UBEM to estimate the heat demand for each building could then be a *simulation*. However, there are technical differences between *simulation* and *calculation*. A simulation replicates the state transitions of a (dynamical) system, or model by computing the state of the system one time step after the other. The result for each time step depends on the preceding one. A calculation is simpler in that each state is computed in isolation from the previous state, hence no real "dynamics"[5]. For what I will do in this thesis, I argue *calculating energy demand* would be the most fitting term for this particular step. Lastly, in the context of this thesis, an *algorithm* is a set of rules, or steps, that a computer follows or performs. An algorithm can obviously be used to do *simulations* and *calculations*, but it can also be used to do *urban building energy modelling* (i.e. prepare the UBEM).

1.2.2 Purpose of UBEMs

The purpose of UBEMs falls under the general term *data driven urban planning*. It relies on the assumption that the more and better data is available, the better policy and planning decisions can be made. For illustrations see (García Kerdan et al., 2017; Martinsen & Krey, 2008; Vásquez et al., 2016). As an example for a national analysis, Sandberg et al. (2017) calculate different scenarios for the Norwegian building stock. In a simulation for the period 2016 to 2050, they obtain a decrease of yearly delivered building energy of 23% under the baseline building renovation scenario. The authors show this can be increased to a 52% reduction but with the integration of more heat pumps and photovoltaics, while increased renovation rate or deeper renovations would only bring very marginal decreases in delivered energy[6]. Such, more subtle conclusions provide important guidance for decision makers, pointing to possibly erroneous commonly held assumptions which stand in the way of effective climate protection. Note that this conclusion is for the Norwegian building stock, neither the authors nor I claim this to be a conclusion for buildings stocks in general.

1.2.3 Types of models, state of the art

Urban building energy modelling has to deal with the basic drawback that data on energy-related building characteristics are generally not available at the individual building level (Sola et al., 2020). The modelling therefore is first concerned with filling the data gaps. Decisions on data structure and richness of detail then follow from what could be accomplished from the first step. This has led to a

[5] Discussing the differences between „simulation" and „calculation" is a rather deep rabbit hole. Different fields of science would have different definitions. UBEMs are in many ways a cross-field, so one may encounter engineers, architects, computer scientists and physicists, all taking part. Because of this, I do not argue that the definition I give in the text is the one single source of truth. Only that it is useful in this context and that it helps explain what I actually do as opposed to what I do not do.

[6] Note that delivered energy is usually the energy at the point of sale or entry in the building/property. Thus, renewable energy integrated in the building usually decreases delivered energy, since it reduces the amount of energy needed "from the outside".

couple of different approaches to urban building energy modelling. Li et al (2017) categorise them broadly as “top-down” and “bottom-up”.

“Top-down” UBEMs are usually prepared (*modelled*) using total building energy consumption at coarse spatial scales – available from national or state-level energy balances, f. ex. in Germany the s. c. “energy balances” (AGEB e.V., 2018). The energy consumption is then distributed to finer-scale spatial units (individual buildings, neighbourhoods or raster cells) based on known characteristics of the latter, like population count or floor area. The upside of this approach is that it is, by definition, valid at coarser scales. The downside is that it has limitations when it comes to the analysis of policy scenarios, as such models are not flexible enough to simulate changes in the building stock in detail.

“Bottom-up” modelling uses characteristics of individual buildings found in digital cadastres[7] to create a model. Building size, shape and location is generally available either through extruded 2D cadastres or measured 3D observations (LiDAR or photogrammetry). However, there are no or extremely few exhaustive[8] data sources with energetic information – heat transmissivity of building shell, types of heating systems, appliances etc. To fill this data gap a building typology approach or “archetype” approach is very often used (Reinhart & Cerezo, 2016). A building typology is derived usually from sample observations at national or regional scales and includes information (building attributes) that can be matched to the ones in the cadastre. It also includes the information that is not in the cadastre, but is energetically relevant and considered *typical* of a certain type (f. ex. transmissivity of building envelope typical for buildings built in 1920s). By matching the objects in the cadastre to the types by their common attributes, each building in the cadastre receives a type and “inherits” the needed energetic information from it. In this way, the bottom-up models are much more flexible, since each building can be represented as detailed as the building types are defined.

Bottom-up models can be further divided into “statistical” and “engineering” (Lim & Zhai, 2017). The differentiation concerns the way the energy demand is calculated. Statistical models use usually a regression model with the building attributes as independent variables and the energy demand as the dependent variable. They require large datasets with building and consumption data to set up the regression model. Engineering bottom-up UBEMs calculate the energy demand from the building attributes via an energy balancing calculation (f. ex. the German DIN 4108 or DIN 18599) or via a dynamic simulation (f. ex. using software like EnergyPLUS, TRNSYS or others).

Here I would argue, it is useful to consider this step (the estimation of the energy demand) as performed <u>with</u> a UBEM, rather than as part of creating the UBEM. One could use the same bottom-up UBEM to perform both a statistical and an engineering energy demand estimation. Therefore, the energy demand estimation should not be part of the definition of the UBEM. Viewing the UBEM preparation and energy demand estimation as two separate steps is also useful when considering the last

[7] Digital ledgers of buildings/real estate.
[8] Including all units within the scope of analysis, f. ex. all buildings in a city, as opposed to a sample of observations

differentiation found in the literature – probabilistic and deterministic UBEMs. The key issue is the general unavailability of data in the face of the data-hungry calculations, simulations and analysis that people need to perform. The archetype approach can only go so far in filling the necessary data gaps and after results stray far from external validation sources, researchers turn to probability to tackle the uncertainties (f. ex see Sokol et al (2017, p. 13)). The uncertainties mentioned in the literature can be summarised in two groups – building related (heat transmissivity coefficients, heat systems, appliances) and occupant related (ventilation, temperature, use of appliances). I would add a third – energy estimation method. With the same input data, a static heat balancing with DIN 4108 would produce different results from an EnergyPLUS simulation. Probabilistic models use distributions, rather than fixed (assumed, estimated or plainly guessed) values for building- and occupant-related model attributes. Uncertainties are thus analysed and the model is calibrated to known external validation sources.

1.2.4 Challenges

Building-related

Building-related challenges in creating UBEMs stem from the performance of the archetype approach – how are types defined and assigned. To quote:

"While the actual division of a building stock into archetypes is obviously of paramount importance for the reliability of the resulting UBEM, the process typically remains ad hoc, relying on generic assumptions" (Reinhart & Cerezo, 2016, p. 198).

"The largest remaining uncertainty for UBEM simulations is associated with the definition and detailed description of archetypes that reliably represent a building stock" (ibid, p. 199)

More concretely, the archetype approach suffers again from data unavailability. Construction epoch usually serves as the main signal for building envelope efficiency, because types of materials and their energetic properties are associated with certain epochs and the introduction of energy efficiency regulation. Even though construction epoch of buildings is not widely available in digital cadastres, this is a manageable problem, some data exists. However, buildings continuously undergo changes: modernisation, exchange of windows, thermal insulation of the façade and installation of new heating and domestic hot water systems. This means that archetypes defined using a matrix of building size and building epoch ("single-family house (EFH)[9] built in the 1920s")[10] are far from homogenous. In reality, buildings of the same archetype can vary greatly in key energetic parameters. A remedy would be to extend the typology with one more dimension – renovation level (as in "SFH built in the 1920s with renovations"). This has two issues. Firstly, German digital cadastres (as most European

[9] I refer to a "single-family house" as EFH, coming from the German *Einfamilienhaus.* This is to preserve consistency with the names of the types the IWU Typology, which is a widely used building Typology in Germany.

[10] "EFH built in the 1940s", "Multifamily house (MFH) built in the 1920s" etc.

cadastres) do not include information on the state of renovation of buildings. Therefore, even if such archetypes are defined, they cannot be matched with the buildings in the cadastre. Secondly, there is high variability in renovations. A building owner could exchange only windows, or add only insulation, or only exchange a gas boiler or only add a PV-panel or any combination of these. Moreover, all of these can have different types that come with different levels of energy efficiency. The number of combinations explode and with them the number of archetypes. However, this conundrum can be tackled with the computing power of even an ordinary modern laptop, as I will show further in this thesis.

All of this refers mostly to residential or, at most, institutional buildings where people do deskwork or simply reside, like offices or schools. The archetype approach is generally not suitable for industry and more specialised buildings. These are also beyond the scope of this thesis.

Occupant-related

Occupant-related challenges concern the assumptions about the behaviour of building occupants in the energy demand estimations. What room temperature do occupants prefer? How do they operate the heating equipment to obtain it? How often do they ventilate? These challenges are generally tackled with probabilistic approaches. My colleague Esteban Muñoz has done great work in advancing a different, more novel approach to this. He used spatial microsimulation (a technique from economics, see Section 6.1) to model buildings together with their occupants in a deterministic way (Muñoz et al., 2016; Muñoz & Peters, 2014). The microsimulation technique has a tradition in health and tax policy analysis and has been extended to account for spatial effects, but had not been applied to buildings by then.

The demand-consumption discrepancy

An additional challenge is the often observed and cited (see f.ex. Yan et el (2015)) discrepancy between demand and consumption. "Demand" is a calculated value based on the properties of a building and, in most cases, an assumed standard ("norm") occupant or building user. Consumption is a measured (metered) value. With complete information on the building and the occupant and a perfect simulation of the thermal behaviour of the building, demand and consumption should be the same.

However, even after an on-site building inspection, complete information on a building's characteristics is close to impossible to obtain. Numeric values are estimated by categorising the type of masonry or by investigating the production year of the windows and checking in catalogues. More precision could be achieved with thermal cameras and generally more precise instruments, but few purposes would allow expending this effort.

Standard energy audits generally do not model occupant behaviour. Applying a simple percentage correction factor derived from metered consumption is possible in some cases, although it is generally considered that a building's energy efficiency should not depend on the user and an energy audit should objectively evaluate the building itself. A problem arises when energy audits are used for calculating returns on investment of energy efficiency measures. If the user is more energy-conscious, actual energy savings would be lower than projected, since status quo is lower to begin with. For this, the mentioned consumption corrections are sometimes applied. Additionally, "rebound" effects play a role in the estimation of savings. With a more energy-efficient building, occupants tend to increase comfort whereby decreasing savings.

Standard computation methods (DIN 18599 or DIN 4108) are "steady-state" or "static" heat balancing methods, which are inferior to dynamic simulations, but considered good enough for the purposes of building planning and retrofitting. More elaborate models and simulations are usually needed for smart building control systems or in highly efficient passive[11] buildings where all aspects of the building design must be extremely optimised.

1.3 Hamburg UBEM Status Quo

Regarding UBEMs for Hamburg, Ecofys Germany GmbH produced a study on the status quo of the Hamburg building stock around 2013 (Ecofys Germany GmbH, 2013). The GEWISS Project (Geographical Heat Information and Simulation System (2019)) created a simulation tool to explore Hamburg building stock scenarios. The *Behörde für Umwelt und Energie* (*BUE,* Hamburg Ministry for Environment and Energy) published an online *Wärmekataster* in the fall of 2017 (Behörde für Umwelt und Energie [Hamburg Ministry for Environment and Energy], 2017). All three of these efforts are "bottom-up" in that they use data from the Hamburg digital cadastre, combined with an energetic building typology. GEWISS and *Wärmekataster* use the *IWU* (*Institut Wohnen und Umwelt*) typology (Loga et al., 2015), while Ecofys used their own typology.

The *Infrastrukturplanung und Stadttechnik* group (Technical Urban Infrastructure Systems Group) at HafenCity University with which I was affiliated at that time was responsible for the "Data Integration" work package in the GEWISS Project. Through the close cooperation between the group and the *BUE*, I have also been the main author[12] of the methodology for the Hamburg *Wärmekataster* (Behörde für Umwelt und Energie [Hamburg Ministry for Environment and Energy], 2017).

2 Objectives and Research Question

One problem of UBEMs in Hamburg is that not one bottom-up model managed to come close to the actual consumption values provided by the local statistical office. All models went around the problem in different ways. The *Wärmekataster* depicted only "useful" heat, since information on energy systems was not available and they were thus not modelled. The problem of the unknown building renovations was worked around by calculating the entire building dataset with three levels of renovation and

[11] "Passive" buildings refer to an unofficial yet widely adopted energy efficiency standard (Passive House Institute, n.d.)

[12] Which would not have been possible without the prior work of my colleague Esteban Muñoz.

displaying all three. Interested parties have to choose which one is more representative of the area they were interested in. A fine work-around, but not solving the problem.

The GEWISS project used a (second[13]) consumption correction to make the model match the totals for Hamburg. For the older buildings built before 1978, which I suspected are the target of most existing renovations, I reduced the heat demand so that the city total moved closer to the values of the statistical office. Not ideal, since this means I averaged over all older buildings, so evaluating renovation policy was more difficult.

Ecofys overestimated the residential heat demand by 20% but stated that the statistical office might have got it wrong (Ecofys Germany GmbH, 2018, p. 13). Although interpreting the numbers of the statistical office is not without its challenges, I highly doubt that this was the main issue.

All models "skirted" around the problem of renovations - both of envelope and heating supply system. I set out to explore if it can be (at least partially) solved. A possibility existed in the integration of other data sources, which although perhaps known, were difficult to integrate[14]. In the end, I present practical use cases for an improved UBEM.

The research question therefore has two parts:

A. ***How can modelling and estimation techniques for data integration be used for improving Urban Building Energy Modelling (UBEM)?***
B. ***What insights for energy policy in Hamburg could a thus improved UBEM provide?***

3 Data

3.1 Availability

The appropriate methods have to come from the concrete task and since data integration and modelling is the task, I will first list what data is available in Hamburg:

1) ALKIS (the acronym for *Amtliches Liegenschaftskatasterinformationssystem*) is the digital cadastre of the city of Hamburg (Freie und Hansestadt Hamburg, Landesbetrieb Geoinformation und Vermessung (LGV) [Hamburg State Office for Geoinformation and Surveying], 2020). It consists of (among many other things) 2D polygon representations of buildings (see Figure 1), their number of storeys and, for some buildings (around 50%), a construction epoch. Comparing samples with other sources, for example energy audits, reveals that the construction epochs in the ALKIS are not without mistakes.

[13] The buildings in the IWU Typology come with a consumption correction themselves.
[14] I worked on the „Data Integration" package in the GEWISS Project. Although I could not finish work on this thesis prior to the project's end, my work there was paramount to my dissertation.

Figure 1. Example of ALKIS buildings data

2) Census Data – 100m raster cells (see Figure 2) for the entire Hamburg area with cross-tabulations[15] of different building characteristics – construction epoch, type of heating and type of ownership (Statistische Ämter des Bundes und der Länder [German Federal Statistical Office], 2018). These, I refer to, as "benchmark" attributes[16].

Figure 2. Example of Census cells

[15] Cross tabulations such as: Number of buildings according to an attribute value, f. ex. cell "X" contains 20 buildings, built in the 1920s, 10 buildings with district heating, 10 buildings with individual boilers, 5 buildings owned by a residential cooperative.

[16] In the British school of Spatial Microsimulation these are referred to as "constraints", while "benchmarks" is the Australian term. I will stick with "benchmarks" to avoid the ambiguity with "constraints" from the realm of mathematical/computer optimisation.

3) Energy audits – approx. 1300 energy audits for different buildings in Hamburg, containing all necessary energy related characteristics (heat transmissivity values, heating systems, etc). This data was provided by the BUE as part of the GEWISS Project (GEWISS Projekt Hamburg, 2019)

Figure 3. Example of buildings with energy audits

4) A distribution of the age of gas boilers in the city, prepared by the local union of chimney sweepers (Landesinnungsverband des Schornsteinfegerhandwerks Hamburg, 2016).

5) Distributions of buildings across Germany according to the presence and type of energy efficiency renovation measures, prepared by Cischinsky & Diefenbach, (2016) from the IWU (*Institut Wohnen und Umwelt)*. The distributions are based on a large (17000) stratified sample of buildings in Germany gathered in 2016. The data provides more advanced cross tabulations – f. ex. multi-attribute (joint) distributions (as in "percent of buildings built before 1970 AND having added insulation"). This data is at the national level, however, I will transfer it to Hamburg. The data shows very low differences between regions and city sizes, so this should not be much of an issue (see Section 5.1). Further, the benchmarking techniques I am going to use can mitigate most of the possible issues with this. They will correct and adjust the data to known Hamburg distributions and thus minimise the possible error of transferring from the national to the city level.

6) District heating grid layout. As part of the GEWISS Project, I had access to the district heating grid layout. I use this to designate all buildings connected to district heating. While being connected does not always mean one is using district heating, it is a rather strong signal for it.

7) Addresses of some housing cooperatives (*Wohnungsgenossenschaften*). Some housing cooperatives in Hamburg are part of an association which publishes a list of their building

stock. I used these to populate some ALKIS buildings with an ownership type. The equivalent benchmarking attribute in the Census counts the addresses of the municipal housing company as well as other housing companies (non-cooperatives) together, so mapping the cooperatives does not exhaust all buildings with this ownership type as defined in the Census. This exemplifies to some extent the type of data integration problem – lots of different data, usually non-exhaustive, usually at different levels of aggregation.

8) Information on renovation measures undertaken, found in construction permits from the years 2014-2018. These are public in the Hamburg "Transparency Portal", however they are also not exhaustive by far. The reason is that not all types of energy efficiency measures require a construction permit. F. ex. adding insulation would not require, in most cases, a construction permit, only if done as part of another type of intervention (adding an extra storey f. ex.). The inverse, however, holds – by using a simple word search for "*WDVS*" (*Wärmedämmverbundsystem* - Exterior insulation and finish system (EIFS)) and similar insulation related German terms, I could filter the construction permits to include only those that relate to insulation.

3.2 Transferability

The data situation is not particularly specific to Hamburg. The ALKIS is a German-wide system. Although in some states it does not include number of storeys, 3D models are now available for all of Germany (see Section 6.4 for a note on 3D data). Census Data is available for the whole of Germany. The distribution of gas boiler age was obtained from a local report, however similar ones probably exist in most cities, since there are such reports at the national level and they must rest on locally gathered data.

The distributions of renovations are from a nation-wide sample survey. Layout of district heating grids are usually moderately easy to come by. If not grid layout, then areas with district heating could be used. List of addresses of cooperatives and renovation measures from building permits are Hamburg-specific, however, equivalents probably exist in many places.

The only Hamburg-specific data source that is probably hard to come by anywhere else in Germany are the energy audits. Although energy audits are prepared in the whole of Germany (and most of Europe), Hamburg has a specific financial support program ("*Hamburger Energiepass*"), through which copies of the audits are gathered centrally and a combined dataset could be obtained.[17]

[17] Nevertheless, obtaining some sample of buildings with energy characteristics is not uncommon. One option would be to inquire with the rent and housing portals (f. ex. https://www.immobilienscout24.de/ or other similar). These obtain energy audit information as part of their offerings. This is because German building regulation stipulates that landlords hold ready a building energy audit (or "pass") to prospective tenants.

3.3 Uncertainty

No data source is perfect. The ALKIS has some known issues with the construction epochs. The Census has a data protection algorithm which changes some values at the cell level in order not to infringe upon personal data protection requirements. I georeferenced the energy audits using an address matching algorithm, which, from sample analysis, is approx. 95% correct, therefore some false geo-reference could be present in the energy audits as well.

4 Defining the Modelling Problem

As noted in the introduction, the archetype approach has its shortcomings. Current Hamburg UBEMs do not mirror building renovations and the total calculated heat demand is 20-30% higher than official statistics, based on the sales of final energy by energy companies. Data on number and types of renovations (buildings with new insulation, buildings with new windows, etc.) is available however, although at an aggregated, Germany-wide level, from the IWU sample study in 2016. Data on construction epochs and ownership (which are important for policy analysis) is also available, but at a different aggregation level (Census cell). Finally, highly detailed data on a sample of buildings (the energy audits) is also available.

The energy audits represent a variety of renovation levels. Some buildings in their "baseline" state (the status quo at the point of carrying out the audit) already have new windows and insulation. This information can be matched with the data on number and types of renovations at the city level. There are 291 unique combinations of energy-related characteristics within the energy audits. These 291 combinations can be viewed as "archetypes" that can describe the building stock. The proper representation of the building stock would then come down to the way the archetypes are matched to Hamburg's buildings[18] and, equally importantly, to the distributions that would arise from this matching at the Census cell and city levels. These distributions have to match the known distributions from the data sources described in the previous section. Note that while the energy audits constitute a biased sample (self-selection bias) I can nevertheless use the information therein and attempt to correct for the bias, see further in Section 4.4.

This points to a set of problems referred to, generally, as "combinatorial optimisation" (CO). CO is concerned with finding an optimal combination (in this case, one that is closest to the aggregated benchmarks, both at the city and Census cell levels) from a finite set of objects (in this case, the energy audit data). There are many concrete problems under the hood of CO, some are known to be NP-complete[19] (the Knapsack problem) and thus can only be approximated or brute-forced, others not – the Minimum Spanning Tree Problem can be solved in (almost) linear time.

[18] I treat the ALKIS objects as "buildings", for more detail see Chapter 5
[19] Refers to computational problems for which no efficient algorithm has yet been found.

4.1 Evaluation criterion - "Good Enough" combinations

Generally, when dealing with optimisation problems, an initial consideration is what solution is satisfactory. Since the aim is to find a realistic UBEM by way of minimising the difference between the resulting model's characteristics and the benchmarks, the satisfactory solution would be, at first glance, one for which the fit is perfect (i.e., the difference is zero[20]). This would be "optimal".

However, given the uncertainties in the benchmarks (stemming from the sources themselves and from the pre-processing of the data) it may well be that a solution with a fit close to the benchmarks could be worse (further from the real building stock) than one with a looser fit. Therefore, an approximate solution to the mathematical optimisation problem would be, for practical purposes, just as good as the mathematically optimal solution. Given the difficulty of finding a mathematically optimal and not just approximated solution to many combinatorial optimisation problems, this is good from a computational standpoint. However, a choice must be made as to where the cut-off for the fit should be. In other words, what are "good enough" combinations? One might consider borrowing fixed numbers, like the 95% confidence level, widely used in statistics. This does not help much. Considering that, for example, for medical purposes, the 99% is more often used, it is evident that "good enough" is in the eye of the beholder or user. Because of this, I would abandon the strictly quantitative approach and not add any predefined cut-off values while modelling. Any model user should evaluate and judge the model fit depending upon the data and the use case. For the concrete Hamburg UBEM that I am preparing for this thesis, I discuss the model fit and what I deem "good enough" in Chapter 7.

4.2 Same-fit combinations and spatial clustering

Provided that the problem concerns optimising counts (numbers of buildings) it is possible that two or more distinct combinations may have exactly the same fit. This either leads to multiple outputs or a choice must be made which one to take. Producing multiple outputs is viable, but I reckon, for practical purposes, it should be avoided if possible. More crucially, it has algorithmic implications, which I will discuss in Section 7.1.

In order to choose between same-fit combinations, I will use one last piece of information contained in the energy audit dataset – the location of each audit. The assumption would be that **if** audits exhibit spatial clustering along the constraining attributes, a combination of *nearest neighbours* would potentially be more realistic than another combination given they have the same fit. The presence of spatial clustering is evidenced in Section 5.4.

[20] Maximising the fit is equivalent to minimising the difference to the benchmarks. I define "fit" and by analogy "difference" in Chapter 7.

4.3 Additional requirements

4.3.1 The need for explainability

The purpose of UBEMs is to support policy analysis and implementation. Urban and regional planning, of which the building sector is a major addressee, has pushed for public participation in policy formation and implementation at least since the Copenhagen Charter (Danish Ministry of the Environment, 2002). In this regard, there is a benefit to and, I would argue, even a requirement for transparent policy based on explainable models and analysis. An example for this is the Hamburg Heat Demand Cadastre (*Hamburgisches Wärmekataster*). It is a tool for policy making and it is public and intended for public use as well.

Although building renovation policies may be driven by an executive branch (for example the Hamburg BUE), building ownership is mostly private. Therefore, the authority to make decisions lies in many cases outside the executive branch and the executive branch has a guiding, mediating and communicating role and (nowadays) more limited regulatory power. Any tools it uses have to be explainable, the more the better.

There are many ways this could translate to modelling methodology – f. ex. an algorithm with less or simpler steps would be better explainable. I argue, it also translates into a preference for deterministic over probabilistic modelling. Of course, obeying the prerequisite that both produce "good enough" models (see Section 4.1). Given the uncertainties in the input data, no solution is preferable to another beyond a certain point. A probabilistic algorithm could find a solution that is preferable based on some other, new criterion. However, this preference has to carry benefits which exceed the costs incurred because the algorithm is less explainable. I have not found such a criterion.

I have to stress that this argumentation is specific to urban planning. For example, it is a different situation if an algorithm has to optimise the operational profiles of electrical power generators of a utility company and finds a solution in a probabilistic manner. The condition of explainability could still apply, but it would be narrowed down to engineers and at most the company board-room members. Provided that the Hamburg Heat Demand Cadastre, an example of the type of UBEM I am generating, is public and intended for public use, it is reasonable to set a higher bar for explainability. A bar higher than in the power generator optimisation example.

Using probabilities and random sampling is a good way to analyse uncertainty and to validate an algorithm. I use these techniques in Section 9.2. However, I aim for a generally deterministic approach to modelling.

4.3.2 Integrating sparse, incomplete building data

UBEMs are primarily prepared by public bodies, which have access to different samples of building information (from local projects, initiatives etc.). This is usually data on only some building characteristics for only some buildings. Nevertheless, it is usually empirical and can be useful. Because of this, it would be advantageous if a modelling algorithm can incorporate such data together with the sources listed in Chapter 3.

4.4 Key Assumptions

There are three key assumptions that underpin the modelling problem:

- A combination of the energy audits (with their energetic characteristics) is enough to create a plausible representation of the energetic characteristics of the building stock. Firstly, the energy audits, per definition, cover the energetically vital characteristics of the buildings. Secondly, they are a sample spanning different locations, building construction epochs, types of heating system and building shell characteristics. They do not cover all possible combinations of building attributes (some are obviously not plausible, f. ex. a building built in 2010 with a wall U-value[21] of 2.0). Also, they are not a representative sample, but I also do not assume that. I assume that the sample contains enough information to allow it to be reweighted so that it can *become* representative.

- A combination of energy audits such that, when aggregated, matches the benchmarks, would be a good representation of the building stock. The key issue here is that it is possible there are a number of very different combinations, that when aggregated lead to the same benchmarks. Therefore, is finding one such combination good enough to represent the building stock? I analyse this in Section 9.2.

- Among all possible combinations of energy audits that match the benchmarks, the one that consists of nearest neighbours is the closest possible to the "true" building stock. I expand on that in Section 5.4.

[21] A measure of thermal transmittance

5 Preparing Data for Modelling

I had to pre-process the three main datasets – buildings (approx. 200000 objects in the electronic cadastre), cells (approx. 20000 100x100m raster cells from the Census) and energy audits (approx. 1300) before modelling. This included choosing appropriate attributes, matching attribute names, and integrating known information at the building level. I will present some of the pre-processing steps and the end results in this section.

5.1 Building attributes

The attributes used for benchmarking and their categories[22] are presented in Table 1. These are the "benchmarking" attributes that apply to all three main data sources – buildings, Census cells and audits. Some of them are self-explanatory, f. ex. "DISTRICT HEAT" means a building is heated with district heating[23]. "BUILD_CENTRAL_HEAT" refers to buildings with a central (to the building) heating system. "APARTMENT_HEAT" and "ROOM_HEAT_STORAGE_HEAT" refer to buildings where each apartment or each room is heated with a separate system. "EFH" stands for "*Einfamilienhaus*" or single-family house. "MFH", respectively for multi-family building. Note that the information on insulations are in the form of joint distributions – "EFH_OTHER_NEW_INS_AW" means single-family house, not owned by a cooperative or public housing company (hence "OTHER") that has new (added) insulation ("NEW_INS") on the outer walls ("AW" stands for "*Außenwände*" – outer walls). Further, "KE" refers to "*Keller*" – basement, and "DA" refers to "*Dach*" or roof). Refer to Appendix I for a detailed description of all categories.

[22] For the purposes of this thesis, I use "category" as one of the possible values that an attribute (which is a categorical variable) can take. F. ex. "district heating" is a category of the "type of heating" attribute. Using the term "categories" in this way might seem counter-intuitive to some readers. There are several ways of denoting these values, f. ex "classes" or "attribute values". However, the computer scientist might object to "class" because that is a reserved term in his discipline. Using "attribute values" would run into issues when one looks at the Census cells, where these "categories" become attributes themselves.

[23] Note that some simplifications are necessary. The Census gathers data by asking people. Therefore, this attribute means, strictly speaking, "the building is said to be heated with district heating". On the other hand, for the ALKIS objects I define as "DISTRICT HEAT" I use the presence of a district heating pipe as the criterion. There are edge cases where these two definitions would diverge and not mean the same thing. For practical purposes I take them as being equivalent.

Attributes	Categories	Level
Type of Heating	DISTRICT_HEAT BUILD_CENTRAL_HEAT ROOM_HEAT_STORAGE_HEAT APARTMENT_HEAT	Census cell (local)
Construction Epoch	(1400,1918) (1919,1948) (1949,1978) (1979,1986) (1987,1995) (1996,2000) (2001,2008) (2009,2011) (2012,2015) (2016,2020)	
Type of Ownership	COOPERATIVE_OR_MUNICIPAL_COMPANY OTHER	
Building Type	MFH EFH	
Insulation outer walls	EFH_OTHER_NEW_INS_AW EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NO_NEW_INS_AW MFH_OTHER_NO_NEW_INS_AW MFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NO_NEW_INS_AW MFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_AW MFH_OTHER_NEW_INS_AW EFH_OTHER_NO_NEW_INS_AW EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_AW	City (global)
Insulation attic	MFH_OTHER_NEW_INS_DA MFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NO_NEW_INS_DA EFH_OTHER_NEW_INS_DA MFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_DA EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_DA EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NO_NEW_INS_DA MFH_OTHER_NO_NEW_INS_DA EFH_OTHER_NO_NEW_INS_DA	
Insulation cellar	EFH_OTHER_NEW_INS_KE EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NO_NEW_INS_KE MFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_KE MFH_OTHER_NEW_INS_KE MFH_OTHER_NO_NEW_INS_KE EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_KE EFH_OTHER_NO_NEW_INS_KE MFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NO_NEW_INS_KE	
Gas- or Oil-Boiler age	GAS_OIL_BOILER_ab1996 NO_GAS_OIL_BOILER GAS_OIL_BOILER_1995 GAS_OIL_BOILER_1987	

Table 1. Overview of building attributes and respective categories used for benchmarking the model

The importance of most of these variables for heat demand calculation are obvious (type of heating, insulation etc.). The construction epoch is generally used as a proxy for building materials and their characteristics. For buildings built after 1978, the second oil crisis, the introduction of energy efficiency standards in Germany led to a notable increase in efficiency. For those built before 1978, one could argue that the older the building, the likelier it was already renovated at some point. This assumption, according to the sample study of the IWU (Cischinsky & Diefenbach, 2016, p. 49) holds. However, the percentages of buildings with insulation is almost the same (~40% for wall, ~65% for roof/ceiling and ~18% for cellar) across buildings built before 1968, therefore having a finer differentiation into construction epochs might not be needed[24]. Nevertheless, I use all epochs available in the data, because the epoch is a defining building characteristic that can be assumed to correlate with a variety of other characteristics (façade decorations for example, that can make the addition of shell insulation impossible or improbable).

There is an additional important detail – the definition of "insulation" in the context of this thesis. Since the construction epoch is a signal of the way a building was built (mostly because of regulation), insulation added at the time of construction is implicitly mirrored by it. Because of this, when I refer to "insulation", I mean <u>additional</u> insulation, added after construction. The said distributions (ibid, p.49) also refer to "added insulation" (*nachträgliche Wärmedämmung*). Looking at the numbers, one notices that, logically, the share of buildings with added insulation falls to 11% for buildings built after 1979 and then further down to 2.8% (after 1995), since the buildings are newer and were built to be more energy efficient. At this point, I simplify the modelling, by neglecting added insulation for buildings after 1978. The main reason for this is that differentiating any given attribute too much leads to more benchmarks, which makes finding a good combination harder. Therefore, I needed to simplify where possible.

It was important to find variables which correlate with the presence of insulation, since insulation is one of the most important attributes that is generally unavailable at the Census cell level. The same IWU study (p. 50-51) suggests that the state, city size and also whether a municipality has population growth or decline do not influence much the shares of buildings with insulation. There are minor differences in the percentages, but they are below or close to the standard error of the sample[25]. However, ownership type does have a large effect. Table 2 (or Nr. 22 in the original report) gives the

[24] The reader might ask about insulation thickness. While it is a valid point that the thickness of added insulation varies across time and in their study IWU gathered data on this, I purposefully ignore the effects of this. Firstly, the added benefit of insulation thickness decreases exponentially, not linearly. The vast majority of reduction in the U-values is gained with the first 5-8 cm of insulation. Everything on top of this has diminishing returns (yes, passive houses require insulation around 30 cm thick to reach the required efficiency, but this does not invalidate the fact that these extra 22 cm or more of insulation bring a lot less than the first 8 cm). Secondly, differentiating insulation thickness rather than modelling a simpler "presence of insulation" would lead to a lot more archetypes and make model fitting a lot harder. Given the first and second point, I believe that, in my case, the informational benefits of modelling insulation thickness do not outweigh the practical costs to the creation of the UBEM.

[25] The only exception is the geographical dimension - the states which were territory of the former GDR exhibit a larger difference when it comes to renovations compared to the rest of Germany. Their share of buildings with insulation is 15 percentage points higher.

joint distribution of added insulation and ownership for multifamily buildings built before 1978 in the "old" states ("*Alte Bundesländer*", the states in the former West Germany). The data exhibits large differences between the shares of buildings with added insulation when split into ownership types – 47% for housing companies, 33% for private owners and only 17% for buildings with multiple owners (individual apartment owners). A plausible explanation for these differences lies in the administrative, planning and organisational capacity of these different ownership types. Housing companies, logically, have the highest share since they can both take decisions and carry them through faster than a building with multiple owners would. In the category "private owners" fall mostly single-family houses[26]. In this case, it is less the ability to take decisions and more the financing that might be a challenge.

Since ownership is available at the Census cell level and the data shows it makes a large difference for the presence of insulation, it was important to use it for benchmarking. For this reason, at the city-level the insulation categories included (f. ex. "EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_AW") actually have the aforementioned multi-attribute (joint) categories – referring to size, ownership and presence of insulation together.

	Added Insulation [% buildings]		
	Outer Walls	Roof/Attic	Cellar/Ground
All multifamily buildings built before 1978	32.7% +/- 2.4%	60.9% +/- 2.0%	15.1% +/- 1.2%
Small cities and municipalities	32.1% +/- 3.2%	62.4% +/- 2.5%	16.8% +/- 1.7%
Large cities and municipalities	33.2% +/- 3.4%	59.8% +/- 2.8%	13.9% +/- 1.7%
Municipalities with growing population	34.9% +/-4.6%	61.9% +/-3.3%	12.3% +/-2.2%
Municipalities with stagnating population	32.2% +/- 2.8%	60.6% +/- 2.3%	15.8% +/- 1.4%
Multi-owner buildings	**17.1% +/- 3.4%**	54.7% +/- 4.0%	10.1% +/- 2.2%
Single private owner	33.0% +/- 3.4%	65.7% +/- 2.3%	15.8% +/- 1.6%
Housing company	**47.5% +/- 4.9%**	58.8% +/- 3.9%	18.6% +/- 3.0%

Table 2. Presence of insulation according to size of the of city (town), population dynamic and ownership. Source (Cischinsky & Diefenbach, 2016, p. 51)

For building construction epoch the mutual exclusion of categories within attributes is obvious. A building cannot be built both before 1900 and after 1900. Other attributes are not per definition mutually exclusive, for example "type of heating". A building can have district heating and a back-up or peak load boiler. However, the Census simplifies this and gives the predominant type of heating. A simplification that I am bound by, since I use the Census.

[26] But not only, some wealthy indviduals might own a building or multiple buildings. While there is no specific data on such, they probably act more like a housing company than like a private owner in terms of their potential for renovations.

Another key step in the pre-processing and analysis of the ALKIS buildings and Census cells is to understand how the two relate to each other. The ALKIS contains cadastral objects (buildings). While the Census measures "buildings" as well, if one sums up the totals for both, the results would differ markedly. That is because the Census actually counts a building entrance as a separate building. The way to match the counts is by counting ALKIS address points and compare those to the number of "buildings" in the Census cells. An example is given in Figure 4. The cell pictured has 12 "buildings" according to the Census. It also contains three ALKIS buildings. The red dots are address points. It is clear then that the Census counts the latter as buildings.

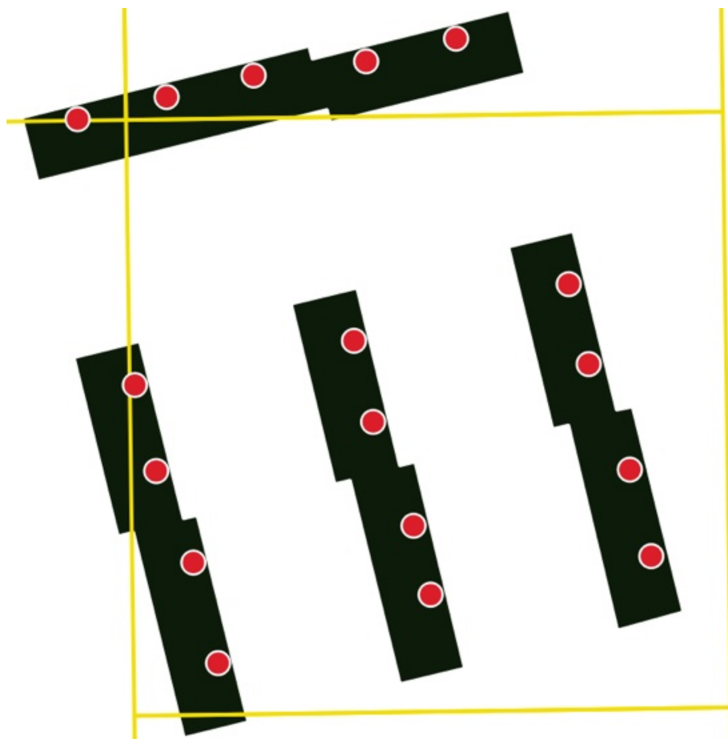


Figure 4. Example of relationship between ALKIS building, Census cell and address points. The black rectangles are buildings in the ALKIS, the red dots are address points, the yellow squares are Census cells.

While understanding this discrepancy allows for its mitigation by simply knowing what counts to sum up, there is also another practical problem. Some ALKIS buildings reside in multiple Census cells. This means the lowest unit of analysis – the ALKIS building – actually has to be broken down into address points. On the other hand – it is unlikely that an address point would have different building characteristics compared to other address points in the same ALKIS object. Therefore, the modelling algorithm has to ensure that a single ALKIS building is not assigned different energetic characteristics from the different address points. This would have further complicated the modelling task and added an additional requirement. I tackled the problem by aggregating all cells until the address points of an ALKIS building always reside in the same aggregated cell. This reduced the number of cells from 26600 raster cells to 19900 cells. An example of this aggregation is in Figure 5. Note that the address point geometry is key here. It is not a problem when a building polygon intersects multiple cells. The problem arises when the address points associated with a single ALKIS building reside in multiple cells. That is what the aggregation tackles.

Figure 5. Example of aggregating Census cells so that address points of a ALKIS building always reside in the same cell.

With this aggregation I can keep the smallest unit of analysis to be the ALKIS building and when Census values are concerned, I do not count each ALKIS building, but its number of address points.

Finally, the data on gas (or oil) boiler age comes in the form of a distribution of boilers according to their age (Landesinnungsverband des Schornsteinfegerhandwerks Hamburg, 2016). While this does not necessarily mean that each boiler corresponds to one building[27], I make this assumption. Thus, I view the distribution of boilers to be equivalent with the distribution of buildings with such boilers. F. ex. if the data on boilers shows 20% of boilers are 20 years old or older, I view that as 20% of buildings not using district heating having a boiler 20 years or older. Single-family houses with multiple boilers would probably be an edge-case, so this equivalence would likely be unproblematic for them. It is more likely that a multi-family building could have multiple boilers. Correcting for this however would be highly problematic. In order to find the connection between building and boiler counts, one has to also pinpoint the end of the supply zone of one boiler in the building and beginning of another zone, together

[27] There could be multiple boilers in one building, or one boiler for two buildings (ALKIS objects)

with the end of one building and the beginning of another. The latter is more or less a given in this case by the ALKIS objects. The former is largely an unknown. The connection between the ALKIS object and the potential supply zones in the building is also unknown. At this point I chose to simplify and assume a one-to-one, one boiler to one building for all buildings. I chose this over attempting a complex correction and estimation that I believe lacks a stable enough footing. Thus, using the distribution of boilers with certain age as equivalent to the distribution of buildings with boilers with certain age is a form of simplification and assumption.

5.2 Excluding Heat demand from benchmarked building attributes

After looking at all attributes above, one might wonder why is heat demand not part of the benchmarking attributes. Each energy audit comes with values for heat demand, which, with a correction for typical consumption, can be benchmarked to known consumption values at the city level. The reason for omitting the heat demand from the benchmarking is that I wanted to use it for external validation. One key argument I make in this thesis is that given a good quality estimation of the building characteristics, a consumption-corrected demand can be calculated that comes very close to metered consumption. It is this argument that I put to the test by not benchmarking the heat demand/consumption itself, but only the building characteristics. If they are well modelled, the consumption should match at the end without explicitly benchmarking it. Refer to Sections 6.2 and 9.3 for evidence for this argument.

5.3 Examples of the datasets used for modelling

The building attributes are the same for all three main datasets (buildings, energy audits and Census cells). I will present examples of these datasets for better understanding of what the modelling algorithm presented in Chapter 8 takes as input. Firstly, an example of the cadastral building dataset after pre-processing and adding known information on the building level is presented in Table 3.

ID	Geometry	Number of storeys	Number of Addresses	Type of Heating	Construction Epoch	Type of Ownership	Build-ing Type	Insula-tion outer walls	Insula-tion attic	Insula-tion cellar	Gas- or Oil-Boiler age
1	POLY-GON (…)	2	1	-	(1949,1978)	-	EFH	EFH_OTHER_NEW_INS_AW	-	-	-
2	POLY-GON (…)	5	3	DISTRICT_HEAT	-	COOPERATIVE_OR_MUNICIPAL_COMPANY	MFH	-	-	-	-
…											
197857	POLY-GON (…)	8	2	-	-	-	MFH	-	-	-	-

Table 3. Example of the input cadastral buildings dataset

The underlying cadastral building dataset in its raw ALKIS form contains the geometry, the number of storeys and the number of addresses. The modelling algorithm should however be able to deal with partially available information at the building level. In the example, one building (ID 1) has a known construction epoch and type of ownership, another (ID 2) has a known type of heating. In the concrete Hamburg case, I estimated the "building type" from the geometry and the buildings with district heating by using information on the location of heating grid pipes. Similarly, using the data on building cooperatives, for some buildings, I could note the type of ownership prior to the algorithmic modelling. Further, here I use the addresses of the energy audits together with the construction permits[28] obtained from the Hamburg "Transparency Portal" to note individual buildings which are most likely already renovated.

The example of the energy audit data looks similar (Table 4). It has the same attributes as the building dataset, however, it is complete (all attributes have values) and there are a number of additional attributes (in italics), which are not used for benchmarking the model, but for calculating the heat demand.

ID	Geometry	…	Gas- or Oil-Boiler age	*Roof U-Value*	*Wall U-Value*	*Cellar U-Value*	*Window U-Value*	*Win-Wall Ratio*	*Domestic Hot Water type*	*Circulation*	*System Supply Temp*	*Pipe Insulation*	*Generator Location*
1	POLYGON (…)	…	GAS_OIL_BOILER_ab1996	1.0	2.5	1.0	2.8	0.22	CENTRAL_WW	mit Zirk.	55	mäßig (Altbau)/halbeEnEV	außerhalb der thermischen Hülle
2	POLYGON (…)	…	NO_GAS_OIL_BOILER	1.7	2.8	1.2	2.8	0.19	DE-CENTRAL_WW	ohne Zirk.	70	unge-dämmt	innerhalb der thermischen Hülle
…													
1502	POLYGON (…)	…	GAS_OIL_BOILER_1995	0.8	0.6	0.8	1.1	0.3	DE-CENTRAL_WW	ohne Zirk	90	EnEV/doppelteEnEV	außerhalb der thermischen Hülle

Table 4. Example of the energy audit dataset. Attributes in italics are not benchmarked, but used for calculating heat demand.

Finally, Table 5 is an example for the Census cell dataset. Each row represents a Census cell, and each column has numeric values – the number of address points in the cell having the respective category of the respective building attribute. F. ex. a building attribute is "type of heating", which can have the categories (values) "DISTRICT_HEAT", "BUILD_CENTRAL_HEAT", etc. These categories then become columns of the cells, since their occurrence in a cell has to be counted.

[28] Note that most renovations do not require a construction permit and not all construction permits relate to energy efficiency measures. However, I took only those construction permits that *do* relate to energy efficiency by filtering based on energy-related keywords. Also, I do not claim that these represent a complete picture of renovated buildings in Hamburg.

ID	Geometry	DISTRICT_HEAT	BUILD_CENTRAL_HEAT	ROOM_HEAT_STORAGE_HEAT	APARTMENT_HEAT	(1400, 1918)	(1919, 1948)	(1949, 1978)	(1979, 1986)	…	EFH_COOPERATIVE_OR_PUBLIC_COMPANY_NO_NEW_INS_KE
1	POLYGON (…)	5	0	0	0	0	5	0	0		4
2	POLYGON (…)	0	10	0	0	0	4	0	6		0
				…							
19900	POLYGON (…)	0	5	4	1	2	4	4	0	…	2

Table 5. Example of the Census cells used as benchmarks

5.4 Analysis of spatial clustering

To test the assumption made in Chapter 4, that using the spatial location of energy audits could be used to choose between same-fit combinations, I will use the *nearest neighbour index* (NNI) (Wilson & Din, 2018) on the energy audits. The NNI is a global measure of clustering and dispersion. It is the ratio between the observed and expected average distance between all pairs of nearest neighbours. In the concrete case, the energy audits are selected based on their category for each attribute. For each category I then compute the nearest neighbour index.

Wilson and Din use bootstrapping to compute the expected distance, while I use a regular pattern as is standard in the ESRI ArcGIS tool Average Nearest Neighbour (Mitchell, 2009). Contrary to Mitchell and in accordance with Wilson and Din, I will use the median rather than the arithmetic mean to compute the observed average distance. This decreases the effect of outliers. Given a regular pattern, the median and arithmetic mean are the same for the expected distance:

Eq. 1.

$$NNI = \frac{med(D')}{D''}$$

Eq. 2.

$$D'' = 2\sqrt{\frac{n}{A}}$$

Where:

NNI is the nearest neighbour index

$med()$ is a function to return the median value from a group of numbers

D' is the set of all nearest neighbour distances of category X in the energy audits

D'' is the expected average distance for category X in the energy audits

n is the number of energy audits

A is the total area of all Census cells

NNI values of above 1 are a signal for dispersion, while below 1 a signal for clustering. The index is very sensitive to the target area via the area parameter. An assumption when using it is that the locations have an unconstrained probability to fall in any location within the target area. Due to this, I used only the area of the Census cells (~262 km^2) instead of the area of the city (~755 km^2)[29]. In this way, areas where a residential building cannot be located (water bodies, industrial zones, harbour, etc.) are excluded and the expected distance is not overestimated. The results for all categories and an interpretation are presented in Table 6.[30]

Attribute	Categories	NNI	Z-Score	Energy Audit Count	Interpretation
Type of Heating	DISTRICT_HEAT	0.2	-28.51	312	Highly clustered, to be expected, given the need for a grid
	BUILD_CENTRAL_HEAT	0.7	-16.13	1084	Somewhat clustered, probably due to the lower count in areas with district heating
	ROOM_HEAT_STORAGE_HEAT	0.9	-2.0	62	More or less neutral
	APARTMENT_HEAT	0.6	-2.8	12	Clustered, but too few audits
Construction Epoch	(1400,1918)	0.4	-17.9	237	Clustered, to be expected, given the natural growth of the city and the re-building after World War II
	(1919,1948)	0.5	-15.87	299	
	(1949,1978)	0.6	-22.88	824	
	(1979,1986)	0.8	-2.89	68	
	(1987,1995)	0.9	-1.34	32	
	(1996,2000)	2.2	5.5	6	Too few audits far plausible interpretation
	(2001,2008)	0.0	-3.82	4	
Type of Ownership	OTHER	0.7	-17.21	1233	Somewhat clustered, probably the result of the cooperatives being clustered
	COOPERATIVE_OR_ MUNICIPAL_COMPANY	0.1	-26.76	237	Clustered, to be expected
Building Type	EFH	0.7	-16.02	736	Clustered, the general city centre vs. outskirts pattern
	MFH	0.3	-36.94	734	
Insulation outer walls	_NEW_INS_AW	0.7	-9.91	283	Somewhat clustered, an important result
Insulation attic	_NEW_INS_DA	0.6	-34.16	2140	
Insulation cellar	_NEW_INS_KE	0.8	-6.41	280	
Gas- or Oil-Boiler age	NO_GAS_OIL_BOILER	0.3	-27.12	374	Clustered, the audits with district heating are counted here
	GAS_OIL_BOILER_1987	0.9	-2.1	159	More or less neutral, rather inconclusive
	GAS_OIL_BOILER_1995	0.7	-12.71	375	Moderately clustered, also an important result
	GAS_OIL_BOILER_ab1996	0.7	-14.81	556	

Table 6. Nearest Neighbour Index as a measure of global spatial clustering

[29] Census cells cover only areas with residential buildings

[30] Note that the number of categories here is 22, as opposed to 46 in Table 1., because here I do not use categories representing joint distributions (f. ex. EFH_OTHER_NEW_INS_KE).

Another way of analysing the clustering is by finding the number of nearest neighbours needed for finding the same audit again ("same" as in "has the same categories"). Running the analysis gives the following values according to percentiles - Table 7:

Percentile	10th	20th	30th	40th	50th	60th	70th	80th	90th
Same nearest neighbour found at neighbour:	1	1	3	6	11	19	34	69	174

Table 7. Percentiles of the distribution of same nearest neighbour

The way to interpret the table is the following: For 20% of the approx. 1351 energy audits[31] an audit with exactly the same categories is actually the nearest neighbour. For 50%, an audit with the same categories can be found among 11 nearest neighbours, etc. These values are referred to again in Section 9.2.

All in all, most attributes exhibit spatial clustering which gives plausibility to the assumption that a nearest neighbour combination would be preferable (potentially more realistic) than any other, given the same fit.

5.5 Supplementing the audits with IWU buildings

As a final step in the pre-processing, I added approximately 151 energy audits that are "synthetic". These are audits that I created for variable combinations that most probably exist, but are not found in the sample. These are for example newly built buildings after 2016. Their attributes I took from the IWU Typology.

[31] 1351 and not 1501, because 1501 are the energy audits together with the 151 "synthetic" audits – see Section 5.5

6 Existing Methods

6.1 Spatial Microsimulation

The problem setup points the search for methods to a field of economics, known as “microsimulation”, pioneered by Guy Orcutt (Orcutt, 1957). At the time, economists were using mostly aggregated economic data for their models. Orcutt argued that models have to be at the “micro” or person/household levels for some concrete simulations, such as health or tax policy. The reason was (still is) that these policy issues have usually multiple, interconnected reasons, the combination of which matters and differs at the individual level. This gets lots in aggregation. F. ex. health is influenced by a large number of factors or personal characteristics – from diet, through genetics and environmental factors to behavioural factors such as smoking and sports. Aggregated models in the form of cross tabulations could not be used to analyse such characteristics in their combinations. In some cases, such data on individual level could be available. The original applications of microsimulation to tax and healthcare used (and still use) “micro-datasets”, f. ex the German *“Mikrozensus”* (a household survey). In some cases, even building “micro-datasets” are available (such as in Peters et al.(2002)). In many other cases however, researchers in the field need to put considerable effort to come up with techniques that create such “micro” datasets from various available data. An example is the more recent “spatial microsimulation”, whereby individual data on, f. ex, country level is combined with aggregated data on, f. ex., regional level to obtain a regional-level “microdataset”. This allows the analysis of differences of the “micro”-units across regions.

The methods used for combining aggregated and individual data can be found in the literature under the general term *population synthesis*, *benchmarking* or *reweighting*. The basic idea is to use the aggregated data as *benchmarks*. Then find such weights for the entries in the individual data that, if aggregated, the variables in the individual dataset match the benchmarks. The reweighted individual data can then be regarded as a disaggregated representation of the population, or “synthetic” population. A known prerequisite for the individual data is to be representative of the general population or at the very least include all or nearly all combinations of individual’s characteristics found in the general population.

If one performs the reweighting for a country, then the model is at the national scale. If the reweighting is performed for regions, districts or some other spatial unit, then a “synthetic” population is created for each unit. Thus, differences between regions and spatial clusters can be analysed.

These techniques have been used in transportation planning (J. Ma et al., 2014) and to a much lesser extent in UBEMs (Chingcuanco & Miller, 2012; Muñoz et al., 2016; Muñoz & Peters, 2014; Nägeli et al., 2018). Tanton (2014) gives an overview of “spatial microsimulation” methods.

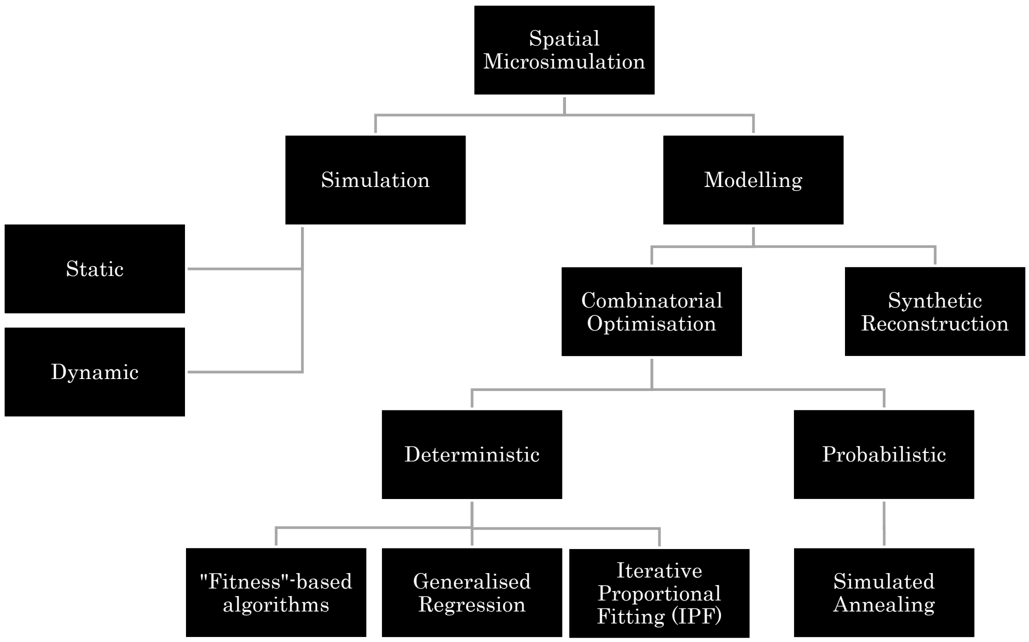


Figure 6. Overview of spatial microsimulation methods[32]. Own elaboration based on Tanton (2014, p. 7)

6.1.1 Static vs Dynamic Microsimulation

Static microsimulation refers to simulations which are static in time - a representation of the population for current policy decisions. In dynamic microsimulation, the population is "aged", its development through time is simulated so that longer lasting effects can be analysed over time. Of course, one can simulate the "ageing" (as in changes over time) of not only people or households, but other micro units as well (f. ex. buildings). An UBEM can be used for both static and dynamic microsimulation, but since the current task is modelling and not simulation, I will not delve further into this differentiation.

6.1.2 Combinatorial Optimisation vs Synthetic Reconstruction

The difference between the two modelling techniques is that *Combinatorial Optimisation* (CO) involves aggregated data and a sample of individual data, while *Synthetic Reconstruction* uses aggregated data and known distributions of characteristics of individuals. In the latter case, probabilistic sampling of the distributions is used until the aggregated benchmarks are matched[33]. Since, in the concrete case of this thesis, a sample is available, the search for appropriate methods is narrowed further to CO.

[32] A side note on terminology: Looking at these methods, however, one finds mostly "modelling" methods (how to get the data?) and much less "simulation" (performing tasks with the data). This is indicative of how much effort is needed (and put) into modelling.

[33] This is essentially very similar to what stochastic bottom-up UBEMs as in Sokol et al. (2017) make use of.

6.1.3 Probabilistic vs Deterministic

All remaining techniques are a form of Combinatorial Optimisation (CO) in that they all search for an optimal combination of a finite set of elements, given some benchmark. Some of them are “stochastic”, which means they make use of a random number generator somewhere along the process. If the provided seed is not constant, the result would be (usually slightly) different each time the algorithm is run. “Deterministic” methods, given the same input, always produce the same result.

6.1.4 Integer vs Floating point results

Another important differentiation is between “*Fitness-based algorithms*” and “*Simulated Annealing*”, on the one hand, and “*Generalised Regression*” and “*Iterative Proportional Fitting (IPF)”* on the other (see Figure 6). The difference lies in the nature of their results. The former produces integer weights, while the latter – floating point weights, which require a further “integerisation” step (Lovelace & Ballas, 2013). In simpler terms – some algorithms produce weights (e.g. number of times a type of person/households is to be found in an area) which are whole numbers – 1, 3, 10, 100 etc. Other algorithms produce decimal weights. Since there cannot be 1,5 households in an area, the decimal weights need to be converted to integers, which becomes an additional modelling step.

6.2 Calculating heat demand

6.2.1 The TABULA[34] Method

Added to the task of modelling the building stock’s properties comes the task of converting building characteristics to energetic values (heat demand). While dynamic simulations of heat demand could provide superior results, they are computationally heavy even for individual buildings, let alone for an entire building stock. Because of this, the most widely used approach in Germany is the heat balancing (“steady-state”) calculation according to DIN 4108 (Deutsche Institut für Normung, 2003) or DIN 18599 (Deutsche Institut für Normung, 2018). The former was the standard for residential buildings until recently, while the latter for non-residential. For the purposes of this thesis however, I will use a modified approach to the DIN 4108 - the “TABULA reference method” designed by the already mentioned IWU (Loga et al., 2015, p. 47; Loga & Diefenbach, 2013). It is based on the DIN 4108, but is a yearly rather than a monthly calculation and more crucially, includes a “consumption-correction”. As described in Subsection 1.2.4, calculated heat demand regularly diverges from measured consumption. However, measured consumption is what is mostly used for UBEM validation. Furthermore, any building model, regardless if for a building stock or for an individual building, that fails to model consumption properly loses credibility with the policy makers and the public.

It is of course essential to understand the reasons for the discrepancy in any individual case and tackle each reason accordingly. When it comes to UBEMs, failing to model renovation levels or, more

[34] “TABULA” stands for „Typology Approach for Building Energy Assessment“ and is the name of an EU project with IWU in the lead that developed energy typologies for European countries - https://episcope.eu/iee-project/tabula/

generally, a bad (as in "unrealistic") estimation of building characteristics is usually the first concern. After that comes user behaviour and lastly the calculation method itself. The bigger part of this thesis is concerned with tackling the first problem – modelling the building stock properly. The other two I would tackle in a simpler, but I reckon effective way, by making use of the TABULA reference method and the aforementioned "consumption correction". The consumption correction is based on a sample of metered data, from which a function was derived. This function is used to convert the calculated heat demand to an estimation of the heat consumption (Loga et al., 2015, p. 77). The benefit of using a function rather than a simpler correction factor is that the demand-consumption discrepancy is not linear. It is usually larger (demand being much greater than consumption) for less energy efficient buildings. Respectively, for moderately insulated buildings, the discrepancy is small and then it reverses for highly energy efficient buildings – demand tends to be lower than actual consumption[35]. This is nicely reflected in the TABULA method (Figure 7).

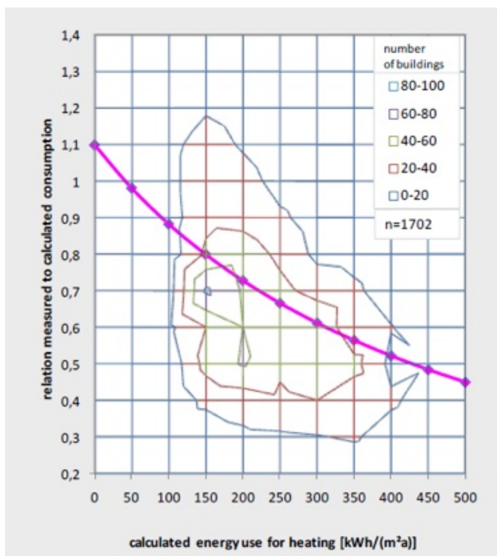


Figure 7. Empirical correction of the demand-consumption discrepancy. Source: Loga et al.(2015, p. 78)

Of course, the sample that IWU used is limited (1702 buildings) and Figure 7 shows that the numbers deviate from the function line. Nevertheless, it is an effective and practical way to describe the connection between demand and consumption.

Still, I performed my own plausibility check of this method. I used the study of the building stock of the state of *Schlesswig-Holstein* by the building construction association ARGE[36] (Walberg et al., 2012). ARGE had gathered a sample of buildings and metered data and created a typology to describe the building stock in the state (similarly, but mostly independently of IWU). The ARGE typology defines

[35] See Subsection 1.2.4 for a further discussion of the demand-consumption discrepancy and possible reasons for it.
[36] ARGE is the "*Arbeitsgemeinschaft für zeitgenössisches Bauen e.V.*", an association based in Kiel, federal state Schleswig-Holstein.

12 building types according to size (single-family and multi-family) and construction epoch (before 1918, 1918-48 etc, see Walberg et al (2012, p. 19)). The 12 types are further subdivided in three baseline states ("not renovated", "moderately renovated" and "well renovated") and three "state of the art" modernisation possibilities at the time (2012). These are the "adequate modernisation", "*EnEV* 2009 modernisation" and "*Effizienzhaus 85*"[37]. For these types and renovation levels the typology gives the building characteristics (U-values, systems etc.) and average metered consumption per square meter. I took two building types and four renovation levels. I matched them to the corresponding IWU buildings and corrected where the building characteristics didn't match. Then I calculated the consumption-corrected heat demand according to the TABULA method and compared it to the values for the average consumption given by ARGE. The result is in Figure 8. It shows that the TABULA method comes generally quite close to the average metered consumption (provided that the buildings have the same characteristics, which I had made sure of).

Given the authority of the IWU (it is a respected organisation) and my own independent plausibility check, I reckon I have enough evidence to adopt the TABULA method as a realistic and practical way of calculating heat demand for the purposes of this thesis.

Lastly, one might ask, if metered consumption exists why do I need a heat calculation method at all? The reason is that metered consumption is always bound by a sample and the building types (or "energetic archetypes") contained in or derived from it. This can greatly limit the flexibility of model creation and use. The model as a baseline state and further analysis of scenarios would both be bound by the building types for which consumption is available. An empirically derived calculation that corrects for the demand-consumption discrepancy serves to lift this limitation.

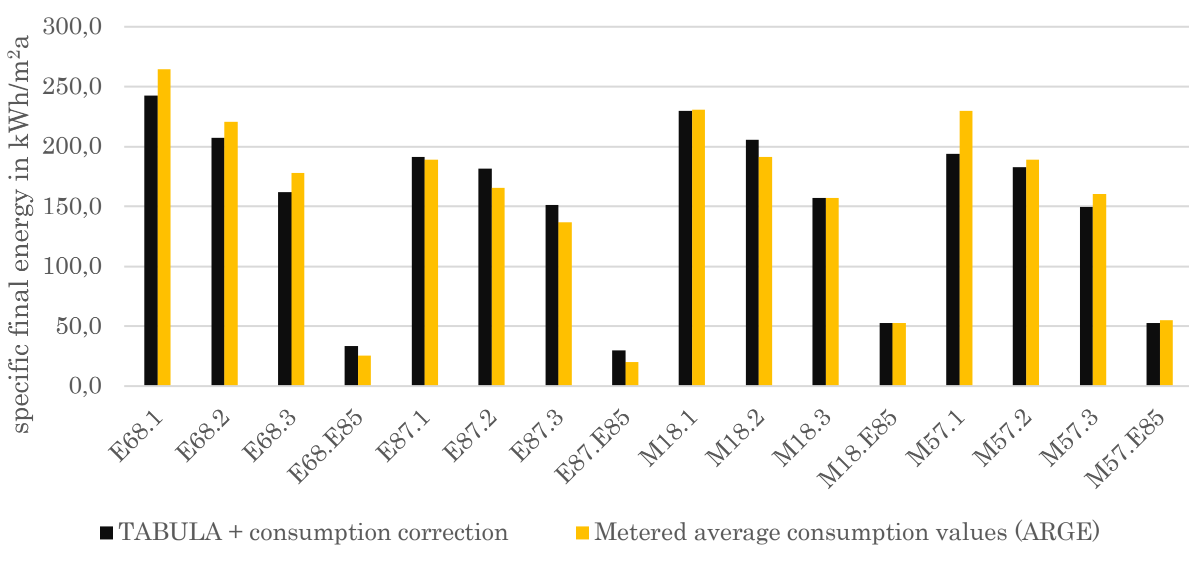


Figure 8. Comparison of TABULA consumption-corrected heat demand and metered average consumption

[37] For the purposes of simplicity, one can view this as simply six renovation levels.

6.2.2 Calculation parameters

The heat demand calculation requires some global[38] input parameters. Firstly, for Hamburg, I will use Climate Zone 2 DIN 4108-6:2003 Hamburg *Fühlsbüttel*, with a long-term average (1970-2019) of 222 heating days, 4.85°C average outside temperature (Institut Wohnen und Umwelt, 2018). This amounts to 3140 heating degree days with heating beginning at 12°C average daily temperature and 19°C indoor target-temperature. Both 19°C and 20°C can be considered standard numbers. I use the lower one for the following reason. Most building energy calculations are for one building, for which heated and unheated spaces are known. In my case, I approximate the heated space via the living space. However, it is implausible that the whole living space of Hamburg is heated to 20°C all day long. To account for that I use 19°C internal temperature. An overview of all parameters is given in Table 8.

Parameter	Value	Units
Heating Days	222	Days
External Temperature	4.85	Degrees Celsius (°C)
Internal temperature	19	
Solar Gains Horizontal	366	kWh/m^2*a
Solar Gains East	252	
Solar Gains South	370	
Solar Gains West	211	
Solar Gains North	130	
Internal Heat Gains	3	W/m^2
Air Change through User	0.4	1/h
Air Change through Infiltration	0.2	1/h

Table 8. Global parameters for the heat demand calculation for Hamburg, see Loga and Diefenbach (2013)

6.3 A note on building typologies

There are many discussions in the field of heat demand estimation[39] of whether a typology and especially the IWU Typology is "good". The reason for the existence of the ARGE Typology is probably that the IWU Typology is German-wide and an assumption was made on part of the ARGE that a regional one would be better, more accurate. However, I looked at the two typologies and noted that the assumed U-values are really not very different. Figure 9 presents an example for a EFH (single-family house) built in the 1960s and one in the 1980s together with a MFH (multifamily) in the 1920s and 1950s, so as to cover the major epochs.

[38] Applied to all data points irrespective of their attributes.
[39] I use "estimation" not in the strictly statistical sense, but in the sense "to arrive at a value for".

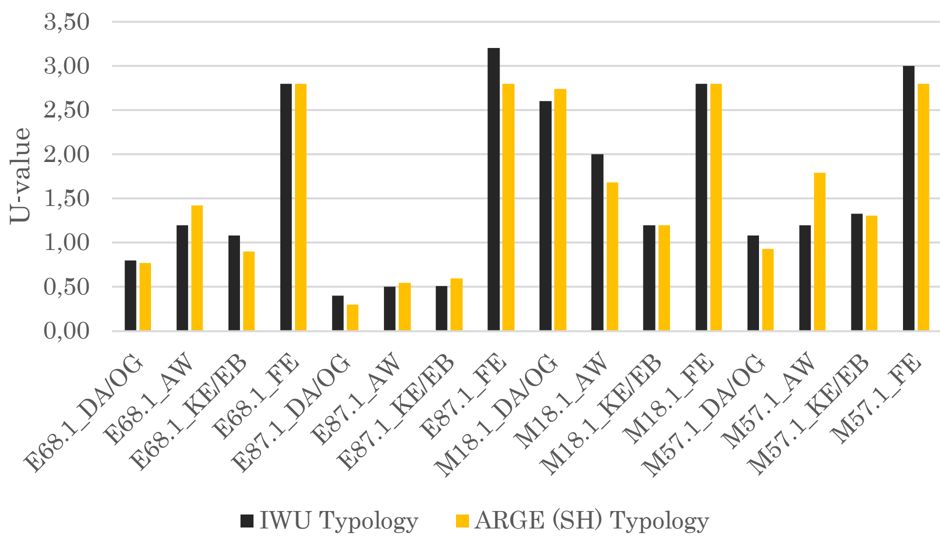


Figure 9. Examples of U-values of the IWU-Typology compared with the ARGE-Typology (SH-Typology). "DA/OG" stands for "roof", "AW" for "outer walls", "KE/FB" for "cellar or ground floor".

I compared the individual building envelope elements (roof, walls, windows and cellar) for each building type. With the exception of the outside walls of the MFH in the 1950s, most values are within 0,1-0,2 W/m²K, which is a small difference.

The ARGE typology however has some superior properties. It distinguishes more finely between the different building types with a total of six renovation levels (three baseline states and three modernisation options on top of them) compared to the IWU's three. This is important and a step in the right direction. All typologies are a form of simplification, where a broad spectrum of real-world objects (in our case, buildings) are grouped together, so that analysis and calculations are easier. Also, in most cases where a typology is not strictly consumption based, but attempts some sort of heat demand calculation, a building type is not an *average* building, but a *typical* building. In other words, not a *mean*, but a *mode* (in the sense of the statistical measures of centrality). If there is high variance of building attributes within a type, its typical value (mode) would not represent it very well, neither at the individual level, nor when scaled up (aggregated for hundreds or thousands of buildings). Averages are different, although wrong at the individual level, when scaled up, they will represent better the group.

According to the sample survey of IWU (Cischinsky & Diefenbach, 2016, p. 50) approx. 30% of buildings built before 1978 have new wall insulation, 60% new roof/attic insulation and 20% cellar insulation. Also, 60% of renovations have been carried out for a single element, while another 26% are for two elements (ibid p. 95). Renovations of all four elements or more (f. ex. together with the exchange of heating) are between 1% and 3% of all buildings. In other words, most buildings have had some form of renovation. Few are in their original state, and few are completely renovated. And this is across building sizes and epochs. This leaves the typical matrix of sizes and epochs vastly insufficient. A “typical” (from an energetic perspective) single family house building in the 1950s does not really exist anymore. The vast majority have had *something* changed or renovated. For this reason, moving to a finer renovation level split is a step in the right direction if one attempts to represent the building stock properly. In this thesis, I have taken this a step further. I consider all energy audits in my sample as a form of building types. If this is an extreme version of the typology approach, where instead of 42 types I have 300 types, or whether it constitutes a different approach altogether, is besides the point. The bottom line is that it better captures the heterogeneity of the building stock, which the data shows is present.

6.4 On using 3D data

It is safe to assume that in the future most cities would probably have and use 3D-data, so at this point the reader should ask where 3D-data comes in? Some researchers in the field are using 3D data explicitly and exclusively, f. ex the group around Ursula Eicker and Volker Coors working on SimStadt in Stuttgart (Nouvel et al., 2015). I have addressed this issue both in a publication on an archetype approach (Dochev, Seller, & Peters, 2020) and also in a publication with the SimStadt group on comparing our methods (Dochev, Gorzalka, et al., 2020). Still, an explanation is in order.

The first thing to clarify is what is 3D-Data and what is 2D-Data. Consider Figure 10, on the left is the footprint of a building. That is 2D building data. In the middle, the same building as a 3D object. This is a Level of Detail 1 (LoD1) 3D building object, a mere extrusion of the footprint.

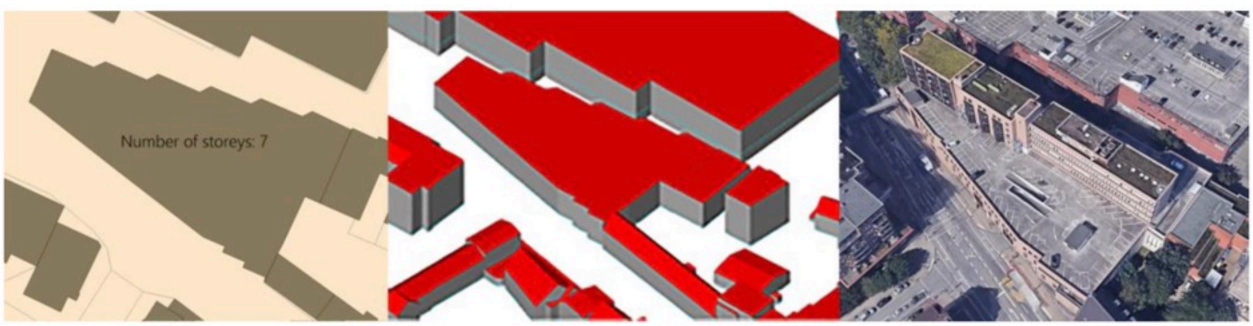


Figure 10. 2D vs 3D building data. Sources: ALKIS (Freie und Hansestadt Hamburg, Landesbetrieb Geoinformation und Vermessung (LGV) [Hamburg State Office for Geoinformation and Surveying], 2020), 3D model (Freie und Hansestadt Hamburg, Landesbetrieb Geoinformation und Vermessung (LGV) [Hamburg State Office for Geoinformation and Surveying], 2017) found in (Dochev, Seller, & Peters, 2020, p. 246)

Using solely the geometry of the footprint of a building will not be enough for almost any UBEM. But that is also **not** what I am using. Note the “number of storeys” label inside the footprint. The ALKIS data I am using has this attribute for all buildings. Consider now the LoD1 object, the difference then between what I am using and LoD1 is only in the way the 3rd dimension is measured or derived[40]. In my case it is via the number of storeys and an assumed height per storey. The LoD1 is derived from a

[40] And how they are visualised, but visualisation is not the issue here.

LiDAR[41] point cloud, which was further processed and cleaned. Therefore, in a way, I am using a form of 3D data, because I have data on the 3rd dimension.

Then the question becomes, is the LoD1 not more precise? Well, yes, but for residential buildings the difference is not very large. The real benefit of LoD1 comes for non-residential buildings[42] where the actual geometry in the 3rd dimension can be complex – f. ex. Figure 10, where in different parts of the footprint the building has different heights However, for Hamburg the LoD1 actually tends to be wrong exactly where it would bring most benefits. See how this non-residential building actually looks like compared to the LoD1 in Figure 10. Neither the ALKIS nor the LoD1 come even close to modelling the actual building volume. Interestingly, the ALKIS as a cadastral system has provisions to tackle such cases. "Building parts" can be defined, which would have different number of storeys and can be then used to model such complex building volumes. The Hamburg ALKIS regrettably does not utilise this feature of the ALKIS, i.e. there is very limited data on building parts. The Berlin ALKIS, however, uses it and that allows a more adequate modelling of such a building (Dochev, Gorzalka, et al., 2020) without formal "3D data" in the form of a LoD model.

Looking at the same Figure 10 one notices that some other buildings have more sophisticated roof forms. That is because the data is actually LoD2, although the building in the middle is presented as LoD1 (the difference between LoD1 and LoD2 is roof form). Residential buildings often have gable or hipped or other non-flat roofs and using only footprint and number of storeys would not capture this. However, the ALKIS actually has information on the type of roof as an attribute and I could factor this in without using the 3D models. The availability of this attribute differs regionally, some cities do not have it. In this case, the Hamburg ALKIS has it, the Berlin ALKIS does not.

Overall, the key issue is to calculate the building volume as close to reality as possible. Having information on the height of buildings is critical. This can be achieved with the ALKIS or with 3D data. The choice depends on the availability and the context. However, merely having something that can be visualised with a 3D-object is by no means a guarantee that this thing is a correct representation of reality. Or that it is better than something visualised on a 2D map with height as attribute.

[41] Light Detection and Ranging - an airborne remote sensing technique
[42] Which I am not covering in this thesis.

7 Formal Representation of Modelling Approach

7.1 Definition of terms

For modelling the Hamburg building stock, the assigning of energy audits to cadastral building has to be benchmarked in a specific way. Let n be the number of (ALKIS) buildings and m be the number of audits. The assigning can be represented with a two-dimensional (2D), boolean array $\boldsymbol{A}$ ($\boldsymbol{A}_{i,j} \in \{0,1\}$) with size nxm where the first dimension (rows (i)) are the buildings and the second dimension (columns (j)) are the audits. A value of 1 for any $\boldsymbol{A}_{i,j}$ indicates that a building with index i has audit with index j assigned. Because a building should get only one, whole audit assigned the sum of all columns for each row of $\boldsymbol{A}$ has to equal 1. Obviously, within the constraints ($\boldsymbol{A}_{i,j} \in \{0,1\}$ and $\sum_j^m \boldsymbol{A}_{*,j} = 1$) there are numerous possible instances of $\boldsymbol{A}$. These represent the search space for the problem.

Example: with n=4 and m=3. I will use the * sign to mean "all indices of a dimension of an array". Additionally, I will introduce the concept of advanced indexing, which is when an array is indexed with another array[43]

$$\boldsymbol{A} = \begin{bmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{bmatrix} \qquad \boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\boldsymbol{A}_{1,*} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{A}_{*,2} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\boldsymbol{A}_{1,2:} = \begin{bmatrix} 0 & 0 \end{bmatrix}$$

$$\boldsymbol{A}_{*,2:3} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}$$

Advanced indexing:

$$\boldsymbol{A}_{1,[1,3]} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

The benchmarks are the values (number of address points) that the available data gives for each category for each Census cell or for the city. Since $\boldsymbol{A}$ relates to buildings, I will define array $\boldsymbol{B}_i \in \mathbb{Z}^{\geq}$ (positive integers, including zero) of size n to store the number of address points per building. In this way, I can relate the buildings to the counts provided in the benchmark data.

Example: n=4, Each value is the number of address points for the row index. The first building has two address points, the second one address point etc.

$$\boldsymbol{B} = \begin{bmatrix} 2 \\ 1 \\ 3 \\ 4 \end{bmatrix}$$

[43] This notation is very similar to the Python numpy array indexing. Introducing it makes describing the algorithm and the code for it more intuitive. However, I avoid using Python's zero indexing and the default right open intervals, because they might be counter-intuitive to the "non-pythonic" reader. So, all indices start from 1, not from 0 as in Python and are both left and right closed. As an example 1:3 means 1 to 3 inclusive.

I am looking for an instance of $\boldsymbol{A}$ that fits the benchmarks. The benchmarks are building counts of (mutually exclusive[44]) categories of building attributes. There are p categories, hence p benchmarks. F. ex., an attribute is "type of heating", a category of this attribute is "district heating". The benchmark would be 39924 – the number of address points that have district heating according to the Census. I can represent the relationship between audits and categories with a 2D boolean array $\boldsymbol{C}$ of size mxp where rows (j) are the audits, columns (k) are all categories. Analog to the above, if $\boldsymbol{C}_{j,k} = 1$, then audit j has ("is in") category k[45]. However, some benchmarks are at the Census cell level, while others at the city level. Let s be the number of Census cell benchmarks. Then, $p - s$ is the number of city benchmarks. Let the columns (k) be ordered so that the Census cell benchmarks are the first ones (indices $1 \leq k \leq s$, and the city benchmarks indices are $s < k \leq p$)[46].

Example: m=3, p=6, s=4 Each audit (row) has some categories, the first four are for the cell benchmarks (indices 1 to s) and the last two for the city level.

$$\boldsymbol{C} = \begin{bmatrix} 1 & 0 & 1 & 0 & & 0 & 1 \\ 0 & 1 & 0 & 1 & & 0 & 1 \\ 0 & 1 & 0 & 1 & & 0 & 1 \end{bmatrix}$$

I will represent the connection between buildings and Census cells with another 2D boolean array $\boldsymbol{D}$. It will have size nxr, where r is the number of cells. Rows (i) represent buildings and columns (v) represent Census cells. A value of 1 for any $\boldsymbol{D}_{i,v}$ indicates that the building with index i resides in cell with index v.

Example: n=4, r=2, first two buildings (i, row) reside in one cell (v, column), the others reside in the second cell

$$\boldsymbol{D} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

Then, I will use a 2D array $\boldsymbol{E}$ with size rxp to store the values for the benchmarks. The rows (v) are the Census cells and columns (k) are the categories. Since these are counts and not relationships as the previous arrays: $\boldsymbol{E}_{v,k} \in \mathbb{Z}^{\geq}$. However, only some benchmarks have values for the cell level[47], therefore there are no values for some elements of $\boldsymbol{E}$. For this, I will use two versions of $\boldsymbol{E}$, one for the Census cell benchmarks ($\boldsymbol{E}'$), and one for all benchmarks at the city level ($\boldsymbol{E}''$). Both have the same number of

[44] See Section 5.1.

[45] Note that in the example array $\boldsymbol{C}$ it is unclear which categories correspond to which attributes so the mutual exclusiveness is not represented in any way. It is unnecessary to represent it, because it is a characteristic of the data (or my data pre-processing) and not a constraint on $\boldsymbol{A}$.

[46] This is for convenience of description. An implementation in a programming language can easily generalise this by indexing an array with another array to extract the different types of benchmarks, so that the order would not matter.

[47] Benchmarks at the Census cell level can easily be summed up to the city level, so all benchmarks have values at the city level.

columns p, but $\boldsymbol{E}''$ has only one row. For $\boldsymbol{E}'_{*,s+1:p}$ there are no values, but I will denote with a placeholder "?", so that the number of dimensions stays the same. This is purely for illustrative purposes.

Example: r=2, p=6, s=4 Each Census cell (v, row) has some total count for each category (k, column)

$$\boldsymbol{E}' = \begin{bmatrix} 2 & 1 & 2 & 1 & ? & ? \\ 3 & 4 & 3 & 4 & ? & ? \end{bmatrix}$$

$$\boldsymbol{E}'' = \begin{bmatrix} 5 & 5 & 5 & 5 & 0^{48} & 10 \end{bmatrix}$$

As discussed in Sections 4.2 and 5.4 the assigning of audits to buildings should also take into account the spatial relationships between them. For this, I will define yet another array $\boldsymbol{F}$ with dimensions (rxm) which will store the order (from closest to furthest) for every cell-audit pair. Note that I simplify and measure the distances from each cell to each audit rather than from each building to each audit. The difference in terms of result is very small. In the vast majority of cases the order of nearest neighbours for each building is the same as the order for the cell centroid. Computationally, however, this decreases the size of $\boldsymbol{F}$ significantly.

Example: Values are spatial orders – a value of 1 would mean the first closest neighbour, 2 the second etc.

$$\boldsymbol{F} = \begin{bmatrix} 1 & 3 & 2 \\ 3 & 1 & 2 \end{bmatrix}$$

I will use $\odot$ for the element-wise multiplication ("Hadamard product" in matrix algebra) of two arrays. A Hadamard product is defined for two matrices with the same dimensions. <u>However, I will include in the definition the concept of broadcasting</u>. It means expanding an array to match the dimensions of an array of higher dimensions by copying it multiple times. This is best explained with an example.

Examples:

$$\boldsymbol{A} \odot \boldsymbol{A}_{2,*} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \odot \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \odot \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{A} \odot \boldsymbol{A}_{*,2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \odot \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \odot \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

[48] The sharp-eyed reader would notice here that if this is a zero, then no building in the entire city should be in this category, which sounds strange – why have the category then? The answer is that this is a very simple example for a city with just four buildings, the characteristics of which I made up only for the purpose of the example. The real dataset does not have such a zero.

Lastly, in Subsection 4.3.2 I define a requirement that sparse information on individual buildings should be integrated and taken into account while assigning. Let $\boldsymbol{G}$ be a 2D boolean array of size nxp where rows are buildings and columns are categories. A value of 1 for any $\boldsymbol{G}_{i,k}$ signals that it is already known that a building has that category. The interpretation of a value of 0 however is different from the previous cases and here it stands not for a negative, but for an unknown, as in if $\boldsymbol{G}_{i,k}$= 0 the building may or may not have that category.

Example: n=4, p=6 Each building (row) has some categories

$$\boldsymbol{G} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Representing the buildings, cells, benchmarks etc. in this form facilitates the understanding of its details, the objective function and the algorithm presented in the next sections. Table 9 provides an overview of the input arrays.

Array	Type	Size	Indices	Rows	Columns
A	Boolean	nxm (197587x1502)	i, j	(ALKIS) buildings	Energy Audits
B	Integer	nx1 (197587)	i	(ALKIS) buildings	Nr. of address points
C	Boolean	mxp (1502x46)	j, k	Energy Audits	Categories (f. ex. “District Heating”)
D	Boolean	nxr (197587x19900)	i, v	(ALKIS) buildings	Census Cells
E	Integer	rxp (19900x46)	v, k	Census Cells	Categories, whereby the first s (=18) are cell level benchmarks, the rest city level.
F	Integer	rxm (19900x1502)	v, j	Census Cells	Energy Audits
G	Boolean	nxp (197587x46)	i, k	(ALKIS) buildings	Categories, whereby the first s (=18) are cell level benchmarks, the rest city level.

Table 9. Overview of input arrays (matrices)

7.2 Objective function(s)

Boolean array $\boldsymbol{A}$ is the unknown (target). Integer array $\boldsymbol{B}$ stores the number of address points for every building. Boolean arrays $\boldsymbol{C}$ and $\boldsymbol{D}$ represent the current relationships between buildings, audits, categories and Census cells as derived from the data. With this I can now calculate the number of building addresses in each Census cell (v) with each category (k) for any given instance of $\boldsymbol{A}$. I will store this in an array $\boldsymbol{U}$[49]. It is effectively an aggregation of the buildings' attributes at the Census cell and city levels. To correspond to the benchmarks, I will again define two versions of $\boldsymbol{U}$. One for benchmarks at the Census cell level: $\boldsymbol{U}'_{v,k}$ a 2D array with size rxp (logically this array has the same shape as $\boldsymbol{E}'$). And one for the city level benchmarks: $\boldsymbol{U}''_{k}$ a 1D array with size p (same shape as $\boldsymbol{E}''$). The two arrays are simply an aggregation of an instance of an assigning ($\boldsymbol{A}$):

Eq. 3.

$$\boldsymbol{U}'_{v,k} = \sum_{i}^{n} \sum_{j}^{m} (\boldsymbol{A} \odot \boldsymbol{B} \odot \boldsymbol{D}_{*,v} \odot \boldsymbol{C}^{T}_{*,k})_{ij}$$

Eq. 4.

$$\boldsymbol{U}''_{k} = \sum_{i}^{n} \sum_{j}^{m} (\boldsymbol{A} \odot \boldsymbol{B} \odot \boldsymbol{C}^{T}_{*,k})_{ij}$$

Integer arrays $\boldsymbol{E}'$ and $\boldsymbol{E}''$ store the values for the benchmarks that have to be matched. Floating-point array $\boldsymbol{F}$ stores the spatial relationships between buildings and audits (using cells as proxy). Boolean array $\boldsymbol{G}$ stores the information on individual buildings where one or more attribute categories are known and the target assigning has to match.

With this apparatus I can now define an objective function which gives a meaningful answer to the question "*does my model fit the benchmarks well?*". Note that the true task is to arrive at the instance of $\boldsymbol{A}$ that is closest to the real building stock. Let this be $\widehat{\boldsymbol{A}}$**.** Finding an instance of $\boldsymbol{A}$ that fits benchmarks would mean that it is likely that $\boldsymbol{A}$ is close to $\widehat{\boldsymbol{A}}$**.** However, this is still an assumption. One I test in Section 9.2. For now, the objective is to fit the benchmarks.

I use the sum of squared differences as the type of metric for optimisation. I have two sets of benchmarks: at the city and Census cell levels. Further, I want to assign audits that are close in geographic space. This turns the problem into a multi-objective one:

[49] i.e. for every assigning of audits to buildings

Minimise with respect to $\boldsymbol{A}$:

Formula 1	$$\sum_{k=1}^{s}(\boldsymbol{E}'_{v,k} - \boldsymbol{U}'_{v,k})^2$$	Minimise the sum of squared differences between estimated counts and benchmarks at cell level
Formula 2	$$\sum_{k=1}^{p-s}(\boldsymbol{E}''_{k} - \boldsymbol{U}''_{k})^2$$	Minimise the sum of squared differences at city level
Formula 3	$$\sum_{i=1}^{n}\sum_{j=1}^{m}(\boldsymbol{A} \odot \boldsymbol{F})_{ij}$$	Minimise the spatial orders (in geographic space) from a building to an assigned audit

where:

Eq. 3. *(defined previously)*	$$\boldsymbol{U}'_{v,k} = \sum_{i=1}^{n}\sum_{j=1}^{m}(\boldsymbol{A} \odot \boldsymbol{B} \odot \boldsymbol{D}_{*,v} \odot \boldsymbol{C}^{T}_{*,k})_{ij}$$	The two aggregation terms - Census cell and city levels
Eq. 4. *(defined previously)*	$$\boldsymbol{U}''_{k} = \sum_{i=1}^{n}\sum_{j=1}^{m}(\boldsymbol{A} \odot \boldsymbol{B} \odot \boldsymbol{C}^{T}_{*,k})_{ij}$$	

subject to:

$$\boldsymbol{A}_{i,j} \in \{0,1\}$$

$$\sum_{j=1}^{m}\boldsymbol{A}_{*,j} = 1$$

Only one, whole audit assigned

$$\boldsymbol{G}_{i,k} = 1 \rightarrow \boldsymbol{A}_{i,j} \odot \boldsymbol{C}_{j,k} = 1$$

If an attribute category is known, make sure the assigning keeps it

8 Designing the Modelling Algorithm

My aim in this thesis is to contribute to the development of UBEMs by finding a good way to integrate widely and regularly available data sources[50]. This has two implications for the modelling algorithm. Firstly, it has to be able to tackle distinct, but similar datasets from other cities, so a certain degree of transferability has to be evidenced. Secondly, the logical similarities of characteristics of the datasets can be used to guide the development of the algorithm.

A simple example for what I mean is the estimation of the search space for the choice of algorithm. The number of possible instances of $\boldsymbol{A}$, given the constraints, are m^n instances. Hamburg, for which I will test the algorithm has a population of approx. 1.8 million with 250 000 residential buildings. The sample of energy audits is approx. 1300. The possible instances of $\boldsymbol{A}$ are then approx. 1300^{250000}, an obviously astronomical number. At this point, it is rather safe to assume that just brute-forcing (going through all possible permutations) to find a solution is not possible. Additionally, it is also safe to assume that there would hardly be a use case where this is not true[51].

However, the energy audits are <u>not</u> a set (every member being unique), when it comes to the permutations of their categories. They are unique with their spatial locations, but for the purposes of maximising the fit to the benchmarks only the unique ones according to their characteristics can be taken – 291. The search space still cannot be brute-forced, but this step can have computational benefits.

Simple brute-force is obviously not possible, a smarter search is required. Given the optimisation methods presented in Chapter 6 – "simulated annealing", "fitness-based", "generalised regression" and "IPF" some logical arguments can be made as to which ones would have most potential.

Firstly, both "generalised regression" and "IPF" produce non-integer results ("weights"). This means that with them a resulting instance of $\boldsymbol{A}$ would be an array of floating points. It could match the benchmarks well and $\sum_m^n \sum_j^m \boldsymbol{A}_{i,j} = n$, but $\boldsymbol{A}_{i,j} \notin \{0,1\}$. In Spatial Microsimulation, this is typically solved with an "integerisation" step (a state of the art method for this "truncate, replicate, sample" (Lovelace & Ballas, 2013)). There are two general approaches to the integerisation step – a deterministic (the simplest of which is rounding) and a probabilistic, where weights are treated as a probability distribution. The "truncate, replicate, sample" method is actually a hybrid of the two, where the decimal weights are split into an integer and decimal part. The "sample" step in the method samples from the probability distribution of the decimal part of the weights and adds the results to the integer part of the weights. Thus, the weights are "integerised".

---

[50] As explained in Chapter 3 most of the datasets I use are not unique to Hamburg, but rather standard in spatial planning – buildings, Census cells etc.

[51] Even a town with 10 audits and 120 buildings would have 10^{120} permutations, roughly the number of possible chess games, still incomputable.

However, here, the nature of the problem at hand renders this approach unsuitable. There are, on average, 12 buildings in each Census cell and 291 unique energy audits (ratio of 0.04). This means that the "candidates" for each cell are on average approx. 25 times more than the targets[52]. After testing IPF, the vast majority of weights I obtained were small decimal numbers (10^{-2}). In other words, in far too many weights, there was no integer part to begin with.

This meant that, in my case, the integerisation step (which came down to sampling) gained far too much on influence and made the whole approach more probabilistic than usually. In Spatial Microsimulation, populations, not buildings are usually modelled and at Census tract or similar spatial level rather than 100m Census cells. At the usual spatial levels, the usual situation is at least a couple of hundred inhabitants per spatial unit (f. ex. Hamburg has on average 2000 inhabitants per Census tract). Then the ratio between the size of the individual sample (energy audits in my case, surveys in Spatial Microsimulation usually) and population size of spatial unit on average is much closer to 1 or above and not 0.04.

For this reason, I concentrated on innately integer-producing algorithms - "fitness-based" and "simulated annealing". Between the two approaches, I adopted the "fitness-based" approach, because it is deterministic. It had produced good results for similar problems in the literature (Ma & Srinivasan, 2015). Note that a fitness-based approach can be transformed into a quasi-annealing approach. Consider taking a window of n top "fitness" scores and choosing among them rather than the single best score. The window can then be reduced after each iteration (similar to what the "temperature" does in simulated annealing). Then the algorithm will be close to a simulated annealing algorithm. I implemented this logic, but do not use it per default, see next section for details.

8.1 Algorithm – simple explanation and pseudocode

The algorithm I use is based on the "fitness-based" approach described by Ma and Srinivasan (ibid) whereby audits are assigned based on a "scoring" ("fitness") value[53]. The basic logic is to iteratively pick energy audits for each building such that the aggregates for the cells and the city converge on the benchmarks. The algorithm orders the buildings according to the number of categories that are known ($\boldsymbol{G}$) and the buildings' size. It then assigns each unique energy audit to the first building in the first cell and chooses the one that best fits (according to the objective function). The tie-breaker, if needed, is the geographical distance from an audit to a building. Note that when an energy audit is assigned to a building this changes the difference between the aggregates (given the concrete assigning) and the cell and city benchmarks. Thus, each assigning influences the scores for the ones that follow. This is not a problem for the cell benchmarks, since each cell's newly calculated aggregates are independent

[52] This is a peculiarity, which is uncommon in the field of spatial microsimulation, where the spatial units comprise of at least a couple of hundred individuals. Therefore, it was not surprising that I did not find examples of similar situations in the literature.
[53] In the field of computer science, the strategy to modify hill-climbing by using some sort of "fit" measure ("Best fit search") to choose the next candidate can be traced back to Pearl (1985).

from those of other cells. For city benchmarks however, every building's assigning influences all others, since they all adhere to a single benchmark for the city. In order to decrease the importance of the building order, the algorithm uses batches of buildings and updates the city benchmarks not after each building, but after each batch. Each batch consists of a number (an algorithm parameter) of buildings each from a different cell. F. ex. batch one would be the first building from cell one, the first building from cell two etc. The algorithm loops over all batches and then starts from the beginning. This gives it multiple chances to "check again" the scores for a building, after other assignings have been changed. The first building in a cell would be checked again after the first iteration of all buildings, but this time its fitting score might change, since the other buildings with assigned audits would influence the totals for the benchmarks. Note that this does not prevent the possibility to get stuck in local optima, but it helps against it. Attempting to (approximately) solve the local optima problem would require a form of stochasticity, similar to the simulated annealing approach. The algorithm I propose allows for this by specifying an "assigning_window_size" and "window_reduction" parameters. The "assigning_window_size" controls how many top scoring audits to consider at each assigning and take a random audit among them. The "window reduction" is by how much to decrease the window size after each batch. I practically turn-off this functionality by setting a window size of 1. The reasons are that even without this stochasticity the algorithm performed well on the Hamburg data and that a deterministic algorithm has benefits when trust and explainability to stakeholders are concerned. Still, the option is available.

The algorithm written with for loops and in pseudocode is given below. A more detailed example using arrays is given in the next section.

```
#prepare inputs and parameters
DECLARE INPUTS

        alkis_buildings,
        census_cells,
        energy_audits

DECLARE PARAMETERS
        max_iterations = 100,
        batch_size = 500,
        scoring_metric = sum_of_squared_differences
        iteration_stop_metric = frobenius_norm,
        assigning_window_size = 1,
        window_reduction = 0

#prepare derived variables
energy_audits_unique = GET UNIQUE energy_audits #unique in terms of their categories

distances_of_audits_to_cells = SPATIAL DISTANCE (energy_audits, cells)

ORDER buildings BY (Census_cell_id, building_known_attributes, building_floow_area)
ORDER buildings BY (index in cell, building_known_attributes, building_floow_area)
                        )
batches_of_cells_with_buildings = GROUP cells_with_buildings IN BATCHES OF batch_size
#Note that the batches are created and ordered in a special way. The buildings are sorted according to their index in each cell,
which is based on the known attributes (categories) and their size. So, at the beginning of the "buildings" list are all "first
buildings". The list continues with all "second buildings", etc until the n-th buildings, where n is the max number of buildings
```

in the Census cells. Within all the first buildings, the order is based again on known attributes and floor area. Then, within all "second buildings" again the same etc. Finally, the list is split into batches of size equal to the batch size.

```
#run fitting
FOR iteration IN max_iterations:
    FOR batch IN batches_of_cells_with_buildings:
        FOR building IN batch:
            IF assigned_audit NOT none:  #if the building has an audit assigned
                current_cell_aggregates -= this_building_cell_attributes #remove it …
                current_city_aggregates -= this_building_city_attributes #… we check if another fits better
            ELSE:
                PASS
            FOR unique_audit IN energy_audits_unique:
                ASSIGN unique_audit TO building
                score_local = CALCULATE scoring_metric (
                                                    building,
                                                    current_cell_aggregates,
                                                    cell_benchmarks
                                                    )
                score_global = CALCULATE scoring_metric (
                                                    building,
                                                    current_city_aggregates,
                                                    city_benchmarks
                                                    )
            END FOR
            SORT unique_audits_assigned ACCORDING TO (
                                                    known_building_benchmarks,
                                                    min_score_local,
                                                    min_score_global,
                                                    distances_of_audits_to_cells
                                                    )
            CHOOSE RANDOM top_unique_audit FROM unique_audits IN window_size #if size is 1: the top score

            UPDATE current_cell_aggregates WITH buildings_with_assigned_audits
        END FOR

        UPDATE current_city_aggregates WITH buildings_with_assigned_audits
        assigning_window_size -= window_reduction
    END FOR

    all_current_aggregates = AGGREGATE buildings_with_assigned_audits
    difference_to_benchmarks = current_aggregates - benchmarks

    convergence = CALCULATE iteration_stop_metric (
                                    difference_to_benchmarks,
                                    difference_to_benchmarks_from_previous_iteration
                                    )

    IF convergence = 0:
        TERMINATE
    ELSE:
        CONTINUE

END FOR

#now all buildings have a unique_audit assigned, next I choose which energy audit to assign, based on the #spatial distance
FOR building IN buildings_with_assigned_audits:
    CHOOSE energy_audit ACCORDING TO distance(
                                    building,
                                    energy_audits_of_assigned_unique_audit
                                    )
```

8.2 Algorithm – more detailed explanation and example with matrices/arrays

While the pseudocode in the previous section describes the major steps, an implementation using arrays is described below. For simplicity I am presenting an example with the "assigning_window_size" and "window_reduction" parameters set to 1 and 0 respectively, which basically turns them off.

Since the benchmarks count the building address points, the algorithm has to account for that and count each building with its respective number of address points. From the examples in Chapter 7:

Example:

$$\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\\1&1&1\end{bmatrix}\odot\begin{bmatrix}2\\1\\3\\4\end{bmatrix}=\begin{bmatrix}2&2&2\\1&1&1\\3&3&3\\4&4&4\end{bmatrix}$$

Notes

The process starts with a fully filled out version of $\boldsymbol{A}$ where every building has every audit assigned and multiply it with $\boldsymbol{B}$. The row and column indices point to buildings and energy audits respectively.

Since the energy audits are not unique, this can be simplified by tacking only the unique and keeping track of which non-unique audits point to each unique audit:

Example:

$$\boldsymbol{C}=\begin{bmatrix}1&0&1&0&0&1\\0&1&0&1&0&1\\0&1&0&1&0&1\end{bmatrix}$$

$$\begin{bmatrix}2&2&2\\1&1&1\\3&3&3\\4&4&4\end{bmatrix}\rightarrow\begin{bmatrix}2&2\\1&1\\3&3\\4&4\end{bmatrix}$$

Notes

Array $\boldsymbol{C}$ holds the energy audits. Since audits at indices 2 and 3 have the same categories, the result from the previous step can be simplified. The rows remain the same, but the number of columns is reduced. A different array is kept which tracks which indices from the non-unique correspond to which in the unique.

Then the now simplified $\boldsymbol{A}$ is split according to the Census cells in which the buildings reside using $\boldsymbol{D}$

Example:

$$\boldsymbol{D}=\begin{bmatrix}1&0\\1&0\\0&1\\0&1\end{bmatrix}$$

$$\rightarrow\begin{bmatrix}2&2\\1&1\end{bmatrix}\quad\begin{bmatrix}3&3\\4&4\end{bmatrix}$$

Notes

Array $\boldsymbol{D}$ describes the connection between buildings (rows) and cells (columns). According to it, the first two buildings are in cell 1 and the other two in cell 2. The original order of the buildings can be arbitrary, from here on it is not anymore.

The algorithm then sorts the buildings within the cells firstly by the number of known categories and then by the floor area (more known categories and larger area go higher up in the order).

Example:

$$\boldsymbol{G} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\sum_{k}^{p} \boldsymbol{G}_{*,k} = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 3 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 3 & 3 \\ 4 & 4 \end{bmatrix}$$

Notes

Array $\boldsymbol{G}$ holds the known categories for individual buildings. Summing up over all categories shows how many known are there for each building (row). In the example, buildings with indices 1 and 3 have one known category, index 2 has two and index 4 has none. The split version of $\boldsymbol{A}$ gets reordered again. According to Array $\boldsymbol{G}$, in the first cell, building with index 1 has one known category, while building with index 2 has two known ones. The order is flipped. Note, the numbers in the array represent the address points, while the reordering is done with via the indices. The second Census cell remains the same, since the building with index 3 (which is index 1 in the cell) has more known categories than the building with index 4.

After that, the algorithm starts assigning audits to buildings for groups of buildings in parallel. The groups are based on the index of the building within a cell, so first all buildings with index 1 in their respective cells, then index 2 etc. These groups are further subdivided with a "batch size" variable, which is a parameter that can be set. The batch size can take values between 1 and the total number of cells. With f. ex. 1000 cells, a "batch size" of 100 would mean the buildings with index 1 in the first 100 cells get an audit assigned first, then the next batch and only after all batches, the algorithm moves to building with index 2. With this, the order of the cells also becomes important. The cells are pre-sorted, similarly to the buildings, based on the known categories and the size of the cell, with cells with more known categories and larger sum of floor areas taken first.

Example:

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 3 & 3 \\ 4 & 4 \end{bmatrix}$$

Notes

In this example, the buildings with index 1 in their respective cells have address point counts of 1 and 3 respectively. A batch size of 2 would mean that they get audits assigned in parallel. A batch size of 1 would mean they get audits assigned in a consecutive fashion. See further down for an explanation into "parallel" vs "consecutive" assigning.

Then, for each building all potential candidate energy audits and their categories are taken.

Example:

$$\begin{bmatrix}1\\1\end{bmatrix} \odot \boldsymbol{C}'$$

$$\begin{bmatrix}1\\1\end{bmatrix} \odot \begin{bmatrix}1 & 0 & 1 & 0 & 0 & 1\\0 & 1 & 0 & 1 & 0 & 1\end{bmatrix}$$

Notes

$\begin{bmatrix}1\\1\end{bmatrix}$ is simply a transposed $[1 \quad 1]$. So, building with index 1 in cell one. It gets multiplied with $\boldsymbol{C}$, which stores the categories of the energy audits, however only the two unique ones instead of all[54].

Then the algorithm considers the benchmarks for the cell and for the city before assigning an audit to the building at hand. For the very first building at the very first iteration, the difference to the benchmarks are the values of the benchmarks themselves. After each building gets an audit assigned, the difference to the benchmarks is changed, because the cells start getting "populated" with audits and thus the modelled counts of categories start to change.

Example:

$$\boldsymbol{E}'_{v=1} = [2 \quad 1 \quad 2 \quad 1 \quad - \quad -]$$

$$\boldsymbol{E}'' = [5 \quad 5 \quad 5 \quad 5 \quad 0 \quad 10]$$

Notes

Before the first building in the first cell gets an audit assigned, the benchmarks have the original values. When the second building gets its turn, the benchmarks are already modified

Then fitting scores are computed for the building in question. They are the squared sum of differences to the benchmarks.

Example:

Census Cell Benchmarks

$$\begin{array}{r} -\,[2 \quad 1 \quad 2 \quad 1 \quad ? \quad ?] \\ [1 \quad 0 \quad 1 \quad 0 \quad 0 \quad 1] \\ \hline [1 \quad 1 \quad 1 \quad 1 \quad 0 \quad 0] \end{array} \rightarrow [4]$$

$$\begin{array}{r} -\,[2 \quad 1 \quad 2 \quad 1 \quad ? \quad ?] \\ [0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1] \\ \hline [4 \quad 0 \quad 4 \quad 0 \quad 0 \quad 0] \end{array} \rightarrow [8]$$

City Benchmarks

$$\begin{array}{r} -\,[5 \quad 5 \quad 5 \quad 5 \quad 0 \quad 10] \\ [1 \quad 0 \quad 1 \quad 0 \quad 0 \quad 1] \\ \hline [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 27] \end{array} \rightarrow [27]$$

$$\begin{array}{r} -\,[5 \quad 5 \quad 5 \quad 5 \quad 0 \quad 10] \\ [0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1] \\ \hline [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 27] \end{array} \rightarrow [27]$$

$$\rightarrow priority \begin{bmatrix}\mathbf{4}\\8\end{bmatrix} \begin{bmatrix}\mathbf{27}\\\mathbf{27}\end{bmatrix}$$

Notes

For both Census cell benchmarks and city benchmarks, both unique candidates get a score. Lower scores are preferred.

After computing the fitting scores, the known categories are considered again.

[54] Obviously having a single address point makes the multiplication unnecessary, but I am giving a generalised presentation of the steps.

Example:

$$\boldsymbol{G} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{C'} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & \mathbf{1} & 0 & 1 \end{bmatrix}$$

$$\rightarrow priority \begin{bmatrix} 2 \\ \mathbf{1} \end{bmatrix}$$

Notes

Array $\boldsymbol{G}$ points to building 2 (which is now taken with index 1) as having category 4. From the potential candidates $\boldsymbol{C'}$ the second one has it, the first does not, so the priority is $\begin{bmatrix} 2 \\ \mathbf{1} \end{bmatrix}$ which means "prefer the second one".

With the two fitting scores and the known categories a choice is made as to which energy audit to assign. The choice is based on a priority of scores which is again a setting that can be changed. I use the known categories first, then the Census cell benchmarks and then the city benchmarks. The final tie-breaker is spatial distance – if all else is equal the closer energy audit gets assigned.

Example:

$$priority \begin{bmatrix} 2 \\ \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{4} \\ 8 \end{bmatrix} \begin{bmatrix} \mathbf{27} \\ \mathbf{27} \end{bmatrix}$$

Notes

For the scores, lower is always better. Energy audit 2 gets assigned, because it has a known category and the known categories are preferred over the other scores. In this case the scores obviously don't change anything, but this is because of the simplicity of the example.

With the current ordering of the scores instead of some compounded index or integral measure, the known categories have a hard and fixed influence. In reality, this means that if a given building is known to have a category, the assigned audit will always also have it. In other words, the known categories are taken as the most accurate of data and the algorithm tries to fit the benchmarks around this known information.

After an audit is assigned, the difference between model (aggregates) and benchmarks is updated to reflect that.

Example:

$$\begin{array}{r} \begin{bmatrix} 2 & 1 & 2 & 1 & ? & ? \end{bmatrix} \\ - \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \\ \hline \begin{bmatrix} 2 & 0 & 2 & 0 & ? & ? \end{bmatrix} \end{array}$$

Notes

Update $\boldsymbol{E}'_{1,*}$. I assigned unique audit with index 2, so now the difference to the benchmarks of the cell changes.

Note that in the example only the Census cell ($\boldsymbol{E'}$) benchmarks were updated. This is an important point – the moment when the city benchmarks ($\boldsymbol{E''}$) are updated is not after each building, but after each batch. If batch size equals 1 (the minimum), then both benchmarks are updated at the same time. If batch size equals the number of cells (the maximum), then the city benchmarks get updated after each building cell index. The necessity for this is that as opposed to Census benchmarks, city benchmarks include all buildings. In other words, an assigning of audits within a cell can be independent of other cells when it comes to the Census cell benchmarks, but not when it comes to the city benchmarks. The following example explains this further:

Example for:	Notes
[3 3] With batch size=1 $\boldsymbol{E}'' = [5 \quad 5 \quad 5 \quad 5 \quad 0 \quad 9]$ With batch size=2 $\boldsymbol{E}'' = [5 \quad 5 \quad 5 \quad 5 \quad 0 \quad 10]$	Consider the second building with index 1 in its cell. It is different from the previous building. For Census cell benchmarks, whichever audit was assigned to the previous building in the batch does not matter, the building is from a different cell. However, the city benchmarks influence all assignings. If batch size is 1, then the city benchmarks would have been updated after the previous building was assigned and would have slightly lower values then if batch size was 2.

The moment at which the city benchmarks get updated is what I referred to previously with "parallel" vs "consecutive" assigning. If two buildings get audits assigned in parallel, the fitting scores used consider the same city benchmarks for both buildings. If they are assigned "consecutively", then the assigning of the first building influences the fitting scores and thus the assigning of the second building. The algorithm continues assigning by looping over all batches of buildings, indices within cells and iterations (Figure 11).

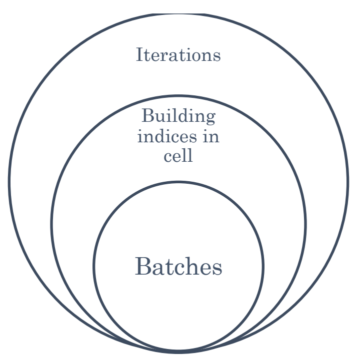


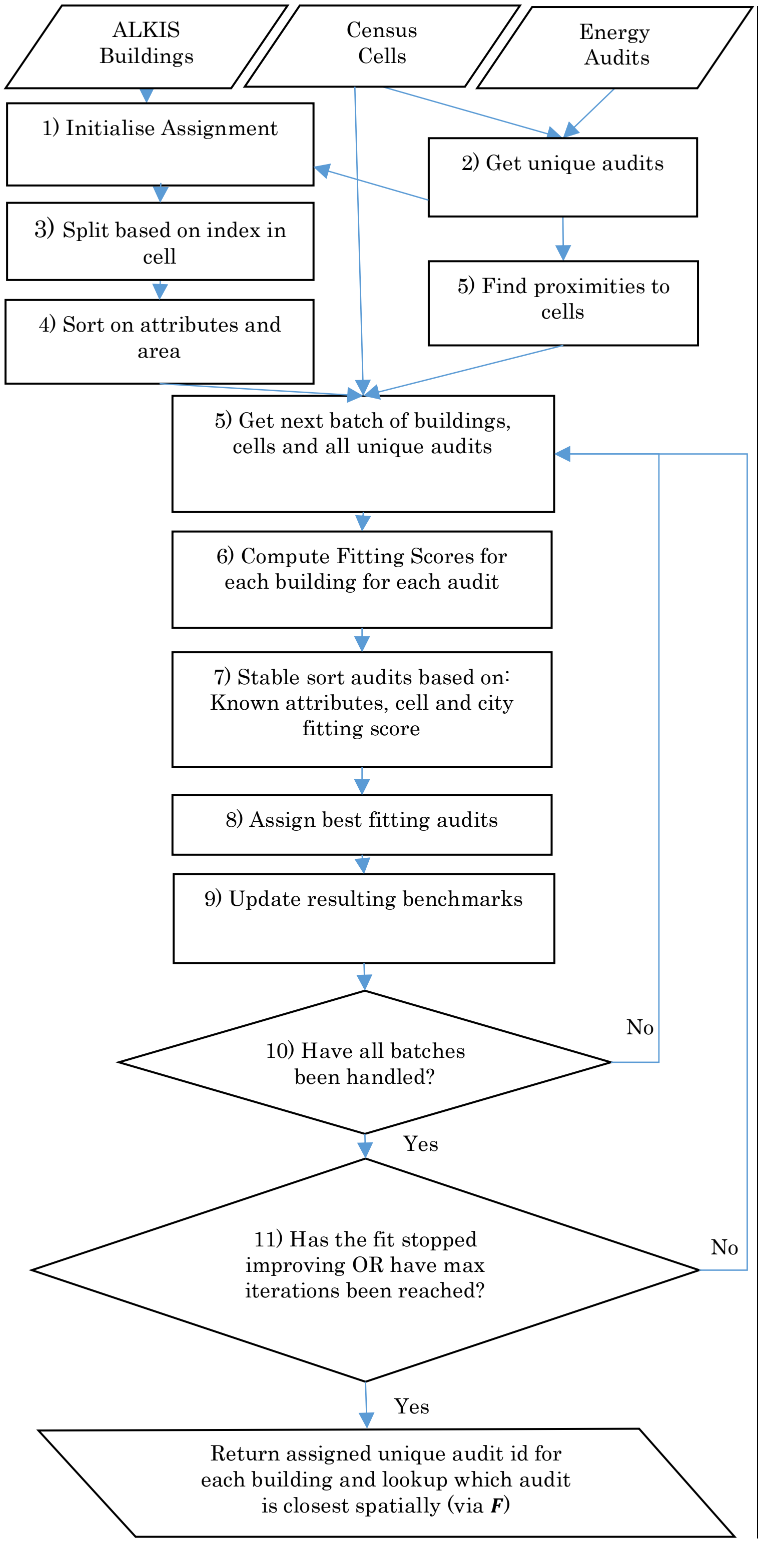
Figure 11. Structure of algorithm loops

After each building from each batch gets assigned an audit an iteration finishes. Then another one begins with the first building. All buildings "remember" their last assigned audit from the previous iterations. Because the benchmarks are calculated at the cell and city levels, for the first building the fitting scores might now be different, because it takes into account the assigned audits to all other buildings which came after it in the previous iteration. Therefore, any building might get a different audit in each subsequent iteration.

The algorithm still needs a termination point. After each iteration a convergence score is calculated. It is the Frobenius norm of the differences to the benchmarks between iterations. If the norms for both Census cell and city benchmarks reach zero, the algorithm terminates. The logic is that if there is no difference between two iterations, then no building got a re-assigned audit. No future iteration would change that, we have reached convergence. Alternatively, a max number of iterations can be set. I used 100 iterations, but the algorithm converged usually sooner.

The assigning until now assigned *unique* audits (in terms of their combination of benchmarking (!) attributes). In a final step, I use $\boldsymbol{F}$ to convert unique audits to the original audit dataset. Basically, for every building multiple non-unique audits can be assigned, so I choose one based on the spatial proximity to the given building.

Figure 12 presents all these steps and how they relate to each other. See next section for runtime and convergence metrics.



1) Initialise	$\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\\1&1&1\end{bmatrix} \odot \begin{bmatrix}2\\1\\3\\4\end{bmatrix} = \begin{bmatrix}2&2&2\\1&1&1\\3&3&3\\4&4&4\end{bmatrix}$
2) Get unique using $\boldsymbol{C}$	$\begin{bmatrix}2&2\\1&1\\3&3\\4&4\end{bmatrix}$
3) Split based on Census cells and index in cell	$\boldsymbol{D} = \begin{bmatrix}1&0\\1&0\\0&1\\0&1\end{bmatrix}$
	[[2 2] [3 3]] [[1 1] [4 4]]
4) Sort based on known attributes ($\boldsymbol{G}$)[55]. Use it to index.	$\sum_{k}^{p} \boldsymbol{G}_{*,k} = \begin{bmatrix}1\\2\\1\\0\end{bmatrix} \rightarrow \text{order} \begin{bmatrix}2\\1\\1\\2\end{bmatrix}$
	[[1 1] [3 3]] [[2 2] [4 4]]
5) First building in first cell and the potential candidates	$\begin{bmatrix}1\\1\end{bmatrix} \odot \begin{bmatrix}1&0&1&0&0&1\\0&1&0&1&0&1\end{bmatrix}$
If after first iteration, take out the previously assigned audits on buildings with this index	$-\begin{array}{l}[2\quad 1\quad 2\quad 1\quad ?\quad ?]\\ [0?\quad 0?\quad 0?\quad 0?\quad ?\quad ?]\end{array}$
6) Cell Fitting $\boldsymbol{E}'_{1,*}$: Squared sum of element-wise differences. Lower is better.	$\begin{array}{c} -[2\quad 1\quad 2\quad 1\quad ?\quad ?] \\ \begin{bmatrix}1&0&1&0&0&1\\0&1&0&1&0&1\end{bmatrix} \\ \hline \begin{bmatrix}1&1&1&1&0&0\\4&0&4&0&0&0\end{bmatrix}\end{array} \rightarrow \begin{bmatrix}\mathbf{4}\\8\end{bmatrix}$
6) City fitting score $\boldsymbol{E}''_{*}$	$\begin{array}{c} -[5\quad 5\quad 5\quad 5\quad 0\quad 10] \\ \begin{bmatrix}1&0&1&0&0&1\\0&1&0&1&0&1\end{bmatrix} \\ \hline \begin{bmatrix}0&0&0&0&0&27\\0&0&0&0&0&27\end{bmatrix}\end{array} \rightarrow \begin{bmatrix}\mathbf{27}\\\mathbf{27}\end{bmatrix}$
7)/8) Known categories. Building 2, which is now taken first has to have category 4. Audit 2 has it, audit 1 does not. The priority order – [known, cell, city] decides that audit 2 gets assigned with local score 8 and city score 27.	$\boldsymbol{G}_{1,2,*} = \begin{bmatrix}0&0&1&0&0&0\\0&0&0&\mathbf{1}&0&1\end{bmatrix}$ $[2\quad 1\quad 2\quad \mathbf{1}\quad ?\quad ?]$ $priority \begin{bmatrix}2\\\mathbf{1}\end{bmatrix} \begin{bmatrix}\mathbf{4}\\8\end{bmatrix} \begin{bmatrix}\mathbf{27}\\\mathbf{27}\end{bmatrix}$
9) Update $\boldsymbol{E}'_{1,*}$ $\boldsymbol{E}''_{*}$ gets updated not after each cell, but after a subset of cells. Subset size is an algorithm parameter.	$\begin{array}{c} -[2\quad 1\quad 2\quad 1\quad ?\quad ?] \\ [0\quad 1\quad 0\quad 1\quad 0\quad 1] \\ \hline [2\quad 0\quad 2\quad 0\quad ?\quad ?]\end{array}$
11) Check Frobenius norm of the differences between iterations (both levels, cell and city)	cell level: $\sum_{k}^{s}\sum_{v}^{r} \begin{bmatrix}1&1&1&1&0&0\\4&0&4&0&0&0\end{bmatrix}^2 = 36$

Figure 12. Algorithm steps

[55] Buildings 1 and 3 have just 1 known category, but 3 goes higher in the order since it is bigger. I omit the floor area calculation based on which the algorithm decides which building is bigger, to keep the example simpler.

8.3 Runtimes and convergence

The following testing was performed using a python numpy implementation of the algorithm within QGIS. The hardware I used was a Core i7, 8GM RAM laptop.

For the real data (~250 000 ALKIS building objects, 1500 energy audits, 291 unique energy audits and ~20 000 Census cells) the algorithm reached convergence after 9 iterations. For the convergence, I use the Frobenius norm of the element-wise difference between the modelled aggregates and the benchmarks. If the difference between the norm in two consecutive iterations reaches zero, the algorithm terminates. Note that this does not mean that the algorithm minimised the difference to the benchmarks, only that it could not do better on the subsequent iteration. Since there are benchmarks on the city level and the cell level, both have to stop improving for the algorithm to terminate.

Iteration	City Benchmarks	Cell Benchmarks
1	25573	4655
2	973	28
3	26.7	0.75
4	13.7	0.027
5	2.08	0.06
6	2.42	0.01
7	1.62	0.005
8	0.71	0.0
9	0.0	0.0

Table 10 Convergence measured with differences in Frobenius norm between two iterations for the Hamburg data

Additionally, I tested the average number of iterations until convergence in a randomised simulation. For the details of how are the random building stocks generated see Section 9.2.

Number of iterations*	Building Stock Size**		
	500	1 000	5 000
Min	3	4	5
Max	17	23	50
Mean	6.21	8.79	22.18
Std. Dev.	2.70	3.62	12.35

* Exploratory statistics of the number of iterations after 100 randomised runs

** Measured in number of Census cells

Table 11. Convergence in 100 simulated modelling cases with building stocks of different sizes

It is worth noting that due to the city benchmarks, the assigning of audits in one cell is not independent from the assigning in other cells[56]. Thus, in some situations the algorithm might diverge and instead terminate when the maximum number of iterations is reached. The divergence would be caused by the attempts of the algorithm to balance the Census cell and city benchmarks.

[56] This is because difference between modelled and city benchmarks is recalculated after each iteration, while for Census cell benchmarks after each batch of cells within an iteration.

9 Validation

The validation of the model has three parts. The internal validation represents how well the algorithm matched the benchmarks. It is "internal", because the benchmarks are used as part of the modelling. The part "Validation through random generation" is where I to attempt to estimate how good the same technique would be for other cities (different input data). For this I use a randomised dataset. The external validation is about comparing the model to known data, which was not part of the modelling. In this case – consumption data.

9.1 Internal validation

For the metric of optimisation, I use the sum of squared differences between category counts in the data (benchmarks) and category counts as computed by my model. While this has some practical benefits while fitting – it penalises larger deviations exponentially[57] – it is not very intuitive for interpretation. Hence, I will use the mean and standard deviation of the absolute percentage difference (APD) for the internal validation of the final result:

Eq. 5.[58]

$$cityMAPD = \frac{\sum_{k=1}^{p}\left(\frac{\left|\boldsymbol{E}_k'' - \boldsymbol{U}_k''\right|}{\boldsymbol{E}_k''}\right)\boldsymbol{E}_k''}{\sum_{k=1}^{p}\boldsymbol{E}_k''} 100\%$$

Weighted mean absolute percent deviation at city level ("city weighted MAPD")

where

k = 1, …, p counts over the categories

Eq. 6.

$$cellMAPD = \frac{\sum_{k=1}^{s}\sum_{v=1}^{r}\frac{\left|\boldsymbol{E}_{v,k}' - \boldsymbol{U}_{v,k}'\right|}{\boldsymbol{E}_{v,k}'}}{sr} 100\%$$

Mean absolute percent deviation at cell level ("cell MAPD")

where

k = 1, …, s counts over the categories that have a value at the Census cell level

v = 1, …, r counts over the Census cells

I would additionally reweight the MAPD based on the number of building addresses with the respective category (Eq. 5.). A high percent difference for a group of buildings that makes up just a small percentage of the building stock (f. ex. buildings built before 1900) is much less of a problem than on a

[57] Of course, it also serves to make all deviations positive, so as to avoid cancelling out.

[58] Obviously, the two $\boldsymbol{E}_k''$ terms in this equation cancel each other out, I leave them in the equation in grey, because it makes it easier to understand the logic. First one takes the absolute difference. Then relativises with totals $\boldsymbol{E}_k''$. Then calculates a weighted mean, by first multiplying each value by a weight ($\boldsymbol{E}_k''$) and then divide by the sum of the weights $\sum_{k=1}^{p}\boldsymbol{E}_k''$.

group of buildings which makes up a larger share. I calculate the absolute percent difference for each benchmark at the city level (for both city and cell level benchmarks, Table 13 and Figure 14)

For cell level benchmarks alone, I calculate the MAPD over each cell and benchmark attribute (cell MAPD, Eq. 6.) and Pearson's R^2. Additionally, for both metrics I compute the standard deviation, in order to track the dispersion.

Figure 13 presents a scatterplot of the model vs benchmark counts per benchmark category at the city level. While the R^2 is very high, a closer look is required. Firstly, looking at Table 12, with 4% weighted MAPD at the city level, the model performs more than reasonably. The standard deviation of 9% shows that for some attributes the performance is worse. Looking at Table 13, most of these categories have low building counts. F. ex. the buildings (address points) built between 2009 and 2011, according to the Census are 1276, while the model has 1655, a 30% absolute percent difference. While this seems like a lot, there are 250 000 address points overall. In other words, large deviation, but on a relatively small group of buildings.

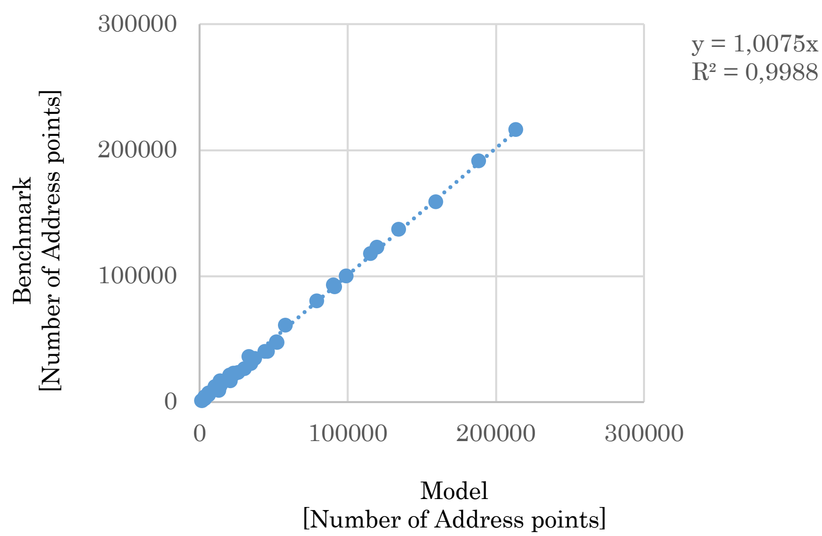


Figure 13. Scatterplot of Model vs Benchmark aggregates

City weighted MAPD	4%
City APD std	9%
Cell MAPD	6%
Cell APD std	26%

Table 12. Model MAPD (Mean Absolute Percent Deviation) on the city and cell levels

As to be expected, the performance drops at the cell level. The cell MAPD[59] is 6%, but with a much higher standard deviation of 26%. This points to limitations in the possible use-cases. Looking at an individual cell from the UBEM in isolation would be a challenge. Then again, if the target buildings are so few, the cost of analysing them simply on-site would also be low. The UBEM is thus suited for use in projects involving multiple 100m raster cells. The mean number of address points per cell is 12, multiple cells would mean approx. 50 address points or more. This is in-line with the scale of the "neighbourhood" (in German: *Quartier).*

[59] Note that the mean MAPD is the mean of a mean. The MAPD is the mean absolute percent difference over the building attributes and the mean MAPD is the mean over the cells and then over the attributes.

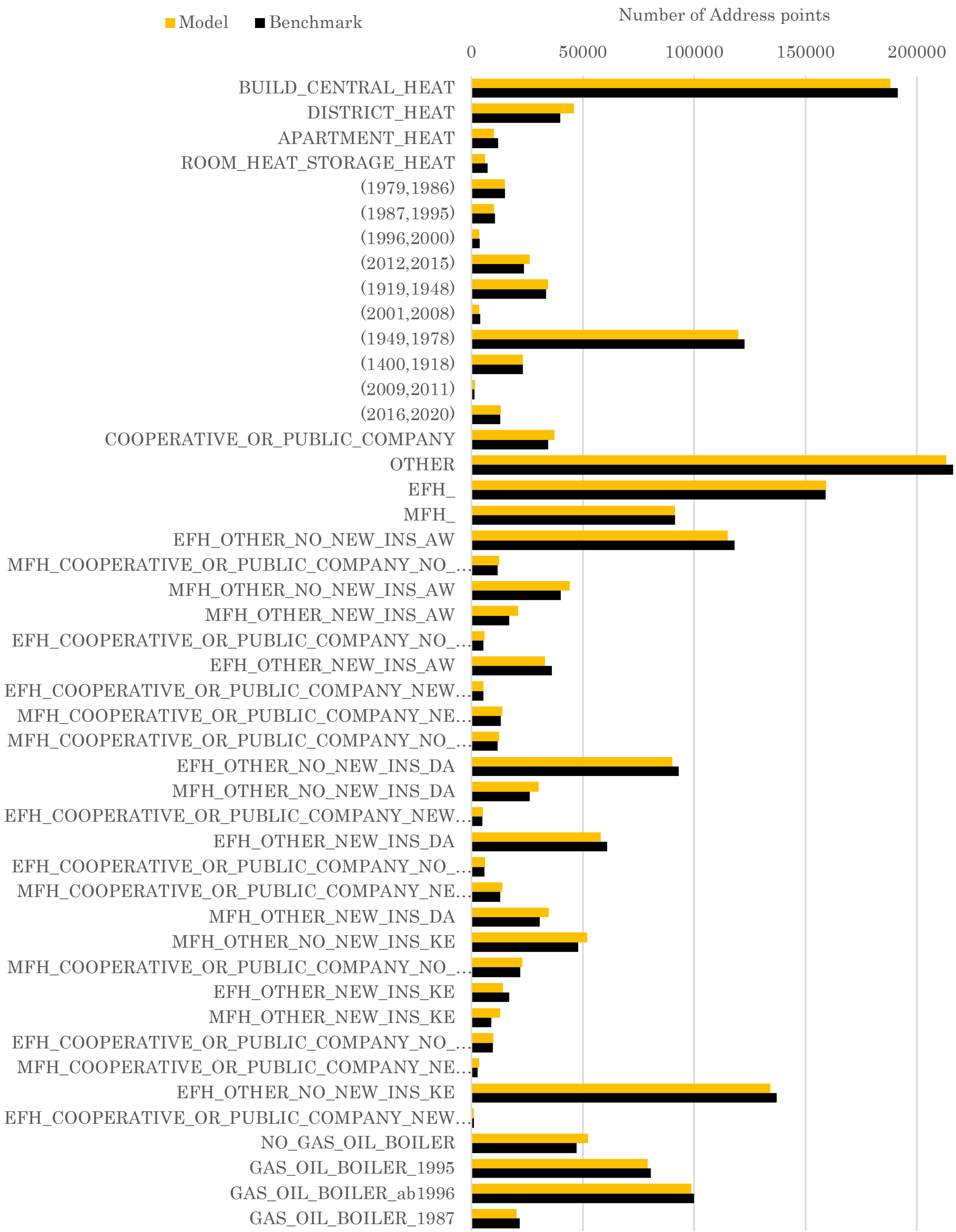


Figure 14. Building categories counts. Refer to Chapter 3 for descriptions of the attributes.

Building Attribute	Model [Number of address points]	Benchmark [Number of address points]	APD
BUILD_CENTRAL_HEAT	188250	191478	2%
DISTRICT_HEAT	45942	39924	15%
APARTMENT_HEAT	10204	12015	15%
ROOM_HEAT_STORAGE_HEAT	6194	7194	14%
(1979,1986)	15157	15131	0%
(1987,1995)	10163	10544	4%
(1996,2000)	3454	3783	9%
(2012,2015)	26063	23678	10%
(1919,1948)	34556	33578	3%
(2001,2008)	3472	3896	11%
(1949,1978)	119746	122723	2%
(1400,1918)	23060	23109	0%
(2009,2011)	1655	1276	30%
(2016,2020)	13264	12893	3%
COOPERATIVE_OR_MUNICIPAL_COMPANY	37345	34405	9%
OTHER	213245	216206	1%
EFH_	159251	159140	0%
MFH_	91339	91471	0%
EFH_OTHER_NO_NEW_INS_AW	115149	118070	3%
MFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NO_NEW_INS_AW	12477	11679	7%
MFH_OTHER_NO_NEW_INS_AW	44095	40058	10%
MFH_OTHER_NEW_INS_AW	20908	16871	24%
EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NO_NEW_INS_AW	5731	5447	5%
EFH_OTHER_NEW_INS_AW	33093	36015	8%
EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_AW	5278	5306	1%
MFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_AW	13859	13061	6%
MFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NO_NEW_INS_DA	12563	11764	7%
EFH_OTHER_NO_NEW_INS_DA	90257	93179	3%
MFH_OTHER_NO_NEW_INS_DA	30296	26259	15%
EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_DA	5045	4918	3%
EFH_OTHER_NEW_INS_DA	57985	60907	5%
EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NO_NEW_INS_DA	5964	5836	2%
MFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_DA	13773	12975	6%
MFH_OTHER_NEW_INS_DA	34707	30670	13%
MFH_OTHER_NO_NEW_INS_KE	51959	47922	8%
MFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NO_NEW_INS_KE	22806	22007	4%

EFH_OTHER_NEW_INS_KE	14013	16935	17%
MFH_OTHER_NEW_INS_KE	13044	9007	45%
EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NO_NEW_INS_KE	9846	9718	1%
MFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_KE	3530	2732	29%
EFH_OTHER_NO_NEW_INS_KE	134229	137151	2%
EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_KE	1163	1035	12%
NO_GAS_OIL_BOILER	52393	47125	11%
GAS_OIL_BOILER_1995	79222	80517	2%
GAS_OIL_BOILER_ab1996	98694	99988	1%
GAS_OIL_BOILER_1987	20281	21576	6%

Table 13. Model Aggregates vs Benchmarks. APD – absolute percent difference

9.2 Validation through random generation

All the techniques presented in Chapter 6 are generally under the hood of combinatorial optimisation, since they are all techniques for finding a certain "optimal" combination. In the context of Spatial Microsimulation and UBEMs, this combination should be the one most similar to the real population (or building stock, in this case). However, all these techniques use benchmarks to arrive at this "optimal" combination. This leads to the notion that maximising the fit (i.e., minimising the difference (MAPD)) between model aggregates and benchmarks is the objective. Hence, I defined my objective function(s) to minimise the squared sum of differences.

However, considering the nature of the problem, the objective is actually to find the combination closest to the real population, while minimising the differences **is the way** (a proxy) to get there. The difference between objective and proxy becomes apparent when one considers that there could be multiple combinations that fit the benchmarks (as noted in Chapter 4). In this situation, whether all of these are equally similar (if at all) to the real population is an assumption. One that is not tested in any way, since it is outside the objective function. Nevertheless, defining the similarity to the true building stock cannot be the objective function for the algorithm, since the true building stock is unknown. Its benchmarks are.

However, I can explore the connection between building stock characteristics (expressed as combinations of energy audits) and benchmarks by simulating the problem. I can randomly create different building stocks, aggregate the building variables, let the algorithm find a combination that matches these aggregates and then compare the found combination to the randomly generated building stock.

For this, first I define a "similarity" score. It is a measure of how similar two combinations of energy audits are. The similarity is the average of the percent of matching benchmark attribute categories between each pair of unique energy audits of the two compared combinations[60].

Then comes the issue of how to generate the random building stocks. Given that I am comparing combinations (not permutations) repetitive sampling from the unique energy audits converges to the same combination, because each audit is picked roughly the same number of times. This means that the algorithm gets tested against the same combination multiple times. This is hardly a good way to validate it. To correct that, I introduce a random bias in the sampling. By randomly excluding between 0% and 50% of the unique audits, the pool from which the building stocks are generated is more different each time and the generated stocks vary much more.

Furthermore, I noted in Section 5.4 that spatial clustering is observed. In order to let the algorithm run against combinations that have spatial clustering, when I assign a random audit from the biased pool I choose among the 20 nearest neighbours for each building, which is the 60th percentile referred to in Section 5.4.

With 200 random combinations of 1000 Census cells, the "similarity" scores and MAPD, I generated Figure 15.:

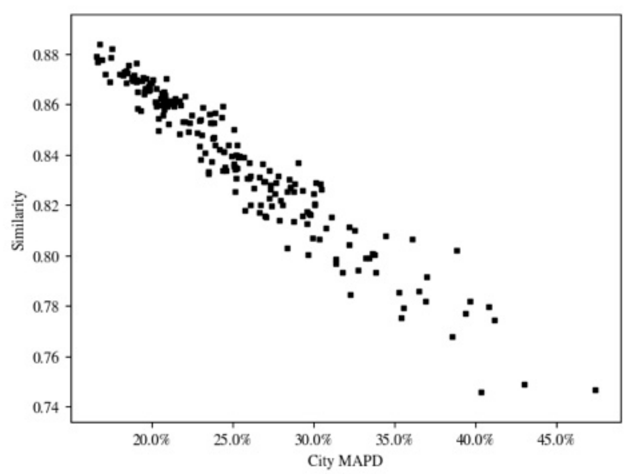


Figure 15. Scatterplot of MAPD vs Similarity for 200 randomly generated building stocks of size 1000 cells.

The figure shows the relationship between the MAPD (city-level) and the similarity exhibited from the randomised simulation. It shows a clear negative correlation between the similarity and the MAPD,

---

[60] For example, a combination consists of 70 assigned unique audits with index 1 and 30 with index 2 Another combination has 50 with index 1 and 50 with index 2. There are 8 categories. The similarity between the two is the average of the pair-wise comparisons: 50 with index 1 are common between the two, so the first 50 comparisons are for a perfect similarity - 1.0. For the other 50 pairs, the number of matching categories divided by the total number of categories gives the scoring. This results in another 50 scores between 0 (no categories are mutual in the pair) and 0.875 (7 out of 8 categories are mutual).

which supports the assumption, that minimising the MAPD leads to building stocks which tend to be similar to the real building stock. This suggests that it is likely that having the algorithm minimise the MAPD would result in a building stock that is similar to the real one.

The same concept in more formal terms:

Let $\boldsymbol{X}'$ be $\boldsymbol{X}''$ be two combinations of audits. A combination can be expressed as an integer array of length m, where the indices j represent a unique energy audit, and the value at index j represents the number of times this audit occurs in the combination.

Then $\boldsymbol{S}$ is the "similarity" score between them:

Eq. 7

$$\boldsymbol{S} = \boldsymbol{f}(\boldsymbol{X}', \boldsymbol{X}'') = \mathbf{1} - \frac{\sum_j^m \left|\boldsymbol{X}_j - \boldsymbol{Y}_j\right|}{\sum_j^m \boldsymbol{X}_j}$$

Example: m=3, Each value is the number of times an audit appears in the combination

$$\boldsymbol{X} = [2 \quad 2 \quad 3]$$

$$\boldsymbol{Y} = [1 \quad 2 \quad 4]$$

$$\boldsymbol{f}(\boldsymbol{X}, \boldsymbol{Y}) = 1 - \frac{2}{7} = 0.71 = 71\%$$

It is clear from the above that $\boldsymbol{f}()$ is simply the total absolute error relative to the size of the arrays[61], and taken out of 100%, so that it constitutes a "similarity" score. It is also clear that the usefulness of $\boldsymbol{f}()$ is under the assumption that such an $\boldsymbol{X}$ exists**,** which represents well the building stock. In other words, there is enough information in the energy audits, so that a combination of them is enough to represent the stock.

Then, let $\boldsymbol{E}$ be the benchmarks[62] and assume that they are 100% correct. There exists a function $\boldsymbol{g}()$ which takes any combination ($\boldsymbol{X}$ or $\boldsymbol{Y}$ or any other) and aggregates it according to the benchmarks. If $\boldsymbol{X}$ is the true building stock, then $\boldsymbol{g}(\boldsymbol{X}) = \boldsymbol{E}$, since aggregating the "true" building stock equals the true benchmarks.

The notion explained previously can now be presented in formal terms. Most methods described in Chapter 6 assume that if a $\boldsymbol{Y}$ is found for which $\boldsymbol{g}(\boldsymbol{Y}) = \boldsymbol{E}$**,** then $\boldsymbol{Y}$ is a good representation of the whole population (is "similar" to the "true" $\boldsymbol{X}$). In other words, if $\boldsymbol{g}(\boldsymbol{Y}) - \boldsymbol{E}$ can be minimised, then $\boldsymbol{Y}$ is a good model. For this reason, one might be inclined to define $\boldsymbol{g}(\boldsymbol{Y}) - \boldsymbol{E}$ as the objective function. However, to

[61] The size of the arrays is per definition equal.

[62] A more concrete definition of $\boldsymbol{E}$ is given in the next section

conclude that if $\boldsymbol{g}(\boldsymbol{Y}) - \boldsymbol{E} = \mathbf{0},$ then $\boldsymbol{f}(\boldsymbol{X},\boldsymbol{Y})$ equals a high number is only an assumption. The problem with $\boldsymbol{f}(\boldsymbol{X},\boldsymbol{Y})$ is obviously that it is unknown, because $\boldsymbol{X}$ is unknown.

9.3 External validation

As opposed to using the data already used for the creation of the model to validate it, in this part I will use external data. It is highly important however to describe the details involved in pre-processing the data for external validation so that false comparisons are avoided (for example comparing values for useful heat with values for final energy or not accounting for weather etc.). Note that in this chapter I will not analyse what the data means for climate policy. Here the goal is to validate the model.

9.3.1 The Hamburg energy statistics

The first way to externally validate the Hamburg UBEM is by using the consumption statistics that the local statistical office prepares (Statistical office for Hamburg and Schleswig-Holstein, 2018a, 2019, 2020). The statistics are gathered from energy providers and give the energy consumption of the city according to sectors and fuel types. Since I am modelling the residential building stock the sector "households" is the one I use for validation. Due to the way the data is gathered, the statistical report lags two or three years behind, so the most current at the time of writing is the 2018 report, published in 2020. Since the local consumption data does not differentiate between heat uses, I use the German-wide split estimated by the Working Group on Energy Balances. It is an association of industry and energy researchers, which is regularly tasked with preparing the official energy balances for the Federal Government (AGEB e.V., 2018). In this way, I estimate the consumption for space heating and domestic hot water (DHW) for Hamburg and correct for weather with degree days (Institut Wohnen und Umwelt, 2018), Table 14)). The assumption I make is that the split of Hamburg does not differ much from the national one. Of course, this is an assumption, therefore the validation should be considered with a grain of salt.

Energy Carrier [GWh]	Hamburg Statistics Households 2018	Country-wide Split		Degree-days 2018/Long Term Average	Result	
		Space Heating	DHW		Space Heating	DHW
Oil	2145	86.1%	13.0%	2850/3140	2035	279
Gas	4165	81.3%	18.4%		3731	766
Electricity	3148	6.6%	12.0%		229	378
District heating	2653	91.9%	8.1%		2686	215
Coal	15	100.0%	0.0%		16	0
Renewables	180	87.3%	12.7%		173	23
Total					8870	1661

Table 14. Hamburg heat consumption 2018, own elaboration based on data from AGEB e.V. (2018), Institut Wohnen und Umwelt (2018) and Statistical Office for Hamburg and Schleswig-Holstein (2020)

In the same way, I calculate and weather-correct the consumption also for the years 2016 and 2017. Then using data from the same statistical office on residential floor area (Statistical Office for Hamburg and Schleswig-Holstein, 2016, 2017, 2018b). I calculate the specific heat consumption for space heating and DHW for the three years (Table 15) and calculate an average for the three years.

		Consumption [GWh]			**Specific Heat Consumption* [kWh/m²]**		**Three year average [kWh/m²]**		
Year	**Residential Floor area [Mill.m²]**	**Space Heating**	**DHW**	**District Heating (Space Heating + DHW)**	**Space Heating**	**DHW**	**Space Heating**	**DHW**	**Total**
2016	71.3	7487	1587	2362	105.4	22.2			
2017	72.0	8060	1572	2576	115.1	21.8	**112.9**	**22.3**	**134.2**
2018	72.8	8870	1661	2570	121.4	22.8			

*The listed external validation sources do not include the floor area supplied with district heating, so the specific heat consumption is only for the total for space heating and DHW

Table 15. Hamburg heat consumption for the years 2016, 2017 and 2018. Own elaboration based on data from: AGEB e.V. (2018), Institut Wohnen und Umwelt (2018), Statistical Office for Hamburg and Schleswig-Holstein (2016, 2017, 2018b, 2018a, 2020)

Table 15 shows that the area, consumption and specific consumption have all increased in the last three-year period. Although this seems like an important and rather worrisome tendency, I would not draw too many conclusions. The year-to-year consumption can fluctuate and even with a weather correction looking at only three years might lead to some spurious observations. Another possible inaccuracy can come from the way the consumption is allocated to the sectors. The data is gathered from energy companies, but the allocation to sectors involves statistical methods and could lead to inaccuracies. For these reasons, for the task of evaluating the UBEM for Hamburg, I consider the three-year average. I will also account for the fact that even the average might be a couple of percent off[63]. Still, with a grain of salt, the tendency of slight increase is observable and the Techem report (on p.17, see next section) confirms that.

9.3.2 The Techem report on energy consumption

Independently of the Hamburg energy balances of the statistical office, the heat accounting firm Techem published a report on heat consumption prepared with their own metering data (Techem Energy Services GmbH, 2019). In this report (2019, p. 193) values are given for the average specific heat consumption of residential buildings in Hamburg. Note that I define the specific heat consumption (SHC) for a building stock as the total consumption over the total floor area. Given that the Techem data comes from a (potentially) non-random sample and that I could not find a mention of reweighting in the report, I assume that the specific heat consumption in the report is actually the average specific heat consumption (aSHC):

[63] In Chapter 2 I quoted the Ecofys report stating that the discrepancy between their UBEM and the consumption data of the energy balances was possibly due to an error in the statistics. In this Chapter, I also argue that the statistics might not be 100% correct. However, I argue the statistics could be off, but a couple of percent, not 20%, as was the case with the Ecofys UBEM.

Eq. 8.

$$SHC = \frac{\sum_{i=1}^{n} Q_i}{\sum_{i=1}^{n} A_i}$$

Eq. 9.

$$aSHC = \frac{\sum_{i=1}^{n} \frac{Q_i}{A_i}}{n}$$

Where:

Q_i is the total consumption of building i

A_i is the residential floor area of building i

n is the number of buildings in the building stock or sample

The difference between SHC and aSHC is subtle[64] and noticeable mostly when the efficiency of smaller buildings differs significantly from that of larger buildings and the sample at hand is highly skewed. Both are possible, so this also has to be considered when comparing the values with the UBEM. The Techem values for the $aSHC$ for Hamburg for 2018 split according to different energy carriers are given in Table 16. The weather correction is again from the (Institut Wohnen und Umwelt, 2018) and equals 2850 degree days divided by 3140 degree days. A complication here is that for buildings with a combined heating and DHW system, the consumption is for the totals. In order to correct for the weather, however, the DHW needs to be subtracted. The report provides values for heating consumption of buildings with decentralised DHW. Simply taking the difference between consumption for heating alone and DHW and consumption for heating in order to estimate the consumption for DHW is, however, not a good way of tackling the problem. This is because when the DHW is centralised and with a circulation system, it acts partially as a secondary heating system. This is mirrored also in the TABULA reference method in that heat losses of distribution for the DHW system are used to offset part of the heating demand. This is the reason why a simple subtraction would result in improbable values for DHW heat demand – 16 and 14 kWh/m^2 final energy. Such values are plausible for useful heat, but not for final energy. The report gives average heat consumption values for DHW of 30 and 31 kWh/m^2 for district heating and natural gas respectively (Techem Energy Services GmbH, 2019, p. 54). I will take these out of the totals, correct for weather and then add them again (Table 16).

[64] SHC is actually equivalent to the area-weighted $aSHC$.

Energy Carrier	aSHC [kWh/m²] Space Heating + DHW for buildings with a combined heating and DHW system	Space Heating for buildings with decentralised DHW	Average DHW consumption	Space Heating for buildings with centralised DHW	Weather Correction	Space Heating for buildings with centralised DHW, weather corrected
Gas	139.0	125.1	31.0	108.0	2850/3140	118.7
District Heating	120.4	104.4	30.0	90.4		99.4

Table 16. Hamburg average specific heat consumption according to the report by Techem (Techem Energy Services GmbH, 2019)

Additionally to the aSHC for Hamburg, the report by Techem provides a frequency distribution of buildings with different energy efficiency at the level of grouped postal codes (ibid, 2019, pp. 69–70). The details of how the buildings are classified is provided in the Appendix. The split for postal codes 20000-29999 (of which Hamburg is part) is given in Table 17. The column names refer to the name and year of energy efficiency regulations in Germany. It is used as a proxy for energy efficiency. “WSVO” and “EnEV” refer to “Wärmeschutzverordnung” and “Energieeinsparverordnung” respectively – the names of past energy efficiency ordinances. The number refers to the year – WSVO77 is the ordinance from 1977. “Altbau” means “older construction”, which in this context means “older than the first energy efficiency ordinance in 1977”. Note that Techem classify into these standards based on the size, energy carrier and yearly SHC, not based on the construction date. Therefore, a building may be built in 1940, but because of energy efficiency improvements fall in the “EnEV02” category.

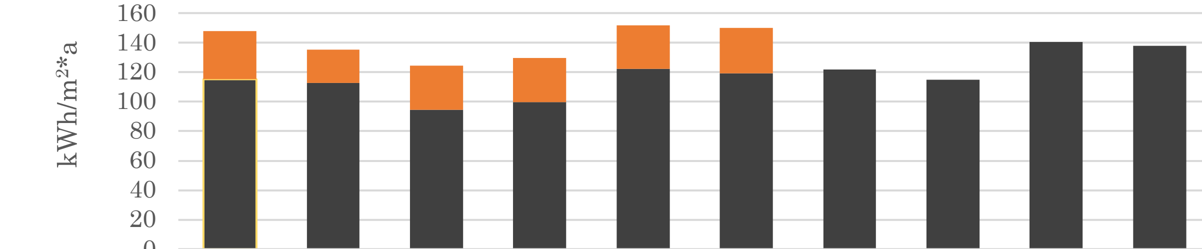
PostalCode Area	Altbau	WSVO77	WSVO95	EnEV02	EnEV09
20000-29999	0.545	0.16	0.136	0.106	0.054

Table 17. Distribution of buildings according to energy efficiency standard in Postal Code Area 20000-29999 (includes Hamburg and the region around). Source: (Techem Energy Services GmbH, 2019, p. 70)

9.3.3 Comparison with the UBEM

A comparison between the energy statistics, the Techem report and the UBEM is given in Figure 16. Firstly, the estimated heat consumption for space heating of the UBEM is very close to the energy statistics. For DHW the difference is more substantial – 32.4 kWh/m² from the UBEM compared with 22.3 kWh/m² calculated from the statistics. Looking at the energy carriers, the numbers are again very close – 94.6 and 99.6 for district heating and 122.4 and 119.0 for Gas. Overall, at this level the model is very close to the actual consumption.

An interesting note here is that the table contains eight pairs of numbers to be compared, all of which could fall under a general term “heat demand”. This means that when communicating and discussing “heat demand” great care must be taken that all parties are aware of which “heat demand” exactly is discussed. Otherwise, it would become an “apples vs oranges” type of comparison.



	UBEM SHC	Energy Statistics SHC	UBEM DH aSHC	Techem DH aSHC	UBEM Gas aSHC	Techem Gas aSHC	UBEM DH aSHC dec. DHW	Techem DH aSHC dec. DHW	UBEM Gas aSHC dec. DHW	Techem Gas aSHC dec. DHW
DHW	32,4	22,3	29,9	30	29,1	31				
Space Heating	115,2	112,9	94,6	99,6	122,4	119,0	121,6	115,0	140,6	137,8

Figure 16. Comparing the UBEM's estimated heat consumption with external sources

Furthermore, Figure 17 gives a comparison of the splits according to the Techem classification. It is clear that while plausible, the model is not perfect. The size of the WSVO77 and the EnEV09 categories are underestimated in the model, while the Altbau, EnEV02 and WSVO95 are overestimated.

The underestimation of the EnEV09 category is compensated by the EnEV02 category. This is expected. The energy audits are biased towards older buildings and the reweighting can only partially compensate the fact that newer, better buildings do not need an energy audit. There simply are not enough examples in the sample with such high energy efficiency.

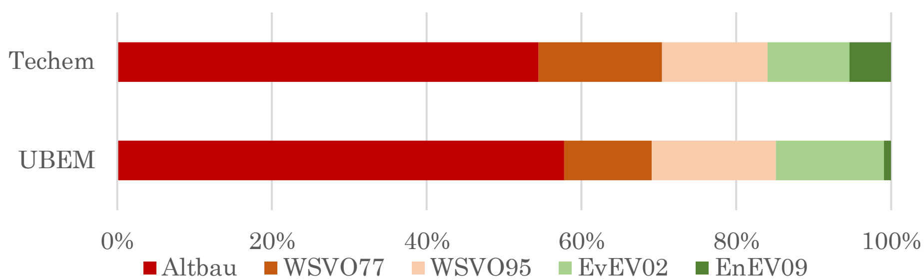


Figure 17. Comparison between the distribution of building energy efficiency in the UBEM and the Techem Report.

10 Use Cases

In the last part of this thesis, I will present some use cases for the UBEM generated in the previous chapters. First, I will give a background to energy policy in Hamburg and then introduce the policy questions and scenarios developed in the GEWISS project. These are to serve as examples of what kinds of questions are important and interesting for the local authorities. I will not delve into simulating the scenarios. Simulating the development of the building stock over a 30-year timeframe would be beyond the scope of this thesis. I will, however, present how these policy questions and the respective parameters relate to my UBEM and why I argue it contains enough information to allow such a simulation.

Finally, I will formulate and analyse a couple of use cases of my own. They are real, and realistic, and constitute applied energy policy (analysis), ranging from the effects of legislation to the socioeconomic conditions relating to building energy efficiency.

10.1 Energy policy in Hamburg

The city-state of Hamburg has taken a number of steps with varying degree of stringency to comply with European, national and global climate protection goals in the past decade. A Climate Protection Concept was approved in 2011 (*Behörde für Stadtentwicklung und Wohnen* [Hamburg Ministry for Urban Development and Housing], 2011). Then, in 2015 a Climate Protection Plan was adopted (*Behörde für Umwelt und Energie* [Hamburg Ministry for Environment and Energy], 2015) and in 2019 this plan was updated (*Behörde für Umwelt, Klima, Energie und Agrarwirtschaft* [Hamburg Ministry for Environment, Climate, Energy and Agriculture], 2019)[65]. Furthermore, an act of the Hamburg Parliament – *Hamburgisches Gesetz zum Schutz des Klimas* - HmbKliSchG [Hamburg Climate Protection Act], 2020 was passed. Additionally, a number of neighbourhood energy plans were prepared and the Hamburg state investment bank (IFBHH) provides programmes and financial incentives for energetic building retrofits and the installation of renewable energy systems.

The goal of all these steps and measures is climate protection. More concretely according to §4 HmbKliSchG, Hamburg has to reduce its carbon footprint by 55% and 95% by 2030 and 2050 respectively, compared with 1990 levels. The legislation explicitly mentions the city energy balances ("*Verursacherbilanz der Freien und Hansestadt Hamburg")* as the data source that is to be used for measuring compliance[66].

10.2 Scenarios of the GEWISS project

As part of the GEWISS Project, I and colleagues from HCU and *Die Hochschule für Angewandte Wissenschaften Hamburg* (HAW) worked together with the BUE on the definition of scenarios (Dochev,

---

[65] All the different names refer to the same entity. The state ministry responsible for energy was reorganised multiple times in the last decade, hence the different names.
[66] This is the same data that I used for the external validation.

Seller, Peters, et al., 2020, p. 346) that the GEWISS simulation tool would simulate. The latter was developed by the HCU and HAW teams as a deterministic rule-based simulation of a given baseline building stock. The time horizon was 2050, or approx. 30 years. The scenarios are defined with a key policy question and corresponding simulation parameter combinations. A key policy question would be "Given the current renovation rates, how would Hamburg's building stock look like (energetically) in 2050"?

An example of a parameter is the renovation rate, defined as a percentage of buildings that are renovated per year. Another parameter would be the share of passive house standard[67] among renovated buildings. The corresponding scenario parameter combination would then be 1% renovation rate and a 5% passive house standard rate. This would mean the simulation will pick 1% of buildings per year for renovation and 5% of those would be renovated to the higher energy efficiency standard. These values are also approximately the current (2020) values, hence this is a "baseline" scenario and reflects current renovation rates.

As in any type of analysis of a future development, one always has to start with an understanding of a "baseline" situation. In other words, in order to understand how the building stock will look in 2050, one needs to know how it looks today. Therefore, any parameter that is not present in the UBEM can, almost per definition, not be analysed for the future. By mapping the scenario parameters to my UBEM attributes, I attempt to show that the UBEM can be used in such a simulation.

The scenarios of the GEWISS project are grouped thematically and vary the parameter values. For the first scenario ("baseline") the associated UBEM attributes are "renovated buildings", "renovations with passive house standard", "type of heating system" and "ownership". Ownership plays a role since cooperatives and housing companies tend to renovate more than private owners. Looking at the UBEM, building ownership is explicitly modelled as such (refer to Table 5 of Section 5.3). By contrast "renovated buildings" and "renovations with passive house standard" cannot be explicitly mapped to attributes. There is however detailed information on insulation, wall and window U-values and heating systems (refer to Table 4) which allows for any such classification.

The renovation rate and depth may imply a change in the heating source and system of a building. In GEWISS this is operationalised with a heating exchange matrix (see Table 18). Similarly to above, while the exact same attributes are not present in the UBEM, it actually has even more detailed data – boiler age f. ex. Therefore, the UBEM can be used here as well, with little to no data crunching.

All the scenarios and respective policy questions are presented in groups and briefly described below (Dochev, Seller, Peters, et al., 2020, p. 346):

[67] "Passive House" refers to a German unofficial yet widely adopted energy efficiency standard

- Scenario Group 1 – "Baseline renovation rates and passive house quotas". The key policy question is the very first example from above: "Given the current renovation rates, how would Hamburg's building stock look like (energetically) in 2050?".
  Scenario 1.1 – "Faster renovation, at current depth". The key policy question is "How would the building stock look like if the renovation rate is increased, but the depth (of the renovation) is retained?".
  Scenario 1.2 – "Same renovation rate, but deeper". The key policy question is "How would the building stock look like if the renovation rate is retained, but the depth is increased?"
  Scenario 1.3. – "Faster and deeper renovation" – "What if Hamburg renovates both faster and deeper?"
  Scenario 1.4. – "Gradual increase in both the rate and depth over time" – The core concern here is whether there is time still to postpone measures.

- Scenario Group 2 – "Heating system exchanges". The scenarios concerning heating system exchanges revolve around different probability matrices, where each cell gives the probability of a switch from the system in the row to the system in the column.
  Table 18 gives an excerpt of the exchange matrix for a baseline scenario. Low temperature boilers (in German, *Niedertemperaturkessel)* are not supposed to be installed anymore, so the scenario gives a 0.00 probability for this outcome. The outcome assumed to be most likely is exchanging a low temperature boiler for a condensing boiler (in German, *Brennwertkessel)*, hence the 0.44 probability in the table. Similarly, to the above scenarios, the UBEM has information on the heating system so this kind of probabilistic simulation can be carried out.

	LOW_TEMPERATURE_ BOILER	**CONDENSING_ BOILER**	**CONDENSING_BOILER_ SOLAR**	…
LOW_TEMPERATURE_ BOILER	0.00	0.44	0.42	…
DISTRICT_HEAT	0.00	0.00	0.00	…
CONDENSING_BOILER	0.00	0.44	0.42	…
…	…	…	…	…

Table 18. Example of a heating system exchange matrix. Source: (Dochev, Seller, Peters, et al., 2020, p. 348)

- Scenario Group 3 – "Impact of the housing companies". The focus here is on giving the housing companies different over- or under-proportional renovation rates, so as to measure their effect on the housing stock. A key question is f. ex. "What would happen if the housing companies do not renovate faster" or "What if they renovated more deeply?".

- Scenario Group 4 – "CO_2-neutral district heating". This scenario analyses the building stock under the assumption that the district heating grid is climate neutral after 2030.

All of these scenarios can be simulated with my UBEM with little data crunching, since it contains the relative attributes or even more detailed ones.

10.3 Heating System Exchanges (§17 (1) HmbKliSchG)

After presenting some example of policy questions that were formulated within the GEWISS project, I will formulate one of my own. A straightforward example of a model use-case is the analysis of the effects of §17 (1) HmbKliSchG, which legislates that after the 30th of June 2021, when a heat generator in a building is exchanged, the owner has to integrate renewable energy in the heat supply. The amount is set at 15% of the building heat demand. The heat demand is defined as the final energy for space heating and DHW (§3 (11) ibid).

The use-case can be viewed from the point of view of the policy maker, as in "how to shape the legislation to achieve an ultimate goal?" or the view of the policy analyst – "what will be the effects of the legislation?". In both cases the questions that need to be asked are similar, since the latter is part of the process of the former. Still, I will be viewing it from an analyst's perspective.

An obvious place to start is to attempt to estimate how big an impact §17 (1) HmbKliSchG could have. Since it comes into force when a boiler is replaced, the question becomes "how many boilers are probably going to be replaced soon?"[68]. Let us assume a typical gas boiler has a lifetime of 20-30 years. If we consider 2021 as the reference year, then 1991-2001 would be the cut-off for construction year for the boilers. Boilers constructed before 1991 would, by 2021, have an age of 30 years or more, so they are likely candidates for replacement.

Attribute	Categories	Level
Gas- or Oil-Boiler Age	GAS_OIL_BOILER_ab1996 NO_GAS_OIL_BOILER GAS_OIL_BOILER_1995 GAS_OIL_BOILER_1987	City (global)

Table 19. UBEM variable related to boiler age. Excerpt from building attributes table.

The UBEM has the "Gas- or Oil-Boiler" variable (see 5.1 Building attributes, Table 19 presents an excerpt relating to only this variable). The cut-off years, as found in the data on boiler ages I used (Landesinnungsverband des Schornsteinfegerhandwerks Hamburg, 2016) are 1987 and 1995. Looking at our boiler lifetime range 20-30, I would take the middle (25 years). Thus, I consider all buildings in the UBEM with the attribute "GAS_OIL_BOILER_1995" and "GAS_OIL_BOILER_1987" as the buildings which likely would need a replacement of their boiler soon and these are the ones which the legislation will probably affect. From the UBEM, I can calculate that this is 42% of address points. For

[68] Theoretically all boilers would be exchanged at some point, but Hamburg (as is Germany and the world for that matter) are in a race against time when it comes to reaching its goals, so it matters when would a legislation start to make a difference.

energy policy however, neither the number of address points, nor the number of buildings are very telling. Buildings vary wildly in their size, address points could have one, two or up to a dozen storeys. For most UBEM-related matters, I argue the better unit to measure is the floor area, as in "how much of the residential floor area is/has attribute X". In the concrete case, the residential floor area of the buildings with boilers 25 years old or older[69] sums up to 31% of the total floor area. This is already an important insight - it shows the scale at which the legislation could affect the building stock. Taking this a step further, from the UBEM, I can calculate how much of the heat demand is in buildings with such boilers – 36%. So, this legislation is likely to target approx. 36% of the heat demand of the city.

Policy relevant question	Answer
How many address points are likely to be affected by §17 (1) HmbKliSchG?	42%
What percent of the residential floor area is that?	31%
What percent of the total heat demand?	36% (4 TWh)

Table 20. Examples for relevant policy question concerning §17 (1) HmbKliSchG and the respective answers according to the UBEM.

Until now the analysis could, in theory, have been carried out without the UBEM, but with the raw benchmark data itself and a bit of calculation. The raw data comes in the form of number of boilers, split into their power capacity in kW and their construction year (before 1987, before 1995 etc). From this data a share of boilers can be calculated directly. Further, based on the kW capacity of boilers one could attempt to estimate the heated area that each boiler supplies and thus estimate my second insight from above (31% of total residential floor area). With area and kW and a bit more data crunching[70], one could approximate also my last insight from above – the percent of the heat demand supplied (the 36%). Of course, this estimation technique comes with its own caveats – everything hangs on the kW and thus the detail of over/under-sized boilers is tacitly ignored.

However, where the UBEM and generally microsimulation techniques shine and definitely outperform other estimation techniques is in the cross-reference of different attributes. In this case, a policy maker might want to investigate which types of owners will be most affected – private owners or housing cooperatives? Given the differences in both financial resources and project management capabilities this is an important question. With the raw data, such analysis would be impossible. On the other hand, the UBEM includes an ownership type attribute, therefore answering the question becomes a trivial table operation – selecting on multiple attributes. The process is similar for all other attributes, present in the UBEM – f. ex. for the presence of wall insulation (Table 21). The implications of this are f. ex. that since over half of the affected heat demand is in buildings with insulation, their boilers could

[69] A sensitivity analysis regarding boiler lifetime age (I assumed 25 years) could be carried out, by computing the target variable for different life spans. However, there are limits to this, due to the boiler age classes found in the data. The original data not provide the construction date of each boiler, so the exact age is not known. It is known that it is "after 1996" or "before 1995" etc.

[70] Assume some typical load curve for the power capacity, integrate under the curve, get the yearly heat demand related to each boiler.

be currently oversized. Since the vast majority (32 percentage points out of 36%) are private owners, it becomes important to inform them of the potential efficiency gains for adjusting the power capacity to the now reduced loads (because of the insulation) when replacing their heating systems.

One could add various other, more advanced analyses to this, f. ex. looking at attributes that are not benchmarked (heating systems, temperatures, pipe insulation etc.).

Policy relevant question	**Answer**
How much of the city's heat demand resides in privately owned buildings which are likely to be affected by §17 (1) HmbKliSchG?	32%
And how much in cooperatives and "SAGA" (the municipal housing company)?	4%
How much of the heat demand resides in buildings with wall insulation which are likely to be affected by §17 (1) HmbKliSchG?	20%
And without?	16%

Table 21. Examples for relevant policy question concerning §17 (1) HmbKliSchG that could not be answered by raw data, but require a UBEM.

Another important benefit of the UBEM is that it is georeferenced. This allows using location to combine it with other data, where the location is used to link the two datasets. For example, in the context of §17 (1) a logical question after knowing how many buildings would be affected is "what options would they have once affected?". Thus asking, f. ex., how many of the affected buildings/area could replace their boilers with district heating?

Combining the UBEM with spatial data on district heating grids (Behörde für Umwelt und Energie [Hamburg Ministry for Environment and Energy], 2018), I can estimate an answer to this. Considering a 200m distance to an existing district heating area as viable for district heating, I produced Table 22. For comparison I will give also a 50m buffer. Figure 18 gives a visual example of what 50 and 200 meters around heating grids look like in an urban setting.

Policy relevant question	**Answer**
How much of the city's heat demand (A) which is likely to be affected by §17 (1) HmbKliSchG (B) resides in buildings within 200 meters of existing heating grids?	1,4 TWh / 12% of A 35% of B
And within 50m?	0,7 TWh/ 6% of A 17% of B

Table 22. Examining policy relevant questions by combining the UBEM with other spatial data

Looking at the table, more than 1/3 of the affected buildings are within the 200m buffer. Looking at Figure 18 and considering the generally expensive and difficult underground construction required by underground heating grid connection, the 200m buffer might be too ambitious. Nevertheless, even the

50m buffer (for which usually only a house connection would be needed) catches circa 17% of the affected buildings. On the whole with 12% and 6% of the total heat demand (depending on the buffer) a potential synergy between §17 and the overall goal of increasing district heating share is possible.

Figure 18. Different buffers (50 and 200m) around a district heating grid. Heating grids source: (Behörde für Umwelt und Energie [Hamburg Ministry for Environment and Energy], 2017), basemap source: (Freie und Hansestadt Hamburg, Landesbetrieb Geoinformation und Vermessung (LGV) [Hamburg State Office for Geoinformation and Surveying], 2022)

10.4 Combining with data on purchasing power

The benefits of the dataset being geo-referenced can be further shown in the following example. Until now, most of the use cases were on the technical side, affected buildings, viable technical options etc. Here, I would add a social aspect. The city of Hamburg has ambitious CO_2 goals and rightfully so, but good policy requires an appreciation of the socioeconomic condition of the affected households. An interesting and, I believe, important question is how the current energy efficiency is distributed among the socioeconomic classes. There are many ways and angles from which one could analyse this. I will attempt to operationalise it by looking at the average specific heat demand (kWh/m^2) in relation to purchasing power.

For this, I would use data on purchasing power obtained as part of the GEWISS project (panadress marketing intelligence/WIGeoGIS GmbH, 2019). Among other things, this dataset contains purchasing power at the address point level for Hamburg (see Figure 19). The data is categorical, split into five categories – average and two classes lower or higher. The dataset metadata does not provide more detail on what the classes mean. I would take them as they are and use them as proxy for the socio-economic status of the households.

Figure 19. Purchasing power at the address level. Data source: (panadress marketing intelligence/WIGeoGIS GmbH, 2019), basemap: (Freie und Hansestadt Hamburg, Landesbetrieb Geoinformation und Vermessung (LGV) [Hamburg State Office for Geoinformation and Surveying], 2022)

In order to integrate this data with the UBEM, some data crunching is necessary. The relationship between the address points and the buildings is in some cases many-to-one. For the simplest case, one-to-one, if only one address point intersects the geometry of a building, I just transferred the point's attributes to the building. For buildings, which did not contain a specific address point I took the attributes of the nearest one to the building as the buildings' attributes. Lastly the many-to-one case, there I took the mode (most common value). In case of a tie, I took the lower value. The result is presented in Figure 20. The data seems plausible with neighbourhoods like Wilhelmsburg and Harburg shown to have lower purchasing power than Blankenese. However, adding Eimsbüttel to the "further below average" class seems strange. Eimsbüttel is a rather expensive, well-off neighbourhood west of the Alster, so this seems less plausible. Therefore, using this data should be taken with a grain of salt.

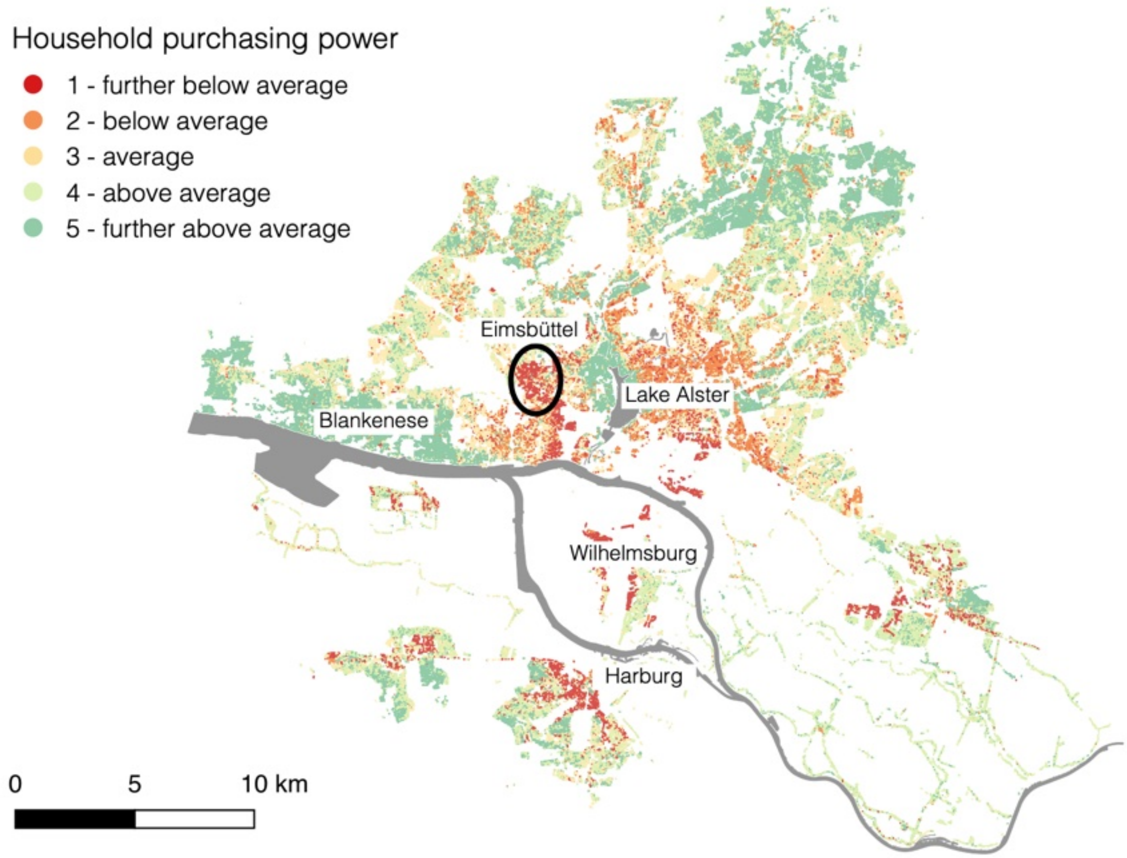


Figure 20. Hamburg household purchasing power at the building level. Data source: (panadress marketing intelligence/WIGeoGIS GmbH, 2019)

With the data on purchasing power now at the building level, I can calculate the specific heat demand split into the five purchasing power classes (Table 23).

Specific Heat Demand of buildings with households with following purchasing power:	kWh/m²
1 – further below average	119
2 – below average	122
3 – average	126
4 – above average	142
5 – further above average	145

Table 23. Specific Heat Demand and purchasing power

The numbers suggest that households with more purchasing power have higher specific heat demand. One could think that is because these households are less inclined to save energy. While this probably plays a role, the main reason lies elsewhere (see Table 24). More higher income households live in single family houses, while less well-off households in large multi-family buildings. The latter are usually owned by cooperatives or the public housing company and actually have large shares of renovations.

Specific Heat Demand of buildings split into building types:	**kWh/m²**
EFH (single-family house)	162
RH (row-house, terraced house)	145
MFH (Multifamily building)	123
GMH (Large multifamily building)	118
HH (Highrise (7+ storeys))	116

Table 24. Specific Heat Demand and building types

Further, due to their geometry, multi-family buildings are more energy efficient than single-family houses, even if all else is equal. The reason is that volume and area scale differently. It does not take twice as much outer wall area to build a building with twice as much volume. This is an old and generally well-known geometric effect which benefits multi-family buildings (expressed with the area-to-volume ratio A/V). Overall, the analysis suggests that well-off households have more potential to increase energy efficiency which is an encouraging conclusion, since they should be the ones who are more able to afford it. Further, given German demographics, such single-family houses are often inhabited by older people whose children have left the house, which sometimes constitutes a barrier for renovation. This could be addressed with policies specifically tailored to this phenomenon.

11 Conclusion and Outlook

In my work, I presented a urban building energy modelling technique that, I believe, can substantially improve UBEMs. Compared to other available models for Hamburg, the presented UBEM has major benefits. It is bottom-up, but also validated with consumption data top-down. It integrates data at various spatial scales. It is georeferenced and is composed of micro-units (buildings) as opposed to neighbourhoods or being an aggregate model at the city level.

Adopting methods from spatial microsimulation – "benchmarking" – allows one to overcome the top-down-bottom-up modelling dilemma. When using top-down, one misses on detail and microdata. When using bottom-up, aggregates often fail to match known aggregate data. Credibility thus diminishes. Benchmarking allows for the creation of a hybrid, whereby microdata and aggregate data can both be used. Of course, one has to take caution not to use all data for modelling and none for validation. By using building attribute data for modelling, and actual energy consumption for validation, I believe I found a good balance.

My work in the GEWISS project gave me the unique opportunity to talk and work with a variety of professionals from the local administration and from energy suppliers. One direct effect was that I understood that being able to integrate known data at every level is important. However good the micro-dataset (in my case the energy audits) and the spatial (aggregate) data are, looking at a single, random building the model might still be wrong. How would one know the model is wrong? In most cities, there would usually be some concrete data from piece-wise renovation initiatives and energy concepts, which, if not integrated into a UBEM, would be at best not optimally utilised, at worst used to contradict the UBEM. This would reduce the UBEM's credibility. Overall, in my experience, there is some understanding among policy makers, analysts and practitioners that a model with 200 000 buildings is bound to be wrong here and there. However, if such sample data can be integrated, the modeller can actually state that what is known is integrated and what isn't, is estimated. It would be the best use of the data, given the many expectations in a policy analysis model, some of which concern the user's perspective and preferences. Therefore, I believe this is an important requirement for the modelling technique.

Nevertheless, there are limitations. Mainly, the availability of aggregate data on added insulation. The overall distribution of buildings with and without insulation that I used comes from an IWU study, which, while respectable, is based on a Germany-wide sample, not a Hamburg-specific one. While the benchmarking technique can correct some of the error, by reweighting to Hamburg-specific benchmarks, having Hamburg-specific data would have been preferable. Further, and perhaps more importantly, the spatial distribution of the renovations would have been more accurate given Census data and not having to depend solely on spatial clustering.

The possible availability of such data is not far-fetched. From personal communication with the local administration, I know there were intentions to add such a question relating to insulation to the 2022

Census. As of May 2022, this does not seem to be the case. That, I consider a huge loss for energy planning. Usually, each question in the Census is carefully picked and of course every planner would like to have it tailored to their needs. However, with energy being such a central topic due to climate change and more recently the war in Ukraine, I believe this omission was a mistake.

While creating a model is all fine, modellers are often confronted with the “so what?” and “what do we use it for?” questions. These questions I attempted to address in the last part of this thesis. First, I presented how one can use the UBEM to analyse the possible effects of legislation. Then, I looked at scenarios for the future development of the building stock and how the UBEM could be used for that. Lastly, I looked at combining the UBEM with socio-economic data to exemplify the benefits of having a geo-referenced micro-level model and what insights can be drawn from that. With the presented use-cases, I tried to cover a broad spectrum and showcase various types of use-cases. Of course, I did not exhaust all possibilities.

In the future, 3D data and BIM (Building Information Modelling) will probably be state-of-the-art. That does not invalidate my work. I explained in Section 6.4 that although not stated as 3D data, I am implicitly using such. As for BIM, an UBEM is a form of a building information model, one that is specialised in energy. An interesting and important development here is the CityGML Energy ADE (“Energy Application Domain Extension”, see Agugiaro et al., (2018)) standard, which attempts to formalise and standardise the way buildings are energetically described. This is definitely the right way to move forward.

All in all, I believe that while my work is by no means the ultimate solution to all challenges facing UBEMs, it does present a notable improvement and contribution to the field.

References

Arbeitsgemeinschaft Energiebilanzen e.V. [Working Group on Energy Balances]. (2023). *Anwendungsbilanzen zur Energiebilanz Deutschland [End-use balances of the German energy sectors]*. https://ag-energiebilanzen.de/wp-content/uploads/2023/01/AGEB_21p2_V3_20221222.pdf

Agugiaro, G., Benner, J., Cipriano, P., & Nouvel, R. (2018). The Energy Application Domain Extension for CityGML: Enhancing interoperability for urban energy simulations. *Open Geospatial Data, Software and Standards*, *3*(1), 2. https://doi.org/10.1186/s40965-018-0042-y

Behörde für Stadtentwicklung und Wohnen [Hamburg Ministry for Urban Development and Housing]. (2011). *Das Hamburger Klimaschutzkonzept [Hamburg climate protection Concept]*. http://www.hamburg.de/contentblob/3155868/1b576a6f0b04dc919c2e8afae1726628/data/download-klimaschutzkonzept-fortschreibung.pdf

Behörde für Umwelt, Klima, Energie und Agrarwirtschaft [Hamburg Ministry for Environment, Climate, Energy and Agriculture]. (2019). *Erste Fortschreibung des Hamburger Klimaplans [Hamburg Climate Plan, first update]*. https://www.hamburg.de/contentblob/13647730/82f1c3fe9959d1d7f70106ae89a80781/data/d-hamburger-klimaplan-2019.pdf

Behörde für Umwelt und Energie [Hamburg Ministry for Environment and Energy]. (2015). *Hamburger Klimaplan [Hamburg Climate Plan]*. https://www.hamburg.de/contentblob/4658414/b246fbfbbf1149184431706972709508/data/d-21-2521-hamburger-klimaplan.pdf

Behörde für Umwelt und Energie [Hamburg Ministry for Environment and Energy]. (2017). *Wärmekataster für Hamburg [Hamburg Heat Demand Cadastre]*. http://www.hamburg.de/energiewende/waermekataster/

Behörde für Umwelt und Energie [Hamburg Ministry for Environment and Energy]. (2018). *Wärmekataster Gebiet mit Wärmenetz Hamburg [Heat Demand Cadastre Areas with district heating in Hamburg]*. https://metaver.de/trefferanzeige?cmd=doShowDocument&docuuid=77450A5E-789A-4E98-B1B1-39C07264CC34

Bundesministerium für Wirtschaft und Energie [Federal Ministry for Economic Affairs and Energy]. (2017). *Energieeffizienz in Zahlen [Energy-efficiency statistics]*. http://www.bmwi.de/Redaktion/DE/Publikationen/Energie/energieeffizienz-in-zahlen.pdf?__blob=publicationFile&v=10

Chingcuanco, F., & Miller, E. J. (2012). A microsimulation model of urban energy use: Modelling residential space heating demand in ILUTE. *Computers, Environment and Urban Systems*, *36*(2), 186–194. https://doi.org/10.1016/j.compenvurbsys.2011.11.005

Cischinsky, H., & Diefenbach, N. (2016). *Datenerhebung Wohngebäudebestand 2016* [*Building data survey for the residential sector 2016*]. https://www.iwu.de/fileadmin/user_upload/dateien/gebaeudebestand/prj/Endbericht_Datenerhebung_Wohngeb%C3%A4udebestand_2016.pdf

Danish Ministry of the Environment. (2002). *Copenhagen Charter 2002: A Statement on the Occasion of the Danish Presidency: European Cities in a Global Era – Urban Identities and Regional Development*. http://naturstyrelsen.dk/media/nst/attachments/92328/era_charter_uk.pdf

Deutsche Institut für Normung [German Institute for Standardisation]. (2003). *DIN V 4108-6:2003-06, Wärmeschutz und Energie-Einsparung in Gebäuden: Teil_6: Berechnung des Jahresheizwärme- und des Jahresheizenergiebedarfs* [*German Industrial Standard 4108-6:2003, Heat preservation and energy savings in buildings*]. Beuth Verlag GmbH. https://doi.org/10.31030/9493310

Deutsche Institut für Normung [German Institute for Standardisation]. (2018). *DIN V 18599-1:2018-09, Energetische Bewertung von Gebäuden* [*German Industrial Standard 18599-1:2018, Energetic assessment of buildings*]. Beuth Verlag GmbH. https://doi.org/10.31030/2874317

Dochev, I., Gorzalka, P., Weiler, V., Schmiedt, J. E., Linkiewicz, M., Eicker, U., Hoffschmidt, B., Peters, I., & Schröter, B. (2020). Calculating urban heat demands: An analysis of two modelling approaches and remote sensing for input data and validation. *Energy and Buildings*, 110378. https://doi.org/10.1016/j.enbuild.2020.110378

Dochev, I., Seller, H., & Peters, I. (2020). Assigning Energetic Archetypes to a Digital Cadastre and Estimating Building Heat Demand. An Example from Hamburg, Germany. *Environmental and Climate Technologies*, *24*(1), 233–253. https://doi.org/10.2478/rtuect-2020-0014

Dochev, I., Seller, H., Peters, I., Tschirschwitz, A.-L., Schiewe, J., Köhler, E., Schäfers, H., & Renz, W. (2020). *Abschlussbericht zum Verbundprojekt: GEWISS - GEografisches WärmeInformations- und SimulationsSystem* [*Final report of consortium project GEWISS - Geographic Heat Information and Simulation System*]. HafenCity Universität (HCU) Hamburg, published in TIB-Technische Informationsbibliothek Universitätsbibliothek Hannover. https://doi.org/10.2314/KXP:1780128819

Ecofys Germany GmbH. (2013). *Flächendeckende Erhebung und Kartierung des energetischen Zustandes des Hamburger Gebäudebestandes* [*Comprehensive survey and mapping of the Hamburg building stock*]. http://www.hamburg.de/contentblob/4313012/6510b8c57c2e9461b1952dda726128e2/data/erhebung-energetischer-zustand-b.pdf

Ecofys Germany GmbH. (2018). *Optionen und Instrumente der Freien und Hansestadt Hamburg zur Reduzierung der CO_2-Emissionen im Gebäudesektor [Tools and instruments for the reduction of CO_2 emissions in the Hamburg building sector]*. https://www.hamburg.de/contentblob/13047276/cfb90c03093a575b2195b9b0d21d5f98/data/d-gutachten-optionen-und-instrumente-ecofys-mit-vermerk.pdf

Freie und Hansestadt Hamburg, Landesbetrieb Geoinformation und Vermessung (LGV) [Hamburg State Office for Geoinformation and Surveying]. (2017). *3D-Stadtmodell LoD2-DE Hamburg [3D City model in Level of Detail (LoD) 2]*. http://datenhamburg.de/geographie_geologie_geobasisdaten/3d_stadtmodell_lod2/LoD2-DE_HH_2017-12-31.zip

Freie und Hansestadt Hamburg, Landesbetrieb Geoinformation und Vermessung (LGV) [Hamburg State Office for Geoinformation and Surveying]. (2020). *Amtliches Liegenschaftskatasterinformationssystem (ALKIS) [Official digital cadastre information system]*. https://daten-hamburg.de/geographie_geologie_geobasisdaten/ALKIS_Liegenschaftskarte/ALKIS_Liegenschaftskarte_ausgewaehlteDaten_HH_2020-01-04.zip

Freie und Hansestadt Hamburg, Landesbetrieb Geoinformation und Vermessung (LGV) [Hamburg State Office for Geoinformation and Surveying]. (2022). *Digitale Orthophotos (belaubt) Hamburg [Digital orthophotos Hamburg]*. https://metaver.de/trefferanzeige?cmd=doShowDocument&docuuid=B3008A4D-BA8E-4FDB-8FDD-559D02975BD7

García Kerdan, I., Raslan, R., Ruyssevelt, P., & Morillón Gálvez, D. (2017). The role of an exergy-based building stock model for exploration of future decarbonisation scenarios and policy making. *Energy Policy*, *105*, 467–483. https://doi.org/10.1016/j.enpol.2017.03.020

GEWISS Projekt Hamburg. (2019). *Geografisches Wärmeinformations- und Simulationssystem Hamburg [Geographic Heat Information and Simulation System]*. http://gewiss.haw-hamburg.de/

Hamburgisches Gesetz zum Schutz des Klimas—HmbKliSchG [Hamburg Climate Protection Act], (2020). https://www.landesrecht-hamburg.de/bsha/document/jlr-KlimaSchGHA2020rahmen

Institut Wohnen und Umwelt. (2018). *Gradtagszahlen Deutschland [Degree days in Germany]:* http://t3.iwu.de/fileadmin/user_upload/dateien/energie/werkzeuge/Gradtagszahlen_Deutschland.xls

Keirstead, J., Jennings, M., & Sivakumar, A. (2012). A review of urban energy system models: Approaches, challenges and opportunities. *Renewable and Sustainable Energy Reviews*, *16*(6), 3847–3866. https://doi.org/10.1016/j.rser.2012.02.047

Landesinnungsverband des Schornsteinfegerhandwerks Hamburg [Hamburg chimney sweeper association]. (2016). *Erhebung des Schornsteinfegerhandwerks für 2016 [Survey of the Hamburg association of chimney sweepers]*.

Li, W., Zhou, Y., Cetin, K., Eom, J., Wang, Y., Chen, G., & Zhang, X. (2017). Modeling urban building energy use: A review of modeling approaches and procedures. *Energy*, *141*, 2445–2457. https://doi.org/10.1016/j.energy.2017.11.071

Lim, H., & Zhai, Z. J. (2017). Review on stochastic modeling methods for building stock energy prediction. *Building Simulation*, *10*(5), 607–624. https://doi.org/10.1007/s12273-017-0383-y

Loga, T., & Diefenbach, N. (2013). *TABULA Calculation Method – Energy Use for Heating and Domestic Hot Water*. https://episcope.eu/fileadmin/tabula/public/docs/report/TABULA_CommonCalculationMethod.pdf

Loga, T., Stein, B., Diefenbach, N., & Born, R. (2015). *Deutsche Wohngebäudetypologie Beispielhafte Maßnahmen zur Verbesserung der Energieeffizienz von typischen Wohngebäuden [German residential building typology. Measures to improve the energy efficiency of typical residential buildings]*. Darmstadt (Germany). https://www.iwu.de/fileadmin/publikationen/gebaeudebestand/episcope/2015_IWU_LogaEtAl_Deutsche-Wohngeb%C3%A4udetypologie.pdf

Lovelace, R., & Ballas, D. (2013). 'Truncate, replicate, sample': A method for creating integer weights for spatial microsimulation. *Computers, Environment and Urban Systems*, *41*, 1–11. https://doi.org/10.1016/j.compenvurbsys.2013.03.004

Ma, J., Heppenstall, A., Harland, K., & Mitchell, G. (2014). Synthesising carbon emission for mega-cities: A static spatial microsimulation of transport CO2 from urban travel in Beijing. *Computers, Environment and Urban Systems*, *45*, 78–88. https://doi.org/10.1016/j.compenvurbsys.2014.02.006

Ma, L., & Srinivasan, S. (2015). Synthetic Population Generation with Multilevel Controls: A Fitness-Based Synthesis Approach and Validations. *Computer-Aided Civil and Infrastructure Engineering*, *30*(2), 135–150. https://doi.org/10.1111/mice.12085

Martinsen, D., & Krey, V. (2008). Compromises in energy policy—Using fuzzy optimization in an energy systems model. *Energy Policy*, *36*(8), 2983–2994. https://doi.org/10.1016/j.enpol.2008.04.005

Martinsen, D., Krey, V., & Markewitz, P. (2007). Implications of high energy prices for energy system and emissions—The response from an energy model for Germany. *Energy Policy*, *35*(9), 4504–4515. https://doi.org/10.1016/j.enpol.2007.03.003

Martinsen, D., Linssen, J., Markewitz, P., & Vögele, S. (2007). CCS: A future CO2 mitigation option for Germany?—A bottom-up approach. *Energy Policy*, *35*(4), 2110–2120. https://doi.org/10.1016/j.enpol.2006.06.017

Mitchell, A. (2009). *Spatial measurements & statistics: Vol. / Andy Mitchell; Vol. 2*. ESRI Press.

Muñoz, M. E., Dochev, I., Seller, H., & Peters, I. (2016). Constructing a Synthetic City for Estimating Spatially Disaggregated Heat Demand. *International Journal of Microsimulation*, *9(3)*, 66–88. https://doi.org/10.34196/IJM.00143

Muñoz, M. E., & Peters, I. (2014). Constructing an urban microsimulation model to assess the influence of demographics on heat consumption. *International Journal of Microsimulation*, *7(1)*. https://doi.org/10.34196/IJM.00096

Nägeli, C., Camarasa, C., Jakob, M., Catenazzi, G., & Ostermeyer, Y. (2018). Synthetic building stocks as a way to assess the energy demand and greenhouse gas emissions of national building stocks. *Energy and Buildings*, *173*, 443–460. https://doi.org/10.1016/j.enbuild.2018.05.055

Nouvel, R., Brassel, K.-H., Bruse, M., Duminil, E., Coors, V., Eicker, U., & Robinson, D. (2015). SimStadt, a new workflow-driven urban energy simulation platform for CityGML city models. *CISBAT 2015*, 889–894. https://doi.org/10.5075/epfl-cisbat2015-889-894

Orcutt, G. H. (1957). A new type of socio-economic system. *Review of Economics and Statistics*, *39(2)*, 116–123.

panadress marketing intelligence/WIGeoGIS GmbH. (2019). *Microbase-Hausdaten Deutschland [Microdata on buildings in Germany]*.

Passive House Institute. (n.d.). https://passivehouse.com

Pearl, J. (1985). *Heuristics: Intelligent search strategies for computer problem solving* (Reprinted with corr). Addison-Wesley.

Peters, I., Brassel, K.-H., & Spörri, C. (2002). A Microsimulation Model for Assessing Urine Flows in Urban Wastewater Management. *Integrated Assessment and Decision Support. Proceedings of the 1st Biennial Meeting of the International Environmental Modelling and Software Society. Ed. by Andrea E. Rizzoli and Anthony J. Jakeman.* https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.125.8448&rep=rep1&type=pdf

Reinhart, C. F., & Cerezo, C. D. (2016). Urban building energy modeling – A review of a nascent field. *Building and Environment*, *97*, 196–202. https://doi.org/10.1016/j.buildenv.2015.12.001

Sandberg, N. H., Sartori, I., Vestrum, M. I., & Brattebø, H. (2017). Using a segmented dynamic dwelling stock model for scenario analysis of future energy demand: The dwelling stock of Norway 2016–2050. *Energy and Buildings*, *146*, 220–232. https://doi.org/10.1016/j.enbuild.2017.04.016

Sokol, J., Cerezo, C. D., & Reinhart, C. F. (2017). Validation of a Bayesian-based method for defining residential archetypes in urban building energy models. *Energy and Buildings*, *134*, 11–24. https://doi.org/10.1016/j.enbuild.2016.10.050

Sola, A., Corchero, C., Salom, J., & Sanmarti, M. (2020). Multi-domain urban-scale energy modelling tools: A review. *Sustainable Cities and Society*, *54*, 101872. https://doi.org/10.1016/j.scs.2019.101872

Statistical Office for Hamburg and Schleswig-Holstein. (2016). *Stadtteil Profile 2016 [Neighbourhood statistics 2016]*. https://www.statistik-nord.de/fileadmin/Dokumente/Datenbanken_und_Karten/Stadtteilprofile/Stadtteilprofile-Berichtsjahre-2013-2016.xlsx

Statistical Office for Hamburg and Schleswig-Holstein. (2017). *Stadtteil Profile 2017 [Neighbourhood statistics 2017]*. https://www.statistik-nord.de/fileadmin/Dokumente/Datenbanken_und_Karten/Stadtteilprofile/StadtteilprofileBerichtsjahr2017.xlsx

Statistical Office for Hamburg and Schleswig-Holstein. (2018a). *Energiebilanz und CO2-Bilanzen für Hamburg 2016 [Energy balance and CO_2 balance for Hamburg 2016]*. https://www.statistik-nord.de/fileadmin/Dokumente/Sonderver%C3%B6ffentlichungen/Energie-_und_CO2-Bilanz_Hamburg/EB_CO2_HH_2016.pdf

Statistical Office for Hamburg and Schleswig-Holstein. (2018b). *Stadtteil Profile 2018 [Neighbourhood statistics 2018]*. https://www.statistik-nord.de/fileadmin/Dokumente/NORD.regional/StadtteilprofileBerichtsjahr2018_1.xlsx

Statistical Office for Hamburg and Schleswig-Holstein. (2019). *Energiebilanz und CO2-Bilanzen für Hamburg 2017 [Energy balance and CO_2 balance for Hamburg 2017]*. https://www.statistik-nord.de/fileadmin/Dokumente/Sonderver%C3%B6ffentlichungen/Energie-_und_CO2-Bilanz_Hamburg/EB_CO2_HH_2017.pdf

Statistical Office for Hamburg and Schleswig-Holstein. (2020). *Energiebilanz und CO2-Bilanzen für Hamburg 2018 [Energy balance and CO_2 balance for Hamburg 2018]*. https://www.statistik-nord.de/fileadmin/Dokumente/Sonderver%C3%B6ffentlichungen/Energie-_und_CO2-Bilanz_Hamburg/EB_CO2_HH_2018.pdf

Statistische Ämter des Bundes und der Länder [German Federal Statistical Office]. (2018). *Gebäude im 100 Meter-Gitter [Census 2011 100-Meter raster cell data on buildings]*. https://www.zensus2011.de/SharedDocs/Downloads/DE/Pressemitteilung/DemografischeGrunddaten/csv_Gebaeude_100m_Gitter.zip?__blob=publicationFile&v=2

Tanton, R. (2014). A Review of Spatial Microsimulation Methods. *International Journal of Microsimulation*, *7(1)*. http://www.microsimulation.org/IJM/V7_1/2-IJM_7_1_Tanton_.pdf

Techem Energy Services GmbH. (2019). *Energiekennwerte 2019 [Values on specific energy use 2019]*. https://www.techem.com/content/dam/techem/downloads/newsroom/studien/Techem-Energiekennwerte-Studie-2019.pdf

United States Energy Information Administration (EIA). (2016). *International Energy Outlook 2016*. https://www.eia.gov/outlooks/ieo/pdf/0484(2016).pdf

Vásquez, F., Løvik, A. N., Sandberg, N. H., & Müller, D. B. (2016). Dynamic type-cohort-time approach for the analysis of energy reductions strategies in the building stock. *Energy and Buildings*, *111*, 37–55. https://doi.org/10.1016/j.enbuild.2015.11.018

Walberg, D., Griechwitz, T., & Schulze, T. (2012). *Gebäudetypologie Schleswig-Holstein [Building typology for the German state of Schlesswig-Hosltein]* (Arbeitsgemeinschaft für zeitgemäßes Bauen e.V., Ed.). https://arge-ev.de/arge-ev/publikationen/schriftenreihe/

Wilson, R., & Din, A. (2018). Calculating Varying Scales of Clustering Among Locations. *Cityscape: A Journal of Policy Development and Research*, *20–1*, 215–231. https://www.jstor.org/stable/26381229

Yan, D., O'Brien, W., Hong, T., Feng, X., Burak Gunay, H., Tahmasebi, F., & Mahdavi, A. (2015). Occupant behavior modeling for building performance simulation: Current state and future challenges. *Energy and Buildings*, *107*, 264–278. https://doi.org/10.1016/j.enbuild.2015.08.032

Appendix I

Description of all building attributes.

Category	Description
DISTRICT_HEAT	Building is heated with district heating
BUILD_CENTRAL_HEAT	Building has a central heating system. “Central” in terms of the building, not district heating.
ROOM_HEAT_STORAGE_HEAT	The rooms in the building are heated independently of one another. F. ex. with a night storage or air conditioner
APARTMENT_HEAT	Each apartment in the building has their own heating system.
(1400,1918)	Building built before 1918.
(1919,1948)	…
(1949,1978)	
(1979,1986)	
(1987,1995)	
(1996,2000)	
(2001,2008)	
(2009,2011)	
(2012,2015)	
(2016,2020)	
COOPERATIVE_OR_MUNICIPAL_COMPANY	Building owned by a housing cooperative or the municipal housing company
OTHER	All other types of ownership
MFH	Multi-family building
EFH	Single-family house
EFH_OTHER_NEW_INS_AW	Single-family house with “other” type of ownership and new (added insulation) of the outer walls
EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NO_NEW_INS_AW	Single-family house owned by a housing cooperative or municipal company and without any new (added insulation) of the outer walls (“AW” – “*Aussenwände*”).
MFH_OTHER_NO_NEW_INS_AW	…
MFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NO_NEW_INS_AW	
MFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_AW	
MFH_OTHER_NEW_INS_AW	
EFH_OTHER_NO_NEW_INS_AW	
EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_AW	
MFH_OTHER_NEW_INS_DA	Multi-family building with “other” type of ownership and new insulation of the roof (“DA” – “*Dach*”).
MFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NO_NEW_INS_DA	...
EFH_OTHER_NEW_INS_DA	
MFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_DA	
EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_DA	
EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NO_NEW_INS_DA	
MFH_OTHER_NO_NEW_INS_DA	
EFH_OTHER_NO_NEW_INS_DA	
EFH_OTHER_NEW_INS_KE	Single-family house with “other” type of ownership and new (added) insulation of the cellar (“KE” – “*Keller*”).
EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NO_NEW_INS_KE	
MFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_KE	
MFH_OTHER_NEW_INS_KE	
MFH_OTHER_NO_NEW_INS_KE	

EFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NEW_INS_KE
EFH_OTHER_NO_NEW_INS_KE
MFH_COOPERATIVE_OR_MUNICIPAL_COMPANY_NO_NEW_INS_KE
GAS_OIL_BOILER_ab1996
NO_GAS_OIL_BOILER
GAS_OIL_BOILER_1995
GAS_OIL_BOILER_1987

Appendix II

Building energy efficiency types according to the Techem study (Techem Energy Services GmbH, 2019)

Tabelle 11: Gebäudeklassen nach Wohnfläche, Beheizung, Heizenergieverbrauch und energetischem Standard

Anzahl Nutzeinheiten	Energieform	Heizenergieverbrauch	Energetischer Standard
2	Fernwärme	>126	Altbau (vor 1977, unmodernisiert)
2	Fernwärme	102 – 126	WSVO 77 (1977 – 1994)
2	Fernwärme	78 – 102	WSVO 95 (1995 – 2001)
2	Fernwärme	56 – 78	EnEV 02 (2002 – 2008)
2	Fernwärme	<56	EnEV 09 (≥ 2009)
2	Öl/Gas	>147	Altbau (vor 1977, unmodernisiert)
2	Öl/Gas	127 – 147	WSVO 77 (1977 – 1994)
2	Öl/Gas	103 – 127	WSVO 95 (1995 – 2001)
2	Öl/Gas	76 – 103	EnEV 02 (2002 – 2008)
2	Öl/Gas	<76	EnEV 09 (≥ 2009)
3 – 4	Fernwärme	>127	Altbau (vor 1977, unmodernisiert)
3 – 4	Fernwärme	105 – 127	WSVO 77 (1977 – 1994)
3 – 4	Fernwärme	86 – 105	WSVO 95 (1995 – 2001)
3 – 4	Fernwärme	65 – 86	EnEV 02 (2002 – 2008)
3 – 4	Fernwärme	<65	EnEV 09 (≥ 2009)
3 – 4	Öl/Gas	>141	Altbau (vor 1977, unmodernisiert)
3 – 4	Öl/Gas	121 – 141	WSVO 77 (1977 – 1994)
3 – 4	Öl/Gas	97 – 121	WSVO 95 (1995 – 2001)
3 – 4	Öl/Gas	66 – 97	EnEV 02 (2002 – 2008)
3 – 4	Öl/Gas	<66	EnEV 09 (≥ 2009)
5 – 7	Fernwärme	>127	Altbau (vor 1977, unmodernisiert)
5 – 7	Fernwärme	93 – 127	WSVO 77 (1977 – 1994)
5 – 7	Fernwärme	82 – 93	WSVO 95 (1995 – 2001)
5 – 7	Fernwärme	62 – 82	EnEV 02 (2002 – 2008)
5 – 7	Fernwärme	<62	EnEV 09 (≥ 2009)
5 – 7	Öl/Gas	>138	Altbau (vor 1977, unmodernisiert)
5 – 7	Öl/Gas	115 – 138	WSVO 77 (1977 – 1994)
5 – 7	Öl/Gas	92 – 115	WSVO 95 (1995 – 2001)
5 – 7	Öl/Gas	64 – 92	EnEV 02 (2002 – 2008)
5 – 7	Öl/Gas	<64	EnEV 09 (≥ 2009)
8 – 16	Fernwärme	>116	Altbau (vor 1977, unmodernisiert)
8 – 16	Fernwärme	84 – 116	WSVO 77 (1977 – 1994)
8 – 16	Fernwärme	76 – 84	WSVO 95 (1995 – 2001)
8 – 16	Fernwärme	61 – 76	EnEV 02 (2002 – 2008)
8 – 16	Fernwärme	<61	EnEV 09 (≥ 2009)
8 – 16	Öl/Gas	>132	Altbau (vor 1977, unmodernisiert)
8 – 16	Öl/Gas	113 – 132	WSVO 77 (1977 – 1994)
8 – 16	Öl/Gas	90 – 113	WSVO 95 (1995 – 2001)
8 – 16	Öl/Gas	66 – 90	EnEV 02 (2002 – 2008)
8 – 16	Öl/Gas	<66	EnEV 09 (≥ 2009)
> 16	Fernwärme	>94	Altbau (vor 1977, unmodernisiert)
> 16	Fernwärme	80 – 94	WSVO 77 (1977 – 1994)
> 16	Fernwärme	70 – 80	WSVO 95 (1995 – 2001)
> 16	Fernwärme	57 – 70	EnEV 02 (2002 – 2008)
> 16	Fernwärme	<57	EnEV 09 (≥ 2009)
> 16	Öl/Gas	>130	Altbau (vor 1977, unmodernisiert)
> 16	Öl/Gas	110 – 130	WSVO 77 (1977 – 1994)
> 16	Öl/Gas	91 – 110	WSVO 95 (1995 – 2001)
> 16	Öl/Gas	66 – 91	EnEV 02 (2002 – 2008)
> 16	Öl/Gas	<66	EnEV 09 (≥ 2009)

Energetischer Standard

- Altbau (vor 1977, unmodernisiert)
- WSVO 77 (1977–1994)
- WSVO 95 (1995–2001)
- EnEV 02 (2002–2008)
- EnEV 09 (≥ 2009)

Die Unterteilung in die verschiedenen Klassen hilft bei der Bewertung unterschiedlicher Gebäude hinsichtlich ihres energetischen Standards.