Building and Environment 149 (2019) 630–639

Contents lists available at ScienceDirect

Building and Environment

journal homepage: www.elsevier.com/locate/buildenv

Automated urban energy system modeling and thermal building simulation based on OpenStreetMap data sets

Jan Schiefelbein*, Jana Rudnick, Anna Scholl, Peter Remmen, Marcus Fuchs, Dirk Müller

RWTH Aachen University, E.ON Energy Research Center, Institute for Energy Efficient Buildings and Indoor Climate, Mathieustr. 10, Aachen, Germany

ARTICLE INFO

Keywords:
City district
Urban energy models
Geographic information systems
Data enrichment
Archetypes
Workflow automation

ABSTRACT

City districts have a large potential to reduce greenhouse gas emissions by usage of energy efficiency measures. Urban energy models (UEM) can be useful to analyze the impact of different energy efficiency actions on city districts. While simulation of demand data with high spatial and temporal resolution is often necessary to evaluate retrofit measures, the city's complex structure and lack of data often prevents a reliable application of such methods. This paper presents an urban energy modeling approach based on open-source geographical information system (GIS) datasets to reduce input data uncertainty and simplify city district modeling. We present a method to automatically extract basic city district data from OpenStreetMap (OSM) and enrich these datasets based on national building stock statistics. Building models with representative geometries and physical properties are automatically generated based on building archetype information. These models enable thermal simulation on urban scale. The approach is demonstrated for a use case in Germany, where a reference city district model has been generated with OSM data extraction and enrichment. The reference city district model has been used to perform a space heating net energy demand uncertainty analysis. The demand values simulated with the reference model show a sufficient fit with measured consumption values. The approach provides a fast and structured methodology to model city districts and simulate space heating energy demand on urban scale.

1. Introduction

Urban areas are large contributors to global greenhouse gas emissions. More than 70% of global CO_2 emissions are emitted by cities [1]. In Germany, more than 40% of final energy is consumed by building structures, with a large share of heating demand [2]. To reach climatic goals of the Paris agreement [3], the overall urban energy efficiency has to be increased by intensified integration of renewable energy, higher process and conversion efficiencies, and reduction in net energy usage.

To analyze and evaluate different energy efficiency measures for cities, urban energy models (UEM) are frequently used. However, city energy modeling is challenging because of two main issues:

- Challenge 1: Lack of data/low data quality
- Challenge 2: City complexity/modeling effort

Elaborated modeling and simulation approaches can hardly perform well without sufficient input data. Many projects on urban scale lack specific knowledge of building data. Gathering and evaluating necessary data sets often requires large manual effort. Even if detailed building geometry data is available, for instance via CityGML files, information about building type, year of construction, or physical properties might still be missing. Moreover, open-access CityGML datasets are limited. This data knowledge gap is challenging for urban energy system analysis.

Cities can be seen as complex structures. They might hold a large number of entities of different kind. The manual modeling requires a lot of effort for data gathering and urban energy model generation. Thus, the question arises how to enable a fast and simple way to access building data sets, enrich them, and generate urban energy models.

According to Keirstead et al. UEM are "formal system[s] that represents the combined processes of acquiring and using energy to satisfy the energy service demands of a given urban area" [4]. UEM can support the demand estimation and energy efficiency analysis of districts. According to Cerezo Davila et al. urban building energy models are physical simulation models of building structure heat flows [5], which can support the estimation of space heating demand requirements of urban areas. They classify urban energy models in top-down and bottom-up models. Top-down models mainly contribute to analyses on macroeconomic scale, such as population dynamics or economic growth analysis on city or regional scale for instance to predict future energy usage. In comparison, bottom-up modeling applies analytical

* Corresponding author.
E-mail address: jan_schiefelbein@gmx.de (J. Schiefelbein).

https://doi.org/10.1016/j.buildenv.2018.12.025
Received 14 July 2018; Received in revised form 2 December 2018; Accepted 11 December 2018
Available online 14 December 2018

methods on single building scale. It can be used to evaluate the influence of single building parameters on the overall urban energy system, for instance as performed by Mostafavi et al. [6] or by Kazas et al. [7].

The modeler has to find a reasonable trade-off between level of detail of each building and runtime requirements for model generation and simulation. While detailed building performance simulation (BPS) is well established and often executable in practical runtime, urban energy modeling and simulation might lead to impractical runtime requirements, if the level of detail per building is high. Moreover, the model parameterization requirements should be low on urban scale. A possible solution approach is the usage of simplified building models such as Reduced-Order Models (ROM) to keep parameterization effort to a minimum and reduce runtime requirements while still being able to account for physical effects. Different building elements, like walls and windows, are merged into lumped resistances and capacitances. ROM are suitable for simulations on urban scale, as shown by Lauster et al. [8], Kim et al. [9] and Fonseca et al. [10].

Parameterization of buildings on urban scale is another important issue for urban energy modeling. Manual parameterization of each building requires effort and might get impractical for larger districts with hundreds of buildings. The archetype approach is an option to simplify the building modeling process. Building archetypes are building models, that hold representative building information for specific regions and construction periods. In this case, representative means that their geometry, wall-window-ratios and their physical properties are derived from statistics about existing building stock. Their usage simplifies the model generation, as many building attributes can be derived from basic building input data, such as year of construction, type of building and ground floor area. Cerezo et al. [11] compared different archetype models in a modeling process for Kuwait city. Dogan et al. choose a shoebox modeling approach [12] to combine building structures into groups of simplified shoebox building models to reduce simulation time on urban scale. Remmen et al. provide the open-source Python tool TEASER to enable building archetype generation and urban energy modeling [13]. TEASER enables the creation of full-scale building construction data sets with low input requirements. TEASER uses archetype information to enrich urban energy models and enable dynamic space heating simulations. Remmen et al. demonstrated the TEASER usage on building, neighborhood and urban scale.

With increasing complexity and size of city districts, the demand for effective data management rises. Thus, many urban energy modeling projects deal with databases and geographical information system (GIS) coupling to manage city district data, such as Monteiro eta al. [14], who describe an urban building database approach. Fuchs et al. describe an automated workflow for modeling of campus districts in combination with a PostgreSQL database, GIS usage and Modelica simulations [15]. Nageler et al. chose a comparable approach for urban energy modeling [16]. These frameworks can be coupled with optimization tools, as done by Fonseca et al. with the City Energy Analyst [10,17].

Besides 2D GIS approaches, 3D applications are used for urban energy modeling. CityGML is an exchange format for 3D city district data, which is also used for energy analyses. Harter et al. [18] performed a life cycle analysis for buildings on city district scale based on CityGML data. Shi et al. developed a three-step simulation-based urban form generation and optimization model to try to tackle the relationship between urban form and urban energy usage [19].

Moreover, urban energy models enable special use cases for city districts. Chen et al. developed a web interface to calculate building energy savings on city district scale [20]. Further research is done on how urban form affects energy demand [21,22]. To increase efficiency, control strategies on city district scale get more important [23]. Furthermore, whole city districts can be used to perform demand side management (DSM), as performed by Müller et al. [24] or Kolen et al. [25].

In summary, urban energy system planners are challenged by two main issues: Data uncertainty and modeling effort. A possible solution to tackle the first issue is the usage of OpenStreetMap (OSM) (https://www.openstreetmap.org/). OSM provides open-access to multiple datasets on urban scale. Moreover, different application programming interfaces (API) exist for OSM. Thus, automated data extraction is possible, which supports a simplified modeling process on urban scale.

OSM data extracts might not hold enough data to parameterize detailed building models. However, basic building information is available, which can be enriched. TEASER can be used to generate building models in situations where only basic building information is available. The archetype model information can directly be used in ROM to enable dynamic simulations on urban scale. ROM are promising to keep runtime requirements for simulation within acceptable limits. This automated modeling methodology is supportive to solve the second issue.

1.1. Contribution of this work

Within this work, we present a simple, GIS-based approach to model and simulate city district building energy systems. This work is an extension of the work presented in Ref. [26]. The novelty of this approach is the automated urban energy system model generation and enrichment based on OSM in combination with a thermal building simulation workflow. In a first step, building and city district data is automatically extracted from OSM with the Overpass API (http://overpass-api.de/) and a Python interface. Thus, large amounts of open-source datasets can easily be extracted. As the data quality and availability might vary for different regions on the OSM platform, data enrichment can be performed. This enables estimation of missing building base data, such as year of construction or height of buildings. Building base data is then used to generate building archetype models with the Python tool TEASER [13]. Within TEASER representative geometries and physical properties can be generated based on basic building input data and building stock statistics. Subsequently, TEASER is used to export each archetype building into ROM to enable dynamic space heating simulations on city district scale. A Python implementation of the VDI6007 [27] building simulation model is used to run all simulations, using the TEASER ROM building models. The approach is tested with a space heating demand uncertainty analysis on city district scale as well as a reference model simulation run. The aim of the case study was to verify the accurate functionality of the modeling and simulation framework by comparing simulation and reference consumption values. The presented modeling framework can be used to simplify building base data extraction, rapidly generate city district models and perform space heating load simulations on urban scale.

1.2. Outline

Section 2 describes the used methodology for urban energy modeling. First, the Python modeling framework is elaborated. Second, the city district data extraction is explained. Third, the data enrichment approach is elaborated. Forth, the building archetype approach and, fifth, the chosen simulation model are described. A case study is performed in section 3. The modeling and simulation approach is used to perform a space heating uncertainty analysis on city district scale. Finally, modeling and case study results are discussed.

2. Methodology

The following section describes the modeling framework for urban energy systems as well as the building modeling and simulation approach. Fig. 1 shows the general modeling and simulation workflow. At the beginning, open GIS data sets are extracted to generate a city model with the urban topology, such as building positions and street routings, and basic building data, such as building types and ground areas. Data enrichment can be used to add basic building data, such as construction years or building heights, as OSM data sets might be incomplete. The

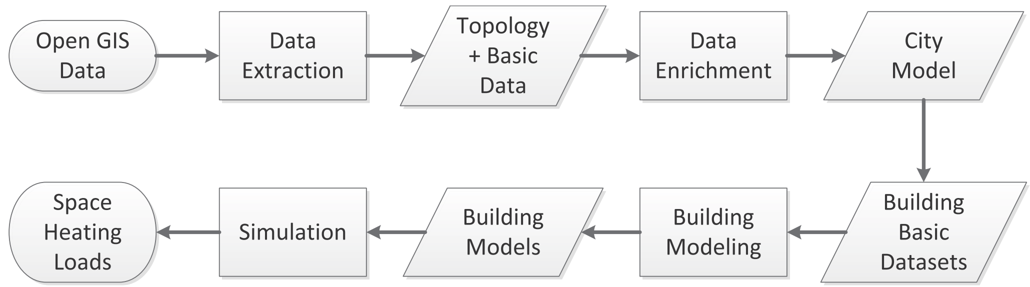

Fig. 1. General modeling workflow.

data enrichment results into the basic city model. Then, building basic datasets are used to automatically generate building models with representative geometry and physical properties. The dynamic simulation of each building leads to space heating load profiles and energy demands on urban scale.

2.1. pyCity modeling framework

UEM are necessary to analyze energy demand requirements and effects of energy efficiency measures on urban scale. Although different UEM frameworks exist, most of them require a lot of effort for manual input data preparation and data management. In order to simplify UEM generation and handling, we designed a modular Python framework. For this framework we identified the following functional requirements:

- Automated open GIS data access
- Data management of city topological information and building data
- Data enrichment of basic city model
- User profile and internal loads generation
- Automated building modeling and simulation

To address these requirements, we designed a framework consisting of five Python packages:

- uesgraphs
- pyCity
- pyCity_calc
- richardsonpy
- TEASER

The package uesgraphs is used to describe urban topologies as graphs with building nodes and street networks (https://github.com/RWTH-EBC/uesgraphs). Moreover, uesgraphs provides an interface to access and extract OSM data.

The package pyCity is used to model city districts and handle city district data. It extends uesgraph models to hold building objects with zones and energy loads as well as building energy systems (BES), such as boilers, photovoltaic systems or thermal storages [28]. pyCity_calc extends pyCity with calculation logic, such as data enrichment for city district models. Fig. 2 shows the dependencies of uesgraphs, pyCity and pyCity_calc.

To enable the generation of user profiles and internal loads for building simulation, a Python version of the Richardson tool [29–31], named richardsonpy, is used (https://github.com/RWTH-EBC/richardsonpy). The package TEASER [13] is used to generate physical building models and perform dynamic space heating load simulations on urban scale (https://github.com/RWTH-EBC/TEASER). Data extraction, enrichment as well as building modeling and simulation will be explained in detail in the following sections.

2.2. OSM city district data extraction

The Python package uesgraphs is used to extract city district datasets from OpenStreetMap. This is done with the Overpass API, which enables OSM data export for a user defined area. Data sets are then parsed to a city object of pyCity. Thus, the basic city object holds building nodes with positions and ground areas as well as street networks. OSM might hold further building data, such as name tags defining the building type and usage or information about construction years, which can also be extracted and parsed to individual building nodes.

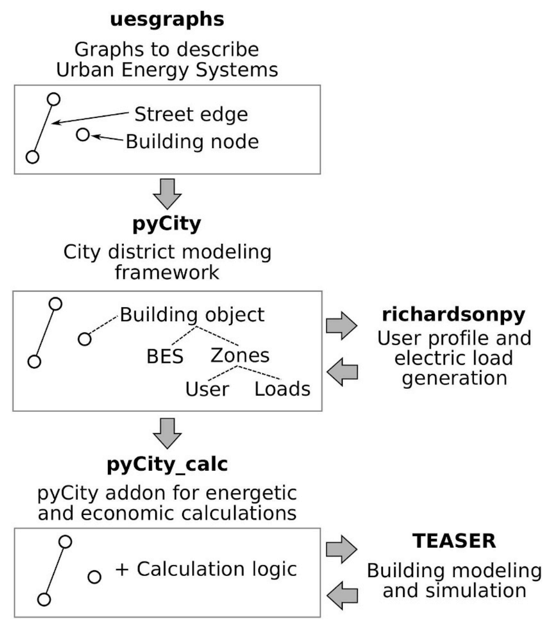

Fig. 2. Schemes and dependencies of uesgraphs, pyCity and pyCity_calc Python packages.

This interface enables a fast basic urban energy model generation with many building entities. However, there is no guarantee for sufficient data availability for the all extracted buildings. Thus, data enrichment might be necessary to estimate missing basic building data.

2.3. Data enrichment on urban scale

The following input data sets are required for TEASER [13] to generate building models:

- Building type (e.g. residential)
- Year of construction (or last year of modernization)
- Basic geometry (e.g. ground area and height)

OSM might hold these basic datasets. However, the OSM data quality varies depending on the area of interest. For instance, the city of Aachen, Germany, holds 2D building information, but lacks height data [32]. Furthermore, a large share of buildings hold building type information, while only a minor share of buildings has year of construction data. If some extracted building lacks these basic data sets, data enrichment can be performed to estimate and add missing information.

The following steps are processed during data enrichment:

1. OSM data analysis
2. Estimate building types
3. Classify building structures into corresponding urban forms
4. Estimate and add missing data

Within step 1, building ground areas and information about adjacent buildings can be extracted from OSM building polygons. In this case adjacent means, that building ground area polygons share common edges respectively polygons are connected. Based on the number of adjacent buildings and the building ground area, the building type is estimated within step 2. For instance, stand-alone buildings with a relatively small ground area are classified as single-family buildings. Interconnected buildings of comparable ground area size and form are identified as terraced houses. Buildings with residential tag and relatively large ground area are classified as multi-family houses. Non-residential buildings are either directly identified with OSM information, such as non-residential building type attributes, or tried to be identified based on ground area and neighbor distance. For instance, relatively large ground areas in combination with relatively high distance to neighbor buildings can be an indicator for commercial or industrial buildings.

In step 3, the data enrichment algorithm is looking for comparable ground areas and structures within a specific area around the chosen building. Comparable to Hegger et al. [33] the search area around the building is limited to 10000 m^2. The area is still large enough to identify common urban structures but small enough to perform the analysis in reasonable runtime. Moreover, the share of paved areas related to the search area around the chosen building is analyzed.

If ground areas have comparable forms or sizes, they might belong to the same urban form, which define building clusters with similar attributes, such as areas with free-standing single-family houses only. Thus, the data enrichment tool assigns buildings to specific urban forms of Hegger et al. [33]. Hegger et al. analyzed different city districts in Germany and classified them into energetic urban forms (EUF) to simplify the estimation of energy demand requirements and renewable energy potential for city districts with comparable urban form. Moreover, they provide data regarding representative construction density, building types and geometry for each EUF. These datasets are used to estimate further missing building data, such as number of floors per building. For instance, residential buildings with relatively large ground area combined with large distances between these buildings can be an indicator for multi-story residential buildings.

In step 4, missing parameters are added to each building. Net floor areas are calculated based on building type, ground area and height. If at least one building within the area of interest holds information about the year of construction, this year is chosen as anchor reference year for the area. Nearby buildings are assumed to have nearly equal years of construction and are, therefore, classified into identical construction year classes. If no information about years of construction exists within OSM, user inputs and German statistics about years of construction [34] can be used. Based on Hoier et al. [35] last years of modernization can be sampled for each building. Depending on building type and net floor area the number of apartments is estimated. Occupants are distributed to each apartment based on Zensus statistics [36]. Further details on the data enrichment approach can be found in Ref. [32].

2.4. Building modeling

Building models are necessary to enable space heating simulation on urban scale. Detailed modeling of each building on urban scale would require large effort and sufficient data availability. To reduce modeling effort and to deal with low data availability, a building archetype approach is chosen for building model generation. The archetype modeling is done with the Python tool TEASER [13]. Based on building type, year of construction and basic geometry, TEASER is able to generate building archetype models with representative geometry and physical properties, such as layer materials and thicknesses. TEASER has the ability to use the archetype building data to parametrize Reduced-Order models (ROM), which enable dynamic space heating demand simulation. ROM describe thermal systems as simplified resistances (R) and capacitances (C) models, comparable to electric circuits. According to Lauster et al. [8] and Kim et al. [9] ROM enable building simulations on urban scale at practical runtime with low reduction of results precision. The TEASER VDI6007 [27] building simulation model implementation is chosen for simulation.

2.5. Dynamic space heating simulation

The following elements are required for dynamic space heating building simulations:

- Physical building model
- Internal gains (such as persons and electric devices as heat source)
- Internal losses (such as user air exchange causing heat losses)
- External gains/boundary conditions (such as outdoor temperature and radiation)

TEASER enables archetype building modeling and ROM parameterization to provide physical building models for dynamic simulations. Infiltration due to building leakages is estimated based on the year of construction, respectively last year of modernization. The possible range of natural infiltration is chosen relative to DIN 18599 [37]. Buildings with younger building age are assumed to have lower infiltration losses.

Internal gains primarily depend on occupant behavior. During presence, persons might use different electric appliances and lighting devices. Human presence and device usage leads to internal heat gains, which can reduce heat demand. To account for occupancy influence in residential buildings, occupancy profiles are modeled with the package richardsonpy, which is a Python implementation of the Richardson tool [29–31]. Human heat flux is defined based on the DIN 18599 [37]. The package richardsonpy uses the number of occupants per apartment as input to generate occupancy and electric load profiles per apartment. As the Richardson tool is based on a UK data survey, the electric load profiles are normalized with a statistic about German electricity usage in residential buildings [38].

Moreover, air exchange profiles due to user activity are generated based on occupancy profiles and outdoor temperature. Therefore, pyCity_calc has a stochastic air exchange rate profile generator, comparable to the methodology of Cali et al. [39]. Window state changes (open/closed) can only happen during occupancy presence and activity. Windows are supposed to have longer periods of being open during periods of higher outdoor temperature, while the probability of windows being closed increases with decreasing outdoor temperature.

External gains are calculated with weather data sets. In this case, Test Reference Year (TRY) datasets of German weather service (including outdoor temperature and radiation) are chosen as input [40]. Each building is simulated individually to enable a parallelization of the calculations. This process can accelerate the overall simulation. However, inter-building radiation is not taken into account. Fig. 3 shows the simulation workflow per building. Fig. 4 summarizes the UEM modeling and simulation framework.

OpenStreetMap is used to extract basic city district data. The extraction leads to a city graph of uesgraphs package with building nodes, street nodes and edges. The building nodes might hold semantic building data extracts of OSM. However, missing building basic datasets can be estimated with pyCity_calc to enrich the city model. Basic

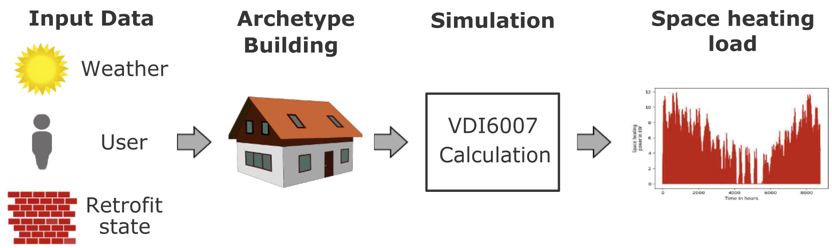

Fig. 3. Workflow for thermal simulation with VDI6007 calculation core.

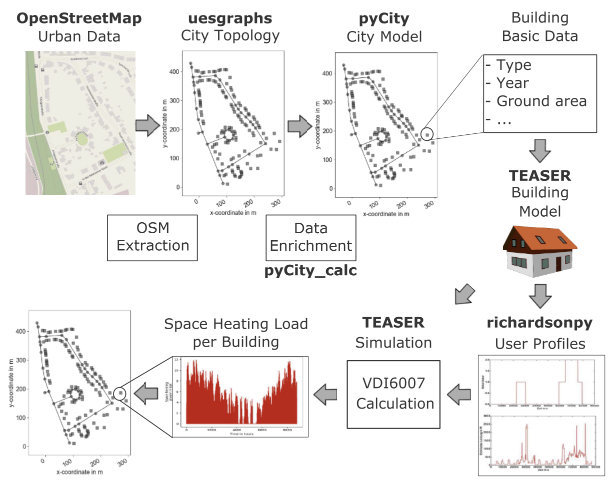

Fig. 4. UEM modeling and simulation framework [Uses data of ©OpenStreetMap contributors].

building data sets, such as building type, year of construction, ground area and building height, are imported into TEASER to generate building archetype models with representative geometry and physical properties. The Python package richardsonpy enables generation of occupancy and electrical load profiles, which serve as inputs for each building simulation. These models can be used to perform thermal simulations on urban scale.

3. Case study

The framework has been used to perform a case study analysis with a city district within Bottrop, Germany. The district has been chosen due to knowledge about yearly energy consumption values. The aim of the case study is to verify the accurate framework functionality by comparing simulated net space heating values of the framework city district model with reference measurement values. Thus, the following steps have been performed:

1. OSM data extraction and enrichment to generate basic city model
2. Dynamic space heating simulation(s):
   (a) Uncertainty analysis: Multiple simulations for samples of uncertain parameters.
   (b) Reference model analysis: Single simulation of reference model with most likely parameters.
3. Comparison of simulated net space heating demand values with reference consumption value

The uncertainty analysis is performed to get a realistic range respectively probability density function of space heating net energy demand values on city district scale.

Fig. 5 shows the case study building block areal view of OpenStreetMap.

The district consists of 55 residential buildings. Reference final energy statistics on building block scale were given in context of the project “EnEff:Stadt - Bottrop, Welheimer Mark” [28,41]. Detailed knowledge about thermal supply units or energy demands per building was not given. However, the majority of buildings is supplied by old gas or oil fired boilers. Thus, we assumed that all reference buildings are boiler-supplied. Cooling demand has been neglected, because of the relatively cold weather in Germany and the lack of cooling equipment in German households.

3.1. Reference city model generation

The city district model has been generated with the framework described in section 2. Therefore, a OSM data extraction has been performed. Extracted building structures with ground areas smaller than 50 m^2 were neglected, as these structures are mainly unheated garages or sheds. The reference city model has then be used as input for the uncertainty analysis and single reference simulation run.

3.2. Uncertainty analysis

The city model has been used within a space heating energy demand uncertainty analysis. Therefore, 10,000 simulation runs have been performed for each building to get a probability density function (PDF) for net space heating energy demand on city district scale. This PDF has been compared to the energy consumption reference value and simulation result of the reference city object.

3.2.1. Input data sampling

To perform a space heating energy demand uncertainty analysis on city district scale, input datasets have to be sampled for every run. Uncertain parameters exist on apartment, building and city scale. The following uncertain parameters are chosen and used within sampling:

- Apartment level:
  – Number of occupants per apartment
  – Electric demand per occupant
  – Desired indoor temperature
  – User air exchange rate
- Building level:
  – Last year of modernization
  – Infiltration rate
- City level: Outdoor temperature profile

10,000 samples have to be selected for each uncertain parameter on each level. Table 1 gives an overview of the selected probability distributions.

The number of occupants per apartment is sampled between 1 and 5 occupants based on statistics of Destatis [44]. Electric demand per occupant is sampled from a uniform distribution related to German Stromspiegel [38].

The desired indoor temperature is sampled from a normal distribution with mean of 20 °C and a standard deviation (STD) of 2.5 °C. The mean has been chosen related to DIN18599 [37]. The relatively large standard deviation has been used to account for extreme user behavior.

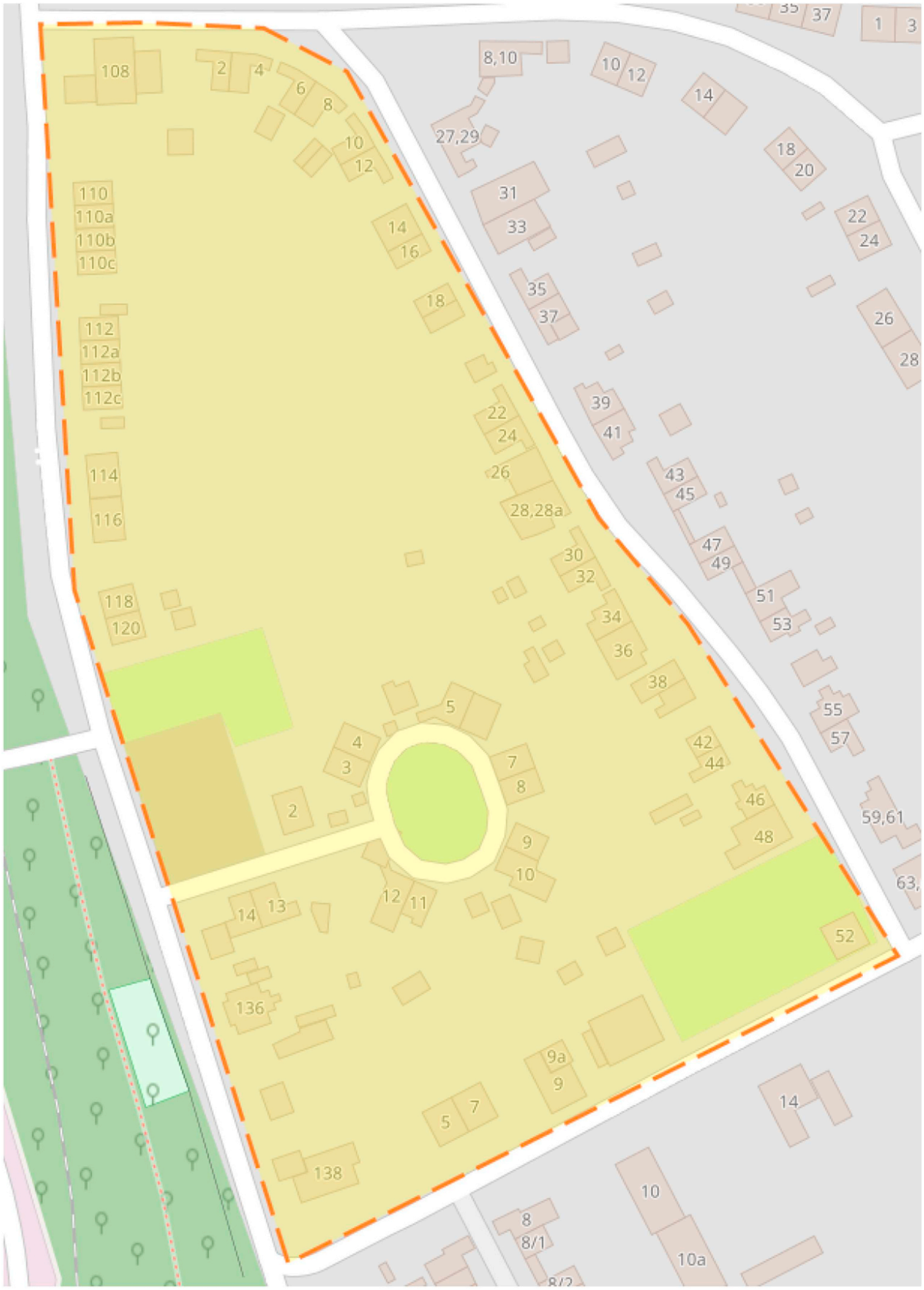

Fig. 5. Case study building block areal view [Uses data of ©OpenStreetMap contributors].

The user air exchange rate in room volume per hour is sampled from a Nakagami distribution with shape 0.6 and spread 0.4, comparable to Stinner et al. [42].

The natural ventilation rate of the building in volume per hour is sampled from a log-normal distribution with mean 0 and standard deviation of 1 divided by 6, related to analyses of Münzenberg [43]. The maximum allowed value for natural ventilation has been set to 2.

The outdoor temperature and radiation curves have been sampled from linear interpolations between warm, regular and cold TRY of German weather service [40], assuming a uniform distribution. Measurement values for outdoor temperature or radiation were not given.

The last year of modernization is chosen from an equal distribution between 1974 and 2014 to account for the majority of German building standards. The sampled years of modernization have been used to generate new archetype building models with TEASER. The archetype geometries remained the same for each building. However, physical properties could change, e.g. through additional insulation for walls and rooftops or through different kind of glazing.

3.2.2. *Simulation runs*

Sample sets for number of occupants, electric usage, user infiltration rate, desired set temperatures and outdoor temperature curves have been used as input for the TEASER VDI6007 calculation core. An hourly timestep for a duration of one year has been chosen for simulation. Thermal simulations have been performed for each building and each data sample set per building to get a distribution of space heating energy demands as result.

Table 1
Overview of probability distributions for uncertainty analysis.

Parameter	Unit	Distribution	Details	Source
Last year of retrofit		Equal	Min: 1974; Max: 2014	
Set temperature	°C	Normal	Mean: 20; STD: 2.5	[37]
User air exchange rate	1/h	Nakagami	Shape: 0.6; spread 0.4	[42]
Natural ventilation	1/h	Lognormal	Mean: 0; STD: 1 (divided by 6); Max: 2	[43]
Outdoor temperature	°C	Uniform	Between cold, regular, and warm TRY	[40]
Radiation	W/m^2	Uniform	Between cold, regular, and warm TRY	[40]
Nb. of apartments		Special	See Table 3	[44]
Electric demand per person	kWh/p	Uniform	See Tables 4 and 5	[38]

3.3. Reference simulation

Additionally, a single simulation of the reference city district model has been performed. Number of occupants and apartments has been taken from the data enrichment approach. The electric demand has been estimated based on average electrical demand in Germany [38]. According to DIN 18599 [37], the desired indoor temperature during daytime is set to 20 °C, while the set temperature during nighttime is set to 16 °C.

Retrofit years have been chosen between 1980 and 1994 for the reference city model. Building natural infiltration rate is estimated based on the year of construction respectively last year of modernization, related to DIN 18599 [37]. Buildings with younger year of construction respectively retrofit year are assumed to have lower natural infiltration through leakages. User air exchange rates are generated with the pyCity_calc stochastic air exchange rate profile generator.

3.4. Results

Due to data confidentiality, the simulated space heating demand values of the uncertainty runs and the reference simulation have been normalized to the reference energy consumption value to enable comparison of consumption and demand values. The thermal final energy demand on building block scale [28,41] has been used to calculate the reference energy consumption value. As we simulated net space heating demand values, the reference final energy consumption had to be converted to be comparable. Therefore, the thermal final energy reference value has been converted into net thermal energy demand value with an assumed boiler efficiency of 0.85 for older boiler supply units. We used heating degree days provided by German weather service to perform a weather normalization of the reference consumption values [45]. The net energy demand value for hot water, based on German statistics about hot water electricity usage [46], has been extracted from the reference value. The conversion led to a net space heating energy consumption reference value for the case study city district.

Fig. 6 shows the normalized space heating net energy demand values of case study city district.

Table 2 lists relevant result parameters of the normalized space heating net energy demand distribution. The distribution of normalized net space heating demands is close to a gaussian distribution with mean of 1.022 and standard deviation of 0.084. Mean, median and reference simulation value are close to the measurement reference value for the given distribution. While mean and median slightly overestimate the space heating demand, the reference simulation led to a small underestimation. The relative interquartile range (RIQR) is chosen for further evaluation. The RIQR is defined as the difference between the 75-quantil and the 25-quantil, related to the median [42], which is 0.111 for the result distribution.

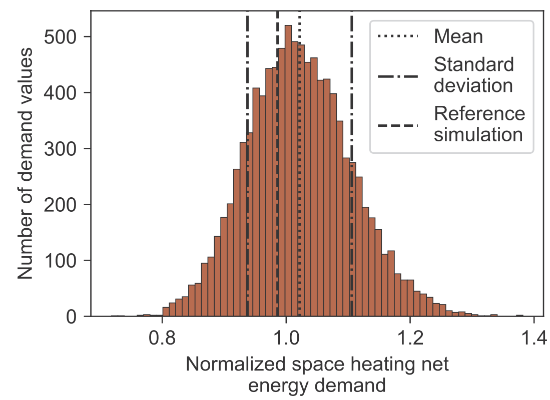

Fig. 6. Histogram of normalized space heating net energy demand values of case study city district.

Table 2
Result parameters of normalized space heating demand distribution.

Parameter	Value
Mean	1.022
Median	1.018
Reference simulation	0.986
Standard deviation	0.084
Minimum	0.718
Maximum	1.382
Relative Interquartile Range (RIQR)	0.111

3.5. Discussion

Mean, median and the reference simulation demand values are close to the reference consumption value (close to 1). Moreover, the RIQR of 0.111 is relatively small. Thus, the reference simulation as well as the uncertainty analysis simulations are good options to estimate net space heating demand values for the case study district. However, the uncertainty analysis has the advantage of providing a possible distribution of demand values for the decision maker.

A possible explanation for the relatively good fit is the effect of averaging out. With increasing number of buildings and apartments, extreme user behavior might average out on city district scale, for instance as shown by Stinner et al. [42]. Thus, the overall user behavior related to desired indoor temperatures, air exchange rates and appliance usage tends to get closer to representative values used by standards and guidelines, such as DIN18599 [37]. However, we assume that the user uncertainty remains high on single building scale, due to strong influence of potentially extreme user behavior.

The retrofit state is another important factor for space heating requirements. The retrofit state of the case study district seems to be representative for the average retrofit state of German building stock, because of the good fit between simulated demand and measured reference value. This will not be the case for every building block or district, as the overall district might have a much better or much worse retrofit state compared to the average German housing stock. However, the uncertainty analysis enables an option to estimate the possible distribution of space heating demands and provide realistic space heating demand ranges on city district scale.

Moreover, the results of the reference simulation show that data extraction and data enrichment of the case study district led to a realistic city district model and plausible net space heating demand on building block scale. Thus, the modeling toolchain enables an easy way to generate city district models based on OSM data, which can be used for further analysis and energy system dimensioning and optimization.

4. Conclusions

Within this paper, we describe a toolchain for urban energy model (UEM) generation. The toolchain provides solutions for the modeling challenges of input data uncertainty and high modeling effort by enabling an GIS-based data extraction and enrichment as well as automated building modeling and simulation. OpenStreetMap (OSM) has been chosen as primary data source for model generation, due to good data availability and OSM interfaces for automated data extraction. The city district topology as well as basic building data, such as ground area or building type, can be extracted from OSM via Overpass API. A statistics-based data enrichment approach has been chosen to estimate missing building datasets, such as years of construction or number of floors.

A building archetype approach has been chosen to simplify the building model generation process. The Python tool TEASER [13] is used to generate representative building models, based on building data about year of construction, usage and rough building geometry. The building information can be used to parameterize Reduced-Order Models (ROM) to enable dynamic thermal simulations on urban scale.

The tool chain has been used to generate a city district model of a building block with 55 houses in Bottrop, Germany. The urban energy model has been used to perform an uncertainty analysis with uncertain input data for space heating demand simulation. Moreover, a single reference simulation has been performed with the reference city model.

The simulated space heating net energy demand values have been compared to a measured consumption reference value on building block scale. There has been a good fit between simulated demand values and reference consumption values. Thus, the modeling toolchain has been able to generate a sufficient reference city district model with realistic space heating energy demand values for the case study district. Real space heating consumption and simulated demand might vary for other districts, especially for districts with very old building structure or modern retrofit standards, when no building age information is given. However, the automated city district model generation is still applicable to automatically generate basic city models, as city district topology, building geometries and usage types can still be extracted from OSM. Moreover, the uncertainty analysis can provide a distribution of possible demand values on building block scale, which can be advantageous if decision makers lack reference consumption values.

The described approach strongly depends on the OSM data quality and availability. Moreover, the data enrichment process is sensitive to the level of heterogeneity of a district. As the enrichment is based on statistics, it can perform well for districts, which are comparable to the reference city district type or urban form, while it might be challenging for districts, which strongly differ from the related city district type or urban form. Another limitation is the lack of inter-building radiation or shading in the model.

In future work, the OSM data gathering and enrichment should include different interfaces and export formats. This enables the definition and exchange of reference city district models. These reference district models could be processed with different simulation tools to enable comparison and benchmarking between diverse simulation engines. As the Python VDI6007 model is able to simulate cooling demands, a validation has to be done.

Overall, the OSM-based modeling tool chain enables an automated way to generate city district models with multiple building entities, which can be used for further analysis, simulation and energy system planning.

Nomenclature

API	application programming interface
BES	building energy systems
BMWi	Federal Ministry for Economic Affairs and Energy
BPS	building performance simulation
C	capacitances
DIN	German Institute for Standardization
DSM	Demand Side Management
EUF	energetic urban forms
GIS	geographical information system
OSM	OpenStreetMap
PDF	probability density function
R	resistances
RIQR	Relative Interquartile Range
ROM	Reduced-Order Models
STD	standard deviation
TRY	Test Reference Year
UEM	urban energy models
VDI	Association of German Engineers

Additional information

The Python package TEASER is available at https://github.com/RWTH-EBC/TEASER.

The Python package uesgraphs is available at https://github.com/RWTH-EBC/uesgraphs.

The Python package richardsonpy is available at https://github.com/RWTH-EBC/richardsonpy.

The Python packages pyCity and pyCity_calc will be published at https://github.com/RWTH-EBC.

Acknowledgements

Grateful acknowledgement is made for financial support by Federal Ministry for Economic Affairs and Energy (BMWi), promotional references 03ET1138D and 03ET1381A. Moreover, we would like to thank our project partners city of Bottrop and Innovation City Management GmbH for the productive cooperation.

Appendix

Table 3
Distribution of occupants per apartment according to Destatis [44].

Number of persons	Share
1	0.414
2	0342
3	0.121
4	0.09
5	0.033

Table 4
Minimum and maximum electric demand per apartment (single family houses) according to German Stromspiegel 2017 [38].

Number of persons in apartment	Minimum electric demand in kWh	Maximum electric demand in kWh
1	1300	4000
2	2100	4400
3	2600	5200
4	2900	5900
5	3500	7500

Table 5
Minimum and maximum electric demand per apartment (multi family houses) according to German Stromspiegel 2017 [38].

Number of persons in apartment	Minimum electric demand in kWh	Maximum electric demand in kWh
1	800	2200
2	1300	3100
3	1700	3900
4	1900	4500
5	2200	5700

References

[1] United Nations, Cities and Climate Change: Global Report on Human Settlements, (2011).

[2] H. Amecke, The impact of energy performance certificates: a survey of German home owners, Energy Pol. 46 (2012) 4–14, https://doi.org/10.1016/j.enpol.2012.01.064.

[3] United Nations, Paris Agreement, (2015).

[4] J. Keirstead, M. Jennings, A. Sivakumar, A review of urban energy system models: approaches, challenges and opportunities, Renew. Sustain. Energy Rev. 16 (6) (2012) 3847–3866, https://doi.org/10.1016/j.rser.2012.02.047.

[5] C. Cerezo Davila, C.F. Reinhart, J.L. Bemis, Modeling Boston: a workflow for the efficient generation and maintenance of urban building energy models from existing geospatial datasets, Energy 117 (2016) 237–250, https://doi.org/10.1016/j.energy.2016.10.057.

[6] N. Mostafavi, M. Farzinmoghadam, S. Hoque, Urban residential energy consumption modeling in the integrated urban metabolism analysis tool (IUMAT), Build. Environ. 114 (2017) 429–444, https://doi.org/10.1016/j.buildenv.2016.12.035.

[7] G. Kazas, E. Fabrizio, M. Perino, Energy demand profile generation with detailed time resolution at an urban district scale: a reference building approach and case study, Appl. Energy 193 (2017) 243–262, https://doi.org/10.1016/j.apenergy.2017.01.095.

[8] M. Lauster, J. Teichmann, M. Fuchs, R. Streblow, D. Mueller, Low order thermal network models for dynamic simulations of buildings on city district scale, Build. Environ. 73 (2014) 223–231, https://doi.org/10.1016/j.buildenv.2013.12.016.

[9] E.-J. Kim, G. Plessis, J.-L. Hubert, J.-J. Roux, Urban energy simulation: simplification and reduction of building envelope models, Energy Build. 84 (2014) 193–202, https://doi.org/10.1016/j.enbuild.2014.07.066.

[10] J.A. Fonseca, A. Schlueter, Integrated model for characterization of spatiotemporal building energy consumption patterns in neighborhoods and city districts, Appl. Energy 142 (2015) 247–265, https://doi.org/10.1016/j.apenergy.2014.12.068.

[11] C. Cerezo, J. Sokol, S. AlKhaled, C. Reinhart, A. Al-Mumin, A. Hajiah, Comparison of four building archetype characterization methods in urban building energy modeling (UBEM): a residential case study in Kuwait City, Energy Build. 154 (2017) 321–334, https://doi.org/10.1016/j.enbuild.2017.08.029.

[12] T. Dogan, C. Reinhart, Shoeboxer, An algorithm for abstracted rapid multi-zone urban building energy model generation and simulation, Energy Build. 140 (2017) 140–153, https://doi.org/10.1016/j.enbuild.2017.01.030.

[13] P. Remmen, M. Lauster, M. Mans, M. Fuchs, T. Osterhage, D. Müller, TEASER: an open tool for urban energy modelling of building stocks, J. Build. Perform. Simulat. (2017) 1–15, https://doi.org/10.1080/19401493.2017.1283539.

[14] C. S. Monteiro, C. Costa, A. Pina, M. Y. Santos, P. Ferrão, An urban building database (UBD) supporting a smart city information system, Energy Build. http://doi.org/10.1016/j.enbuild.2017.10.009.

[15] M. Fuchs, J. Teichmann, M. Lauster, P. Remmen, R. Streblow, D. Müller, Workflow automation for combined modeling of buildings and district energy systems, Energy 117 (2016) 478–484, https://doi.org/10.1016/j.energy.2016.04.023.

[16] P. Nageler, G. Zahrer, R. Heimrath, T. Mach, F. Mauthner, I. Leusbrock, H. Schranzhofer, C. Hochenauer, Novel validated method for GIS based automated dynamic urban building energy simulations, Energy 139 (2017) 142–154, https://doi.org/10.1016/j.energy.2017.07.151.

[17] J.A. Fonseca, T.-A. Nguyen, A. Schlueter, F. Marechal, City Energy Analyst (CEA): integrated framework for analysis and optimization of building energy systems in neighborhoods and city districts, Energy Build. 113 (2016) 202–226, https://doi.org/10.1016/j.enbuild.2015.11.055.

[18] H. Harter, V. Weiler, U. Eicker, Developing a roadmap for the modernisation of city quarters – comparing the primary energy demand and greenhouse gas emissions, Build. Environ. 112 (2017) 166–176, https://doi.org/10.1016/j.buildenv.2016.11.031.

[19] Z. Shi, J.A. Fonseca, A. Schlueter, A review of simulation-based urban form generation and optimization for energy-driven urban design, Build. Environ. 121 (2017) 119–129, https://doi.org/10.1016/j.buildenv.2017.05.006.

[20] Y. Chen, T. Hong, M.A. Piette, Automatic generation and simulation of urban building energy models based on city datasets for city-scale building retrofit analysis, Appl. Energy 205 (2017) 323–335, https://doi.org/10.1016/j.apenergy.2017.07.128.

[21] A. Hargreaves, V. Cheng, S. Deshmukh, M. Leach, K. Steemers, Forecasting how residential urban form affects the regional carbon savings and costs of retrofitting and decentralized energy supply, Appl. Energy 186 (2017) 549–561, https://doi.org/10.1016/j.apenergy.2016.02.095.

[22] L. Chen, J. Hang, M. Sandberg, L. Claesson, S. Di Sabatino, H. Wigo, The impacts of building height variations and building packing densities on flow adjustment and city breathability in idealized urban models, Build. Environ. 118 (2017) 344–361, https://doi.org/10.1016/j.buildenv.2017.03.042.

[23] J. Reynolds, Y. Rezgui, J.-L. Hippolyte, Upscaling energy control from building to districts: current limitations and future perspectives, Sustain. Cities Soc. 35 (2017) 816–829, https://doi.org/10.1016/j.scs.2017.05.012.

[24] D. Müller, A. Monti, S. Stinner, T. Schlösser, T. Schütz, P. Matthes, H. Wolisz, C. Molitor, H. Harb, R. Streblow, Demand side management for city districts, Build.

Environ. http://doi.org/10.1016/j.buildenv.2015.03.026.
[25] S. Kolen, C. Molitor, L. Wagner, A. Monti, Two-level agent-based scheduling for a cluster of heating systems, Sustain. Cities Soc. 30 (2017) 273–281, https://doi.org/10.1016/j.scs.2017.01.014.
[26] J. Schiefelbein, A. Javadi, M. Diekerhof, R. Streblow, D. Mueller, A. Monti, GIS supported city district energy system modeling, in: System Simulation in Buildings - SSB2014 (Ed.), Proceedings of the 9th International Conference on System Simulation in Buildings, Liege, 2014, https://doi.org/10.13140/RG.2.1.5036.1444.
[27] Verein deutscher Ingenieure (VDI), VDI 6007: Calculation of Transient Thermal Response of Rooms and Buildings.
[28] D. Müller, A. Monti, J. Schiefelbein, M. Diekerhof, A. Javadi, H. Pilick, Abschlussbericht: EnEff:Stadt – Bottrop, Welheimer Mark: Integrale Planung einer energetischen Aufwertung und neuen Versorgungsstruktur der Sektoren Arbeit, Wohnen und Infrastruktur.
[29] I. Richardson, M. Thomson, D. Infield, A high-resolution domestic building occupancy model for energy demand simulations, Energy Build. 40 (8) (2008) 1560–1566, https://doi.org/10.1016/j.enbuild.2008.02.006.
[30] I. Richardson, M. Thomson, D. Infield, A. Delahunty, Domestic lighting: a high-resolution energy demand model, Energy Build. 41 (7) (2009) 781–789, https://doi.org/10.1016/j.enbuild.2009.02.010.
[31] I. Richardson, M. Thomson, D. Infield, C. Clifford, Domestic electricity use: a high-resolution energy demand model, Energy Build. 42 (10) (2010) 1878–1887, https://doi.org/10.1016/j.enbuild.2010.05.023.
[32] J. Rudnick, Modelling of a Generator for Urban Districts in PyCity.
[33] M. Hegger, J. Dettmar, Energetische Stadtraumtypen. Strukturelle und energetische Kennwerte von Stadträumen, Fraunhofer IRB Verlag, Stuttgart, 2015.
[34] Institut für Wohnen und Umwelt - IWU, Deutsche Gebäudetypologie: Beispielhafte Maßnahmen zur Verbesserung der Energieeffizienz von typischen Wohngebäuden.
[35] A. Hoier, H. Erhorn, A. Pfnür, N. Müller, Energetische Gebäudesanierung in Deutschland.
[36] Statistische Ämter des Bundes und der Länder, Gebäude- und Wohnungsbestand in Deutschland: Erste Ergebnisse der Gebäude- und Wohnungszählung, (2011).
[37] Deutsches Insitut für Normung e.V., DIN V 18599-1: Energy Efficiency of Buildings – Calculation of the Net, Final and Primary Energy Demand for Heating, Cooling, Ventilation, Domestic Hot Water and Lighting.
[38] co2online gemeinnützige GmbH, Stromspiegel für Deutschland 2017, http://www.die-stromsparinitiative.de/fileadmin/bilder/Stromspiegel/broschuere/Stromspiegel_2017_web.pdf, (2017).
[39] D. Calì, R.K. Andersen, D. Müller, B.W. Olesen, Analysis of occupants' behavior related to the use of windows in German households, Build. Environ. 103 (2016) 54–69, https://doi.org/10.1016/j.buildenv.2016.03.024.
[40] Deutscher Wetterdienst, Testreferenzjahre, http://www.dwd.de/TRY, (2015).
[41] J. Schiefelbein, A. Javadi, M. Fuchs, D. Müller, A. Monti, M. Diekerhof, Modellierung und Optimierung von Mischgebieten, Bauphysik 39 (1) (2017) 23–32, https://doi.org/10.1002/bapi.201710001.
[42] S. Stinner, R. Streblow, D. Mueller, Dynamic energetic uncertainty analysis for city districts, in: International Building Performance Simulation Association (Ed.), Proceedings of BauSIM2014, Aachen, 2014.
[43] U. Münzenberg, Der natürliche Luftwechsel in Gebäuden und seine Bedeutung bei der Beurteilung von Schimmelpilzschäden, Umwelt, Gebäude & Gesundheit: innenraumhygiene, Raumluftqualität und Energieeinsparung, Ergebnisse des 7 (2004) 263–271.
[44] Statistisches Bundesamt, Bevölkerung in deutschland, https://www.destatis.de/DE/ZahlenFakten/Indikatoren/LangeReihen/Bevoelkerung/lrbev05.html;jsessionid=4AACC10D2225591EC88C40EDEFB5EDAC.cae2, (2017).
[45] Deutscher Wetterdienst, Klimafaktoren (KF) für Energieverbrauchsausweise, htttp://www.dwd.de/klimafaktoren, (2013).
[46] co2online gemeinnützige GmbH, Stromspiegel für Deutschland 2017: Klimaschutz zu Hause.