Accepted manuscript. To cite this article:
Ghiassi, N., Tahmasebi, F. & Mahdavi, A. Building Simulation (2017) 10: 1005-1021.
https://doi.org/10.1007/s12273-017-0356-1

Harnessing buildings' operational diversity in a computational framework for high-resolution urban energy modeling

Neda Ghiassi, Farhang Tahmasebi, Ardeshir Mahdavi

Abstract

To achieve computational efficiency, efforts toward developing urban-scale energy modeling applications frequently rely on various domain simplifications. For instance, heat transfer phenomena are captured using reduced order models. As a consequence, specific aspects pertaining to the temporal dynamics of energy load patterns and their dependency on transient phenomena (e.g., weather conditions, inhabitants' presence and actions) cannot be realistically represented. To address this circumstance, we have conceived, implemented, and documented a two-step urban energy modeling approach that combines cluster analysis and sampling techniques, full dynamic numeric simulation capability, and stochastic methods. The paper describes the suggested urban energy modeling approach and the embedded cluster analysis supported sampling methodology. More particularly we focus on the aspects of this approach that explicitly involve the representation of inhabitants in urban-scale energy modeling. In this regard, the potential to recover lost dynamic diversity (e.g., in computation of temporal load patterns) due to the deployed reductive sampling is explored. Parametric runs based on stochastic variations of underlying building use profiles facilitate the generation of highly realistic load patterns despite the small number of buildings selected to represent the simulation domain. We illustrate the utility of the proposed urban energy modeling approach to address queries concerning the energy efficiency potential of behaviorally effective instruments. The feasibility of the envisioned scenarios concerning inhabitants and their behavior (high-resolution temporal load prediction, assessment of behavioral variation) is presented in detail via specific instances of district-level energy modeling for the city of Vienna, Austria.

1. Introduction

The Paris environmental summit of 2015 concluded that global net zero emissions must be achieved before 2070 to avoid catastrophic levels of global warming (COP 2015). Governments across the world have set up ambitious plans for the reduction of emissions and energy use. Urban areas are considered as the primary human habitat since over half of the global population now lives in cities (World Bank 2015). Building stock as a major component of the urban environment, contributes generously to urban energy use and greenhouse gas emissions. Therefore, development of energy efficiency and emission mitigation strategies focused on the urban building stock is indispensable to a systematic shift towards emission-free cities. Towards this end, various demand-side and supply-side energy management approaches can be adopted. These include but are not limited to the improvement of the thermal quality of building envelopes, employment of more efficient systems and appliances, exploitation of the renewable energy resources through centralized and distributed plants, influencing the microclimatic conditions through introduction of vegetation, water bodies, and shading elements, encouraging sustainable building operation routines either directly via informational campaigns or indirectly via energy pricing strategies, etc. However, given the severity and urgency of the current situation and the substantial financial resources, time and effort required for the deployment of such strategies, trial and error is not an option. Integrative urban-level decision support environments, which allow for the investigation and comparative analysis of the implications of various energy and emission management plans, can help ensure the effectiveness of the envisaged strategies and an efficient allocation of the available resources.

The computational method or the energy model is the core component of these environments, determining the scope of their utility. Over the past years, efforts towards development of energy and emission models of the urban building stock for the assessment of various urban change and intervention scenarios, and their consequences have been steadily increasing. These models vary substantially in view of the general approach, scenario modeling capabilities, disaggregation level, required input parameters, and temporal and spatial resolution of the results. However, to achieve computational efficiency, efforts toward developing urban-scale energy modeling applications frequently rely on various domain simplifications. As such, they cannot accommodate queries pertaining to transient phenomenon such as microclimate and occupant behavior with appropriate level temporal and spatial resolution.

The present contribution reports on the development of a bottom-up space-heating energy demand model of the urban building stock. The model is intended as an extendable core computational component of an integrative urban decision support environment. The environment is targeted at comparative analysis of various urban change and intervention scenarios pertaining to the following aspects:

- Physical interventions: Thermal retrofit, densification, etc.
- Technological advancements: Use of solar technologies, efficient heating systems, etc.
- Climatic changes: Urban Heat Island Studies, etc.
- Occupant behavior changes: Induced by demographic changes, lifestyle changes, etc.

This development enables the employment of dynamic performance simulation for urban level energy assessments through a reductive process. Full-fledged simulation is well suited for representation of transient phenomenon with appropriate temporal resolution. For various urban scale queries (e.g., pertaining to the energy implications of microclimate variations or incorporations of small-scale renewable energy technologies), high spatial resolution is also required. As such, the model is integrated with Geographic Information Systems (GIS), which serves as an appropriate interaction, data-processing and visualization platform.

The following sections provide an over view of some previous efforts in this domain, describe the main components of the developed computational schema, and demonstrate its utility for investigation of behavioral change scenarios through simple illustrative examples.

2. State of the art

Bottom-up engineering (physical) urban energy models (Swan and Ugursal 2009) rely on heat transfer principles to arrive at the value of the energy demand of the entire population or a number of representative buildings in the study domain, in the latter case using the representative assessment results to extrapolate the energy performance of the entire population. Due to their computational method and their independence from historical demand data, these models are generally deemed most suitable for evaluation and investigation of various urban scenarios (Kavgic et al. 2010). The range of the observable scenarios depends on the versatility and resolution of the underlying building level energy computation routine and its representation of various energy-relevant building aspects (such as internal and external boundary conditions, thermal properties of the envelope, geometry, etc.). The predictive performance of the model on the other hand depends not only on the modeling capabilities of the computational engine, but also on the inevitable domain simplifications and the reliability of the model input parameters.

Few former urban energy modeling efforts have attempted a detailed representation of the entire building population under study. This is in part due to the scarcity of informational resources and in part due to the computational cost and effort required for such extensive modeling activities. To address this circumstance, several approaches have been adopted by former efforts: 1) adopting simplified computation methods (e.g., Li et al. 2015, Glawischnig

2016); 2) focusing on a smaller area of investigation (e.g., Beatens and Saelens 2015); 3) archetypical representation of the entire population based on statistical data (e.g., Sansregret and Millette 2009); and 4) employing reductive procedures. Most frequently reductive procedures such as sampling or archetyping are used to limit the number of necessary computations in proportion to the demandingness of the integrated energy assessments routines. However, loss of diversity is a natural consequence of the reductive methods. The severity of this loss depends on the efficiency of the building segmentation or classification schema, based on which the representative buildings are selected or generated. The most common approach towards stock segmentation relies on definition of critical thresholds for various energy-influential building characteristics such as envelope quality, geometry, operational parameters, etc.

Snäkin (2000) developed a non-dynamic numerical bottom up engineering model of the province of North Karelia in Eastern Finland. Through a stock segmentation by building usage, built form, construction/retrofit period, primary heat energy source, and heat distribution type, 4163 building types were identified. The study does not consider heat loads from solar energy and users and focuses on annual demand estimations.

Jones et al. (2001) analyzed data on heated ground floor area, façade, window to wall ratio, and exposed end area to identify 20 typical built forms for the assessment of the energy performance of the building stock of the city of Cardiff (UK). Along with 5 construction periods, this led to the definition of 100 building typologies. The buildings selected to represent these types, were subjected to assessments in a building performance benchmarking tool, UK Standard Assessment Procedure (SAP), which does not represent the temporal distribution of demand or capture the intricate effects of occupant presence and activity on energy demand. This stock energy model is a component of an integrative modeling environment aimed at facilitating urban level policy making, covering various aspects of the urban domain including traffic and industrial processes.

Hens et al. (2001) classified the Belgian building stock according to age, total floor area, built form, primary energy source and heating system type (central heating vs. dispersed heating units), to investigate the effectiveness of several CO2 emission reduction strategies. The envisaged CO2 reduction measures included shifting to low emission fuels, installation of heat pumps, conversion to renewable sources and improvements to the energy efficiency of buildings. The identified classes were represented by synthetic archetype buildings, which underwent a steady-state single zone monthly energy demand assessment procedure.

In a study on the energy demand of the Canadian building stock, Parekh (2005) segmented the building stock into classes of buildings with similar usage, vintage and climate region, generating 56 building types. Other building parameters required for energy assessments were statistically determined in every class. The resulting archetypical buildings were simulated in the HOT2000 energy assessment tool (Natural Resources Canada 2016), used to estimate annual energy consumption of low-rise residential buildings (single-family houses, semi-detached houses, and row houses).

The TABULA project (Episcope 2016), aiming towards development of residential building typologies for energy assessments across 11 European countries also relies on a building stock classification per climate zone, vintage and dwelling type. Real buildings representing the various characteristics of the buildings in each class are suggested as references, for performance computation purposes. The resulting typology has been used in various reductive urban energy assessment models (e.g., Dascalaki et al. 2011; Ballarini et al. 2014).

Huang and Broderick (2000) used a segmentation scheme by usage, vintage and location, developing a total of 120 commercial and 144 residential prototype/location combinations to characterize the US building stock. Simulation input files for the DOE-2 (Hirsch and Associates 2016) program were developed for each combination. Results were up-scaled and employed for cogeneration potential studies. Similar simulation-supported reductive approaches have been followed by Heiple and Sailor (2008), Caputo et al. (2013), Tuominen et al. (2014), and Orehounig et al (2014).

Most previous efforts adopt age and built form to express the thermal quality of the building envelope and its exposure to elements respectively. Operational parameters are captured through buildings' primary usage, which may be suboptimal when dealing with multi-usage buildings. Building orientation and transparency, as well as the effect of mutual shading are frequently ignored in various stock segmentation schemes. Despite adoption of reductive routines, very few studies focus on employment of more versatile dynamic performance evaluation procedures. Due to limitations in computational and informational resources, reduced order models are more frequently used. Although beneficial to provide an overview of the energy behavior of an urban domain, these simplified models are incapable of representing the temporal dynamics of energy load patterns and their dependency on transient phenomena (e.g., weather conditions, inhabitants' presence and actions) with appropriate detail. Consequently, specific queries, particularly those regarding the urban microclimate variance or occupant behavioral aspects, cannot be accommodated with suitable levels of resolution.

Sansregret and Millette (2009) developed software that automatically generates archetypical simulation models for the representation of the Quebecois building stock using building usage, floor area, construction period, location and main heating source as input. Relying on an extensive repository of building information data, for various characteristics of buildings probability distributions are statistically derived. These characteristics include aspect ratio, number of floors, thermal properties of envelope components, as well as operational parameters such as control settings and occupancy, lighting and appliance load schedules, and ventilation rates. The stock model is generated through stochastic determination of these variables according to the underlying probability distributions. As such, this method efficiently represents the diversity of the urban building stock, while minimizing modelling effort. However, a reliable spatial disaggregation of demand, necessary for proximity-dependent queries (e.g., pertaining to grid optimization and integration of distributed generation plants) cannot be expected since the method does not allow for the consideration of contextual parameters such as adjacency and shading relations.

In an article declaring Distributed Generation (DG) as the "future power paradigm", Manfren et al. (2011) count the "direct customer's involvement in energy demand and peak power reduction programs" among the manifold strengths of a DG scheme. The identification of potential sites through analysis of existing customers' distribution, energy demands and load patterns is, however, considered the preliminary phase towards efficient adoption of DG in urban areas. Investigation of the energy implications of behavior change scenarios, caused by demographic changes (e.g., ageing society, elevated birthrate, etc.), or new social developments (e.g., changes in official weekly work hours, increase in part-time employment) can enhance the utility of the urban energy predictive models.

Emphasizing the significance of the occupancy related variance in energy demand in view of the increasingly stringent thermal codes, Munoz and Peters (2014) question the efficiency of the reference operational schedules for assessments pertaining to DG schemes. They argue that even though social or behavioral diversity may not play a major role in the current centralized grids, development of decentralized energy generation paradigms requires reliable data on the dynamics of energy demand at a higher spatial resolution. To address this issue, they use the TABULA building typologies for Germany to characterize Hamburg's existing stock (Episcope 2016). Focusing on the residential buildings, they analyze micro-census data, defining household types and their occurrence likelihood, in each statistical area. These households are allocated to the prototypical buildings within the designated spatial domains, but the resulting configuration is assessed using a simplified standard heat balance model, which is ill suited to incorporate this elaborate representation of inhabitants towards performance assessment.

Baetens and Saelens (2015) identify various occupant types (full-time employed, unemployed, minor) for the Belgian context using time-use and household budget survey data. These typical profiles are then used as a basis for the stochastic modeling of occupant presence and activity schedules. The results are used to examine the uncertainties associated with user behavior in neighborhood scale energy assessments. The resulting model efficiently represents diversity of buildings and inhabitants, however, due to the extensive modelling effort required its applicability to larger building assemblies is limited.

In a demand model for the residential building sector in the city of Osaka, Japan, Shimoda et al. (2003) incorporate a detailed survey on household demographics and activities to determine 23 household types. The data provides the probability distribution of each living activity such as sleep, meal, work, etc. in 15-minute time intervals for weekdays, Saturdays and Sundays for each family member's category (classified by gender, age, and occupation: employed or not). With 20 dwelling typologies of detached and apartment houses (distinguished by size) and the defined household types, 460 building typologies were defined for which hourly energy consumption was simulated. The results were extrapolated to the entire city. The solid empirical basis allows for a detailed representation of the occupant-dependent aspects including heating, cooling, lighting and appliance use schedules. However, the physical aspects of buildings may have been over-simplified as the only criteria for the segmentation of dwelling types is area.

An overview of the various characteristics of the presented models is offered in Table 1. In this table the term “Authentic building representation” refers to modeling approaches, in which actual buildings from the study domain have been selected to represent the area, as opposed to “Synthetic building representation”, where archetypical buildings are adopted. These models address various aspects of urban energy computing. However, there is a requirement for versatile computational frameworks that can handle high resolution representations of occupants, and physical and contextual building parameters to support integrative urban decision support environments.

Table 1 An overview of the consulted models

Computation Method	Modeling Extent	Building Representation	Inhabitant Representation	Examples	Comments
simplified	Entire population	Authentic	Simple aggregate representation	Li et al. (2015) Glawischnig (2016)	• Low informational requirements • Low computational cost • Loss of diversity in case of reduced modeling extent • Inefficient representation of contextual parameters in case of reduced modeling extent • Low temporal resolution of results • Limited modeling and representation capabilities (in particular with regard to transient phenomena)
		Synthetic	Simple aggregate representation	Snäkin (2000) Hens et al (2001) Jones et al. (2001) Parekh (2005) Dascalaki et al. (2011) Ballarini et al. (2014)	
			Detailed diversity representation	Munoz and Peters (2014)	
Dynamic performance simulation	Entire population	Synthetic	Simple operational schedules	Sansregret and Millette (2009)	• High informational requirements • High computational cost • High temporal resolution of results • limited application domain due to reliance on a statistical approach towards model generation
		Authentic	Stochastic representation	Baetens and Saelens (2015)	
	Reduced	Authentic	Simple operational schedules	Orehounig et al. (2014)	• High temporal resolution of results • Loss of diversity • Inefficient representation of contextual parameters • Representation of diversity with regard to inhabitants
		Synthetic	Detailed diversity representation	Shimoda et al. (2003)	
			Simple operational schedules	Tuominen et al. (2014) Caputo et al. (2013) Huang and Broderick (2000) Heiple and Sailor (2008)	

3. Approach

3.1. Overview

Figure 1 illustrates the overall structure of the developed framework. The model relies on GIS information and standards to develop an energy relevant representation of the building stock. It employs well-known data-mining methods to effectively reduce the computational domain to a set of representative buildings, enabling the deployment of detailed multi-zonal dynamic performance simulations. The simulation models pertaining to these representatives are diversified with a two-fold method to recover part of the lost diversity, due to the reductive procedure as well as the utilization of average or typical operational representations (schedule-based average trends of occupant presence, use of appliances and lighting, typically provided in standards such as ASHRAE 2013). In this context, two distinct computational modules are developed: 1) The reductive module, and 2) The diversification module.

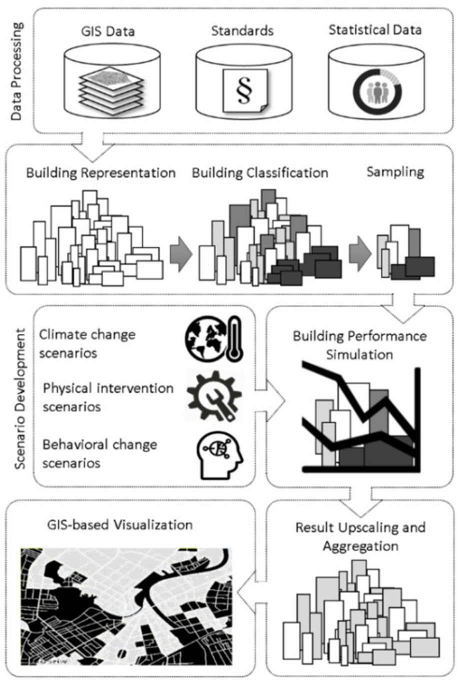


Figure 1 Structure of the proposed framework

Since the model reduces diversity in one step and enhances it in a second step, the authors adopted the term “hour-glass model” to describe the developed method. The present contribution provides a brief introduction to the methods adopted in the first module, followed by a more detailed description of the second. The utility of the model for comparative analysis of the energy implications of behavioral change scenarios is demonstrated through an illustrative example on a neighborhood in Vienna, Austria.

3.2. Case study

The case study is a neighborhood in the center of the city of Vienna covering a morphologically diverse area of about 1.3 square kilometers. The selected neighborhood, including some 750 buildings of various construction periods and usages, well represents the historical building stock of the Austrian capital. However, new Viennese buildings (constructed after 1945) are underrepresented due to their low count in the central districts. A less central location may have better captured the age diversity of the stock, however, due to their architectural and historical quality, the central districts have been better documented in the official GIS data. Figure 2 and Figure 3 illustrate the selected neighborhood and the distribution of buildings by construction period and primary usage. Buildings with uncommon usages, such as kiosks in a permanent market place, underground station entrances and a church were eliminated from the analysis.

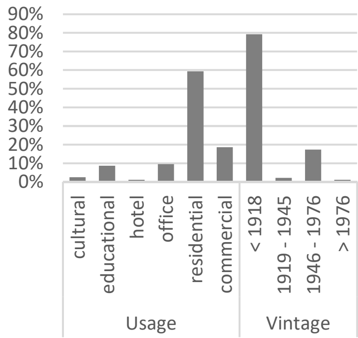


Figure 2 (Left) Selected case study

Figure 3 (Right) Distribution of buildings in the study area by usage and age

4. The reductive module

4.1. Tools and prerequisites

The first module is developed as a plug-in for the open-source QGIS platform (2016). QGIS is a Geographical Information Systems (GIS) environment based on the programming language Python (2015), enhancing it with specialized packages for performing geometrical analyses on GIS data. The plug-in, also developed in Python language following an object-oriented approach, is embedded with functions from the R Program for Statistical Computing (2015), which enable the implementation of intricate data-mining routines. The current version of the plug-in is adapted to the Viennese context. This adaptation pertains to the format and the content of the available GIS data, as well as the incorporated codes and statistical information. However, the method can be applied to other geographical locations with minor modifications to the code, provided that the necessary input data can be procured. This data includes land-use vector data of the urban area under study containing the footprint geometry of the buildings and eaves height information, Digital elevation raster data containing punctual values for building height, and Geo-referenced information on construction period, building usage, and number of floors.

The above data can be utilized to generate a Sky View Factor (SVF) raster imagery of the area using the DEMTools QGIS plug-in (Hammerberg 2014). This data has been used for solar exposure analyses. Optionally, the official land-use data can be refined using other available data sources, such as the crowd-sourced Open Street Maps (2015) GIS data repository. Austrian standards providing use-profiles, age-based thermal performance of building components, etc. have been used to represent various characteristics of the building stock (Table 2).

Table 2 Incorporated urban data

	Source	Data type	Contained data
GIS DATA	ViennaGIS 2015	Land Use Plan (vector layer)	Building Footprint Polygons Construction type (main/annex) Relative eaves height Elevation from ground
		Digital Elevation Model raster layer)	Relative height of every point
		Building Inventory (vector layer)	Building construction period/year Main building usage Number of Floors
	OpenStreetMap 2015	Land Use Plan (vector layer)	Building Footprint Outline Building usage
	Hammerberg 2014	Sky View Factor Map (raster layer)	Sky View Factor of every point at ground level
STANDARDS	OeNorm B8110-5 2011	Thermal insulation in building construction: Model of climate and user profiles	Usage-based internal gains Usage-based infiltration rate Usage-based use hours Reference weather data
	OeNorm B8110-6 2014	Thermal insulation in building construction: Principles and verification methods, Heating demand and cooling demand	Average window to wall ratio Average frame to window ratio Average net to gross floor area ratio
	OIB-RL 6 2015	Guidelines: Energy-technical behavior of buildings	Age-based component U-values Age-based window solar transmittance

4.2. Data processing

In the official GIS data of Vienna, buildings are identified with a unique reference number, which also establishes a link between the land-use data and the geo-referenced building inventory data. The developed plug-in browses through the land-use plan identifying features with a unique reference number as “Building Parts” constituting the footprint geometry of a single “Building” object. The construction type information is used to determine if the feature is associated with a main building or an annex utility building. The latter buildings are considered unconditioned. The eaves height information associated with every “Building Part” is used to generate “Wall” objects. “Buildings Parts” are examined for touching perimeter segments. Heights of the “Wall” objects associated with these segments are compared to determine the parts of the colliding walls, which form the outer boundaries of the building envelope. These “Walls” are maintained in the “Building” object. The aggregated area of the lowermost envelope components as well as the enclosed volume is computed. Through geometric analysis of footprint outlines a list of neighbors is identified for each building. The “Walls” are investigated for adjacency with neighboring buildings, determining their boundary condition (outside, adiabatic, or adjacent to unheated building). For “Walls” adjacent to outside air, angle to north, total area, as well as the glazed area are computed, using the average standard-suggested window to wall ratio. The Sky View Factor (SVF) for a point in the middle of the wall’s baseline is extracted from the SVF raster layer as an indicator of the shading effect of the surrounding buildings. A simple rule-based logic based on the difference between the average height of every “Building Part” calculated from the Digital Elevation Model, and the eaves height given by the land-use layer, determines the shape of the roof (flat/sloped) and the condition of the attic space in case of a sloped roof. If the attic space is assumed to be used, an approximation for the additional conditioned volume and the roof area is made.

The building inventory data point associated with the building is identified by the reference number, thus providing information on the main usage of the building and its construction period. If Open Street Map data is available, it is browsed for data points associated with each building, for additional information on other functionalities present in the building (e.g., banks, shops, restaurants, etc.). Based on simple rules, the entire volume of the building is distributed among the various usages. Operational parameters such as annual day-time and night-time use hours, daily HVAC operation hours, average area-related hourly internal gains and air change rate is extracted from the standards for the present usages. Average values of the operational parameters, weighted by their designated share of the total volume are calculated for the building. According to the construction period/year, U-values of various building components are extracted from the national guidelines. The resulting representation of the urban area is then used to generate a matrix of energy-influential descriptive indicators, which are used to partition the building stock into groups of buildings with similar thermal behavior. Figure 4 summarizes the data processing routine.

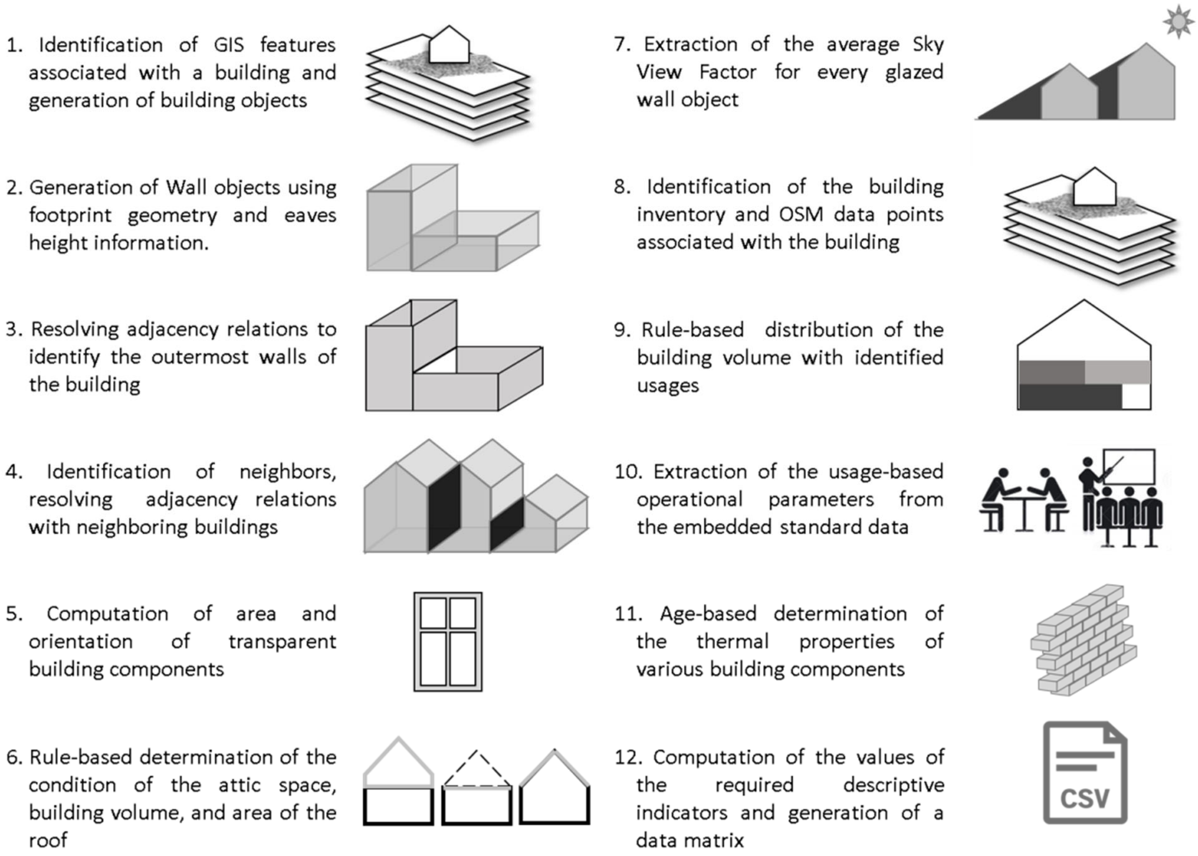


Figure 4 Overview of the steps of the data processing routine

4.3. Stock classification

Rather than relying on construction period, usage and built form as classification criteria, a set of descriptive indicators were defined to capture the various energy-relevant aspects of the buildings, including physical, operational and contextual characteristics. This allows for the development of a more generic approach towards stock classification in view of the dynamic nature of cities. Table 3 provides an overview of the adopted indicators as well as the corresponding computational methods. Note that the contextual parameters (the reductive impact of mutual shading on solar gains, as well as the influence of adjacencies on conductive heat loss) have been involved in the definition of several indicators (thermal compactness, effective glazing ratio, and effective envelope U-value). For every building, the values of the above indicators are calculated by the plug-in. The results are stored in a CSV file, in which every building is represented by a row, and the columns contain the values of the descriptive indicators. The resulting data matrix is subjected to Multivariate Cluster Analysis (MCA), a well-known data-mining method for unsupervised data classification.

There are major differences in the magnitude of the values of these indicators. For instance, the values associated with net volume are much greater than those expressing the effective U-value of the envelope. Prior to cluster analysis, the dataset is standardized to prevent the differences in the magnitude of values, from influencing the clustering outcome as an unintentional weighting. The performance of three clustering algorithms, K-means clustering (MacQueen 1967), Hierarchical agglomerative clustering (Hair et al. 2010), and Model based clustering (Fraley and Raftery 2002) towards efficient partitioning of the data space was evaluated. K-means clustering partitions the data space into k clusters, such that the data points in each cluster are closer to the center of their own cluster than to the centers of any other clusters, based on the Euclidean distance.

Table 3 Adopted classification criteria

	Abbr.	Variable	Description/ Comments	Formula	Parameters	
Geometry	V_n	Net Volume [m^3]	An indicator of the size of the building	$V_n = \sum(A_{feat, i} \cdot h_{feat, i}) \cdot f_n$	$A_{feat, i}$ $h_{feat, i}$ f_n	Area of footprint feature [m] Height of foot print feature [m] Net to gross volume ratio
	h_e	Effective floor height [m]	Ratio of the building volume to the floor area	$h_e = V_n / (A_{f, i} \cdot n_f)$	$A_{f, i}$ n_f	Total floor area [m] Number of floors
	C_t	Thermal compactness [m]	Ratio of the net building volume to the thermally effective envelope area	$C_t = V_n / A_e$	A_e	the thermally effective envelope area [m]
Solar gains	GR_e	Effective glazing ratio	Average glazing to wall ratio weighted by orientation and corrected for the shading effect of the surroundings Weights associated with orientations were based on reference climate data	$GR_e =$ $WWR \,.\, GWR \,.\, g \,.\, \sum(A_{ow, i} \cdot f_{o, i} \,.\, SVF_i) /$ $\sum A_{ow, i}$	WWR GWR $A_{ow, i}$ $f_{o, i}$ g SVF_i	Window to wall ratio Glass to window ratio Area of outside wall [m] Corresponding orientation correction factor Solar factor of glazing Sky View Factor in the vicinity of the wall
Thermal Quality	U_e	Effective average envelope U-value [W.m^{-2}.K^{-1}]	Average U-value of the envelope corrected for adjacency relations and weighted by the corresponding areas	$U_e = \sum(U_i \,.\, A_i \cdot f_{t, i}) / A_e$	U_i A_i $f_{t, i}$	U-value of building element [W.m^{-2}.K^{-1}] Area of building element [m] Corresponding temperature correction factor
Operation Parameters	O_u	Fraction of the year used	Fraction of time the building is used annually	$O_u = t_{use, a} / t_a$	$t_{use, a}$ t_a	Annual use hours [h] Total hours in a year[h]
	Ig_d	Daily area related internal gains [Wh.m^{-2}.d^{-1}]	Daily internal heat gains per unit of area during the heating season	$Ig_d = (q_{i, h} \cdot t_{use, d})$	$q_{i, hd}$ $t_{use,}$	Usage-based internal gains rate [W.m^{-2}] Daily use hours [h]
	Ac_d	Daily air-change rate [d^{-1}]	Daily air-change rate	$Ac_d = n_v \,.\, t_{use, d}$	n_v	Usage-based hourly air-change rate [h^{-1}]

In Hierarchical agglomerative clustering, each data point forms a cluster with a single member. The closest clusters are consecutively joined until only one cluster remains. The steps are traced back to arrive at the desired number of clusters. In this project, the Ward's method with squared Euclidean distance (Ward 1963) has been used as the linkage criterion. Model-based clustering recasts the problem of partitioning the data space as a statistical model choice problem. The set of multivariate Gaussian components best approximating the data space are identified, each component representing a cluster (For more detail on the adopted clustering methods see Ghiassi and Mahdavi 2016). The K-means and Hierarchical agglomerative methods require the number of the desired clusters as a prerequisite to the partitioning. To solve this problem, the nbClust package for R (Charrad et al. 2014) was utilized. This package receives a range of values for the number of clusters, computes over 20 clustering performance indicators for the resulting clustering schemas, and delivers the optimal number of clusters by majority vote. Note that each cluster has to be represented by at least one building. Therefore, the upper limit of the range of cluster numbers considered depends on the available computational and informational resources. In this project a range of 6 to 30 clusters was considered at each run. The center of every cluster is determined as a virtual data point, the dimensions of which assume mean values of the indicators across the cluster. The building closest to this virtual center is selected as the cluster representative.

Obviously, there is no unique or best set of parameters that can cover various energy-relevant building characteristics. For instance, thermal quality of the envelope can be expressed with a lower level of aggregation through effective wall, ceiling and floor U-values. Operational parameters can include more detail, such as day-time and night-time use fractions. Initially a larger set of variables was considered, which in addition to the variables described in Table 3 included thermally effective envelope area, effective wall, ceiling/roof, and floor U-values, daytime use intensity, annual daytime and nighttime use fractions, average area-related internal gains and average hourly air change rate. Eleven scenarios with regard to the clustering criteria were developed. These scenarios involved between 7 to 12 indicators for the representation of the energy influential building characteristics with various resolutions (e.g., average envelope U-value versus effective U-values of major building components to represent constructions). The combination of these 11 scenarios (with regard to input variables) and 3 clustering algorithms resulted in 33 clustering schemas. A simplified evaluation process was adopted to identify the clustering schema that best represented the energy diversity of the neighborhood.

For this purpose, a simplified annual heating demand calculation was carried out on the entire data-set using the previously generated building data representation and Standard Austrian Weather data (ÖNORM 2011). In a second step, for every clustering schema, the volume-related heating demand of the cluster representatives and the net volume of the buildings in the cluster were used to approximate the heating demand of the buildings. The clustering-based predicted values were compared to the originally calculated demand to identify the clustering schema that resulted in the smallest errors in the prediction of the aggregated demand of the neighborhood as well as the smallest average deviation of the disaggregated (building level) results. The evaluation led to the selection of the aforementioned set of indicators (described in Table 3) with the k-means clustering algorithm, which were integrated into the plug-in. The resulting clustering schema involves 7 clusters. It predicts the aggregated heating demand of the neighborhood with a relative error below 1%. Mean relative error in the prediction of building level demand is 11%, with 12% of the buildings featuring a prediction error of above 20%. Of course, the implemented evaluation can only roughly indicate the performance of the model in the prediction of annual demand. The quality of the sample in representing the temporal diversity of load patterns must ultimately be evaluated by means of high resolution performance assessments or empirical data.

5. The diversification module

5.1. Tools and prerequisites

The simulation program EnergyPlus (2016) was used for the generation of the reference buildings. Multi-zonal simulation models for the selected sample were generated based on the detailed drawings procured from the Viennese municipality. For various usages present in reference buildings, occupancy, HVAC, lighting, and equipment schedules provided by ASHRAE (2013) (for Weekdays, Saturdays and Sundays) were adopted as reference schedules. The non-residential schedules were readjusted such that the overall hours of HVAC operation matched the average daily operation hours provided by the Austrian Standards (ÖNORM 2014). No internal gains were assumed for unconditioned zones such as corridors, basements and unheated attic spaces.

Layered building constructions were defined according to the common practice of the period of construction. For this purpose, base case assumptions of the Austrian Handbook for Building Thermal Retrofit (Schöberl et al. 2012) were used. The material properties were adjusted such that the component U-values matched those suggested by national guidelines for performance assessment purposes of historical buildings (OIB 2015). In case of newer buildings, the details available on the plans were used to define constructions, while matching the U-values with the standard guidelines.

All required information for Energy Plus simulations is structured and stored in a text-based object-oriented Input Data File (IDF). The Eppy (2016) package for Python offers the possibility to browse and modify IDF objects through Python codes. This package was utilized to automate the diversification process.

The csv file of the descriptive indicators matrix generated by the first module is used as a basis for diversification. More specifically, the following parameters calculated in the previous phase (some of which are not involved in the current classification scheme) are utilized:

- Effective U-values of the external wall, uppermost enclosure (Floor of the unheated attic space or roof of the heated attic space), and basement ceiling. These values are calculated according to the following equation (Equation 1):

$$U_{e,c} = U_c \,.\, f_c \qquad (1)$$

$U_{e,c}$: Effective U-value of the construction [W.m^{-2}.K^{-1}]

U_c : U-value of the construction as suggested by the standard [W.m^{-2}.K^{-1}]

f_c : Share of the construction from the effective envelope area given by Equation 2:

$$f_c = \sum(A_{c,i} \,.\, f_{t,c,i}) \;/\; \sum(A_j \,.\, f_{t,j}) \qquad (2)$$

$A_{c,i}$, $f_{t,c,i}$: Area and corresponding temperature correction factor of an element associated with the particular construction

A_j , $f_{t,j}$: Area and corresponding temperature correction factor of an arbitrary building element

- Daily internal gains (see Table 3)
- Daily air change rate (see Table 3)

5.2. Computational logic overview

As suggested by the simple evaluation performed in the previous phase, the selected buildings provide a fair overall representation of the energy demand of the neighborhood. However, much of the diversity of the building stock has been lost in this representation. This loss of diversity is in part due to the reductive process, and in part attributable to the use of reference schedules for the representation of the operational parameters. The idea behind the diversification module is to use the developed reference simulation models as a basis for generation of a diverse set of models that better reflect the various characteristics of the building stock.

Acquisition of relevant building information and generation of the geometric model of a building is the most time and effort intensive activity in building performance simulation (Mahdavi and El-Bellahy 2005). Assuming that the identified sample of buildings represents the geometric features of the stock with acceptable fidelity, the diversification module readjusts the non-geometric parameters of the reference simulation models to recapture part of the lost diversity. The building parameters currently subjected to diversification are the thermal properties of the main components of the building envelope (uppermost and lowermost enclosures, external walls), number of occupants, area-related equipment and lighting power, as well as occupancy, activity level (metabolic rate), lighting, and equipment schedules. Diversification of some geometry-dependent aspects (such as effective glazing ratio) is part of the intended future research.

For every building in the neighborhood, a duplicate of the reference simulation model pertaining to the relevant representative building is created. In the new model, the values of the above-mentioned parameters are modified according to the information available on the target (represented) building (acquired by the first module). This procedure results in the generation of a unique simulation file associated with a unique set of schedules for every building in the study domain. These models share all geometric properties of the reference models, but emulate the thermal and operational properties of the target buildings more closely. The resulting simulation files are batch processed from the Energy Plus launcher. Hourly simulation results for volume-related energy loads of these diversified models are used to arrive at the hourly demand of the target buildings according to the following equation (Equation 3):

$$Q_{i,h} = Q_{reference,h} / V_{reference} \cdot V_i \quad (3)$$

$Q_{i,h}$: Heating demand of building *i* in time step *h* (kWh)

V_i : Net volume of building *i* (m^3)

$Q_{sim,h}$: Heating demand of the simulated model in time step *h* (kWh)

$V_{reference}$: Net volume of the reference building based on which the simulation model was generated (m^3)

5.3. Capturing the behavioral diversity

The reference schedules represent the temporal distribution of internal loads in aggregate terms. However, use of average profiles for demand assessments will result in identical peak load hours and unrealistically monotonous profiles across building classes. In order to stochastically represent occupancy-related factors, for each building based on the reference schedules for various days of the week, a set of randomized schedules were created and stored as schedule files compatible with Energy Plus requirements. Each schedule file has 8760 rows (for every hour of the year), and 5 columns corresponding to occupants' presence, lighting use, equipment use, HVAC operation and activity level. The first three variables assume real values in the range of 0 to 1, and the fourth is a Boolean variable (0 or 1). HVAC schedules are not diversified. In case of residential and gastronomy usages, schedules for occupancy, activity level, lighting and equipment use were diversified separately. For this purpose, for every hour of the year, the value suggested by the reference schedule as the mean value of the time step, as well as a default coefficient of variance (CV) were used to generate a probability distribution. The identification of the specifically appropriate CV value is an open research question. Former studies suggest that for certain applications (e.g., the stochastic generation of presence patterns), CV might display a distinct value range (Mahdavi and Tahmasebi 2015). Accordingly, a CV value of 0.2 was deployed in the present study. A random value was then generated based on this distribution and assigned to the time step. Rules were set to ensure that the selected value remained within the acceptable range. In case of office spaces, where equipment and lighting use are very strongly correlated to occupancy presence, the rate of lighting and equipment use in every time step of the reference schedules was expressed as a function of the occupancy rate in addition to a minimum value. In this case, occupancy schedule was subjected to diversification. Equipment and lighting rates were computed based on the occupants' presence rate at every time step. The same method was applied to generate diverse activity level schedules. Natural ventilation was assumed to follow the occupancy schedule. The diversification code links the IDF file associated with each building to a unique set of diversified schedules. Figure 5 illustrate the standard schedules as well as examples of the data generated for the gastronomy and office usages for two buildings for the second week of January.

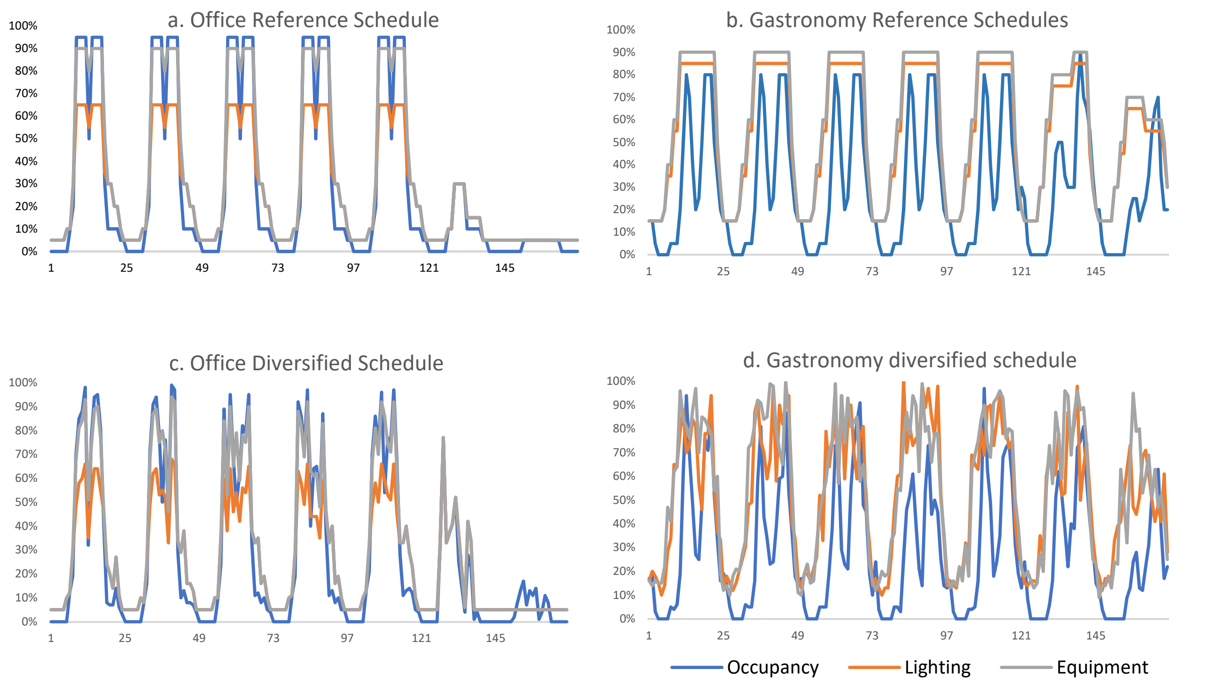


Figure 5 a: Office reference schedules according to ASHRAE (2013); b: Office reference schedules according to ASHRAE (2013); c: A week's data of the diversified schedules generated for the office zones of one building; d: A week's data of the diversified schedules generated for the gastronomy zones of one building

5.4. Readjustment of internal gains and ventilation rates

Once the schedules were generated, the reference values for equipment and lighting power, number of occupants, as well as hourly air change rates were computed such that the aggregated internal gains and ventilation rates matched the values computed for each building in the first phase. For this purpose, the annual internal gains were computed based on the average daily values (see Table 3) and the use days provided by the standard. Based on the same logic, average hourly air change rate across the year was determined for each building. The annual value of area-related internal gains was disaggregated into people, lighting and equipment gains using average ratios provided by literature. For instance, for residential spaces, 58%, 19%, and 23% were assumed for gains pertaining to equipment, lighting and people respectively (Kemna and Moreno Acedo 2014). For every time-step, Energy Plus computes the values of lighting and equipment gains by multiplying the applicable usage rate (given by the schedule) by the reference power value. The following equations (Equation 4 and Equation 5) deliver this reference value (in this case for lighting):

$$\mathrm{Power}_{\mathrm{Lighting}} = \mathrm{Q}_{\mathrm{lighitng,a}} / \mathrm{FLH}_{\mathrm{Lighting}} \quad (4)$$

$$\mathrm{FLH}_{\mathrm{Lighting}} = \sum(\mathrm{HR}_{\mathrm{Lighiting}}) \quad (5)$$

$\mathrm{Power}_{\mathrm{Lighting}}$: reference area-related lighting power [W.m^{-2}]

$\mathrm{Q}_{\mathrm{lighitng,a}}$: Annual internal gains from lighting [Wh. m^{-2}.a^{-1}]

$\mathrm{FLH}_{\mathrm{Lighting}}$: Aggregated annual full load hours of lighting [h. a^{-1}]

$\mathrm{HR}_{\mathrm{Lighiting}}$: Hourly rate of lighting use provided by the schedule

The same logic was applied to acquire the reference air change rates, such that the annual average matches the expected value. In the case of occupants, the hourly gains depend not only on hourly occupancy presence rates, but also on the metabolic rate or activity level assumed for the time step. In this case, the following equation is applied (Equation 6):

$$\mathrm{Number}_{\mathrm{People}} = \mathrm{Q}_{\mathrm{People,a}} / \sum(\mathrm{HR}_{\mathrm{People}} \,.\, \mathrm{H}_{\mathrm{ActivityLevel}}) \quad (6)$$

$\mathrm{Number}_{\mathrm{People}}$: maximum area-related number of occupants [Person.m^{-2}]

$\mathrm{Q}_{\mathrm{People,a}}$: Annual internal gains from people [Wh. m^{-2}.a^{-1}]

$\mathrm{HR}_{\mathrm{People}}$: Hourly presence rate of occupants provided by the schedule

$\mathrm{H}_{\mathrm{ActivityLevel}}$: Hourly activity level provided by the schedule [W. Person^{-1}]

5.5. Readjustment of thermal properties

The diversification of the thermal properties of the envelope relies on the diversity of effective element U-values. The effective element U-value is not only a measure of the thermal quality of the element's construction, but also a function of the significance of the said construction in the overall thermal performance of the building. This significance is determined by the share of the elements associated with a particular construction in the total thermally effective area of the envelope (corrected for adjacencies). The diversification process modifies the pertinent constructions in such a way as to replicate the effective construction U-values computed for every principle building element in the first step (see Equation 1). Since the geometry of the diversified model is identical to that of the reference model, any deviations from the effective U-values of the reference building must be accounted for by manipulating the U-values of the constructions in the new model. For instance, if the walls of a building have a more significant share in the building's effective envelope than is the case for the reference building, and both buildings have the same effective wall U-value, the wall element of the diversified model will assume a higher U-value, such that the effective U-value of the wall remains the same.

As mentioned before, for every reference simulation file, building component constructions were defined according to the common practice of the period of construction of the reference building. In every relevant construction, only the most thermally effective layer was subjected to modifications. (e.g., the massive masonry layer, or the ceiling timber in the case of historical buildings, and the insulation layer in case of newer buildings). Since a modification of the thermal mass of the building was not intended, only the thermal conductivity of the layer was changed to reach the target construction U-value. Rules have been applied to prevent the layer from assuming unreasonably large or negative values. This may happen in rare cases where the effective U-value of the components vary significantly from those of the reference building.

6. Modeled scenarios

6.1. Diversification scenarios

To investigate the impact of the diversification module on the model outcomes, three scenarios were deployed. In the first run, the hourly simulation results for volume-related energy loads of the non-diversified reference buildings with standard assumptions for thermal quality of components, schedules, area-related internal gains, and air change rate were used to arrive at the hourly demand of the buildings in the neighborhood. In a second step, only the schedules were diversified. The third case involved all the diversification steps described in the previous chapter (Table 4).

Table 4 Overview of the diversification scenarios

Abbr.	Schedules	Thermal properties	Internal gains	Number of simulations
NDS	Not diversified	Not diversified	Not diversified	7
DS-S	Diversified	Not diversified	Not diversified	744
DS-A	Diversified	Diversified	Diversified	744

6.2. Illustrative behavior change scenarios

To demonstrate the utilities of the developed computational method for investigation of behavioral change scenarios, the fully diversified model was considered as the reference point. Then, two simple scenarios with different assumptions on occupant behavior with regard to HVAC system temperature set points were defined, simulated, and compared to the reference values. The first scenario considers a setback value for non-residential spaces, which represents more realistically the general operation patterns of these buildings. The second scenario maintains the setback values, but seeks to emulate the behavior of a more energy aware population. It is based on the assumption that occupant presence rates correlate with the fraction of the spaces conditioned. In other words, when occupant presence rate is low, some spaces are unoccupied and thus unheated. Therefore, the average indoor temperatures in the building drop below the standard assumptions. Note that the developed behavior change scenarios are intended as illustrative examples to demonstrate the modeling possibilities of the developed framework and are not presumed to realistically capture the behavior of inhabitants. Table 5 describes the rules defining the HVAC operation in each scenario.

Table 5 Overview of the base case and scenario assumptions for HVAC operation

<table>
<tr><td></td><td></td><td colspan="2">Residential</td><td colspan="2">Office</td><td colspan="2">Gastronomy</td></tr>
<tr><td rowspan="2">DS-A (Base Case)</td><td>Set point assumptions [°C]</td><td colspan="2">20</td><td colspan="2">20</td><td colspan="2">20</td></tr>
<tr><td>HVAC Availability</td><td colspan="2">24 hours a day</td><td colspan="2">14 hours on weekdays</td><td colspan="2">14 hours a day</td></tr>
<tr><td rowspan="2">BS-1</td><td>Set point assumptions [°C]</td><td colspan="2">20</td><td colspan="2">20 during work hours
14 at other times</td><td colspan="2">20 during work hours
14 at other times</td></tr>
<tr><td>HVAC Availability</td><td colspan="2">24 hours a day</td><td colspan="2">24 hours a day</td><td colspan="2">24 hours a day</td></tr>
<tr><td rowspan="2">BS-2</td><td>Set point assumptions [°C]</td><td>16
16
20
Interpolate</td><td>Night hours
Occupancy rate <25%
Occupancy rate > 55%
Other times</td><td>14
16
20
Interpolate</td><td>Not working hours
occupancy rate <25%
Occupancy rate > 75%
Other times</td><td>14
16
20
Interpolate</td><td>Not working hours
occupancy rate <25%
Occupancy rate > 75%
Other times</td></tr>
<tr><td>HVAC Availability</td><td colspan="2">24 hours a day</td><td colspan="2">24 hours a day</td><td colspan="2">24 hours a day</td></tr>
</table>

7. Results and discussion

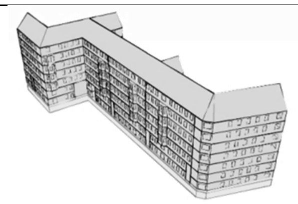
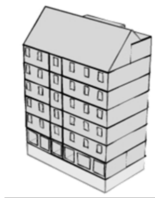
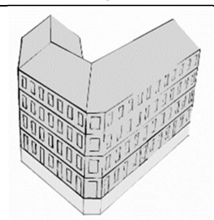
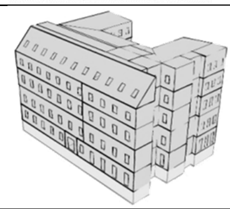
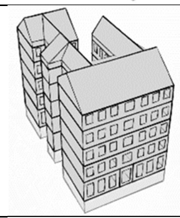
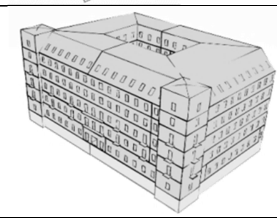
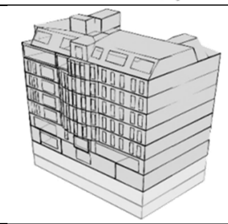
7.1. Neighborhood representation

The final outcome of the operation of the reductive module plug-in is a map of the investigated urban area, in which clusters are identified by colors (Figure 6), the CSV file containing the values of all descriptive indicators, as well as the list of buildings representing the clusters. On an average commercial PC, the application of the reductive plug-in on the current case study requires about 15 minutes.

Figure 6 The GIS maps of the neighborhood featuring the clustering schema, generated by reductive plug-in. Buildings belonging to various clusters are identified by colors.

Table 6 describes the particular characteristics of each cluster and the associated representative building. Three different usages are present in the sample: Residential, Gastronomy, and Office.

Table 6 Description of clusters and representative buildings

	Cluster Description	Representative Description	
Cluster 1	Prominently sized buildings of mainly residential use with high ceilings, and an L-shaped or U-shaped foot print resulting in high exposure to outside air Number of buildings in cluster: 125	Construction Year: 1914 Usages: Residential + Gastronomy	
Cluster 2	Residential buildings of varied sizes, less affected by shading. Most post 1945 buildings are grouped here Number of buildings in cluster: 94	Construction Period: After 1945 Usage: Residential	
Cluster 3	A collection of medium sized residential buildings with higher ceilings Number of buildings in cluster: 201	Construction Year: 1846 Usages: Residential + Gastronomy	
Cluster 4	A mix of mainly educational, cultural, commercial and office buildings of medium size Number of buildings in cluster: 179	Construction Year: 1868 Usage: Office	
Cluster 5	Smaller sized residential buildings with low compactness Number of buildings in cluster: 109	Construction Period: 1848-1918 Usage: Residential	
Cluster 6	A collection of the largest non-residential buildings on site, with high exposure to elements Number of buildings in cluster: 30	Construction Year: 1872 Usages: Office	
Cluster 7	A cluster of all buildings constructed after 1976, whose thermal performance is significantly superior to that of all other classes of building Number of buildings in cluster: 6	Construction Year: 2002 Usages: Residential	

7.2. Diversification scenarios

Table 7 summarizes the results of the modeled cases. As seen in the table, both levels of diversification result in minor changes in the model outcome in terms of the aggregate energy demand of the neighborhood, as well as annual peaks. The use of diversified schedules alone barely modifies the results at all. The cumulative effect of the diversified schedules resembles closely, that of the impact of the average schedule, which was to be expected. The physical and operational adjustments introduced in scenario DS-A, result in a slightly more substantial, yet still not so severe change in the predicted value of annual demand. Since the aggregate results remain rather close to the NDS scenario, the overall integrity and representativeness of the sampling model has not been compromised by the diversification process. This being said, the ultimate reliability of the predictions with regard to overall demand can only be assessed and validated based on real data, currently unavailable. The generation of diversified simulation files for all buildings in the case study, on an average PC, requires about an hour.

Table 7 Summary of the results of modeled scenarios

	Maximum hourly load [MWh]	Relative deviation from reference scenario [%]	mean hourly demand [MWh]	Standard deviation	Total annual space heating load [MWh]	Relative deviation from reference scenario [%]
NDS	153.13	0	22.64	26.83	198354.47	0
DS-S	154.84	1.11	22.4	26.34	196406.17	-0.98
DS-A	151.39	-1.13	21.88	25.97	191659.24	-3.38

At the level of the clusters, the deviations from the NDS are more substantial in DS-A (Table 8). The non-residential building clusters are more sensitive to schedule changes.

Table 8 Deviations in the annual demand of clusters as predicted by diversified models from the non-diversified one

		Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
NDS	Aggregated annual heating demand [MWh]	45584.23	22397.46	38435.705	41929.35	16690.24	32003.70	1313.78
DS-S		45489.35	22343.98	38334.817	41001.84	16663.19	31264.58	1308.41
DS-A		43109.15	25443.04	32301.631	37494.41	17984.79	33739.19	1587.04
DS-S	Relative deviation from NDS [%]	0	0	0	-2	0	-2	0
DS-A		-5	14	-16	-11	8	5	21

The percentage deviation in the volume-related annual heating demand of the buildings acquired from the DS-A, from the reference (non-diversified) values are shown in Figure 7. For the majority of the buildings, the DS-A volume-related demand remains within 20% of the reference value.

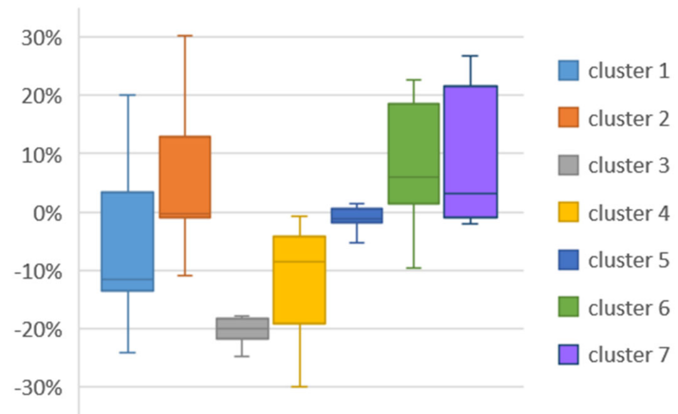


Figure 7 Percentage deviation of the volume-related annual demand across clusters in DS-A from NDS (outliers are excluded from this representation)

Figures 8 displays the diversity caused by the two diversification scenarios in the hourly heating demand values of the buildings in a single time step (11 am, January 8th). In this graph the percentage deviation from the hourly demand value predicted by the non-diversified model has been shown. The two diversification scenarios, DS-S and DS-A, result in 2% and 8% average deviation from the hourly predictions of the non-diversified model respectively. In the design and deployment of DG systems, and net-zero energy neighborhoods, due to the smaller spatial scale of considerations, these variations can be of great significance.

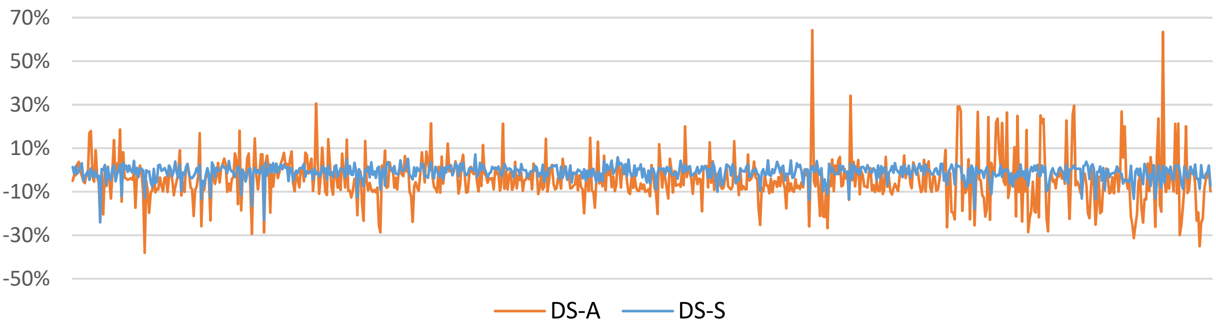


Figure 8 Relative deviation of hourly demand results of all buildings as predicted by the DS-A and DS-S from NDS predictions for a single time step (11 am, January 8th)

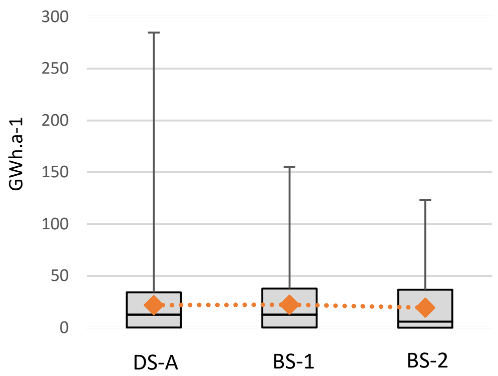
7.3. Behavior change scenarios

The results of the investigated illustrative behavior change scenarios are summarized in Table 9. The consideration of a setback value for the operation of non-residential spaces leads to a minor increase (2%) in the overall energy demand of the neighborhood. However, it leads to effective reduction of (annual) peak load by 18%. Lower peak loads are advantageous for grid stability particularly in Distributed Generation energy systems. The occupancy sensitive control of heating set points during daytime in all buildings and the reduction of set points during the night in residential buildings can lead to an 11% decrease in the overall energy required for space heating. Peak loads are reduced by an impressive 26%.

Table 9 Summary of the results of behavior change scenarios

	Maximum hourly load [MWh]	Relative deviation from Base case [%]	mean hourly demand [MWh]	Standard deviation	Total annual space heating load [MWh]	Relative deviation from Base case [%]
DS-A	151.39	0	21.88	25.97	191659.24	0
BS-1	124.48	-18	22.29	25.41	195220.55	2
BS-2	111.71	-26	19.44	24.39	170303.13	-11

Figure 9 displays the range of hourly loads (of the entire domain) for various scenarios in terms of boxplots. In Figure 10, the cumulative frequency graph for hourly demand of the neighborhood in all scenarios has been plotted. Note that even though BS-1 and BS-2 result in decreases in the magnitude of the peak load, they increase the frequency of hours with higher demands. Such information can support an efficient deployment of supply side energy management strategies.



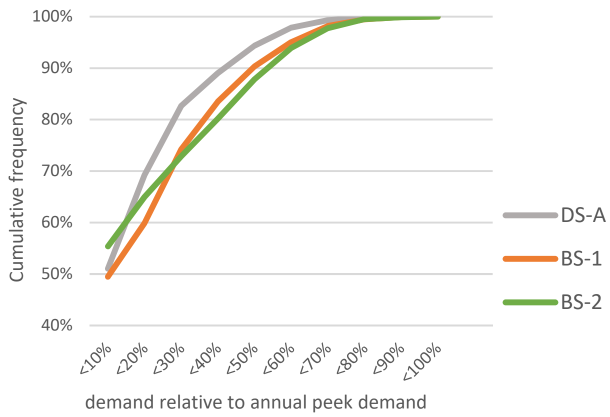


Figure 9 (Left) Comparative analysis of the hourly demand predicted in various scenarios

Figure 10 (Right) Cumulative relative frequency of the hourly heating load in relation to annual peak load in various senarios

8. Conclusion and future research intentions

Addressing the requirement for more versatile urban energy decision support environments, we have reported on the development of a two-step hourglass energy computation model based on detailed performance simulation. The first component of the model is tasked with the reduction of the computational domain, to enable the utilization of dynamic performance simulation for urban level energy assessments. This is done through an automated GIS-based sampling schema, enhanced with Multivariate Cluster Analysis. The second component incorporates stochastic methods as well as the available large scale building data to reintroduce some of the lost diversity back into the model. It generates permutations of the representative simulation models, which emulate the specific properties of the represented buildings more closely. To demonstrate the modeling potential of the developed framework, it has been deployed on a case study to investigate a number of diversification and illustrative behavior change scenarios. In the diversification of the reference schedules, due to the nature of the stochastic method adopted, in aggregate terms, the tendencies of the original standard schedules are maintained. However, the diversification results in a more realistic representation of the temporal as well as spatial distribution of energy demand. The data-oriented diversification of internal gains and thermal properties leads to more significant changes in the overall demand predictions of the model. However, it is expected to improve the performance of the model for comparative analysis of the energy impact of various scenarios due to the better representation of urban stock diversity.

Future research intentions include exploring the potential of diversifying other building characteristics such as solar exposure. Also, the utility of the diversification process for calibration of reductive urban energy demand models based on detailed monitored data from selected buildings will be investigated. The relative reliability of the sampling schema in view of the representation of bulk annual demand has been ascertained through simplified evaluations. However, the performance of the reference buildings in representing the temporal demand tendencies as well the predictive performance of the diversified model require detailed analyses based on monitored demand data currently unavailable.

References

ASHRAE. (2013). Standard 90.1: Energy Standard for Buildings Except Low-Rise Residential Buildings. American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc. U.S. Department of Energy. Atlanta. USA.

Baetens, R., & Saelens, D. (2015). Modelling uncertainty in district energy simulations by stochastic residential occupant behaviour. Journal of Building Performance Simulation, 1493(September), 1–17.

Ballarini, I., Corgnati, S. P., & Corrado, V. (2014). Use of reference buildings to assess the energy saving potentials of the residential building stock: The experience of TABULA project. Energy Policy, 68, 273–284. http://doi.org/10.1016/j.enpol.2014.01.027

Caputo, P., Costa, G., Ferrari, S. (2013). A supporting method for defining energy strategies in the building sector at urban scale. Energy Policy, 55, pp. 261–270. DOI: 10.1016/j.enpol.2012.12.006

Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A. (2014). NbClust: An R package for determining the relevant number of clusters in a data set, Journal of statistical software, Vol.61, Issue 6.

COP. (2015). Paris Summit 2015. www.cop21.gouv.fr

Dascalaki, E. G., Droutsa, K. G., Balaras, C. a., Kontoyiannidis, S. (2011). Building typologies as a tool for assessing the energy performance of residential buildings – A case study for the Hellenic building stock. Energy and Buildings, 43(12), pp. 3400–3409.

Energy Plus. (2016). www.energyplus.net (cit. 15.11.2015)

Episcope. (2016). www.episcope.eu/building-typology (cit. 11.15.2016)

Eppy. (2016). www.pypi.python.org/pypi/eppy/0.4.0 (cit. 15.11.2015)

Fraley, C. and Raftery, E. (2002). Model-based clustering, discriminant analysis, and density estimation. Journal of the American statistical association, Vol. 97, No. 458, pp. 281-297.

Ghiassi, N., & Mahdavi, A. (2016). Urban energy modeling using multivariate cluster analysis. In Proceedings of BAUSIM 2016. Dresden. Germany

Glawischnig, S. (2016). An urban monitoring system for large-scale building energy assessment. Dissertation. Department of building physics and building ecology, Faculty of architecture, TU Vienna. Vienna, Austria.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E. (2010). Multivariate Data Analysis – a Global Perspective. Pearson Global Editions, New Jersey, USA.

Hammerberg, K. (2014). DEMTools. QGIS plugins repository. https://plugins.qgis.org/plugins/DEMTools (cit. 12.03.2016)

Heiple, S., Sailor, D. J. (2008). Using building energy simulation and geospatial modeling techniques to determine high resolution building sector energy consumption profiles. Energy and Buildings, 40(8), pp. 1426–1436. DOI: 10.1016/j.enbuild.2008.01.005

Hens, H., Verbeeck, G., Verdonck, B. (2001). Impact of energy efficiency measures on the CO 2 emissions in the residential sector, a large scale analysis. Energy and Buildings, 33, pp. 275-281.

Hirsch, J.J. and Associates. (2016). DOE2. www.doe2.com (cit. 15.11.2015)

Huang, Y.J., & Brodrick, J. (2000). A Bottom-Up Engineering Estimate of the Aggregate Heating and Cooling Loads of the Entire US Building Stock Prototypical Residential Buildings. In 2000 ACEEE Summer Study on Energy Efficiency in Buildings, pp. 135–148.

Jones, P., Lannon, S., Williams, J. (2001). Modelling building energy use at urban scale. Proceedings of the 7th IBPSA Conference, pp 175–180. Rio de Janeiro, Brazil.

Kavgic, M., Mavrogianni, A., Mumovic, D., Summerfield, A., Stevanovic, Z., Djurovic-Petrovic, M. (2010). A review of bottom-up building stock models for energy consumption in the residential sector. Building and Environment, 45(7), pp. 1683–1697.

Kemna, R. and Moreno Acedo, J. (2014). Average EU building heat load for HVAC equipment. VHK. Delft. Netherlands.

Li, Q., Quan, S.J., Augenbroe, G., Yang, P.P.J., Brown, J. (2015). Building Energy Modelling at Urban Scale: Integration of Reduced Order Energy Model with Geographical Information. Proceedings of the 14th IBPSA Conference. Hyderabad. India.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. Proceedings of the 5th Berkeley symposium on mathematical statistics and probability. University of California Press, Berkeley, USA.

Mahdavi, A., & El-Bellahy, S. (2005). Effort and effectiveness considerations in computational design evaluation: a case study. Building and Environment, 40. 1651 - 1664.

Mahdavi, A., Tahmasebi, F. (2015). The Inter-individual variance of the defining markers of occupancy patterns in office buildings: A Case Study. In Proceedings of the 14 IBPSA Conference, V Garg et al. (ed.); issued by: IBPSA; IBPSA, Hyderabad, India.

Manfren, M., Caputo, P., & Costa, G. (2011). Paradigm shift in urban energy systems through distributed generation: Methods and models. Applied Energy, 88(4), 1032–1048.

Munoz H, E. M., & Peters, I. (2014). Constructing an Urban Microsimulation Model to Assess the Influence of Demographics on Heat Consumption. International Journal of Microsimulation, 7, 127–157.

Natural Resources Canada. (2016). www.nrcan.gc.ca/energy/efficiency/housing/home-improvements (cit. 11.15.2016)

OIB. (2015). RL-6: Energy behavior of buildings. Austrian Institute of Construction Technology, Vienna Austria.

ÖNORM. (2011). B 8110-5 Thermal insulation in building construction, Part5: Model of climate and user profiles. Austrian Standards Institute, Vienna, Austria.

ÖNORM. (2014). B 8110-6 Thermal insulation in building construction, Part6: Principles of verification methods – Heating demand and cooling demand – National application, national specifications and national supplements to ÖNORM EN ISO 13790. Austrian Standards Institute, Vienna, Austria.

Open Street Map. (2015). www.openstreetmap.org (cit. 18.03.2015).

Orehounig, K., Mavromatidis, G., Evins, R., Dorer, V., Carmeliet, J. (2014). Predicting energy consumption of a neighborhood using building performance simulations. Proceedings of BSO Conference. UCL, London, UK.

Parekh, A. (2005). Development of archetypes of building characteristics libraries for simplified energy use evaluation of houses. Proceedings of the 9th IBPSA Conference, pp. 921–928. Montreal, Canada.

Python. (2015). www.python.org (cit. 11.18.2015).

QGIS. (2016). www.qgis.org (cit. 10.03.2016).

R Project for Statistical Computing. 2015. www.r-project.org (cit. 11.18.2015).

Sansregret, S., & Millette, J. (2009). Development of a functionality generating simulations of commercial and institutional buildings having representative characteristics of a real estate stock in Québec (Canada). Proceedings of the 11th IBPSA Conference, pp. 1437–1443. Glasgow, Scotland.

Schöberl, H., Hofer, R., Lang, C. (2012). Handbuch thermische Gebäudesanierung Optimale Ausführungsvarianten. Wirtschaftskammer Niederösterreich Landesinnung Bau. Österreich.

Shimoda, Y., Fujii, T., Morikawa, T., Mizuno, M. 2003. Development of Residential Energy End-Use Simulation Model at City Scale. Proceedings of the 8th IBPSA Conference, pp. 1201–1208. Eindhoven, Netherland.

Snäkin, J.P.A. (2000). An engineering model for heating energy and emission assessment: The case of North Karelia, Finland. Applied Energy, 67.

Swan, L. G., Ugursal, V. I. (2009). Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. Renewable and Sustainable Energy Reviews, 13(8), pp. 1819–1835.

Tuominen, P., Holopainen, R., Eskola, L., Jokisali, J., Airaksinen, M. (2014). Calculation method and tool for assessing energy consumption in the building stock. Building and Environment, 75, pp. 153-160. DOI: 10.1016/j.buildenv.2014.02.001

Vienna GIS. (2015). www.wien.gv.at (cit. 18.03.2015)

Ward, J.H. (1963). Hierarchical grouping to optimize an objective function, Journal of the American statistical association, Vol. 58, No. 301, USA.

World Bank. (2015). www.data.worldbank.org