An Algorithm for Efficient Urban Building Energy Modeling and Simulation

Orçun Koral İşeri¹, İpek Gürsel Dino²

¹ Ph.D. Student, METU
Ankara, Turkey
koral.iseri@metu.edu.tr

² Assoc.Prof.Dr., METU
Ankara, Turkey
ipekg@metu.edu.tr

ABSTRACT

The urban population increases continuously since the industrial revolution, and the residential buildings have the primary responsibility for the total energy demand. There is a need for the analysis of the residential building stock for energy efficiency and sustainable planning. However, energy modeling and simulation in urban scale is expensive in computational complexity and time, due to various building geometries and occupancy types. This research proposes a method to increase the efficiency of the simulation process by reorganizing the building geometries with functional clustering and radiation analysis scaling. In order to accelerate the urban building energy modeling (UBEM) process, the building geometries are modified based on energy simulation standards, then, clustering is determined based on radiation analysis and outside boundary conditions. The candidates are selected according to the selection percentage that has been identified before the process to simulate in building energy software. Three different simulation types are compared to validate the performance of the proposed algorithm with complete model simulations in terms of the error rate of the objectives and the simulation runtime.

Author Keywords

Urban energy simulation; Building energy flows; Energy management; Automation; Residential building stock

ACM Classification Keywords

I.6.4 [Simulation and Modeling]: Model Validation and Analysis; I.6.5 [Simulation and Modeling]: Model Development – Modeling methodologies.

1 INTRODUCTION

Sustainable, energy-efficient solutions are a priority for cities towards decreasing carbon emissions and increasing user comfort [24]. 70% of the global CO₂ emissions are attributed to the urban areas, due to their population density, high rates of economic activity, and associated energy and resource consumption. As 92% of the total population of Turkey lives in cities, and the residential urban areas are responsible for 41% of total energy demand, it is crucial to understand the energy consumption profiles of residential buildings [6,20,32]. However, energy simulation of a wide variety of residential units requires comprehensive models with many parameters, i.e., building volumes, user types, and layout. Due to close interactions with the environment of the residential building stock, the analyzed models are organized as a build-up from the neighborhood level [18,24]. Therefore, simulating an urban region is expensive in terms of time and computing power. This study proposes a bottom-up method for energy modeling of the residential building stock to address this problem. The method aims to decrease the total simulation time for new design projects or retrofit analysis of the urban building model.

1.1 Literature Review

Urban building energy models (UBEM) have the potential to support energy policy decision-making processes for cities to form effective design strategies for building sustainable urban environments [5]. City-scale building data sets are essential for UBEM, which demand different level of building properties for modeling, e.g., footprint, floor area, date of construction, space conditioning type, heating and lighting load, internal load [12]. There are two classes of modeling methods to analyze and estimate overall building stock energy performance, such as top-down, bottom-up approaches [15].

1.1.1. Top-down Approach

The top-down approach evaluates urban models while ignoring household energy demand. Generally, top-down models use collected historical energy data instead of using physical features of individual units and predict end-use energy demand of the building stock by top-level variables, e.g., energy cost, climate effect, macroeconomic indicators such as inflation, gross domestic value [13,24]. The main goal is to find a correlation between macro patterns of the past and the future.

1.1.2. Bottom-up Approach

Bottom-up models consist of the building geometry compositions that are defined as ‘archetypes’ [19,20]. Comprehensive models enhance the description of each building unit in terms of how the unit energy efficiency can be improved and CO₂ emissions can be reduced. These models present cost-effective options for energy demand estimations and CO₂ reduction strategies [22]. This study
adopts a bottom-up approach by processing information from the smallest unit to the top neighborhood level.

1.1.3 Residential Sector and Energy Demand
As the residential sector is one of the leading sectors for energy usage and the environmental impact, residential building stock should be analyzed in detail. However, energy usage profiles of residential types differ widely compared to industrial or commercial buildings due to the variety of building types, areas, or materials. Even, the neighborhood planning composes of a holistic planning strategy; individuals could change their units from years to years for different purposes. Different occupant behaviors and schedules complicate the metering of household energy demands [16]. The situation becomes complex in the aspect of energy demand types of residential units because there are multiple types, e.g., heating load \( Q_h \), cooling load \( Q_c \), domestic hot water (DHW), appliances, and lighting \( Q_l \) [25]. Each demand type is interacting with others from season to season based on user behavior and interior unit layout. Therefore, building energy modeling demands a comprehensive study to reach accurate simulation results for residential buildings.

1.2 BEM and Neighborhood Models
UBEM requires a wide range of information on buildings, such as geometric and non-geometric characteristics (constructions, appliance systems, schedules), and meteorological features of the environment [4]. For geometrical data, Geographic Information Systems (GIS) databases can supply valuable information that contains building age, user type, shape; however, when GIS data is missing, local municipality databases can present reasonable solutions [5,11]. Nevertheless, the computing cost of the model does not only increase depending on the geometry of the structure, the surface heat transfer (u-value, airtightness) and infiltration rates also contribute to the cost [9]. Therefore, there is a need to simplify the modeling process.

Building energy simulation (BES) is an informative model for building energy performance analysis in terms of presenting accurate performance indicators. Energy simulations could serve as feedback at the early design stage, in the way of comparing design alternatives, analyzing the problematic area for the evaluation of the architectural design. The method developed in this study can calculate the energy demand patterns in the neighborhood, and it can sustain valuable information for decision-making on the neighborhood level in terms of energy efficiency and sustainability to architects, planners, or policy-makers.

1.3 Neighborhood Models and Occupant Behavior
Mostly, building energy performance accuracy in simulations decreases when variances in occupant behavior is ignored as a model parameter [14]. However, the occupants have a significant influence on energy demand [26]. The lack of a realistic model that captures occupant behavior creates a demand gap due to the difference between estimated energy performance and actual energy demand influenced by occupant behavior, e.g., daily user schedules, interaction with lighting, and appliances [10,29]. As an important feature, a methodology that can be associated with occupant types can contribute to reducing the variance between calculated and actual energy demand levels [3,23]. In this study, occupant types are generated based on government statistical data repository that represent the actual occupancy profiles, instead of using standard libraries for simulations [27].

2 MATERIALS

The study area is the Kültür neighborhood in Izmir, Turkey (Figure 1). This neighborhood mostly contains retail units on the ground level and residential units on the upper floors. Based on the GIS information, the neighborhood contains 726 residential units with approximately 76,344 m\(^2\) floor area. Approximately, 200 buildings were eliminated as they were atypical in terms of building footprint area for residential function; consequently, 525 residential units were simulated in the study area. The threshold value of the floor area can be changed according to the district. The information on buildings' total number of floors was derived from the Turkish Statistical Institute (TUIK) [27]. On the other hand, the floor height values were not specified in the in the Open Street Map file (.osm). Therefore, researchers determined the height of the building and units measurements based on in-situ observations. Accordingly, the maximum height of the buildings is set to approximately 24 meters, corresponding to 6-7 floors. However, if the height information can be accessed, the model should be constructed based on these values.

![Figure 1. Selected Urban Area in Izmir, Turkey – Red Border Area (2019)](image)

The .osm urban models contain different data types, such as roads and connection points. This information is derived from the GIS, and the dataset converted from 2 dimensional to 3-dimensional model. All layout curves are transformed into four-edged convex geometry for energy simulation tool (i.e. EnergyPlus [17]), in accordance with the modeling restrictions. Except for building geometries and ground surfaces, all other elements in the urban model are ignored, such as roads, urban street elements.

Several parameters are taken from TUIK to generate the model precisely, such as occupant types by ages, the ratio of
space conditioning types [27], while others are estimated, e.g., window to wall ratio, building height, number of floors, residential unit zone division. In addition, the residential block is located in a dense urban area, which can eliminate possibilities of natural daylighting and ventilation. Objects found outside the analysis area are introduced as context geometries as shade-making elements. The context geometry is defined by the ray-casting process, which includes only visible surfaces for simulations. This method decreases total simulation runtime by using only visible elements instead of complete, intricate geometries [8].

3 METHODOLOGY

This chapter presents the method developed for bottom-up neighborhood energy and occupant comfort modeling workflow. The method consists of five steps: data input (as explained previously in Section 2), radiation analysis, model development, simulation, comparative analysis (Figure 2).

3.1 Two-Phased Sampling

This phase involves the statistical sampling technique that helps reduce the computing cost of simulations due to the high number of thermal zones in the urban model. A two-phase stratified sampling method is used, which aims to reduce the number of unit zones to a smaller set that is representative of the whole population by dividing the whole number of members into strata (subgroups) that are have similar thermal characteristics. Stratified sampling works in two steps. Firstly, the algorithm splits the whole data set into different groups with similar characteristic elements according to radiation analysis results, then to distribute the units under the subgroups according to the floor based clustering, which is the division based on the outside boundary conditions. Two-phase grouping leads to an increase in the possibility to select the right distribution for the samples. Finally, randomly sampling within each strata by selecting representative members according to the selection ratio [31]. This process is based on the total solar radiance (SR) incident on envelope surfaces for each unit, which is an adaptation of Dogan and Reinhart’s method [8].

![Figure 3. Solar Radiation-Based Clustering and Vertical Position Sub-Clustering](Image)

3.1.1. Surface Discretization

All vertical building surfaces in the urban model are discretized in vertical and horizontal (each floor) directions (approx. 3m). To form an equal radiation surface area for solar radiation analysis, the division numbers are proportioned based on the façade dimensions of the building, i.e., width, length (Figure 3). The roof surfaces are excluded.

3.1.2. Solar-Radiation Analysis

After discretization, solar radiation analysis (SRA) is conducted for all surfaces. SRA aims to systematically sort residential and retail units according to their incident solar radiation. The radiation results are assigned to the discretized surfaces. The radiation value of a unit is calculated based on the average value of the radiation analysis surfaces found adjacent to the unit. Based on the façade dimensions, each unit could have different number of radiation analysis surfaces. Then, units are sorted according to the weighted radiation values (Figure 3). For this study, a 10-level radiation scale is set. Residential and retail units are categorized by the level of solar radiation.

3.1.3. Clustering based on Vertical Position of Unit

SRA values help the sampling of the units into ten different clusters. Units are sorted based on their weighted radiation values; then, each cluster is divided into three sub-clusters based on the vertical position of each unit in the building (Chapter 3.1). This is because the ground temperature or exposed roof surfaces have different levels of heat transfer compared to adjacent horizontal surfaces in the middle floors. Therefore, three groups are formed based on solar exposure surface type, e.g., top floors, middle floors, ground floors, under the radiation analysis clustering as sub-clusters. The 5%, 10%, 20% sample size are applied for each sub-cluster to execute a uniform selection. This second sampling step is for the equally-distributed selection of the units based on similar thermal characteristics.

3.1.4. Selection of Units for Energy Simulations

Yearly solar-radiation simulations were carried out on the 3D urban model, and the results were sorted into ten different radiation level groups. Then, the units are divided into subgroups with floor-based clustering. The units for energy simulations are extracted from these clusters. In terms of
efficiency, the 5% sample size was 95 minutes, the 10% sample size was 149 minutes, and the 20% sample size was 258 minutes, while the full model lasted 1134 minutes. These simulation types are generated based on the selection ratio in Table 1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Residential elimination</td>
<td>Based on the footprint area</td>
<td>Multiple values, {top, bottom}</td>
</tr>
<tr>
<td>Radiation level</td>
<td>Radiation scale</td>
<td>Single value, 10</td>
</tr>
<tr>
<td>Extracting candidates for simulation</td>
<td>Division of total number of units</td>
<td>Single value, 10</td>
</tr>
<tr>
<td>Number of floors</td>
<td>Based on building height limit [30]</td>
<td>Multiple values, {a,b,c}</td>
</tr>
<tr>
<td>Floor height</td>
<td>Based on building type</td>
<td>Multiple values, {a,b,c}</td>
</tr>
<tr>
<td>Space Conditioning Type</td>
<td>Zone Conditioning</td>
<td>Selection (heat, cool, mixed)</td>
</tr>
</tbody>
</table>

Table 1. Parameters of the UBE Model

3.2 Model description

This chapter describes the space conditioning type and indoor thermal characteristics of each unit. The model consists of a cooling space conditioning system for some units based on the usage ratio of the total number of units, if the climate of the region demands it [27]. For this purpose, residential units are divided into two different clusters, that are mixed-mode (cooling and heating both exist, %20 of all units) and heating-only mode (only heating exists, %80 of all units). This ratio is parameterized in the model, and can be changed with users’ preferences (Table 2). For mixed-mode residential units, there are different zones; living room and bedroom with cooling, and service areas without cooling. For heating-only residential units, the whole unit is considered as one single zone (Figure 4). All retail units are considered as heating-only.

Each unit zone contains different surface types and it could change due to outside boundary conditions, e.g., ceiling or roof surface. Therefore, each unit zone is distributed under different groups based on boundary condition properties on a vertical scale. Besides, the window openings and thermal heat transfers from the surfaces are organized based on the same methodology (Figure 5). The adjacent surfaces were determined as adiabatic surfaces to increase the simulation performance in terms of efficiency.

The space conditioning division, occupant characterization based on age and number, construction definition based on building construction period are identified from the government statistics bureau dataset [27]. The building constructions and schedules is selected according to TS-825 Turkish Standards and ASHRAE standards [1,2,28] (Table 2).

Table 2. Object Properties Used in the Model

<table>
<thead>
<tr>
<th>Surface Heat Transfer</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-value, Wall (before, after 1980)</td>
<td>0.60, 1.88 W/m²-K</td>
</tr>
<tr>
<td>U-value, Roof (before, after 1980)</td>
<td>1.88, 3.12 W/m²-K</td>
</tr>
<tr>
<td>U-value, Floor (before, after 1980)</td>
<td>0.93, 1.92 W/m²-K</td>
</tr>
<tr>
<td>U-value, Window (before, after 1980)</td>
<td>5.1, 2.1 W/m²-K</td>
</tr>
<tr>
<td>Cooling Set Point (Mixed-mode)</td>
<td>25.0 °C</td>
</tr>
<tr>
<td>Heating SetPoint</td>
<td>20.0 °C</td>
</tr>
<tr>
<td>Heating Set Back</td>
<td>10.0 °C</td>
</tr>
<tr>
<td>Natural Ventilation Type</td>
<td>One-sided</td>
</tr>
<tr>
<td>Infiltration Rate Per Area</td>
<td>0.0003 m³/s-m²</td>
</tr>
<tr>
<td>The fraction of Glazing Area</td>
<td>0.25</td>
</tr>
<tr>
<td>Natural Ventilation Limits</td>
<td>21.0 – 24.0 °C</td>
</tr>
<tr>
<td>Number of People Per Area</td>
<td>0.0078 – 0.0394 ppl/m²</td>
</tr>
</tbody>
</table>

Figure 4. Mixed-mode (left) and Heating-only (right) Zone Unit Division

In the second development phase, internal loads are defined. Both population characteristics and occupancy habits are generated under nine occupant types based on the dataset of the TUIK [27]. These types are clustered according to occupant age that represents the population of the urban model, i.e., 0-65 adults from 0-65 age range and adolescent from 0-65 age range, 65+, in accordance with the national statistical datasets [27].

According to occupation types, schedules are generated to represent daily occupancy profiles, e.g., number of people, lighting, equipment and ventilation schedules. Additionally, natural ventilation is introduced in the model to reduce the cooling load in summer when the zone is occupied.
3.3 Simulation
For this chapter, the simulation process for the selected residential and retail units are described. Each unit in the buildings are modeled as described above, and is simulated separately. Annual energy simulations are performed using EnergyPlus. For radiation analysis, RADIANCE is selected [30]. Both simulation tools were operationalized using in the Grasshopper visual coding platform that contains both geometry formation, and energy calculation possibility with extra plug-ins, i.e., Honeybee, Ladybug [21].

3.4 Comparative Analysis
A second UBEM was built for a systematic comparative analysis and validation. For this, all zones (11972 units in 545 buildings) are constructed and simulated. This approach is commonly known as exhaustive search. The comparison metrics of energy demand types (i.e., heating, cooling, lighting) are arithmetic mean (\(\bar{x}\)), standard deviation (\(\sigma\)), mean absolute error (MAE). The success rates of the simulation types with the proposed algorithm are compared based on the 5% confidence interval statistical metric. A confidence interval is the range of elements in a group that evaluates under a degree of confidence [7].

4 RESULTS
This chapter presents the results of the simulation types with the proposed algorithm compared to the complete-model. Four urban building energy models were generated, the first three models are the proposed algorithm and the last one is the complete model of the whole neighborhood.

4.1 Modeling and simulation with proposed algorithm
The proposed algorithm consists of three different simulation types that are differentiated based on unit selection ratio 5%, 10%, 20% sample size. Each simulation process starts with the radiation analysis of the whole unit zones and distribution of the zone units for grouping the zone units. Then, zone units are selected from these groups randomly based on the selection ratio parameter. The proposed algorithm aims to provide efficiency for the total simulation process and reliability for the success rate of the process.

4.1.1. Radiation Analysis Results
The radiation analysis step is only included in the proposed algorithm simulation types. It has resulted in the sampling of the residential units took approximately 40 minutes for each simulation. 162275 surfaces and 11972 units were analyzed for their solar radiation value on vertical surfaces (Figure 6). According to the radiation results, the selected units were sorted for their radiation values. Following, each radiation scale level was divided in terms of floor-based clustering as sub-groups, i.e., top floors, middle floors, ground floors. In total, 30 different groups were formed for each simulation type with the proposed algorithm. Based on the radiation scale parameter or the floor-based clustering parameters, the total group number could increase or decrease to increase the reliability.

4.2 Modeling and simulation of complete model
Unlike the proposed algorithm, the complete-model simulation does not include radiation analysis and floor-based clustering processes. All unit zones are simulated individually with the brute-force method to provide validation for measuring the performance of the simulation types with different sample sizes. For this reason, although the duration of radiation analysis has been extracted from the simulation time, the total time is considerably longer than the proposed algorithms since the all unit zones are put into the structure energy simulation individually.

4.3 Comparative analysis between the proposed model and the complete model
Three different selection ratio numbers were simulated for testing the performance of the proposed algorithm (Table 1). For the Kültür Neighborhood of Izmir case, the ratio of space conditioning type were determined as 20% of a mixed-mode,
80% of the heating-only. The comparative analysis process is shown as:

- 20% sample size; heating-only (residential, retail), mixed-mode (residential)
- 10% sample size; heating-only (residential, retail), mixed-mode (residential)
- 5% sample size; heating-only (residential, retail), mixed-mode (residential)
- Complete model; heating-only (residential, retail), mixed-mode (residential)

In Table 5, the comparison of the units with mixed-mode simulation results is presented in the aspect of Heating Load (QH), cooling load (QC), and lighting load (QL) with average and standard deviation. The reliability for mixed-mode units is lower, due to their lower sample size as compared to the heating-only units. For QH, the 10% and 20% sample size results are similar to each other as 33.6 and 33.7 kWh/m², but the 5% sample size is lower than the other types as 30.8 kWh/m². In parallel, a similar trend is observed in the standard deviation values of QH. From 5% to 20%, as the selection rate increases, the reliability ratio also increases for QC and QC. On the other hand, QL and QL values result in approximately the same as 8.0 to 8.2 kWh/m² and 1.8. Lastly, although some comparisons had similar results with other model types, the 5% sample size performed worse than the 10% and 20% sample sizes in all comparisons.

![Figure 7](image.png)

**Figure 7.** Error Ratio for Heating-Only (HO) and Mixed-Mode (MM) units

Table 4. Heating-only units (%80 of total simulations) simulation comparison

<table>
<thead>
<tr>
<th>kWh</th>
<th>5% sample</th>
<th>10% sample</th>
<th>20% sample</th>
<th>Complete model</th>
</tr>
</thead>
<tbody>
<tr>
<td>QH</td>
<td>35.7</td>
<td>36.6</td>
<td>38.9</td>
<td>38.7</td>
</tr>
<tr>
<td>σQH</td>
<td>16.1</td>
<td>16.9</td>
<td>15.7</td>
<td>17.0</td>
</tr>
<tr>
<td>QL</td>
<td>17.2</td>
<td>17.3</td>
<td>17.3</td>
<td>17.1</td>
</tr>
<tr>
<td>σQL</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
</tbody>
</table>

Table 5. Mixed-mode units (%20 of total simulations) Simulation Comparison

<table>
<thead>
<tr>
<th>kWh</th>
<th>5% sample</th>
<th>10% sample</th>
<th>20% sample</th>
<th>Complete model</th>
</tr>
</thead>
<tbody>
<tr>
<td>QH</td>
<td>30.8</td>
<td>33.6</td>
<td>33.7</td>
<td>34.8</td>
</tr>
<tr>
<td>σQH</td>
<td>9.4</td>
<td>12.8</td>
<td>12.9</td>
<td>12.7</td>
</tr>
<tr>
<td>QC</td>
<td>29.2</td>
<td>30.8</td>
<td>31.7</td>
<td>32.4</td>
</tr>
<tr>
<td>σQC</td>
<td>12.8</td>
<td>13.3</td>
<td>13.5</td>
<td>14.1</td>
</tr>
<tr>
<td>QL</td>
<td>8.2</td>
<td>8.1</td>
<td>8</td>
<td>8.0</td>
</tr>
<tr>
<td>σQL</td>
<td>1.8</td>
<td>1.8</td>
<td>1.8</td>
<td>1.7</td>
</tr>
</tbody>
</table>
The aim of the algorithm is to increase efficiency in terms of simulation period by means of a sampling method that implements radiation analysis and floor-based clustering for candidate selection. Table 6 presents the time comparison between consecutive simulation processes. As all simulations with proposed algorithm include radiation analysis for the whole urban geometry model, the radiation analysis is also added in the simulation time. Floor-based clustering process does not contain any analysis. Therefore, there is no time addition for the simulation types due to floor-based clustering.

In conclusion, the proposed algorithm performed successfully for the %10 and %20 sample sizes in terms of $Q_H$, $Q_c$, $Q_b$ based on 5% confidence interval comparison. The error ratio proportionally increased when the sample size was lowered. On the other hand, the simulation runtime was significantly lower than the complete model simulation process. If the number of buildings to be simulated increases, the time difference between the proposed algorithm and complete model simulation in expected to increase.

5 DISCUSSION

In this study, a bottom-up approach for UBEM is presented. There are some comparative steps for different selection levels of the units for the energy simulation to calculate the success rate of the algorithm. The 10% and 20% sample size simulation types were successful, resulting in lower than the 5% error rate for the selected objectives. However, the %5 sample size simulations did not show complete success for heating and cooling loads. More reliable simulations should be conducted by increasing radiation scale levels for unit extraction for energy simulations or increasing the sample size. Moreover, the simulations were executed in the urban context, but for this process, there was no context element differentiation as a constraint between three different levels of simulation groups. The context geometries are highly effective for solar gain and natural ventilation objectives, which have an impact on both heating and cooling demands. Similarly, user types provided important input values for energy simulations by organizing the occupancy, heating/cooling set point, and setback schedules. As future work, additional parameters can be added related to context geometries and occupant types to increase the precision of the simulation results.

6 CONCLUSION

This study proposed an algorithm that decreases the computing cost of the UBEM and its simulations with a bottom-up approach. The proposed framework categorizes the building geometries as residential and non-residential geometries before energy simulation. By implementing solar-radiation analysis at the beginning, the units of the buildings are sorted based on weighted radiation values on the façade. Then, clustered units are categorized according to their positions in the building under three different sub-clusters such as top, middle, and ground floor clustering. Finally, the selected units are categorized according to their space conditioning. Based on the unit selection rates, three simulation levels were compared with the complete model simulations. 10% and 20% sample sizes error ratio was lower than 5% for all objectives. The 5% sample size resulted in increased errors due to the insufficient number of sample sizes. For future work, it is necessary to develop an inclusive UBEM framework with more precision for objectives, both to facilitate the simulation process and increase the content of the model in terms of other urban context components such as landscape elements (i.e. greenery, water) and urban infrastructure.

ACKNOWLEDGMENTS

This research was partially supported by the Newton – Katip Celebi Fund, Grant No. 217M519, by the Scientific and Technological Research Council of Turkey (TUBITAK) and British Council, UK.

REFERENCES

8. Dogan, T. and Reinhart, C. Shoeboxer: An algorithm


18. MURE, O.- Monitoring of EU and national energy efficiency targets. 2012.


27. TÜİK. *TÜRKİYE İSTATİSTİK KURUMU Turkish Statistical Institute*. 2010.


