A Mobility-Driven Approach to Modeling Building Energy

Andreas Berres
Computational Urban Sciences
Oak Ridge National Laboratory
Oak Ridge, Tennessee
berresas@ornl.gov

Piljae Im
Building Technologies Research & Integration
Oak Ridge National Laboratory
Oak Ridge, Tennessee
imp1@ornl.gov

Kuldeep Kurte
Computational Urban Sciences
Oak Ridge National Laboratory
Oak Ridge, Tennessee
kurtekr@ornl.gov

Melissa Allen-Dumas
Computational Urban Sciences
Oak Ridge National Laboratory
Oak Ridge, Tennessee
allenmr@ornl.gov

Gautam Thakur
Geoinformatics Engineering
Oak Ridge National Laboratory
Oak Ridge, Tennessee
thakurg@ornl.gov

Jibonananda Sanyal
Computational Urban Sciences
Oak Ridge National Laboratory
Oak Ridge, Tennessee
sanyalj@ornl.gov

Abstract— **Buildings are one of the primary energy consumers in any city's energy use [12]. Presence and absence of humans is a major contributing factor to the energy use in a building. In this paper, we present an approach to generating a realistic model of human building occupancy throughout a typical work week. We use the Toolbox for Urban Mobility Systems (TUMS) to generate a synthetic population based on population distribution estimates, we schedule the population's daily commute based on National Household Travel Survey (NHTS) survey data, and we simulate their daily travel patterns using an agent-based transportation simulation (TRANSIMS). We process and fuse the simulation output to produce a list of the first and last seen location of each agent in the simulation. Based on the arrival at the last destination, we map each agent to one of the nearby buildings. Using these agent arrivals, as well as NHTS data, we create an customized hourly occupancy schedule for each building, which replaces the typical generic occupancy schedule that is usually used for building models [23]. We successfully demonstrate this workflow at the example of the Chicago Loop, a major business district in Chicago, Illinois.**

This manuscript has been authored [in part] by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

I. Introduction

Urban environments are complex systems in which social factors, mobility, building energy, and urban climate interact with each other. The interactions between buildings and urban microclimate are driven by physical processes for which there are existing models. However, modeling the interaction between cities and people has largely depended on general assumptions about transportation mode choices and building use. Transportation and building energy are two of the top consumers of energy use in the United States. Transportation accounts for 29% of energy use, whereas buildings account for 38-40% of energy use (combined residential and commercial) [1], [12]. There are many factors that influence energy use in any particular building, including climate, building shape, building materials, and number of occupants. In this paper, we introduce an agent-based model (ABM) approach to modeling the population's behavior and decision-making, and obtain a more accurate representation of energy use in buildings. We generate a realistic model of human building occupancy throughout a typical work day. The model is based on the population's daily commute patterns, using a transportation simulation that is tailored to average measured traffic volumes in the area. Several hundred thousand vehicle occupants are then assigned to tens of thousands of buildings, to generate building occupancy schedules.

Energy use in an urban environment is tightly linked with human mobility. The purpose of this work is to create a workflow to create a realistic building energy model based on a data-driven model of mobility for a population in an urban scenario. Traditionally, occupancy is represented by a generic, piecewise linear curve (seen in Figure 8), which does not account for differences between buildings or individual people's schedules [23].

In this paper, we present a workflow for a data-driven coupling from daily commute to building energy. We

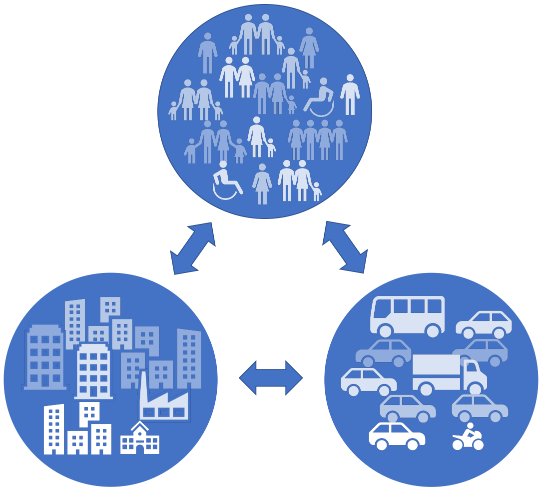


Fig. 1: In an urban environment, there are several major components that are interdependent: population, transportation, and buildings.

1) create a realistic commute scenario based on population models and observed travel patterns,
2) compute agent arrival and departure schedules for a typical work week from simulated vehicle traces, and
3) produce a building occupancy schedule for each building and simulate building energy use.

II. Related Work

A. Transportation

ABMs have been used for a variety of urban mobility applications. Many of such models have been used to model evacuation times and path choices [9], [13], however, some have been used for migration analysis between rural and urban areas [30], or commute patterns of school children based on demographic data [7]. Some works feed into these migration analysis works, such as an approach to assessing population dynamics at high temporal resolution based on LandScanUSA [27], [31] population distribution data by Bhaduri et al. [6]. They consider two approaches to refining the temporal resolution: occupancy curves by business type to enhance daytime and nighttime population distributions, and integrating LandScanUSA as initialization for transportation modeling. Others, like Xue et al. [34], build their work on top of ABMs to evaluate children's exposure to air pollutants on their commute to school, or use ABMs to model urban residential choice [17].

There are a wide variety of ABM transportation models, which each have their unique advantages and disadvantages [28]. The TRansportation ANalysis SIMulation System (TRANSIMS) [32], [33] is a well-established open source ABM transportation model that has been used to model larger regions such as Dallas [3] and Maryland [14]. It can work at macroscopic, mesoscopic, and microscopic levels. For the purpose of this paper, we need microsimulation. Required inputs include Origin and Destination Matrices (O/D Matrices) which determine origins and destinations for each trip, a road network, and a trip schedule. While TRANSIMS is by far not the only traffic simulation available, it is open source, it supports microsimulation, which is required to get vehicle traces, and there is a tool to generate vehicle trips. The Toolset for Urban Mobility Simulations (TUMS) [5] unifies transportation simulation and population models into a single tool. It consists of three components: data processing, traffic simulation models (TRANSIMS and MITSIM), and web-based visualization. It integrates Open Street Map (OSM), LandScan population distributions, and other open data into a single tool to generate evacuation and commute scenarios. It has been demonstrated for the city of Cleveland, Tennessee.

B. Building Energy

Building energy simulations have been an important topic for the past 50 years, and there are vast differences between different simulations' capabilities to model complex building geometries, HVAC, human thermal comfort, solar analysis (effect on solar radiation on building climate and lighting), insulation analysis, advanced fenestration, and more [10]. EnergyPlus [11], [22], is one of the most complete [10] whole building hourly simulation programs. It provides energy analysis and thermal load simulation program based on heat balance-based solutions. A number of previous works have used building occupancy information to simulate running utilities such as HVAC on flexible schedules dependent on building occupancy instead of fixed schedules. These works use sensors to provide real-time occupancy information to the simulation [2], [15] or statistical modeling for human tasks throughout the day [26]. The simulated results indicate potential energy savings from 10-15% [2], [15] through occupancy-driven building control.

III. Methodology

A realistic model for urban energy use has to consider the interconnection between population, transportation, and buildings (see Figure 1). Each individual in a population has a variety of trips they make to move between all the activities they have planned for a given day. Our example person (Figure 2) begins their typical day at home, walks their child to the school bus, and then takes the car to commute to work. They might walk somewhere for lunch, and at the end of the work day, they stop at the gym and the grocery store on their way home. At any given point in time, this person will have an impact on energy use. Driving a vehicle consumes energy, and depending on weather, heating or air conditioning may be required to keep the car at a pleasant temperature. The person's presence or absence from buildings also impacts energy use. In the workplace, a person requires a computer or other equipment. While at home, they may use large appliances, watch television, shower, etc. In either place, people require lighting, as well as a comfortable room temperature. Lighting accounts for almost 25% of commercial building energy use and 11.6% of residential energy use [12]. Heating and cooling account for almost 25% of commercial

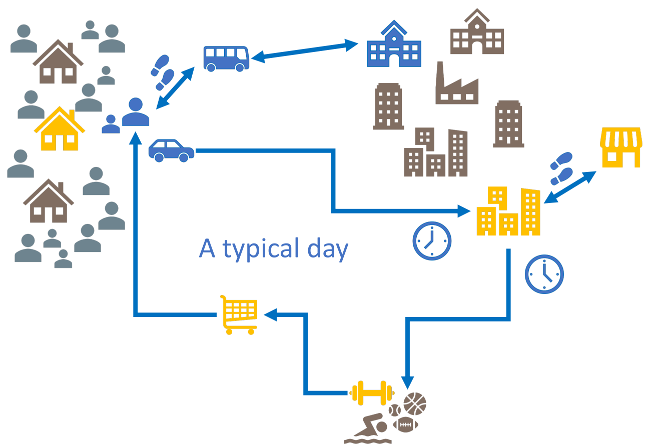


Fig. 2: A typical day in the life of an individual includes a variety of trips to move between the different activities throughout any given day.

and almost 40% of residential energy use [12]. These can be adjusted based on outdoor temperatures, time of day, and –for more efficient energy use– building occupancy.

In this paper, we will focus on building occupancy in office buildings based on a synthetic population's daily commute. For the purpose of building occupancy, we focus this work on the Chicago Loop area, a large business district in Chicago, Illinois.

A. *Workflow*

Figure 3 presents an overview of the workflow presented in this paper. The workflow begins with the initialization, execution, and postprocessing of TRANSIMS to produce vehicle traces with coordinates for each time step a vehicle is in the system (Section IV). TRANSIMS is initialized with vehicle trips that are generated by TUMS based on LandScan population distribution data (Section III-C1), and a trip schedule that is based on commute times that are derived from travel surveys (Section III-C3). It is then executed for an entire work week. The output of the simulation includes a snapshot of vehicle (or *agent*) locations with respect to road network links (sections of roads with identical numbers of lanes, speed limit, etc) for every given time step. In a post-processing step, the agent locations are fused from TRANSIMS output and OSM road network geometries.

In the next set of steps of the workflow (Section V), we determine each agent's exact first and last seen locations as one can reasonably assume that individuals will park their vehicles in the vicinity of buildings they live or work in. Before determining the mapping of individuals to buildings, we first simplify the building models by flattening the file structure, and then augment relevant information such as overall building area. Then, we map agents to the nearby buildings based on their last seen locations. This feeds into agent arrival schedules for each individual building. To determine agent departures, we utilize commute times that are derived from travel surveys (Section III-C3) directly as time in traffic does not directly affect departure times. From the agent arrival and departure schedules, we can derive a building occupancy schedule for each building. Finally, this occupancy schedule, along with the building geometries (Section III-C2) are used as inputs to EnergyPlus. The resulting outputs then give information about each buildings energy use by type of end use, type of energy source, floor, and other criteria.

B. *Area of Interest: Chicago Loop*

For the purpose of this paper, we focus on the Chicago Loop, which holds a large portion of the central business district of Chicago. The central business district houses a large population and is site to hundreds of thousands of jobs [16]. It is also a fast-growing area for high-rise buildings [18], which makes it an interesting study area for building energy use. Figure 4 shows a rendering of the Chicago Loop buildings for which all required information required for building energy modeling is available. However, for benchmarking purposes, we extend the evaluated region to a larger area of about 200 square kilometers, for which we have just the building footprints without energy-relevant metadata. The Chicago Loop area includes 343 buildings, whereas the larger area includes 86,093 buildings. As we are interested in commute, we consider a typical work week for our simulation.

C. *Data Sources*

For this coupling, we are incorporating a range of different data sources. Some of these data are observational or community standard, other are simulated using different simulations, and finally some are created as part of this work. While the second and third type is discussed at greater detail in the later sections, we would like to provide some background on the first type (observational or community standard data).

1) Population Distribution: The workflow begins by generating trips for a synthetic population. The basis for this generation is the LandScanUSA dataset [27], [31], which provides daytime and nighttime population distributions. It is a well-respected standard in the population modeling community. In the United States, the dataset has a resolution of 90 meters, which is much higher than the global LandScan dataset which has a resolution of about 1 kilometer, and only represents ambient population distribution averaged over 24 hours.

2) Building Geometries: The building geometries used in this work are informed by LiDAR and provided as proprietary ESRI shapefiles. They were converted to GeoJSON for easier processing with standard libraries. We use two separate building geometry datasets which are displayed in Figure 4a: The first one encompasses the entire City of Chicago, which contains 820,598 buildings. From this dataset, we have derived a smaller dataset of 86,093 buildings which fall within the traffic simulation area which covers about 200 square kilometers. The second dataset covers only the Chicago Loop, which contains 343 buildings. This dataset contains additional properties for the buildings which are required to run EnergyPlus, the building energy simulation.

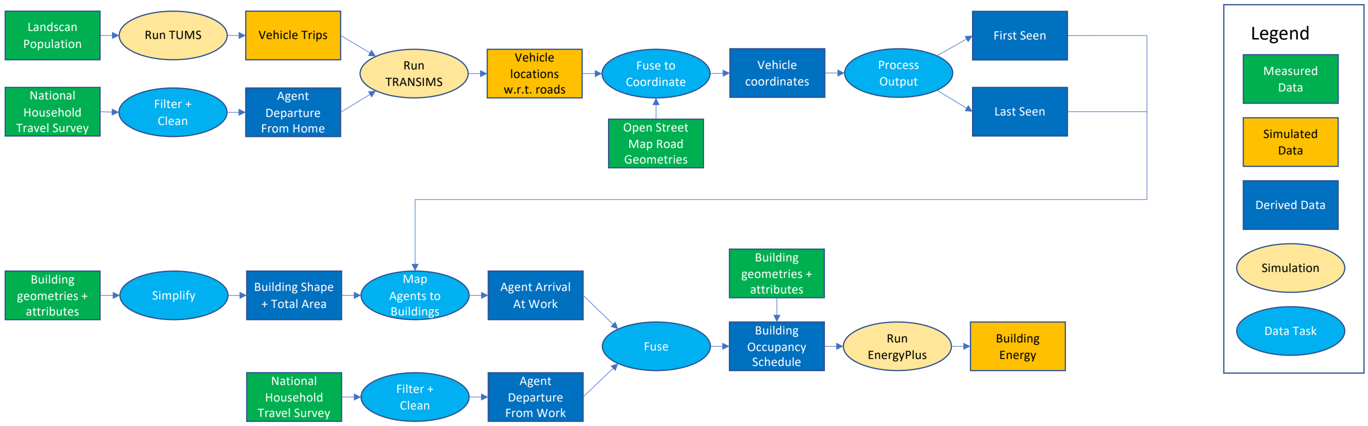
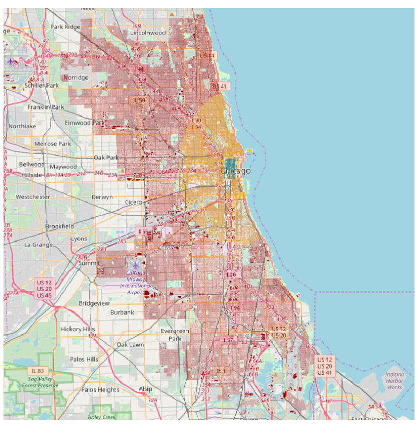
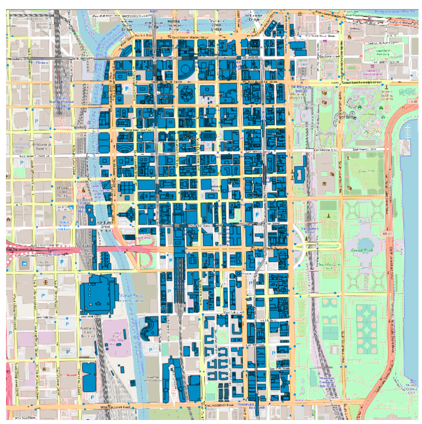


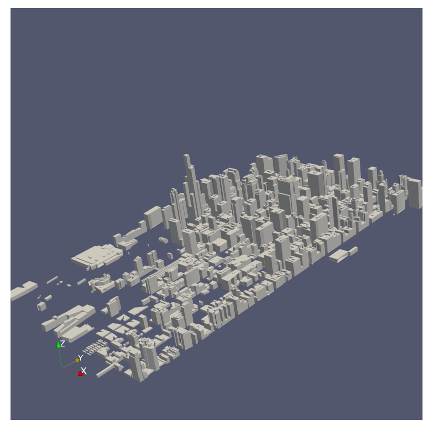
Fig. 3: Overall workflow for the mobility-driven approach to modeling building energy. The workflow includes measured data (green boxes), simulated data (yellow boxes), and derived data (blue boxes). The different data products are used as inputs for simulations (pale yellow ellipses) and tied together by data transformation and algorithmic tasks (teal ellipses).



(a) View of the entire Chicago area.



(b) Close-up of the Chicago Loop.



(c) 3D rendering of the Chicago Loop.

Fig. 4: Building geometries used in this paper: Chicago Loop buildings (blue), buildings within the traffic simulation area (yellow), and buildings for the whole Chicago area (red).

3) Commute Times: To accurately model commute times, we need to know at what time the population departs for work, and at the end of the day, departs from work. One possible data source are the average traffic volumes provided by the Illinois Department of Transportation (IDOT) [24]. These traffic volumes provide hourly schedules for different weekdays for all daily traffic. However, since we are specifically interested in energy use in Chicago's business district, which are largely dependent on commute. The National Household Travel Survey (NHTS) [20] is a valuable source of information for trip purposes. The NHTS provides national-level statistics on personal travel in the United States. It includes detailed information on households, people, vehicles, travel mode (e.g. car, bus, walk, bike), trip purpose (e.g. home, work, school). While for much of the United States, the published data is aggregated to regional level, large metropolitan areas, such as Chicago, are treated separately. From an overall of 923,572 responses in the 2017 dataset [21], we were able to filter out 6,955 responses for the City of Chicago.

IV. Transportation Model

We use TRANSIMS, an agent-based transportation model to simulate a realistic commute scenario. In order to create the required urban mobility scenario, we employ a mixture of simulation and data fusion techniques to initialize TRANSIMS with the required inputs, beginning with the synthetic population. Next, we tune the transportation model to adapt the number of agents (vehicles) in the system to observed travel patterns (Section III-C3). Once the model is run, we process it to create agent arrival and departure schedules (Section V-A). For the purpose of this paper, we focus on passenger vehicles for daily commute in Chicago, which account for the majority of commuters in Chicago [8]: 57.1% in Chicago, 69.8% in Cook County, and 85.6% nation-wide in 2017.

A. Generating Realistic Trips

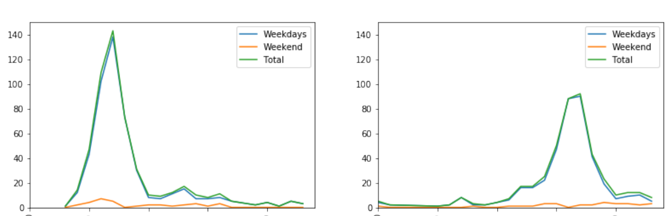
We create a synthetic population using the TUMS [4], an established tool chain for population modeling. TUMS integrates OSM, LandScan [27], [31], as well as other open data (such as Census data), with the TRANSIMS and MIcroscopic Traffic SIMulator (MITSIM) traffic simulation engines. As such, it provides a basic traffic simulation that is informed by daytime and nighttime population distributions.

We feed the traffic volumes, and the OSM road network into TRANSIMS. TRANSIMS takes a matrix of origin/destination pairs, and finds the best route for each of these trips with respect to other traffic. However, these origin/destination pairs alone do not suffice: the simulator also needs a schedule to determine at which times it will send off each vehicle.

1) Initialization with Hourly Schedules: One area in which TUMS falls short is temporally distributing individual trips across the simulation time frame based on departure times. As a first step to tuning TRANSIMS to real-world volumes, we provide such a schedule for each day of the week, based on the average hourly traffic volumes for each work day. As TUMS forces a choice between morning and evening commute, we generate two scenarios for each work day. The first scenario uses the commute-to-work setup and the second scenario uses the commute-from-work setup. To achieve a realistic distribution, we obtained 2017 NHTS [21] data (Section III-C3), which provides nation-wide travel survey information. To generate a trip departure schedule for TRANSIMS, we execute the following steps:

1) Filter the dataset for responses from City of Chicago (6955 responses).
2) Filter for all responses indicating the trip purpose was commute:
   - ”travel to work” (791 responses)
   - ”travel from work” (799 responses)
3) Normalize relative traffic volume with respect to total daily volume by hour of day as TRANSIMS requires all relative traffic volumes to proportionally add up to 100%.
4) Write out in TRANSIMS trip schedule file format.

Figure 5 visualizes the responses for departure times to travel to work or from work, prior to normalization.



(a) (b)

Fig. 5: NHTS responses for Chicago area trip departure times from home to work (left) or work to home (right) [21] aggregated to hourly intervals.

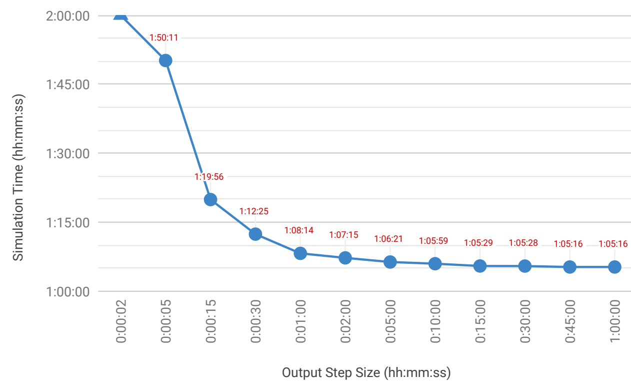


Fig. 6: TRANSIMS runtimes for different output step sizes from hourly outputs to 2 second outputs.

B. Running the TRANSIMS Simulation

We simulate traffic for commute to work and from work for each work day, using the observed travel patterns and origin/destination pairs. The simulation runs at a granularity of 0.1 seconds, however, such a high resolution is neither necessary nor reasonable for the purpose of this study. The parameter to adjust the write step size can be changed in the simulator configuration. For a given step size, e.g. 1 hour, the simulator pauses its compute at each step to write out a snapshot all agents that are in the system at that given time. Therefore, the step size affects not only write frequency but also write volume. At a coarser resolution, a lot of agents that are on shorter trips, are missed. For instance, if an agent were to leave at 8:01 a.m. and arrive at 8:59 a.m., it would neither appear in the simulation snapshot for 8:00 a.m. nor in the one for 9:00 a.m. – this is an important consideration for the choice of step size as we do not want to lose any data.

We ran some benchmarks on a simulation with about 285,000 agents to determine the best output step size. As Figure 6 demonstrates, we found that TRANSIMS is compute bound up to a step size of 15 minutes. At smaller output steps, the main limiting factor in simulation speed is writing to disk – the simulation is input/output (I/O) bound. At 30 second outputs, the I/O overhead already hits 11%, and at 5 second outputs, it is at almost 70% extra compute time.

We choose a write frequency of 30 seconds. This implies the assumption that the last 30 seconds (or less) of any trip are spent within reasonably close proximity to the final destination. The compute overhead for 30 second outputs is acceptable, which makes this choice a reasonable compromise between accuracy and computational demands.

C. Fusing Coordinates to Agent Locations

The snapshot files from the previous step contain, among other information, agent ID, timestamp, link ID, direction on link, and position on link for each agent at each time step it is active. It lacks coordinates for the actual agent location. The OSM road network file contains a collection of links – sections of road (line strings) which share the same properties – for which it provides link ID, main direction (0 or 1), coordinates for the link (one or more line segments), and other parameters. By fusing the two sources, we are able to append coordinates to the snapshot files.

1) Flatten the road network file to tabular format (coordinates are maintained as a linestring).
2) Compute overall arc length of each link and append it to the table.
3) For each agent location:
   a) Identify link and, if the agent is traveling in opposite link direction, reverse the order.
   b) Determine which link segment the agent is traveling on by comparing the relative arc length at each node of the link to the agent's relative position (arc length on the link).
   c) Compute coordinates for agent's position on that line segment.
   d) Convert coordinates from latitude/longitude to Universal Transverse Mercator (UTM) as UTM is locally cartesian and facilitates distance computations.
   e) Append the agent's coordinates to the snapshot file.

V. Assigning Agents to Buildings

The goal of this paper is to create a building energy model based on a realistic occupancy schedule. For the purpose of this paper, we assume that agents travel from a specific building near their origin location to a specific building near their destination location, and that they spend the majority of their time between their arrival and their subsequent departure inside this building. In this Section, we present the different steps required to create EnergyPlus input data from given TRANSIMS output.

A. Compute First and Last Seen Agent Locations

For building occupancy, the most relevant part of the simulation output are the departure and arrival times and locations for all vehicles. Therefore, we want to produce agent arrival and departure schedules as a more compact representation of the simulation output. At a write resolution of 30 seconds and a simulation area of 35 square kilometers, a typical simulation day snapshot has about 25 million lines of agent movement data. Processing such large amounts of information sequentially is rather slow. Therefore, we apply a divide-and-conquer approach that splits up the data into chunks (as outlined below), and processes the chunks in parallel.

1) We read in chunks of the file separately. For each chunk, we determine the first and last seen time and location for every agent in the chunk.
2) Then we recursively merge adjacent chunks and update the arrival information for every agent as needed, until we have one cohesive list for the entire day.

The chunk size should be chosen such that it is larger than the maximum number of agents in the system at any given time, but substantially smaller than the total number of lines. For a snapshot size of 25 million lines, we found that a chunk size of 500,000 worked well.

B. Pre-Processing Buildings

The provided buildings files are in a complex JSON data structure. To facilitate faster processing, we simplify the buildings files by flattening them into a tabular format that only contains

- building ID: unique identifier for each building
- total area in square meters: this property is provided for some but not all building geometries. Where unavailable, we approximated the total area as
$$\frac{\text{footprint area} \cdot \text{avg height}}{2.5}$$
to approximate the relationship between total area, footprint area, and average height for the buildings which included total area in their properties.
- primitive geometry: We reduce the building to its bounding box and centroid as this is a sufficiently accurate approximation for the vast majority of buildings in the simulated area, and because it facilitates faster processing.

Through this simplification, we observed that while tools such as QGIS [25] provide an inaccurate building count: complex building geometries are represented by *multi polygons* which consist of multiple polygons, as the name implies. While they still belong to the same building, each ob the sub polygons is counted as a separate feature. For instance, for the Chicago Loop geometry, QGIS cited 2575 features where there were only 343 buildings. This was not the case for the 200 square kilometer area data and the difference was marginal (less than 10 buildings difference) for the City of Chicago data which have a slightly different format for building geometries.

C. Mapping Agents to Buildings

Provided the last seen location for each agent, one can assign each agent to a nearby building. On average, we have over 640,000 agents in a TRANSIMS simulation for a single day of simulation within an area of 34 square kilometers around the loop. We begin by culling agents that are not within a reasonable distance of any Chicago Loop Buildings, which reduces the number of agents to be considered to about 400,000. At 343 buildings in the loop, this still constitutes over 137 million comparisons.

To address this, we employ a quadtree-based [29] approach with flexible split criteria, given a set of agents and buildings. The algorithm follows these steps:

1) Count agents and buildings
2) Determine midpoint of the given area
3) Check if the number of agents and the number of buildings are greater than the split criterion
   a) Yes: subdivide into four quadrant, using the midpoint as a split point. This step can duplicate buildings if their corners fall into different quadrants.

This is intentional as these buildings may still be the best choice for an agent from neighboring quadrants.

   i) For each quadrant, jump back to step 1.

b) No: Within the given quadrant, map each agent to a building.

For split criterions, we experimented with a range from 5 to 100 buildings (or agents) per leaf quadrant. We took a sample of 1000 buildings and 1000 agents, which would produce 1 million comparisons when executed serially. Figure 7 shows the resulting average times to map all agents to buildings. Unsurprisingly, smaller split criteria performed much better as they only have to perform a fraction of the comparisons. For this work, we have used a split criterion of 10. The Chicago Loop has 343 buildings that are grouped in about 100 city blocks. 10 buildings correspond to about 3 blocks surrounding the agent location.

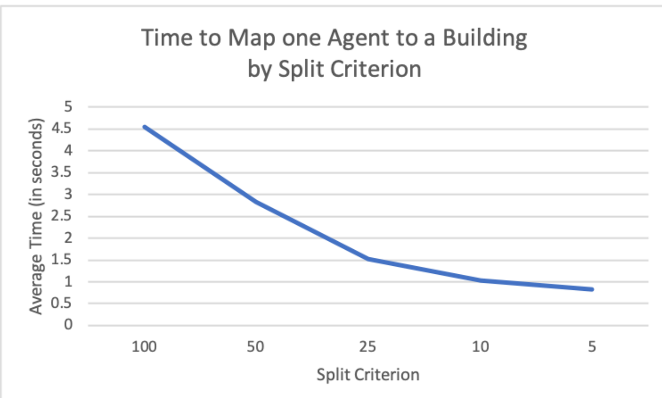


Fig. 7: Benchmark of mapping agents to buildings using Euclidean Distance for differently sized split criteria.

For the purpose of this paper, we restricted ourselves to two simple distance metrics that are based on building centroids and total areas. Alternative mapping methods could include polygon-based distance metrics (which we avoided in favor of faster compute), or stochastic assignment to buildings.

1) Euclidean Distance: We compute the Euclidean Distance between each agent and each building's centroid within a given quadrant. The agent is assigned to the building with the smallest distance.

2) Weighted Euclidean Distance: We compute a Weighted Euclidean Distance using weights that are based on the building's total area.

$$\text{relative area} = \frac{\text{total area}}{\sum_{\text{buildings}} \text{total area}}$$

We use the building's inverse relative areas as weights, such that larger buildings will decrease the distance measure whereas smaller buildings will increase the distance measure. The agent is assigned to the building with the smallest weighted distance.

D. Creating a Building Occupancy Schedule

EnergyPlus requires two inputs to define occupancy: an building occupancy schedule for a given time interval (we use one hour per step) which contains relative building occupancy throughout the day, as well as the maximum occupancy for each building.

To create this schedule, we first need the number of agents arriving and departing the building throughout the day. Given the agent-to-building mapping we can easily create an agent arrival schedule for each building. We aggregate the agent arrivals to an hourly schedule, such that we obtain an hourly schedule of agents arriving at the building for each building.

For the agent departure schedule, we need the times of departure before we can proceed. While we simulate the commute home from work for all agents for other purposes, such as emissions modeling, we have reliable survey information from the NHTS [21] on the average times individuals in Chicago usually leave work. To determine the agent departure schedule, we apply the same transformations to the commute-from-work data that we applied to the commute-to-work data to obtain a schedule of relative volume throughout the day.

1) Compute the relative proportion of individuals leaving their workplace during any given hour (see Section IV-A1 step 3).
2) For each building:
   a) Obtain agent arrival schedule.
   b) Determine cumulative hourly arrivals throughout the day.
   c) Compute number of hourly departures for each building based on its maximum occupancy during the day and the hourly proportion of individuals departing the building.

Finally, we can compute hourly building occupancy by taking the difference of the cumulative arrival and departure times. We then convert it to the required format for EnergyPlus by storing the maximum occupancy for all buildings in a separate file, and normalizing building occupancy to [0,1] before saving it to one file per building.

The simulated occupancy is compared to the generic office schedule in Figure 8. It is slightly faster than the generic occupancy in the mid-morning hours, but quite similar in the afternoon and during off-peak hours. As lunch breaks are not represented in the transportation model, the lack of a drop during lunch hours is expected; however a lunch hour could be simulated if needed. Overall, the simulated schedule is a more useful input than the generic one for the building energy model as it is based on a statistical representation of the activity of actual population.

E. EnergyPlus Building Energy Output

Finally, we initialize EnergyPlus [22], a building energy simulation, with the building occupancy schedules from the previous step. Like other whole building energy simulation programs, users are required to define building geometry, envelope thermal characteristics, HVAC system specification

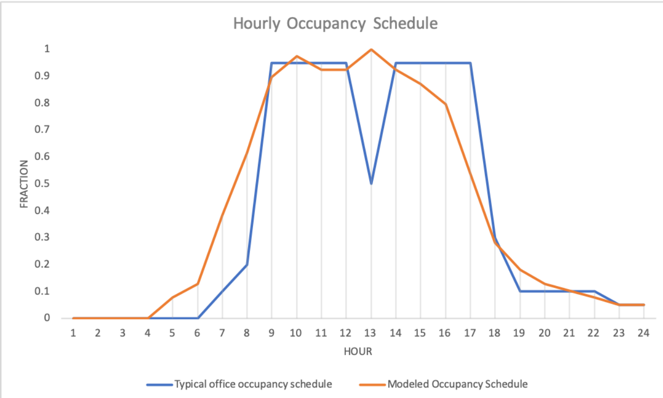


Fig. 8: Comparison of a typical office occupancy schedule vs the modeled occupancy schedule.

as well as operation schedules such as thermostat schedule, occupancy schedule, lighting and plug load schedules. Once input file is ready to be run along with a weather file, the simulation is performed for a whole year or part of year in hourly or sub hourly time resolution. The simulation outputs provide the simulated energy consumption for heating, cooling, ventilation, lighting and plug and process loads as well as indoor conditions such as temperature and relative humidity. In this study, generic occupancy density and schedules in a building model will be replaced with the occupancy schedules generated from transportation simulation.

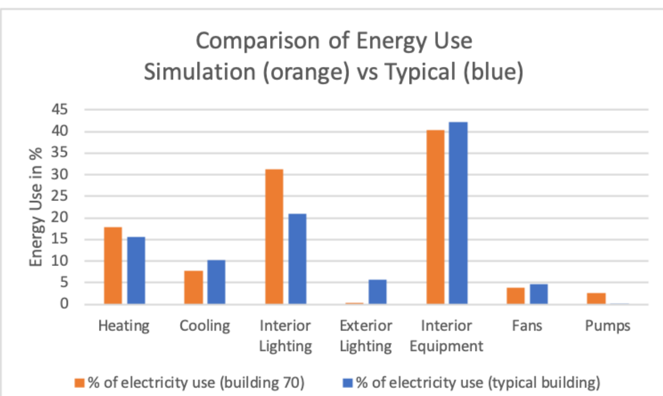


Fig. 9: Average daily electricity use by end use for building 70 as simulated with EnergyPlus compared to a typical medium-sized office building.

In the following, we present results for one of the buildings which is located on the corner of South Franklin St and West Adams St. The occupancy model inputs for this building was created from the previous step, and the maximum occupancy for the building was estimated as 3,000. According to a prototype building model input [23], a typical office building in this size would have around 4,600 of maximum occupancy. Therefore, the occupancy model estimated slightly lower (i.e., about 35% lower) occupancy for this building. The selected building is a typical office building installed with packaged Direct Expansion (DX) cooling unit with electric reheating Variable Air Volume (VAV) system. The model input values for the building and HVAC characteristics except occupancy input were defined per the prototype office building model [23]. Figure 9 displays the average daily electricity use by type of end use. Heating uses 17.7% of all electricity which is used for VAV reheating system, and cooling uses 7.8%. Interior lighting is third at 31.2%. The majority of electricity use (40.3%) is consumed by interior equipment. The remaining energy use types share the remaining 3%. We compare this energy use to that of a typical medium-sized office building at a similar latitude (Buffalo, New York) [23]. As seen in the Figure 9, heating consumed slightly more energy and cooling consumed slightly less. Interior equipment and fans consumed about the same. The larges difference lies in interior lighting (higher for simulated results), exterior lighting (higher for typical building), and pumps (higher for simulated results). Overall, the results are within the expected range.

VI. Conclusion and Future Work

We have presented an end-to-end workflow for a mobility-driven approach to modeling building energy, from transportation simulation setup to building energy output. For every step of the workflow, we have discussed required data inputs and outputs, and explained design choices that were made based on domain-specific reasons. To extend this work, we will replace the census tract TUMS inputs with finer-scale population inputs. Moreover, we would like to create a model for full year of travel, including non-commute travel, and include multiple modes of transportation, based on daily traffic volumes from IDOT and NHTS data. This provides a significant challenge as it will require many more data sources to create such a scenario. We would further like to explore more complex mapping methods for the assignment of agents to buildings, such as the ones outlined in Section V. Finally, we would like to consider weather as an input to the transportation simulation, as it has a profound effect on not only building climate but also driving behavior and transportation mode choice [19].

Acknowledgments

The authors would like to than the Exascale Computing Project (ECP) and the U. S. Department of Energy (DOE) for funding this work. We would furthermore like to thank the ORNL Leadership Computing Facility (OLCF) for supporting this work with compute time on the Titan supercomputer. Moreover, we would like to thank Cheng Liu on his guidance on using TUMS, Srinath Ravulaparthy for his suggestion to use NHTS data, and Steve Peterson for his help with QGIS. Finally, we would like to thank IDOT for providing us with detailed traffic data in addition to their publicly available annual reports [24].

References

[1] U.S. Energy Information Administration. Use of energy in the united states - energy explained. https://www.eia.gov/energyexplained/index.php (retr. 2019-08-14).

[2] Yuvraj Agarwal, Bharathan Balaji, Rajesh Gupta, Jacob Lyles, Michael Wei, and Thomas Weng. Occupancy-driven energy management for smart building automation. In *Proceedings of the 2Nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building*, BuildSys '10, pages 1–6. ACM, 2010. event-place: Zurich, Switzerland.

[3] Richard Beckman, Chris Barrett, K. Berkbigler, K. Burris, Brian Bush, S. Hull, J. Hurford, P. Medvick, Deborah Kubicek, Madhav Marathe, J. Morgeson, K. Nagel, D. Roberts, L. Smith, M. Stein, Paula Stretz, S. Sydoriak, K. Cervenka, and Rick Donnelly. Transportation analysis simulation system (transims) - the dallas-ft. worth case study. Technical report, Los Alamos National Laboratory, 12 2002.

[4] B. L. Bhaduri, W. Lu, C. Liu, G. Thakur, and R. Karthik. Toolbox for Urban Mobility Simulation: High Resolution Population Dynamics for Global Cities. *AGU Fall Meeting Abstracts*, pages PA42A–05, December 2015.

[5] BL Bhaduri, W Lu, C Liu, G Thakur, and R Karthik. Toolbox for urban mobility simulation: High resolution population dynamics for global cities. In *AGU Fall Meeting Abstracts*, 2015.

[6] Budhendra L Bhaduri, Edward A Bright, Amy N Rose, Cheng Liu, Marie L Urban, and Robert N Stewart. Data driven approach for high resolution population distribution and dynamics models. In *Proceedings of the Winter Simulation Conference 2014*, pages 842–850. IEEE, 2014.

[7] Budhendra L Bhaduri, Cheng Liu, James J Nutaro, and Lauren A Patterson. Geospatial modeling and simulation based approach for developing commuting patterns of school children. *CSI Communications*, 32(9), 2008.

[8] U. S. Census Bureau. 2013-2017 american community survey 5-year estimates. https://factfinder.census.gov (retr. 2019-04-11).

[9] Xuwei Chen and F Benjamin Zhan. Agent-based modelling and simulation of urban evacuation: relative effectiveness of simultaneous and staged evacuation strategies. *Journal of the Operational Research Society*, 59(1):25–33, 2008.

[10] Drury B Crawley, Jon W Hand, Michaël Kummert, and Brent T Griffith. Contrasting the capabilities of building energy performance simulation programs. *Building and environment*, 43(4):661–673, 2008.

[11] Drury B Crawley, Linda K Lawrie, Frederick C Winkelmann, Walter F Buhl, Y Joe Huang, Curtis O Pedersen, Richard K Strand, Richard J Liesen, Daniel E Fisher, Michael J Witte, et al. Energyplus: creating a new-generation building energy simulation program. *Energy and buildings*, 33(4):319–331, 2001.

[12] DOE Office of Energy Efficiency & Renewable Energy. Energy efficiency trends in residential and commercial buildings.

[13] Marco D'Orazio, Luca Spalazzi, Enrico Quagliarini, and Gabriele Bernardini. Agent-based model for earthquake pedestrians' evacuation in urban outdoor scenarios: Behavioural patterns definition and evacuation paths choice. *Safety science*, 62:450–465, 2014.

[14] Sevgi Erdogan, Fredrick W Ducca, Krishna Patnam, Zuxuan Deng, and Xiang Wang. Inclusion of time-dependent networks in maryland statewide transportation model. Technical report, National Center for Smart Growth Research and Education at the University of Maryland, 2014.

[15] Varick L. Erickson, Yiqing Lin, Ankur Kamthe, Rohini Brahme, Amit Surana, Alberto E. Cerpa, Michael D. Sohn, and Satish Narayanan. Energy efficient building environment control strategies using real-time occupancy measurements. In *Proceedings of the First ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, BuildSys '09, pages 19–24. ACM, 2009. event-place: Berkeley, California.

[16] Greg Hinz. Downtown jobs hit a record high. https://www.chicagobusiness.com/article/20141129/ISSUE05/31-1299947/downtown-chicago-employment-hits-a-record-high (retr. 2019-08-18).

[17] Qingxu Huang, Dawn C Parker, Tatiana Filatova, and Shipeng Sun. A review of urban residential choice models using agent-based modeling. *Environment and Planning B: Planning and Design*, 41(4):661–689, 2014.

[18] Blair Kamin. The growth of chicago's super loop: So much building, so little architecture. https://www.chicagotribune.com/columns/blair-kamin/ct-chicago-building-boom-kamin-met-0903-20170908-column.html (retr. 2019-08-18).

[19] Kuldeep Kurte, Srinath Ravulaparthy, Andreas Berres, Melissa Allen-Dumas, and Jibonananda Sanyal. Regional-scale spatio-temporal analysis of impacts of weather on traffic speed in chicago using probe data. In *9th International Conference on Sustainable Energy Information Technology (SEIT 2019)*, 2019.

[20] N McGuckin and A Fucci. Summary of travel trends: 2017 national household travel survey. Technical Report FHWA-PL-18-019, US Department of Transportation, Federal Highway Association, 2018.

[21] NHTS. National household travel survey data. https://nhts.ornl.gov/downloads (retr. 2019-08-18).

[22] United States Department of Energy. EnergyPlus. https://energyplus.net/ (retr. 2019-08-16).

[23] US Department of Energy. Commercial prototype building models | building energy codes program, 2013. https://www.energycodes.gov/development/commercial/prototype_-models (retr. 2019-08-19).

[24] State of Illinois Department of Transportation Office of Planning and Programming. Illinois travel statistics 2017, 2017.

[25] QGIS Community. QGIS. https://qgis.org/en/site/ (retr. 2019-08-18).

[26] Ian Richardson, Murray Thomson, and David Infield. A high-resolution domestic building occupancy model for energy demand simulations. *Energy and Buildings*, 40(8):1560–1566, 2008.

[27] Amy N Rose and Eddie A Bright. The landscan global population distribution project: Current state of the art and prospective innovation. Technical report, Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States), 1 2014.

[28] Mustapha Saidallah, Abdeslam El Fergougui, and Abdelbaki Elbelrhiti Elalaoui. A comparative study of urban road traffic simulators. In *MATEC Web of Conferences*, volume 81, page 05002. EDP Sciences, 2016.

[29] Hanan Samet. The quadtree and related hierarchical data structures. *ACM Computing Surveys (CSUR)*, 16(2):187–260, 1984.

[30] Jaylson J Silveira, Aquino L Espíndola, and TJP Penna. Agent-based model to rural–urban migration analysis. *Physica A: Statistical Mechanics and its Applications*, 364:445–456, 2006.

[31] Gautam S. Thakur, Kevin A. Sparks, Robert N. Stewart, Marie L. Urban, and Budhendra L. Bhaduri. Curating transient population in urban dynamics system. *CoRR*, abs/1606.08488, 2016.

[32] TRANSIMS Community. TRANSIMS studio. https://sourceforge.net/projects/transimsstudio/ (retr. 2019-08-18).

[33] Michael D Williams, Gary Thayer, and LaRon Smith. Development of the TRANSIMS environmental model. Technical Report LA-UR-9782, Los Alamos National Laboratory, 1997.

[34] Jianping Xue, Thomas McCurdy, Janet Burke, Budhendra Bhaduri, Cheng Liu, James Nutaro, and Lauren Patterson. Analyses of school commuting data for exposure modeling purposes. *Journal of exposure science and environmental epidemiology*, 20(1):69, 2010.