The Macro Impacts of Micro-Climates on the Energy Consumption of Urban Buildings

Samantha Inneo, Daniel Wadler, Jack Schneiderhan, and Ronald Estevez

Stevens Institute of Technology, Hoboken NJ 07030, USA

Abstract. This paper addresses Challenge 3 of the SMC data challenge by leveraging data-driven tools to understand the relationships between our built environment and nature, and how this relationship impacts energy consumption. It presents detailed results to the research questions posed, along with the rationale for the tools used and limitations of the developed solutions.

Keywords: urban micro-climate, energy analysis, spatial visualization

1 Introduction

From the urban heat islands (UHI) of New York City [1], to Chicago's notorious wind-tunnel effect [2], the adverse impact of urban morphology on local climate is undeniable. Recent studies [3] have uncovered links to extreme weather events such as severe thunderstorms, and heat/cold waves. Despite growing concern [4,5], scientific understanding of this phenomenon - both cause and effect - remains limited [6]. This is in part due to the modeling complexities involved, sometimes requiring exascale computing resources [7]. By leveraging recent advances in data science and machine learning, this paper presents a computationally inexpensive, data-driven understanding of the coupling between nature and urban infrastructure. A survey of the literature is discussed in the following section to highlight previous advancements in this research area.

2 Literature Review

The coupling between urban climate and energy has long been the subject of research interests. Vallati et al. [8] for example explored the impact of urban climate on the heating/cooling demand of standalone vs. urban building types. Their findings indicated that urban buildings required notably less energy for heating. This may be explained by the urban heat island (UHI) effect - a byproduct of the pervasive use of heat-absorbing materials such as steel and concrete. This idea served as the basis for work conducted by Arifwidodo et al.[9] which found strong correlations between the presence of UHIs and increased electricity expenses. Subsequent research focused on the impact of urban morphology [11], climatic variations [10], and population trends [12], all of which found correlations between urban morphology, local climate gradients and energy consumption patterns.

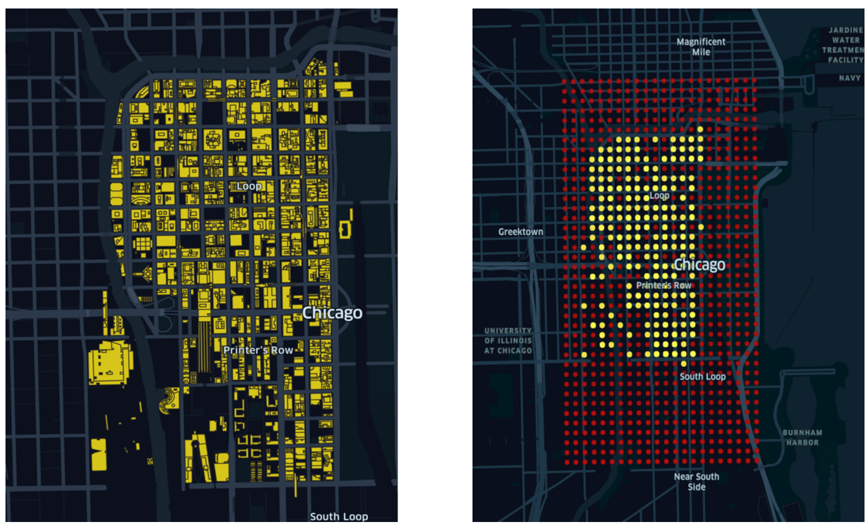
From a modeling perspective, researchers at the Oak Ridge National Laboratory [13] conducted studies focused on how variations across weather datasets impact micro-climate simulations, and ultimately energy consumption forecasts. New et al. [7] then leveraged discrete building energy modeling at urban-scale to inform on the dynamics between urban form and climate. While representing significant progress, the computationally intensive nature of these research efforts (some requiring the world's most powerful supercomputer at the time [7]) deter from widespread use. This presents an opportunity to develop computationally inexpensive alternatives. To this end, this paper leverages building energy consumption as an alternative lens through which to investigate the complex multi-scale coupling between nature and urban infrastructure. Specifically, it makes a contribution by investigating the role urban buildings play in shaping the climate around them through the data-driven analysis of variations in their annual energy consumption.

The remainder of the paper details how this was achieved and is structured as follows. Section 3 begins with an overview of the scope of study and an understanding for the type of data used. Section 4 follows with detailed solutions to the research questions posed. It details the exact approaches used, discusses the results and highlights their limitations. Section 5 concludes with a discussion of the broader impacts and future potential of the presented work. Source code and a comprehensive set of all generated visualizations and animations have been open-sourced, and is accessible via our Github page (link in Appendix).

3 Initial Data Collection and Processing

Generated by the Oak Ridge National Laboratory (ORNL) [14], the primary dataset describes the weather, buildings, and energy consumption of the "Chicago Loop" - the second largest commercial business district in the US, located in downtown Chicago. The dataset is composed of three sub-datasets, the first of which (herein called *Building-Data*) contained building IDs, longitude and latitude coordinates and structural height for 334 buildings (see Figure 1a). The second dataset (herein called *Energy-Data*) contained the annual energy consumption of each building. Finally, the third dataset (herein called *Weather-Data*), provided high resolution, 90-meter simulated weather data for the year 2015, at 15-minute intervals (with known gaps toward the end of each month). In addition to the provided datasets, our team sourced additional open datasets to strengthen our data pipeline. Namely, we obtained a building footprint dataset [15] to aid with building visualizations, and an energy bench-marking dataset from the Chicago Data Portal [16] which provided Energy-Star ratings of each building.

Combining our external datasets to *Building-Data* proved rather challenging. Despite having similar building references, the long/lat coordinates differed significantly between datasets. By spatially analyzing our data using the Folium python library [17], we made a fascinating finding. We discovered that the building coordinates provided in *Building-Data* utilized a Lambert Conformal Conic projection (with 90-meter resolution grids). Consequently, the coordinates in *Building-Data* referenced grid cell centroids, as opposed to real world building locations. To address this problem, our team scripted a nearest neighbor algorithm which computed and assign the nearest real world building (using Euclidean distance) to each centriod location. Our approach was validated by randomly sampling nearest neighbor assignments and visually verifying them using two popular web tools, namely Google Maps [18] and Koordinates [19].



(a) Building footprint visualization **(b)** Building overlaid with weather sensors

Fig. 1: *Data exploration of building (yellow) and weather (red) dataset.*

Exploring the *Weather-Data* dataset revealed a total of 880 sensor points. By combining this data with *Building-Data*, we discovered that the weather data covered a land area about twice the size of that occupied by our buildings (see Figure 1b). We narrowed the scope of our weather data to match that of our buildings in terms of land area (please refer to Appendix for more details). We created a local SQLite database for the resulting dataset to enable faster data retrieval. With respect to *Energy-Data*, we found data parsing to be particularly challenging because the dataset contained extraneous data in the form of null values. This was addressed using standard data cleaning practices.

3.1 Data Visualization Pipeline

We developed a visualization pipeline (see Figure 2) to streamline the spatial visualization of our data and results. For any visualization task, we first extracted the necessary WKT data contained our SQLite database and transformed it into a shapefile using Mapshaper, an online GIS service. Next, we converted the shapefile into a Python friendly GeoJSON file format and visualized it using the Kepler.gl python library. In terms of hardware, our team relied on personal laptops to execute the pipeline, the most capable of which had a Core i7 processor and 16GB of RAM. Having briefly discussed the dataset, the next section presents the research questions asked of this dataset, along with our detailed responses for each.

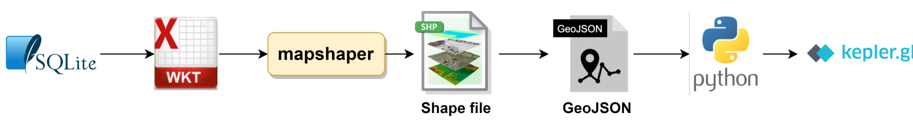


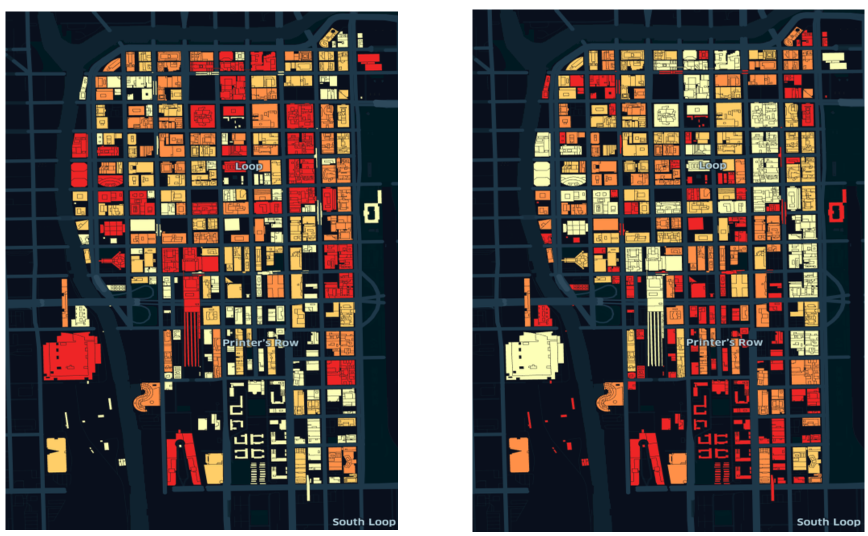
Fig. 2: *Data visualization pipeline*

4 Research Questions

4.1 Are there interesting variations in the weather and building energy use data for the geographic area?

***Approach*:** Using our visualization pipeline, we mapped all building footprints against multiple energy and weather variables. Color was used as a relative measure for comparisons between buildings, with lighter color shades representing lower use/intensity, and darker shades representing high use/intensity. This gradient-based, color-to-intensity relationship is maintained across all visualizations presented in this paper. For visualizations involving weather, we developed a multithreaded Python script that efficiently summarized *Weather-Data* into daily descriptive statistics (including standard deviation, variance, and mean) and saved the output as a JSON file. We then narrowed the features of our weather dataset using personal intuition. Although approaches such as Principal Component Analysis (PCA) are best suited for feature selection tasks, we did not feel comfortable using it owing to a limited understanding of how it works. This represents an area for improvement in future work. Based on our intuition, we selected temperature, wind speed, long-wave radiation, and relative humidity as the features of interest for this work. To visualize potential micro-climatic effects, we further narrowed our scope to only days which displayed significant variance across our selected features. This was motivated purely by computational expense as an exhaustive exploration would have proved computationally infeasible given our limited computational resources.

***Variations in Building Energy Use*:** Our analysis revealed several variations across both energy use and weather. An interesting trend - which we termed "*two halves of Chicago*" - emerged when we compared both electricity usage and intensity across the entire building stock. According to Figure 3a, buildings in the northern half generally used more electricity compared to the south. This relationship is, however, reversed with respect to electricity intensity (see Figure 3b). This implies the existence of two unique urban morphologies, with the north having predominantly tall buildings (hence lower electricity intensity), and the south having relatively shorter, less efficient buildings. Visualizing building heights across the dataset confirmed our suspicions. Please refer to our Github page (link in Appendix) for a more comprehensive set of all generated visualizations for this problem.



(a) Total electricity use **(b)** Total electricity intensity

Fig. 3: *Comparing relative electricity use and intensity - darker shades represent greater use/intensity*

***Variations in Weather*:** Figure 4 illustrates weather gradients observed across Temperature (F), Wind speed (m/s), Long-wave radiation (W/m2) and Relative humidity (%). We discovered that weather gradients typically emerged from one predominant direction and spread across the map. From these plots, it is evident that micro-climatic effects exist and it would be interesting to apply machine learning algorithms in future work to recognize patterns in these variations and begin to correlate them directly to features of the urban morphology.

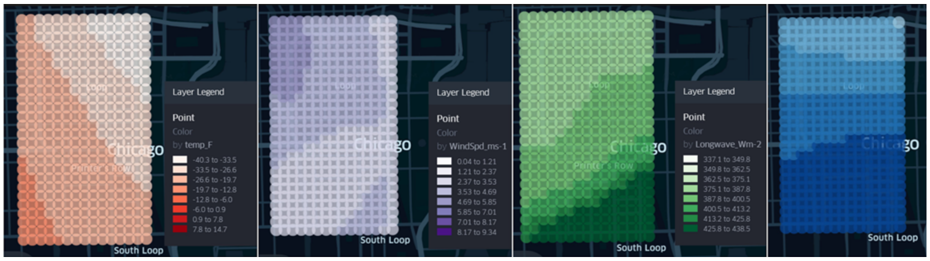


Fig. 4: *From left to right - temperature, wind speed, long-wave radiation, and relative humidity visualizations. Darker shades represent greater intensity.*

***Limitations of solution*:** Our analysis relied heavily on physical building variations (e.g building height) owing to a lack of adequate granularity in our energy dataset. Additionally, the reduced scope of our weather analysis limited our findings. Future iterations will look to leverage high-performance computing (HPC) resources along with advanced data structures such as HDF5 to enable efficient, detailed analysis. We hypothesize that leveled mappings of annual weather data at 15-minute intervals with energy use data of equal resolution will yield very interesting results regarding the connection between building energy use and micro-climates.

4.2 Which buildings in the study are most sensitive to weather effects?

***Approach*:** Addressing this question required an understanding of energy use variation across short term events (i.e rain) and seasons (i.e. winter). However, a major shortcoming of our dataset is that it aggregated energy use over an annual cycle, hence it lacked the granularity we needed. This presented an opportunity for "out of the box" thinking. Our resulting approach made one key assumption - that the ratio between heating and cooling energy for all buildings is constant, irrespective of building size. While improbable, we concluded that for any building to violate this relationship, it must be because its sensitivity to weather forced it to use more cooling or heating energy than the "normal ratio" requires.

To test this hypothesis, we divided each building's HVAC consumption based on end use (heating and cooling) and plotted them against each other. Surprisingly, our assumption was indeed correct. Figure 5 (leftmost plot) shows a strong linear relationship between heating and cooling across the entire building stock. Note that this accounts for all building sizes. It also revealed some weather-sensitive outliers. By plotting the building footprints of the outliers, we noticed that the majority of the outliers were located in the bottom half of our map (see middle plot of Figure 5). This underscored the impact of urban morphologies on energy use. For deeper insights, we then combined our findings with the energy audit dataset we had sourced externally. This revealed that several of the buildings in our outlier set had low Energy-Star ratings (lighter color shades), further strengthening our conclusions that these buildings are most sensitive to weather (see rightmost plot in Figure 5).

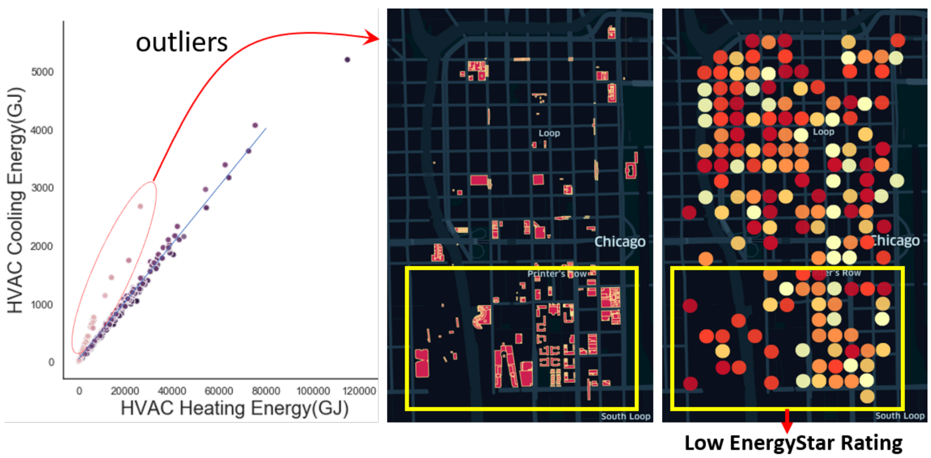


Fig. 5: *Identifying buildings most sensitive to weather*

***Limitations of solution*:** Although weather sensitivity is indeed one of the potential drivers behind the presented results, it is also likely that our findings are the result of inefficient/faulty HVAC systems in these buildings. Should granular energy data become available in the future, our approach to this problem will be to identify days with extreme weather conditions and compare each building's energy usage on these days to a derived building baseline. This will enable us to perform sensitivity analysis to identify buildings with strong sensitivities.

4.3 How can the data best be divided into subsets for meaningful analysis and visualization?

***"Two Halves of Chicago"*:** As highlighted earlier, the Chicago Loop can be separated into two halves based on urban morphology, with the north end comprising mainly of high rise buildings and the south having a comparatively shorter cityscape. Buildings in the north had lower energy intensity compared to the south. We hypothesize that given granular building energy use data, meaningful micro-climate analysis can be performed to compare the performance of similar buildings across the geographic divide.

***Unsupervised Hierarchical Clustering*:** Another idea we tested was to perform unsupervised learning to discover natural clusters within the dataset. We were confident that this will enable us to uncover hidden dynamics between buildings that went beyond simple spatial relationships. To this end, we leveraged the HDBscan algorithm to perform hierarchical clustering on our dataset. One of the main motivations for using this algorithm was the fact that it treats cluster size as a hyper-parameter. Using this algorithm, we were able to segment our dataset into 8 main clusters as illustrated in Figure 6. Although we did not explore further with this approach, we found it to have a lot of potential, especially as we introduce additional datasets to the data pipeline.

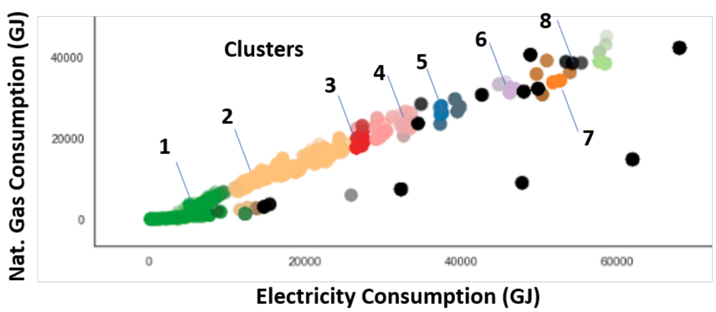


Fig. 6: *Building clusters generated using unsupervised learning*

***Extracting Workdays and Types of Buildings*:** During our preliminary data analysis, we explored the idea of segmenting based on building function. Moreover, we were interested in sub-setting energy output according to workdays, weekends and holidays. We believed this to be a good line of inquiry given that the Chicago Loop is a CBD. This idea motivated additional interest into further sub-setting based on building zoning type (i.e residential or commercial). Although we managed to source relevant datasets to enable all of these, we eventually abandoned these efforts owing to the lack of energy data granularity. The possibility of revisiting this idea using detailed energy data, however, presents a promising avenue for future efforts.

4.4 How does energy use in each building change throughout the year?

***Approach*:** Our analysis revealed a clear variation between the energy consumed during the winter months (requiring heating) and the summer months (requiring cooling). It should be noted that Fall and Spring seasons were neglected in this study. On average, buildings in the dataset were found to use 15 times more energy for heating than cooling. Given that heating primarily occurs in winter, we conclude that energy usage peaks during the colder months. We followed this by calculating the normalized weather energy use intensity (EUI) as well as the average energy demand per heating/cooling day. We also computed the average energy demand per heating/cooling day, illustrated in Figure 7.

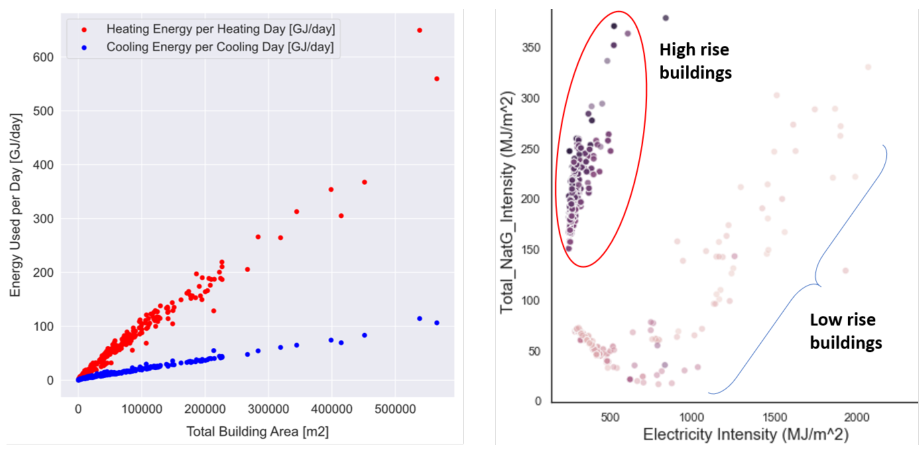


Fig. 7: *Energy demand per heating/cooling day(left) and EUI plot (right)*

***Analysis*:** Figure 7 reveals a trivial, yet important trend - the larger the building, the more energy it uses for heating/cooling (left plot). Additionally, we compared energy intensities of heating (natural gas) and cooling (electricity) (see right plot). From this, we discovered that not only do high rise buildings use comparatively more natural gas than electricity (measured in gigajoules), but they tend to be more efficient (hence the tighter cluster of dark colored points in Figure 7). In contrast, not only did low rise buildings rely more on electricity than natural gas, but most importantly, they exhibited irregular trends with respect to efficiency (see larger spread of lighter colored circles). This was a very interesting finding and potentially points to high rise building benefiting from newer construction and HVAC systems.

4.5 How is energy use different during the coldest/hottest months as compared to during those of less extreme temperature?

Based on the findings presented in Figure 7, we can assume that the coldest months should have the highest energy usage, while the hottest months should have comparatively lower energy usage (measured in gigajoules). Despite this, the months of extreme temperature will still likely exhibit higher energy usage than in the more temperate months. An analysis of energy used for heating and cooling compared to the overall energy used revealed that some buildings used as much as 69% of their annual energy on heating and cooling (see Figure 8). These particular buildings will see a significant decrease in the least extreme months.

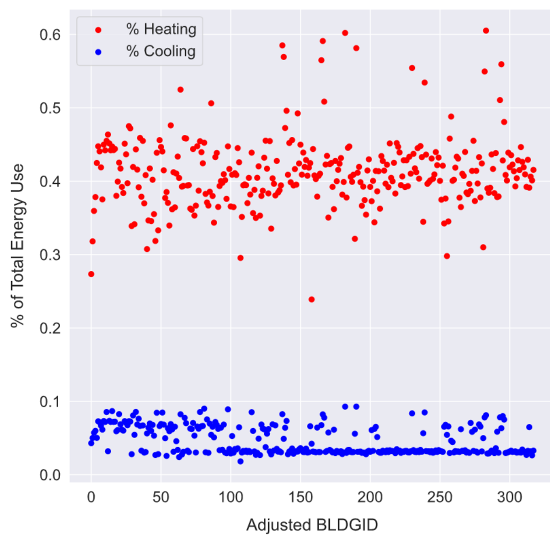


Fig. 8: *Breakdown of HVAC energy use over total energy expenditure*

4.6 Are there any interesting visualizations that illustrate the changing dynamics of the simulated urban environment?

Figure 9 illustrates the urban heat island effect projected over a 3D representation of the Chicago Loop. This figure provides a number of key takeaways. It visually confirms the morphological differences between the two halves of Chicago. But more importantly, it highlights man-made heat zones. Each building is shaded based on its heat rejection energy data. As we can deduce, a heat-zone emerges between the high rise buildings in the north, in stark contrast to the south buildings. Future work will look to leverage wind directional patterns to observe impacts on temperature gradients across the Loop.

5 Conclusion

Using detailed data analysis and visualization, the results presented in this paper shed light on some of the complex relationships between nature and urban infrastructure in the downtown Chicago area. Computationally inexpensive in

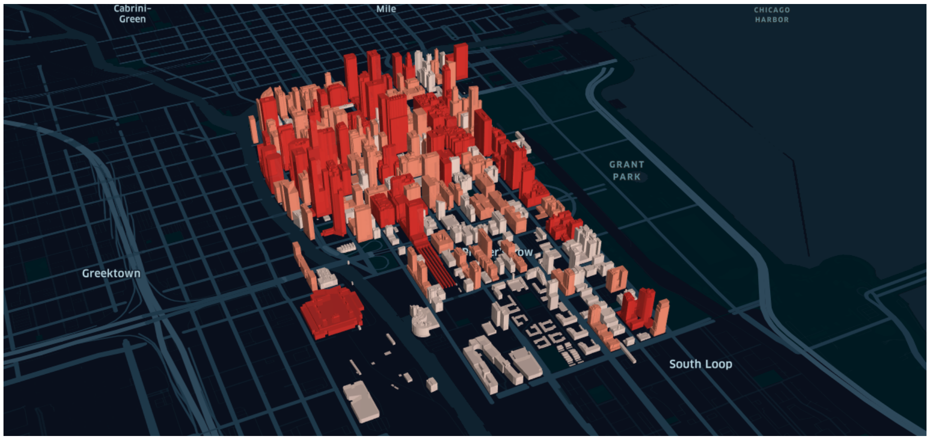


Fig. 9: *Visualizing the urban heat island effect using heat rejection data*

nature, the approach used, along with its findings serve to provide an exploratory step prior to more detailed energy modeling and analysis. It also sets the stage for use of high performance computing resources with more granular data to uncover even more hidden insights and relationships.

Acknowledgments

The authors would like to acknowledge the Pinnacle Scholar Program at the Stevens Institute of Technology for their support of this work. The authors would also like to thank Dr. Philip Odonkor for his mentorship and support of this work.

Support for DOI 10.13139/ORNLNCCS/1619243 dataset is provided by the U.S. Department of Energy, project SMC2020 under Contract DE-AC05-00OR22725. Project SMC2020 used resources of the Oak Ridge Leadership Computing Facility at Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725

APPENDIX

This bounding box used to graphically bound our weather data was developed using the following coordinates: (41.858452, -87.641479), (41.858452, -87.617188), (41.891693, -87.641479), (41.891693, -87.617188).

Source code and comprehensive set of visualizations and animations have been open-sourced and can be accessed via our Github page.

Github Link - https://bit.ly/3hGEwo0

References

1. Bornstein RD.: Observations of the urban heat island effect in New York City. Journal of Applied Meteorology. (1968) Aug;7(4):575-82.
2. Kijewski-Correa T, Kilpatrick J, Kareem A, Kwon DK, Bashor R, Kochly M, Young BS, Abdelrazaq A, Galsworthy J, Isyumov N, Morrish D.: Validating wind-induced response of tall buildings: Synopsis of the Chicago full-scale monitoring program. Journal of Structural Engineering. (2006) Oct;132(10):1509-23.
3. Brookhaven National Laboratory: Predicting urban and coastal microclimates. (2019) . Retrieved July 23, 2020, from https://www.bnl.gov/newsroom/news.php?a=213156
4. Jiang Y, Han X, Shi T, Song D.: Microclimatic Impact Analysis of Multi-Dimensional Indicators of Streetscape Fabric in the Medium Spatial Zone. In: International Journal of Environmental Research and Public Health. (2019) Jan;16(6):952.
5. Chapman S, Watson JE, Salazar A, Thatcher M, McAlpine CA.: The impact of urbanization and climate change on urban temperatures: a systematic review. Landscape Ecology. 2017 Oct 1;32(10):1921-35.
6. Yang, Q., Huang, X. and Li, J.: Assessing the relationship between surface urban heat islands and landscape patterns across climatic zones in China. Sci Rep 7, 9337 (2017). https://doi.org/10.1038/s41598-017-09628-w
7. New JR, Adams MB, Im P, Yang HL, Hambrick JC, Copeland WE, Bruce LB, Ingraham JA.: Automatic Building Energy Model Creation (AutoBEM) for Urban-Scale Energy Modeling and Assessment of Value Propositions for Electric Utilities. Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States); 2018 Jun 1.
8. Toparlar Y, Blocken B, Maiheu B, van Heijst GJ.: Impact of urban microclimate on summertime building cooling demand: A parametric analysis for Antwerp, Belgium. Applied Energy. (2018) Oct 15;228:852-72.
9. Arifwidodo S, Chandrasiri O.: Urban heat island and household energy consumption in Bangkok, Thailand. Energy Procedia. (2015) Nov 1;79(1):189-94.
10. Yang J, Jin S, Xiao X, Jin C, Xia JC, Li X, Wang S.: Local climate zone ventilation and urban land surface temperatures: Towards a performance-based and wind-sensitive planning proposal in megacities. Sustainable Cities and Society. (2019) May 1;47:101487.
11. Allegrini J, Dorer V, Carmeliet J.: Influence of morphologies on the microclimate in urban neighbourhoods. Journal of Wind Engineering and Industrial Aerodynamics. 2015 Sep 1;144:108-17.
12. Fu KS, Allen MR, Archibald RK. Evaluating the relationship between the population trends, prices, heat waves, and the demands of energy consumption in cities. Sustainability. 2015 Nov;7(11):15284-301.
13. Bhandari MS, Shrestha SS, New JR, Allen MR.: Comparison of Microclimate Simulated weather data to ASHRAE Clear Sky Model and Measured Data. Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States); (2017) Jul 19.
14. Dumas, Melissa, and New, Joshua.:Chicago microclimate and building energy use data. United States: N. p., (2020). Web. doi:10.13139/ORNLNCCS/1619243.
15. Oak Ridge National Laboratory: Chicago Loop. (2020) . Retrieved July 23, 2020, from https://evenstar.ornl.gov/autobem/chicago/
16. Chicago Data Portal. Retrieved July 23, 2020, https://data.cityofchicago.org/
17. Folium. Retrieved July 23, 2020, https://python-visualization.github.io/folium/
18. Google Map. Retrieved July 24, 2020, https://www.google.com/maps
19. Koordinates. Retrieved July 24, 2020, https://koordinates.com/