Exploring the Empirical Relationship Between Urban Form and Building Energy Use

Chi On Ho, Marco Miotti and Rishee Jain

September 22, 2023

Exploring the Empirical Relationship Between Urban Form and Building Energy Use

Chi On Ho[1], Marco Miotti[2], and Rishee Jain[3]

[1] Urban Informatics Lab, Stanford University. Email: eleanor.ho@stanford.edu
[2] Urban Informatics Lab, Stanford University. Email: mmiotti@stanford.edu
[3] Urban Informatics Lab, Stanford University (corresponding author). Email: rishee.jain@stanford.edu

ABSTRACT

Buildings account for 30% of global energy use and is expected to continue to increase as urbanization progresses. Among the various drivers, urban form is known to have a non-negligible effect on building energy consumption. Past studies have focused on modeling the physical obstructions and microclimate around the building. However, large-scale modeling is usually computationally prohibitive. In this paper, we aim to empirically evaluate how urban form affects energy use of residential and commercial buildings in Santa Clara, CA. We combined a large-scale building energy dataset with information on building properties and urban environment. We build various regression models and analyze the non-linearity and interaction effects between variables. Results show that the relationship between occupancy density and building energy use is non-linear and the monthly energy use intensity could drop by up to 50%, 18% and 17% for single-family residential, multi-family residential and commercial buildings respectively if we increase the density in surrounding areas. These results can assist urban infrastructure planning and policy-making professionals to make informed decisions regarding land-use and building decarbonization goals.

1 INTRODUCTION

Currently the buildings sector represents 30% of global final energy consumption and 28% of energy-related CO_2 emissions worldwide (Agency 2019). With the prospective rapid urbanization, global urban population is expected to increase from 56.2% in 2020 to 68.4% in 2050 (Nations 2018), resulting in a significant increase of total energy consumption over the globe. Drivers of building energy use can be categorized as internal (intra-building) and external (inter-building) factors. Internal factors include active energy consuming activities, including HVAC operation to satisfy human need for space heating and cooling, daily use of electric appliances, etc. and passive design of the building including building envelope, interior, materials, etc. (Ourghi et al. 2007; Smith et al. 2010; Sadineni et al. 2011; Chen and Hong 2018). External factors include climatological conditions such as solar radiation, ambient temperature, relative humidity, etc., and urban form, which entails surrounding physical characteristics such as floor coverage ratio, heights of adjacent buildings, street depth, etc. (Ko and Radke 2014; Silva et al. 2017; Li et al. 2018).

These physical characteristics of urban form impact building energy consumption by placing physical obstructions and interfering with the heat flow. It is of importance to understand the extent of its impact from the perspective of urban planning and policy-making given the long-term decarbonization and sustainability goals. Urban form has been defined and characterized in different ways, for example, urban morphology, urban configuration, urban texture, urban structure, etc. Past studies have shown that urban form plays a non-negligible role in building energy consumption (Krüger et al. 2010; Pisello et al. 2012; Han et al. 2017; Strømann-Andersen and Sattrup 2011) by posing physical obstruction and altering the microclimate around the building including airflow patterns, wind speed, solar radiation and convection, etc. (Ko and Radke 2014; Silva et al. 2017; Li et al. 2018; Cheng et al. 2006; Steemers 2003; Van Esch et al. 2012; Tereci et al. 2013; Asfour and Alshawaf 2015; Deng et al. 2016; Chen et al. 2017; Salvati et al. 2019).

Pisello et al. quantified the inter-building effect on building energy consumption by simulating the energy performance of a network of 20 single-family homes compared to a single building in Energy Plus (Pisello et al. 2012). Results demonstrate that depending on the climatological contexts and season, inter-building effect could substantially affect the accuracy of energy demand prediction (up to 42% in summer and up to 22% in winter) for space heating and cooling. Han et al. disaggregated and quantified the inter-building effect in terms of mutual shading and mutual reflection on building energy consumption and conducted case studies for a hypothetical network of a commercial building and two calibrated residential buildings in Italy (Han et al. 2017). They found that shading has a relatively larger individual impact compared to reflection within the building's microenvironment. Strømann-Andersen et al. used dynamic thermal and daylight simulations to model passive solar gains under different urban canyon scenarios (Strømann-Andersen and Sattrup 2011). The authors found that geometry of urban canyon, i.e., height/width ratio between adjacent buildings, has an impact on total energy consumption up to 30% for offices and 19% for housing compared to unobstructed sites in a north-European setting. Krüger et al. came to a similar conclusion by simulating the cooling load of a residential building in Israel that higher building and narrower street helps lower cooling demand (Krüger et al. 2010).

In this work, we aim to empirically evaluate the effect of density on typical building energy use using a large-scale dataset for both residential and commercial buildings in Santa Clara, CA. We consider the linear and non-linear effects and examine whether high density zoning may lead to non-linear increase in energy use due to conditions of the urban microclimate, such as heat island effects. We also examine if interaction effects exist between building floorspace and density.

2 METHODOLOGY

2.1 Description of the Datasets

We receive energy data from our collaborator SVCE. And we combine the dataset with publicly available data including tax assessor data and census data to enrich the variable space. We use the energy data as the predicted variable and all other building properties as predictor variables.

Energy Consumption Data. The energy dataset contains monthly energy consumption of 170,224 SVCE customers located in Santa Clara County, CA in 2018.

Tax Assessor Data and Census Data. Tax assessor data and census data are retrieved and simple calculations are performed to obtain the predictor variables. There are three building use types: 1) single-family residential, 2) multi-family residential, and 3) commercial. Continuous variables include: building floorspace in sqft, number of parcel units, year built, number of floors, building coverage ratio, land value, building value, residential and commercial occupancy density (number of people per km^2). Boolean variables include: AC, heat, BEV (battery electric vehicle), PHEV (plug-in hybrid electric vehicle), solar installation, and storage installation, indicating whether the customer possesses the respective items or not.

2.2 Complying with Privacy Regulations

Due to privacy regulations preventing raw energy consumption data at the building level being shared by SVCE for every building in Santa Clara county, we apply a data transformation algorithm to group the similar customers into clusters of size 20. The building properties and energy consumption are then averaged across all customers in each cluster. The detailed procedure is described in (Miotti and Jain 2022). Although the spikes in energy consumption data are smoothed, the result represents the average energy use for a typical customer of similar building properties.

2.3 Feature Engineering

Monthly energy consumption is divided by building floorspace to yield monthly energy intensity, which is the predicted variable of interest. Besides the abovementioned building properties, an additional predictor variable, month, is dummy encoded into 12 indicator variables to fit to the data so that the effect of each month could be quantified independently.

After examining the distribution of each feature, outliers are removed, and logarithmic transformation is applied to: 1) floorspace, 2) land value, 3) building value, 4) residential and commercial density, and 5) monthly energy intensity, in order to bring the feature distribution closer to normal distribution as well as to be able to better interpret the regression model coefficients in subsequent steps. Additionally, the feature domain of log(density) (and optionally log(floorspace)) is cut into 5 equally spaced buckets. And it is further converted as dummy variables by one-hot encoding so that we are able to test the non-linearity of the effect. At last, interaction effects between floorspace and density are included.

2.4 Modeling

Monthly data of the three building use types (single-family residential, multi-family residential, commercial) are separated and fitted to different models respectively. For each dataset, models of different feature selections are examined and compared based on the root mean squared error on the testing set (30% of all data). A final model per building use type is selected and presented in the following sections. Table 1 shows the variables used for the three monthly models.

The Python package **Statsmodels** (Seabold and Perktold 2021) is used to build the linear regression models in order to quantify the effects of the various variables that attribute to monthly energy intensity. Min-max scaling is applied to bring all predictor variables to the same scale from 0 to 1. Using the variables shown in Table 1 as predictors, the predicted variable

$$log(energy\ intensity) = const + \sum_{i \in U} \theta_i x_i$$

where θ_i is the coefficient for variable x_i and U is the full set of variables specified in Table 1.

3 RESULTS AND DISCUSSIONS

Due to space constraints, the fitted coefficients are not shown in this paper. Interpretations of the coefficients are provided below.

3.1 Single-Family and Multi-Family Residential Models

For the two residential models where log(floorspace) is modeled as continuous and log(density) is modeled as dummy variables, deriving from the fitted linear regression model, for a customer that

Table 1. Variables for monthly models. x = variable included. r = reference level excluded.

variable	variable type	building use type		
		single-family residential	multi-family residential	commercial
number of parcel units	continuous		x	x
year built	continuous	x	x	x
number of floors	continuous	x	x	x
building coverage ratio	continuous	x	x	x
log(land value)	continuous	x	x	x
log(building value)	continuous	x	x	x
AC	boolean	x	x	x
heat	boolean	x		
BEV	boolean	x	x	x
PHEV	boolean	x	x	x
solar	boolean	x	x	x
storage	boolean	x	x	x
month				
January	boolean	r	r	r
February	boolean	x	x	x
...	boolean	x	x	x
December	boolean	x	x	x
log(floorspace)	continuous	x	x	
log(density) residential / commercial	continuous			
log(density) group 0	boolean	r	r	
log(density) group 1	boolean	x	x	
...	boolean	x	x	
log(density) group 4	boolean	x	x	
interaction between floorspace and density				
log(floorspace) * log(density) group 0	continuous	r	r	
log(floorspace) * log(density) group 1	continuous	x	x	
...	continuous	x	x	
log(floorspace) * log(density) group 4	continuous	x	x	
log(floorspace) group 0 * log(density) group 0	boolean			r
log(floorspace) group 0 * log(density) group 1	boolean			x
...	boolean			x
log(floorspace) group 4 * log(density) group 4	boolean			x

log(density) of which falls under group n,

$$log(energy\ intensity) = const + \sum_{i \in U} \theta_i x_i = const + \sum_{i \in S} \theta_i x_i + (\alpha + \gamma_n) \cdot log(floorspace) + \beta_n$$

where U is the full set of variables specified in Table 1, S is the set of variables excluding log(floorspace), log(density) group 0...4, and the interaction effects between them, α is the coefficient for log(floorspace), β_i is the coefficient for log(density) group i, γ_i is the coefficient for interaction log(floorspace) * log(density) group i.

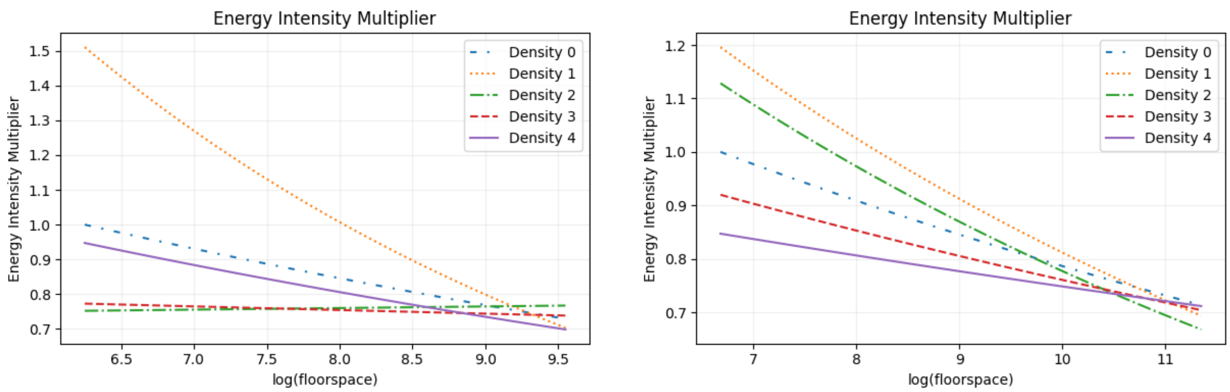
If all other variables in S are kept constant,

$$energy\ intensity = e^{const + \sum_{i \in S} \theta_i x_i + (\alpha + \gamma_n) \cdot log(floorspace) + \beta_n} = C \cdot e^{(\alpha + \gamma_n) \cdot log(floorspace) + \beta_n}$$

where $C = e^{const + \sum_{i \in S} \theta_i x_i}$.

Figure 1 shows the energy intensity multiplier $e^{(\alpha + \gamma_n) \cdot log(floorspace) + \beta_n}$ as a function of floorspace for five levels of residential density for single-family and multi-family residential models. Note that for the single-family residential model, each density level increase corresponds to the actual density increase by a factor of 4.25, while for the multi-family residential model, each density level increase means that the density increases by 2.59 times.

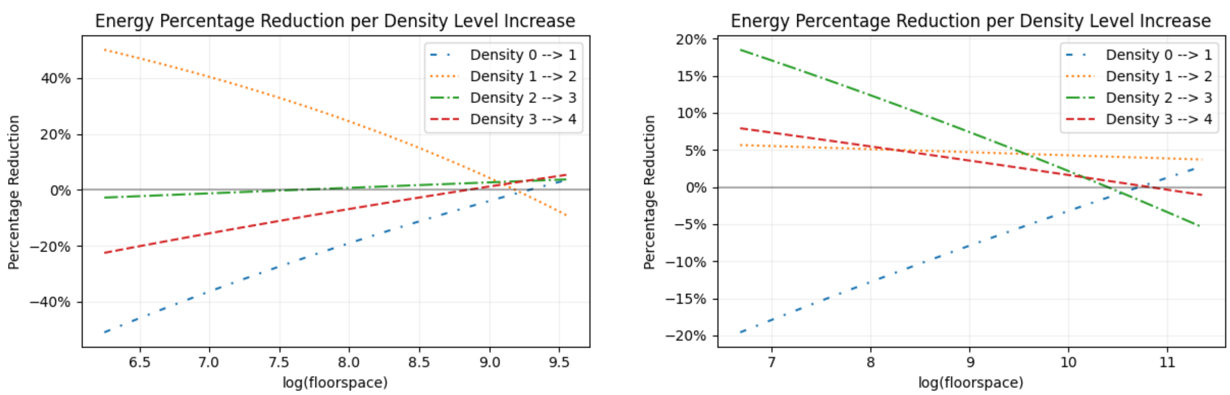
In general, energy intensity decreases as floorspace increases. For the largest density levels 3 and 4, the trend is less salient. For small floorspace, different density levels have high variations in energy intensity. From Figure 1 we can read the optimal level of density for any given floorspace.



(a) **(b)**

Figure 1. Energy intensity multiplier. (a) Single-family residential: density ranges are level 0 (8-36), level 1 (36-152), level 2 (152-648), level 3 (648-2757), level 4 (2757-11731) people per km^2; (b) Multi-family residential: density ranges are level 0 (100-260), level 1 (260-675), level 2 (675-1748), level 3 (1748-4528), level 4 (4528-11731) people per km^2.

Hypothetically, if we increase the density by one level, the percentage change in energy intensity is shown in Figure 2.



(a) **(b)**

Figure 2. Energy intensity percentage reduction. (a) Single-family residential: density ranges are level 0 (8-36), level 1 (36-152), level 2 (152-648), level 3 (648-2757), level 4 (2757-11731) people per km^2; (b) Multi-family residential: density ranges are level 0 (100-260), level 1 (260-675), level 2 (675-1748), level 3 (1748-4528), level 4 (4528-11731) people per km^2.

Single-family residential customers living in less dense areas (density level 1) could reduce energy by up to 50% through a densification process. Considering the energy reduction per capita, the benefit is even more drastic. As the density level increases from 1 to 2, the actual density increases by 4.25 times. Energy reduction per capita then becomes $1 - \frac{1-50\%}{4.25} = 88\%$. As shown in Figure 2a, this energy reduction diminishes as floorspace increases. For all other density levels 0, 2 and 3, it is only worth increasing density for very large building floorspace houses.

As shown in Figure 2b, for multi-family residential customers living in areas of density level 1, it makes sense for all customers to increase density to achieve an energy reduction of about 5% in average. When floorspace is approximately below $e^{10.5} = 36316$ sqft, customers of density level 2 and 3 should consider increase density to achieve a non-trivial energy reduction of up to 18% and 8% respectively. And the corresponding energy reduction per capita is $1 - \frac{1-18\%}{2.59} = 68\%$ and $1 - \frac{1-8\%}{2.59} = 64\%$ respectively.

3.2 Commercial Model

For the commercial model where both log(floorspace) and log(density) are modeled as dummy variables and their interactions are included, deriving from the fitted linear regression model, for a customer that log(density) of which falls under group n and log(floorspace) of which falls under group m,

$$log(energy\ intensity) = const + \sum_{i \in U} \theta_i x_i = const + \sum_{i \in S} \theta_i x_i + \gamma_{mn}$$

where U is the full set of variables specified in Table 1, S is the set of variables excluding interactions between log(floorspace) group 0...4 and log(density) group 0...4, γ_{ij} is the coefficient for interaction log(floorspace) group i * log(density) group j.

Therefore, if all other variables in S are kept constant,

$$energy\ intensity = e^{const + \sum_{i \in S} \theta_i x_i + \gamma_{mn}} = C \cdot e^{\gamma_{mn}}$$

where $C = e^{const + \sum_{i \in S} \theta_i x_i}$.

Table 2 shows the energy intensity multiplier $e^{\gamma_{mn}}$ for different levels of floorspace and density for the commercial model. The energy intensity multiplier for the reference level log(floorspace) group 0 * log(density) group 1 is 1. The emtpy fields indicate that there is no customer falling into those groups. For the commercial model, each density level increase corresponds to the actual density increase by a factor of 2.93, and each floorspace level increase corresponds to the actual floorspace increase by a factor of 3.67.

Table 2. Energy intensity multiplier. Commercial: density ranges are level 0 (42-123), level 1 (123-361), level 2 (361-1056), level 3 (1056-3090), level 4 (3090-9045) people per km^2. Floorspace ranges are level 0 (330-1210), level 1 (1210-4429), level 2 (4429-16220), level 3 (16220-59397), level 4 (59397-217510) sqft.

	log(floorspace)				
log(density)	group 0	group 1	group 2	group 3	group 4
group 0			0.7695	0.7740	
group 1	1.0000		0.6779	0.7634	
group 2	0.8286	0.8164	0.7362	0.7118	0.7470
group 3	0.9981	0.8601	0.7320	0.7896	0.7819
group 4	0.9897	0.8344	0.7177	0.7451	0.7210

If the level of floorspace stays the same, percentage change in energy intensity per density level increase is shown in Table 3.

From the interaction coefficients we can conclude that the relationship between log(floorspace) and log(density) is nonlinear. For example, for the lowest floorspace level 0, increasing density from level 1 to 2 results in a 17% energy reduction which is equivalent to $1 - \frac{1-17\%}{2.93} = 72\%$ energy decrease per capita, but increasing density from level 2 to 3 would have an adverse effect.

Table 3. Energy intensity percentage reduction. Commercial: density ranges are level 0 (42-123), level 1 (123-361), level 2 (361-1056), level 3 (1056-3090), level 4 (3090-9045) people per km^2. Floorspace ranges are level 0 (330-1210), level 1 (1210-4429), level 2 (4429-16220), level 3 (16220-59397), level 4 (59397-217510) sqft.

	log(floorspace)				
log(density)	group 0	group 1	group 2	group 3	group 4
group 0 → 1			11.9%	1.4%	
group 1 → 2	17.1%		-8.6%	6.8%	
group 2 → 3	-20.5%	-5.4%	0.6%	-10.9%	-4.7%
group 3 → 4	0.8%	3.0%	2.0%	5.6%	7.8%

In general, if the density level is relatively low for the corresponding floorspace level, increasing density by one level could help reduce energy.

4 CONCLUSIONS AND FUTURE WORK

In this paper, we empirically evaluate the impact of urban form on building energy intensity for single-family, multi-family residential and commercial buildings in Santa Clara, CA. A large-scale monthly energy data is combined with the publicly available tax assessor data and census data to quantify the effect of density.

Linear regression models are built for each building use type. Results show that density affects building energy intensity nonlinearly and there is synergistic interaction between density and building floorspace. Increasing density solely while keeping all other variables contant could result in energy reduction of up to 50%, 18% and 17% for the three building use types respectively. This could possibly incentivize urban planning professionals and policy-makers to make informed decisions regarding land-use and density regulations to achieve energy saving and decarbonization goals.

In terms of future work, one possible avenue is to further explore the relationship between building energy and density and possibly other building properties on a more granular temporal scale, for example, on an hourly basis. The proposed method could be applied on an hourly dataset to deepen our understanding of density on building energy use.

5 ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (NSF) under Grant No. 1941695, the UPS Endowment and the Center for Integrated Facility Engineering (CIFE). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF, UPS and/or CIFE. We would like to thank Silicon Valley Clean Energy (SVCE) for their assistance in curating and providing the data utilized in this study.

REFERENCES

Agency, I. E. (2019). "The critical role of buildings, <https://www.iea.org/reports/the-critical-role-of-buildings> (Apr).

Asfour, O. S. and Alshawaf, E. S. (2015). "Effect of housing density on energy efficiency of buildings located in hot climates." *Energy and Buildings*, 91, 131–138.

Chen, L., Hang, J., Sandberg, M., Claesson, L., Di Sabatino, S., and Wigo, H. (2017). "The impacts of building height variations and building packing densities on flow adjustment and city breathability in idealized urban models." *Building and Environment*, 118, 344–361.

Chen, Y. and Hong, T. (2018). "Impacts of building geometry modeling methods on the simulation results of urban building energy models." *Applied Energy*, 215, 717–735.

Cheng, V., Steemers, K., Montavon, M., and Compagnon, R. (2006). "Urban form, density and solar potential." *Report no.*

Deng, J.-Y., Wong, N. H., and Zheng, X. (2016). "The study of the effects of building arrangement on microclimate and energy demand of cbd in nanjing, china." *Procedia engineering*, 169, 44–54.

Han, Y., Taylor, J. E., and Pisello, A. L. (2017). "Exploring mutual shading and mutual reflection inter-building effects on building energy performance." *Applied Energy*, 185, 1556–1564.

Ko, Y. and Radke, J. D. (2014). "The effect of urban form and residential cooling energy use in sacramento, california." *Environment and planning B: Planning and Design*, 41(4), 573–593.

Krüger, E., Pearlmutter, D., and Rasia, F. (2010). "Evaluating the impact of canyon geometry and orientation on cooling loads in a high-mass building in a hot dry environment." *Applied energy*, 87(6), 2068–2078.

Li, C., Song, Y., and Kaza, N. (2018). "Urban form and household electricity consumption: A multilevel study." *Energy and Buildings*, 158, 181–193.

Miotti, M. and Jain, R. (2022). "A computationally efficient algorithm to enable privacy preserving urban energy data sharing under the "15/15" rule." *Energy*, 2004, 2965.

Nations, U. (2018). "World urbanization prospects, <https://population.un.org/wup/>.

Ourghi, R., Al-Anzi, A., and Krarti, M. (2007). "A simplified analysis method to predict the impact of shape on annual energy use for office buildings." *Energy conversion and management*, 48(1), 300–305.

Pisello, A. L., Taylor, J. E., Xu, X., and Cotana, F. (2012). "Inter-building effect: Simulating the impact of a network of buildings on the accuracy of building energy performance predictions." *Building and environment*, 58, 37–45.

Sadineni, S. B., Madala, S., and Boehm, R. F. (2011). "Passive building energy savings: A review of building envelope components." *Renewable and sustainable energy reviews*, 15(8), 3617–3631.

Salvati, A., Monti, P., Roura, H. C., and Cecere, C. (2019). "Climatic performance of urban textures: Analysis tools for a mediterranean urban context." *Energy and Buildings*, 185, 162–179.

Seabold, S. and Perktold, J. (2021). "Statsmodels: Econometric and statistical modeling with python." Statsmodels development team, <https://www.statsmodels.org/stable/index.html>.

Silva, M., Oliveira, V., and Leal, V. (2017). "Urban form and energy demand: A review of energy-relevant urban attributes." *Journal of Planning Literature*, 32(4), 346–365.

Smith, A., Luck, R., and Mago, P. J. (2010). "Analysis of a combined cooling, heating, and power system model under different operating strategies with input and model data uncertainty." *Energy and Buildings*, 42(11), 2231–2240.

Steemers, K. (2003). "Energy and the city: density, buildings and transport." *Energy and buildings*, 35(1), 3–14.

Strømann-Andersen, J. and Sattrup, P. A. (2011). "The urban canyon and building energy use: Urban density versus daylight and passive solar gains." *Energy and buildings*, 43(8).

Tereci, A., Ozkan, S. T. E., and Eicker, U. (2013). "Energy benchmarking for residential buildings." *Energy and Buildings*, 60, 92–99.

Van Esch, M., Looman, R., and de Bruin-Hordijk, G. (2012). "The effects of urban and building design parameters on solar access to the urban canyon and the potential for direct passive solar heating strategies." *Energy and Buildings*, 47, 189–200.