International Conference on Applied Energy 2021
Nov. 29 - Dec. 5, 2021, Thailand/Vatual
Paper ID: 828

The nonlinear influence of urban and building factors on residential building energy use: An empirical study with quantile regression in Seoul

Sujin Lee [1], Steven Jige Quan [1, 2,*]

1 City Energy Lab, Graduate School of Environmental Studies, Seoul National University, Seoul, 08826, South Korea

2 Environmental Planning Institute, Graduate School of Environmental Studies, Seoul National University, Seoul, 08826, South Korea

*Corresponding Author, E-mail: sjquan@snu.ac.kr | stevenchuan6656@gmail.com

ABSTRACT

The relationship between urban and building factors and building energy use is important for low-carbon city development. However, previous empirical studies often encountered the nonlinearity and non-normality issues common in complex urban datasets. This study examines the nonlinear influence of urban and building factors on building energy use under complex distribution conditions using the quantile regression model. This study focuses on Seoul's residential building electricity use in August 2017 and compares the quantile regression model and the ordinary least squares (OLS) model. The quantile regression results are generally in line with OLS results. However, considering the energy use distributions, the quantile regression results show that nonlinear influences of the urban and building factors on building energy use are strong at the right tail of the energy use distribution. Specifically, the positive relationship between coverage ratio and building energy seems to change rapidly above quantile 80%. The influence of distance to water body on building energy is insignificant in 25% and 50% quantile models, but it turns to a significantly negative effect at quantile 75% and above. Unlike the OLS results, no significant difference between the older adults ratio is found in all quantile models. This study suggests that the quantile regression reveals the nonlinear relationship between urban and building factors and building energy use, providing more detailed evidence for policymaking.

Keywords: building electricity use, quantile analysis, building density, urban factors, conditional distribution

1. INTRODUCTION

Urbanization increases energy demand and generates more emissions. Worldwide, cities account for over 75% of all primary energy usage and over 80% of direct and indirect greenhouse gas emissions [1].

Energy used in cities can be categorized into buildings, traffic, and industrial energy. Building energy use has a large share in the total urban energy use, e.g., 54% in South Korea [2]. To cope with buildings' increasing energy use, Seoul enacted the "Law of low-carbon green growth [3]" and implemented policies such as "Standard of building energy-saving design [4]." Urban planning and design have been receiving more attention in such an effort.

In urban planning and design, the relationship between urban and building factors and building energy use has been a rising topic for sustainable city development [5, 6, 7]. Empirical studies were conducted to reveal the relationship. Among those factors, urban and building density factors are increasingly examined [8, 9, 10], but there's still a debate on their influences [10]. Socio-economic factors such as income, population density, and education level were found to influence building energy in different ways [11, 12, 13]. Influences of environmental factors, such as green areas and water body, were also widely considered in the literature [14, 15, 16]. Many previous empirical studies examined the influence of urban and building factors based on ordinary least squares (OLS). The OLS regression model is efficient

Selection and peer-review under responsibility of the scientific committee of the 13th Int. Conf. on Applied Energy (ICAE2021).

and relatively straightforward for explaining the relationships between dependent and independent variables [17]. This model estimates minimizes squared errors to find the best-fit line [18, 19]. However, the OLS model has its strong assumption upon the joint distribution, which may be way more complex in real-world urban building energy data [20]. In practice, studies often used the transformation of variables to approximate the complex and nonlinear relationship, which is still far from enough in complex cases.

To better identify such complex relationships, there were a few explorations using other statistical models to examine determinate factors of building energy in the field. One type of effective model of such is the quantile regression model. The quantile regression model examines a dependent variable with different quantile levels and draws inferences about the values that rank below or above the conditional median and quantiles [21]. The use of quantile regression allows for examining questions of problems associated with assuring a normal distribution or filtering of skewness points [20]. Therefore, the quantile regression model seems to provide a complete picture of which part of building energy use is more susceptible to which factors [17, 20]. Studies using quantile regression stressed the importance of nonlinearly distributed errors, robustness against outliers, and the ability to detect heterogeneity of building energy use [22, 23]. However, few studies applied the quantile regression model to building energy at the urban scale to understand the nonlinearity in the influence of urban and building factors together.

This study aims to fill this gap by examining the influence of urban and building factors on residential building energy use in Seoul, South Korea, using the quantile regression model. The study estimated those influential factors' complex and varying influences considering different quantile levels. Unraveling such complexity and nonlinearity in the relationship between urban and building factors and building energy use can provide more detailed evidence for policy toward energy-efficient city development.

2. METHODOLOGY

2.1 Analysis scope and data collection

The study area is Seoul Fig. 1, the capital city of South Korea. Seoul has better data availability than other cities. The study period is August, a typical summer month in Seoul, in 2017. The data used in this study include buildings' energy, urban form and building, socio-economic, and temperature data. They were collected

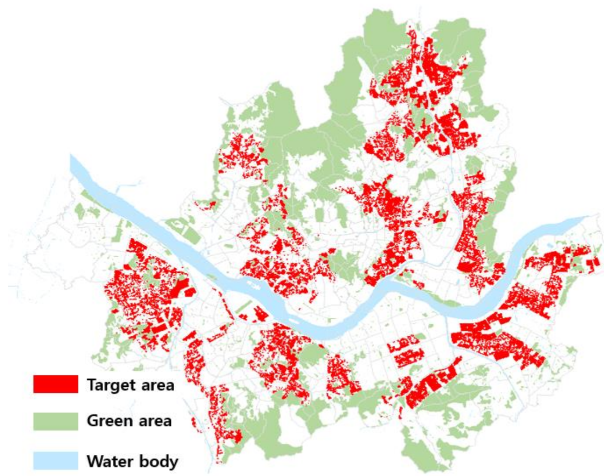

Fig 1. Target area

from the Ministry of Land, Infrastructure and Transport (MOLIT) [24], the Ministry of Interior and Safety (MOIS) [25], the city government website; and the United States Geological Survey (USGS) dataset [26].

2.2 Model specifications

In this study, quantile regression model is used to figure out the relationships with distinguishing 25%, 50%, and 75% quantiles of energy use. The OLS regression model is compared with the quantile regression results. The analysis uses the “qreg” command in STATA 17.

The dependent variable is defined as residential building energy use intensity (EUI) for electricity. To measure it, this study divided the electricity energy use by the total floor area at the parcel level.

The independent variables include the floor area ratio (FAR), the coverage ratio (CR), surface volume ratio (SVR), and building age as urban form and building indicators. Economic factors, namely the land price, were included. Demography factors were population density and older adults ratio. The normalized difference vegetation index (NDVI), the smallest distance to green space, and the smallest distance to a body of water were also included as environmental factors. The details are shown in Table 1.

3. RESULTS

The results of the quantile regression and OLS model are presented in Table 2. In the result of the quantile regression models, the relationships between dependent and independent variables were similar to those in the OLS results. For factors regarding urban form and building, FAR has negative effect whereas CR and SVR have positively affected the building energy use

Table 1. Variables

Division		Variables	Measurement	Source
Dependent variable		EUI	Ln(Electricity consumption / total floor area) (unit: kWh/m^2)	MOLIT
Independent variables	Urban form and Building	FAR	Total floor area / parcel area (unit: N/A)	MOLIT
		CR	Building footprint area / parcel area (unit: N/A)	MOLIT
		SVR	Surface Volume Ratio (unit: m^{-1})	MOLIT
		Building age	Average of building age within parcel (unit: Year)	MOLIT
	Economic	Land Price	Land price / parcel area (unit: 100,000 KRW/m^2)	Seoul
	Demography	Average occupant density	Total population / total floor area (unit: person/m^2)	Seoul
		Older adults ratio	Older adults population (over 65)/parcel area (unit: person/m^2)	Seoul
	Environment	Dis Water body	Distance to the nearest water body (unit: m)	MOIS
		Dis Green area	Distance to the nearest green area (unit: m)	MOIS
		NDVI	Normalized difference vegetation index: (Near infrared-red)/(Near infrared+red) (unit: N/A)	USGS

throughout all quantile regression models. However, building age has an insignificant effect in quantile 25%, whereas 50% and 75% have significant positive effects. In the economic part, the land price has a positive effect, which is already identified in the previous studies. In the demography part, population density has a positive effect on 50% quantile and older adults ratio is insignificant at all levels of the model. In the environmental part, the distance to a body of water was not significant in quantile 25% whereas in quantiles 50% and 75%, the distance to a body of water is negatively significant. The distance to a green area was positively significant, and NDVI was negatively significant for all levels. Pseudo R^2 is 5.84%, 8.81%, and 12.84% in quantile 25%, 50% and 75%, respectively.

As a result of the OLS model, all factors are shown to be significant. The results are similar to quantile regression models except for CR, the older adults ratio and distance to a body of water. CR has negative effects in all quantile models, whereas CR has a positive effect in the OLS model. The distance to a body of water has a significant relationship only in quantile 75%, however, it was shown to be significant in OLS models. The older adults ratio has an insignificant relationship in all quantile models but it was negatively significant in OLS models. The R^2 of the OLS model is 14.98%.

4. DISCUSSION

The graph results show nonlinear relationships across all quantiles of the building energy use (Fig. 2). This indicates that it is necessary to analyze and consider the differential impact of each factor by quantiles for building energy reduction.

In general, energy use and building density have a negative relationship. However, some studies argue that energy and urban form factors such as FAR and CR have positive relationships. The negative effect of an increase in the FAR on the building energy use is strong at the right tail of the energy use distribution. CR in the quantile regression models have positive effects, whereas CR in OLS have negative effects. Table 3 shows an additional result in the high quantiles. Fig. 2 and Table 3 show that CR has a negative effect on building energy use in both 90% and 95% quantiles. In addition, the coefficient decreases rapidly from around quantile 80%. The absolute value of the coefficient at quantile 95% is 0.245, which is a large difference when compared to that at quantile 75% (6.021). This shows that the influences are nonlinear according to the quantile level of energy use. However, this study cannot find the significance between CR and energy use in quantile models.

The older adults ratio shows a nonlinear relationships where the quantile 25% and 75% shows negative and the quantile 50%, 90% and 95% shows positive relationships.

Table 2. Result of quantile regression model

Variables	Quantile 25%		Quantile 50%		Quantile 75%		OLS	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
FAR	-.638***	.074	-1.687***	.107	-3.643***	.300	-4.746***	.183
CR	4.030***	.594	5.056***	.639	6.021***	.995	-4.119***	1.177
SVR	1.659***	.215	1.996***	.375	3.621***	.601	7.351***	.485
Building age	-.008	.005	.025***	.005	.085***	.014	.027**	.011
Land price	.100***	.003	.166***	.004	.240***	.007	.209***	.005
Population density	2.520	2.124	6.619*	3.883	47.294	33.217	11.793***	.920
Older adults ratio	-1.048	.938	.614	1.485	-.326	2.039	-5.510**	2.207
Dis Water body	.000	.000	-.000	.000	-.001**	.000	-.001**	.000
Dis Green area	.003***	.000	.005***	.001	.007***	.001	.007***	.001
NDVI	-10.480***	.811	-13.109***	1.059	-17.917***	.995	-26.523***	2.402
cons	1.154**	.401	1.891**	.796	4.817**	1.607	11.448***	1.106
Pseudo R^2	0.0584		0.0881		0.1284		R^2 0.1498	

* p<0.1, **p<0.05, ***p<0.01

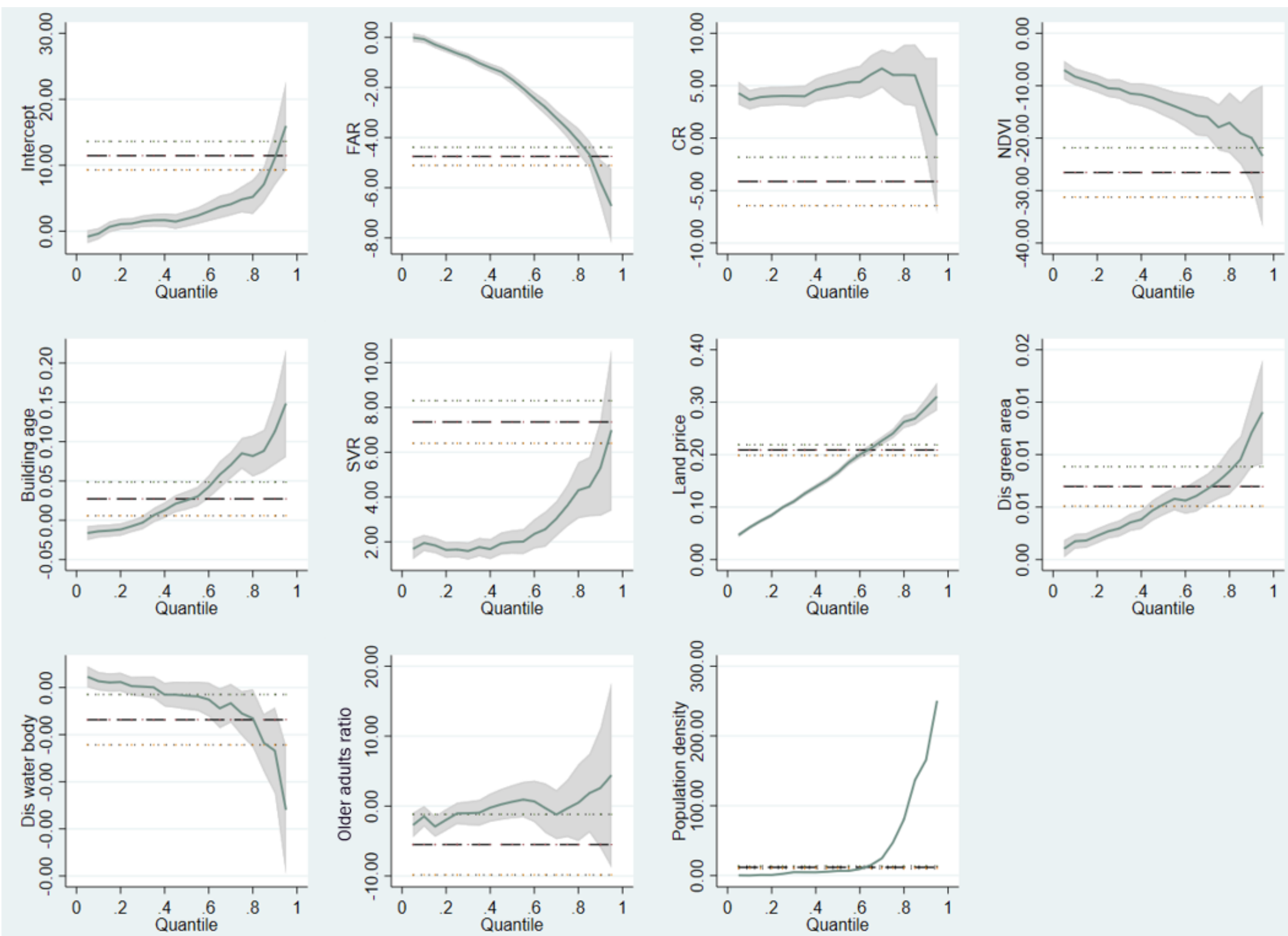

Fig 2. Graph of quantile regression model

Thus, the difference of coefficient is not patterned for each quantile. In addition, the older adults ratio is insignificant in all quantile models, whereas it is negatively significant in the OLS model. These results can be explained by the idea that older adults practice more energy-efficient behavior than younger people and are more vulnerable to heat in general. Nevertheless, to understand and explain the difference between energy use distributions is challenging. In population density, the coefficient increases rapidly from quantile 75%. The value of the coefficient at quantile 95% is 250.434, about 100 times larger than of quantile 25% (2.520).

Distance to a body of water has an insignificant relationship with building energy use in 20% and 50% quantile models, whereas quantiles 75%, 90%, and 95% models and OLS results are negatively significant. According to the results shown in Fig. 2, the distance to a body of water has stronger negative effect as it drops rapidly at the right tail of the building energy use distribution. Table 2 and Table 3 show the significantly negative effect in quantile 75% that switched the result from insignificant to significant. Residents prefer to live in areas with accessible environmental facilities such as parks and waterfronts [27, 28], and Seoul residents prefer residential locations near the Han River [29]. Residential houses close to environmental facilities have higher housing prices and use more energy because of high income residents.

Table 3. Additional result of high quantile part

Variables	Quantile 90%		Quantile 95%	
	Coef.	**S.E.**	**Coef.**	**S.E.**
FAR	-5.767***	.383	-6.727***	.550
CR	3.039	2.109	.245	3.685
SVR	5.300***	1.085	6.991***	1.645
Building age	.113***	.027	.149***	.040
Land price	.290***	.012	.311***	.018
Population density	165.507***	31.136	250.434***	40.026
Older adults ratio	2.591	3.787	4.418	5.559
Dis Water body	-.001**	.001	-.003***	.001
Dis Green area	.012***	.002	.016***	.002
NDVI	-19.938***	2.976	-23.403***	6.529
cons	11.105***	2.473	15.967***	4.026
Pseudo R^2	0.1839		0.2243	

* $p<0.1$, **$p<0.05$, ***$p<0.01$

5. CONCLUSION

The relationships between urban and building factors and building energy use are complex and nonlinear. Previous empirical studies mostly used the OLS or similar regression models to understand this topic. Still, they often suffered from the strong assumptions upon the distribution of variables in those models that may not apply to the urban building energy data. This study aims to fill this research gap by adopting the quantile regression model to examine the nonlinear influence of urban and building factors on building energy use under complex distribution conditions. This study focuses on the residential building electricity use in Seoul in August 2017 and compares the quantile regression and OLS models. In general, the relationships found in the two models are generally in line with each other; however, considering the energy use distributions, the results show that nonlinear influences of the urban and building factors are strong at the right tail of the energy use distribution.

Specifically, the positive relationship between coverage ratio and building energy seems to change rapidly above quantile 80%. The influence of distance to water body on building energy is insignificant in 25% and 50% quantile models, but it turns to a significantly negative effect at quantile 75% and above. Unlike the OLS results, no significant difference between the older and young adults groups is found in all quantile models. In sum, the findings show the effects of factors on building energy use more comprehensively and with more detail.

As Seoul plans to improve building energy efficiency to reach carbon-neutrality, this study can provide more detailed and helpful evidence for policymaking in urban development. The findings reveal the nonlinear relationships between urban and building factors and building energy use that depends on different quantile levels, which helps better understand those complex relationships and develop comprehensive urban energy policy.

This study still has a few limitations, including the incomplete data coverage and the lack of considerations of some influential variables. These issues will be addressed in future studies.

ACKNOWLEDGEMENT

This work was supported by the Creative-Pioneering Researchers Program through Seoul National University (SNU), the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science and ICT) (No. 2018R1C1B5043758), the National Research Foundation of Korea (NRF) grant funded by the

Korea government (Ministry of Education) (No. 5120200113713), and the Chungbuk Research Institute of Korea.

REFERENCE

[1] United Nations. World's population increasingly urban with more than half living in urban areas; 2014. Available:http://www.un.org/en/development/desa/news/population/world-urbanization-prospects-2014.html

[2] Korea Energy Economics Institute (KEEI). Report of energy consumption survey, MOTIE; 2015.

[3] Korean Law information center, accessed Nov 11. 2021,https://www.law.go.kr/ordinInfoP.do?ordinSeq=1275315

[4] Korean Law information center, accessed Nov 11. 2021,URL:https://www.law.go.kr/%ED%96%89%EC%A0%95%EA%B7%9C%EC%B9%99/%EA%B1%B4%EC%B6%95%EB%AC%BC%EC%9D%98%EC%97%90%EB%84%88%EC%A7%80%EC%A0%88%EC%95%BD%EC%84%A4%EA%B3%84%EA%B8%B0%EC%A4%80

[5] Steemers K. Energy and the city: density, buildings and transport. Energy and Buildings 2003;35:3-14.

[6] Ko Y. Urban form and residential energy use: A review of design principles and research findings. Journal of Planning Literature 2013;28:327-351.

[7] Hui S. Low energy building design in high density urban cities. Renewable Energy 2001;24(3-4):627-40.

[8] Quan SJ, Economou A, Grasl T, et al. Computing Energy Performance of Building Density, Shape and Typology in Urban Context. Energy Procedia 2014;61:1602-1605.

[9] Barley D, Deru M, Pless S, Torcellini P. Procedure for measuring and reporting commercial building energy performance. US department of Energy; 2005.

[10]Quan SJ and Li C. Urban form and building energy use: A systematic review of measures, mechanisms, and methodologies. Renewable and Sustainable Energy Reviews 2021; doi: 10.1016/j.rser.2020.110662.

[11] Strømann-Andersen J, Sattrup PA, The urban canyon and building energy use: Urban density versus daylight and passive solar gains 2011;43(8):2011-2020.

[12] Silva M, Oliveira V, Leal V. Urban forma and energy demand: A review of energy-relevant urban attributes. Journal of Planning Literature 2017;32(4):346-365.

[13] Wiedenhofer D, Lenzen M, Steinberger JK. Energy requirements of consumption: Urban form, climatic and socio-economic factors, rebounds and their policy implications. Energy Policy 2013;63:696-707.

[14] Aram F, García EH, Solgi E, Mansournia S. Urban green space cooling effect in cities. Heliyon 2019;5(4):e01339

[15] Farhadi H, Faizi M, Sanaieian H. Mitigating the Urban Heat Island in a Residential Area in Tehran: Investigating the Role of Vegetation, Materials, and Orientation of Buildings. Sustainable Cities and Society 2019;46:101448.

[16] Norton BA, Coutts AM, Livesley SJ, Harris RJ, Hunter AM, Williams NS. Planning for cooler cities: A Framework to Priorities Green Infrastructure to Mitigate High Temperatures in Urban Landscapes. Landscape and Urban Planning 2015;134:127-138.

[17] Roth J, Rajagopal. Benchmarking building energy efficiency using quantile regression. Energy 2018;152(1):866-876.

[18] Kontokosta C, Tull C. A data-driven predictive model of city-scale energy use in buildings. Applied Energy 2017;197:303-317.

[19] Currit N. Inductive regression: overcoming OLS limitations with the general regression neural network. Computers, Environment and Urban Systems 2002;26(4):335-353.

[20] Meng Q, Xiong C, Wu M, Ren X, Wang W, Li Y, Song H. Chang-point multivariable quantile regression to explore effect of weather variables on building energy consumption and estimate base temperature range. Sustainable Cities and Society 2020;53:101900.

[21] Geraci M. Modelling and estimation of nonlinear quantile regression with clustered data. Computational Statistics and Data Analysis 2019;136:30-46.

[22] Hannoudeh S, Nguyen DK, Sousa R. Energy prices and CO2 emission allowance price: A quantile regression approach 2014;70: 201-206.

[23] Kaza N. Understanding the spectrum of residential energy consumption: A quantile regression approach

[24] MOLIT, accessed Nov 11. 2021, https://open.eais.go.kr/main/main.do

[25] MOIS, accessed Nov 11. 2021, https://www.juso.go.kr/openIndexPage.do

[26] USGS, accessed Nov 11. 2021, https://www.usgs.gov/

[27] Schroeder H. Variations in the Perception of Urban Forest Recreation Sites. Leisure Sciences 1983;5(3):221–230.

[28] Feijten P, Hooimeijer P, Mulder CH. Residential Experience and Residential Environment Choice over the Life-Course. Urban Studies 2008;45(1):141-162.

[29] Fotheringhan AS, Park B. Localized Spatiotemporal Effects in the Determinants of Property Prices: A Case Study of Seoul. Applied Spatial Analysis and Policy 2018;11:581-598.

