Big-data analysis on energy consumption of office buildings in Seoul

K.U.Ahn[1], H.S.Shin[1] and C.S.Park[1,*]

[1]School of Civil, Architectural Engineering and Landscape Architecture,
SungKyunKwan University, Suwon, 16419, South Korea

ABSTRACT

Due to substantial energy consumptions in the building sector, a variety of political actions, policies (e.g. energy certifications, carbon emission regulations) and technical supports (e.g. passive and active technologies) have been conducted. One way to achieve energy efficiency for existing buildings is objective analysis on the features of actual energy consumption on the building sectors. This paper aims to investigate the actual energy consumption of office buildings (4,625 buildings) in Seoul based on national building energy database including location, building age, use, total floor area, the number of story and elevator, monthly energy consumption. In order to give an intuitive understanding of the features on the large building stock, this study conducted big-data analyses in three ways: (1) energy consumption of buildings according to their location, (2) relevance of building age to energy consumption, and (3) the correlations between energy consumption and given building information. The analyses were made based on the scatterplot matrix and density plot. The results of this study show that the degree of energy consumption was irrelevant to the building age. In addition, it can be concluded that the database should be supplemented to identify the dominant factor on the energy consumption.

KEYWORDS

Big-data, Data mining, Energy performance, Energy Consumption, Existing building

INTRODUCTION

The energy consumed by existing buildings accounts for more than 30% of global energy use (Wu et al. 2016). With the prospect that the building sector is and will continue to be a major energy consumer in the years ahead (Liu et al. 2015), the replacement rate of existing buildings by the new buildings is only around 1.0-3.0% per annum (Wu et al. 2016). Therefore, there has been awareness that retrofitting existing buildings is important to reduce the building energy consumption.

During the last several decades, a number of energy saving policies and certifications have been developed to improve energy efficiency in existing buildings: LEED and Energy Star in U.S.; BREEAM in U.K.; Energy Performance Certificate in E.U.; HK-BEAM in H.K. The International Energy Agency (IEA) also has conducted projects to improve energy efficiency of existing buildings: Annex 46, Annex 55, and Annex 56.

* Corresponding author email: cheolspark@skku.ac.kr

In addition, research studies based on technical measures (e.g. simulation for high-performance windows or insulations, model predictive controls, measurement and verification of energy efficiency) have been conducted to reduce and optimize the energy consumption in the existing buildings.

The aforementioned efforts have quantified the building energy performance using engineering methods based on their appropriate criteria such as normative condition to predict the energy consumption, benchmarking rating and labelling for energy certificates. However, Menezes et al. (2012) and de Wilde (2014) have reported the increasing concerns with regard to performance gap between predicted and actual energy consumption. On the contrary to the energy simulation model, using the actual empirical data may explain the factors (e.g. occupant behaviors, inappropriate controls, and uncertain indoor and/or outdoor environments) that are difficult to account for in the simulation based energy consumption.

Meanwhile, the government of South Korea released the national building energy dataset including monthly energy consumption and building information to the public. This paper aims to investigate the actual energy consumption of office buildings (4,625 buildings) in Seoul based on national database that includes location, building age, use, total floor area, the number of story and elevator, monthly energy consumption.

BUILDING ENERGY DATABASE WITH DATA MINING

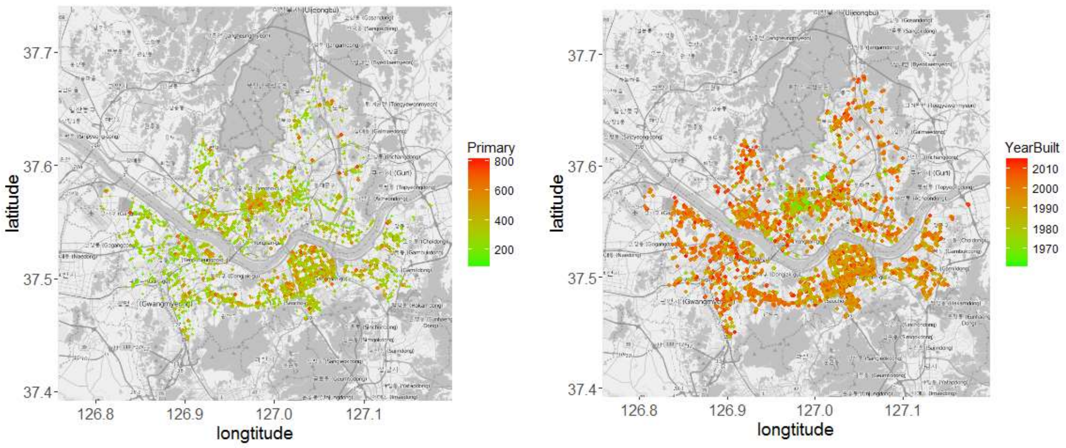
Data mining for big-data is an interdisciplinary field bringing together techniques from machine learning, pattern recognition, statistics, databases and visualization to address the issue of information extraction from large databases (Cabena et al. 1997). Several kinds of techniques can be used for data mining tasks such as classification, clustering, association rules, regression, generation and summarization, pattern based similarities (Han et al. 2012). It is worth noting that visual data exploration, which is to present the data in a visual form, allows the human to get intuitive insight or understanding in the data (Keim 2002). Therefore, the visualization for the big-data of existing buildings can be an effective way to capture the relationships between energy consumption and the building information.

The South Korean government released the a database (https://open.eais.go.kr, http://open.greentogether.go.kr/) which include information on building address, monthly electric and/or gas energy consumption, building name, building use, total floor area, building height, the number of households, the number of floors, the number of elevators, the number of parking lots, and building age. The database include 4,625 office buildings located in the Seoul.

ACTUAL BUILDING ENERGY CONSUMPTION DEPENDING ON DISTRICTS AND BUILDING AGES

Figure 1 shows the location of buildings in the database, their primary energy consumption per unit area, and year built. The energy consumption of buildings vary regardless of the location or districts (Figure 1 (a)). The districts which 75th quantile of yearly primary energy consumption exceeds the 400 kWh/m^2•year are Dongjak, Gangnam, Jongro, Junggu, Nowon, Seocho, and Songpa districts (Table 1). Meanwhile, green dotted circles, which mean that a building is built at about 1970, are mainly concentrated in certain district named Junggu (Figure 1 (b)). The statistic result (Table 1) shows the average of year built in Junggu district is 1986, which is 11 year older than those of Seoul.

According to the results shown in Figure 1 and Table 1, it is difficult to identify a relationship between the building age and energy consumption. For example, while the 25th quantiles are 1,971 (kWh/m^2•year) in the Junggu district and 1,991 (kWh/m^2•year) in the Gangnam district, yearly primary energy consumptions are similar 446.9 kWh/m^2 and 444.8 kWh/m^2 respectively (Table 1). Figure 2 shows the scatterplots to describe the relationship between year built and energy consumption. As shown in Figure 2, the age of a building and a level of energy consumption of a building seem to be irrelevant.



(a) Primary energy consumption (kWh/m^2•year) (b) Year Built

***Figure 1**.Distributions of energy consumption and year built*

***Table 1.** Statistic results of the yearly primary energy consumption and year built*

Name of district	Yearly primary energy consumption [kWh/m^2·year]					Year built				
	quantile			avg.	std.	quantile			avg.	std.
	25th	50th	75th			25th	50th	75th		
Dobong	242.2	277.8	335.2	294.5	89.5	1996	2003	2005	2001	7
Dongdaemun	205.6	277.1	341.5	286.7	106.5	1987	1994	2004	1996	11
Dongjak	263.5	324.6	**457.7**	368.4	135.9	**1990**	1997	2003	1997	10
Eunpyeong	231.9	276.0	327.2	296.6	99.0	2002	2003	2004	2002	8
Gangbook	266.2	327.8	377.0	333.1	105.7	1990	1997	2003	1996	10
Gangdong	250.5	302.1	386.9	337.5	134.5	1992	1998	2004	1998	8
Gangseo	225.7	274.2	324.5	288.1	91.3	1997	2002	2004	2001	7
Gangnam	290.7	361.3	**444.8**	374.7	122.5	**1991**	1995	2004	1997	9
Geumcheon	219.1	273.8	361.2	291.2	110.4	1996	2004	2007	2001	10
Guro	217.9	269.5	339.8	292.2	106.5	1991	2003	2008	2000	10
Gwanak	260.7	297.4	359.6	319.6	93.6	2002	2003	2004	2002	5
Gwangjin	209.3	272.2	348.4	300.4	117.2	1992	2000	2004	1999	9
Jongro	280.3	331.1	**424.3**	352.2	111.0	**1983**	1991	2002	1991	12
Junggu	286.3	363.8	**446.9**	373.7	136.2	**1971**	1986	1999	1986	15
Jungnang	257.1	284.4	343.2	312.8	99.8	1996	2003	2004	2000	7
Mapo	254.0	306.6	386.4	330.4	113.1	1991	2002	2004	1998	9
Nowon	269.2	345.1	**414.1**	366.8	135.2	1990	2000	2004	1998	7
Seocho	270.8	339.6	**416.6**	356.3	117.8	**1991**	1994	2003	1996	8
Seodaemun	261.2	309.6	369.1	325.0	110.9	1992	2003	2005	1999	10
Seongbuk	245.0	317.2	369.0	326.3	118.1	1988	2000	2004	1996	13

Seongdong	225.3	288.9	386.0	315.8	115.4	1990	1998	2006	1998	11
Songpa	259.6	332.3	**402.2**	347.3	120.5	**1991**	1994	2003	1996	7
Yangcheon	206.2	264.6	304.8	270.8	96.1	1996	2002	2004	2001	7
Yeongdeungpo	232.3	284.6	371.6	314.6	114.9	1991	2001	2005	1998	10
Yongsan	247.4	297.5	394.5	330.2	113.9	1988	1994	2004	1995	11
Total (Seoul)	256.8	318.1	400.3	339.5	119.1	1991	1998	2004	1997	10

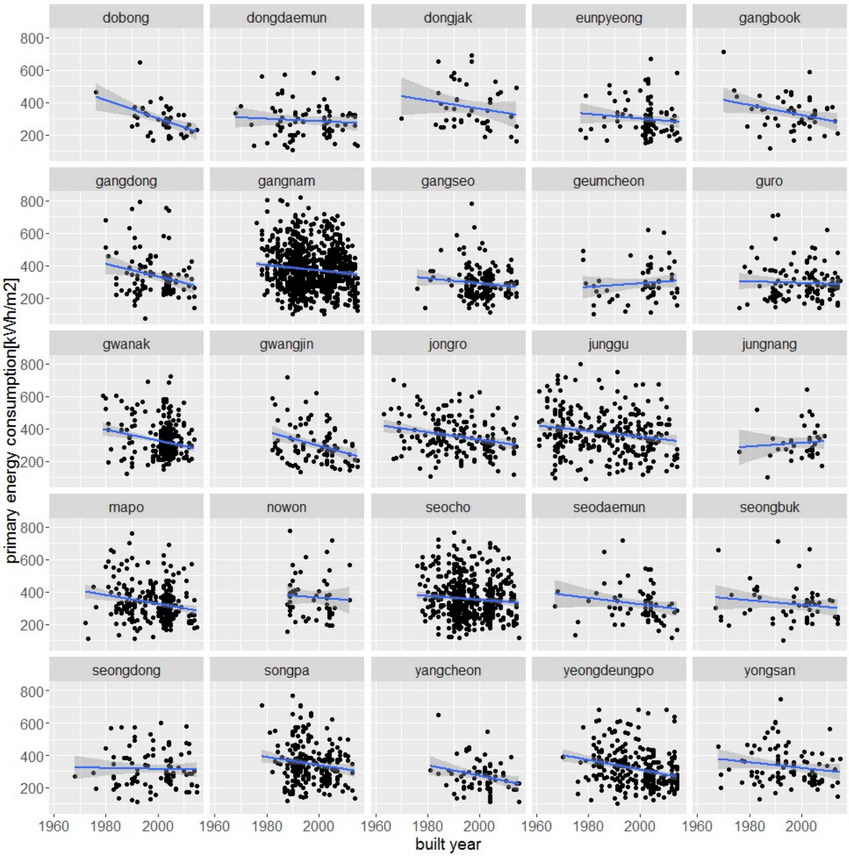


Figure 2*. Scatterplots for year built and primary energy consumption of the districts*

CORRELATION BETWEEN ENERGY CONSUMPTION AND BUILDING INFORMATION

Figure 3 shows the scatter plots among yearly primary energy consumption and building information in the constructed database. The data of yearly primary energy consumption are classified into 4 levels (0~250, 250~500, 500~750, over 750 kWh/m^2) and mapped to the points shown in the each plane of scatterplots. As shown in Figure 3, it is difficult to find a relationship among energy consumption, location, year built and building scales. In addition, 4 levels of energy consumption are mixed and/or overlaid in the scatterplot planes among the axis. It means that the degree of energy consumption could not be explained by the building information (District, YearBuilt, NumofStory, NumofElev, FloorArea)

Figure 4 shows a similar result mentioned above. The four probability density graphs are drawn in each plane, and each graph represents the probability density of the building information that correspond to the energy consumption classified into the four levels. It can be found that the degree of energy consumption is not relevant to the building information.

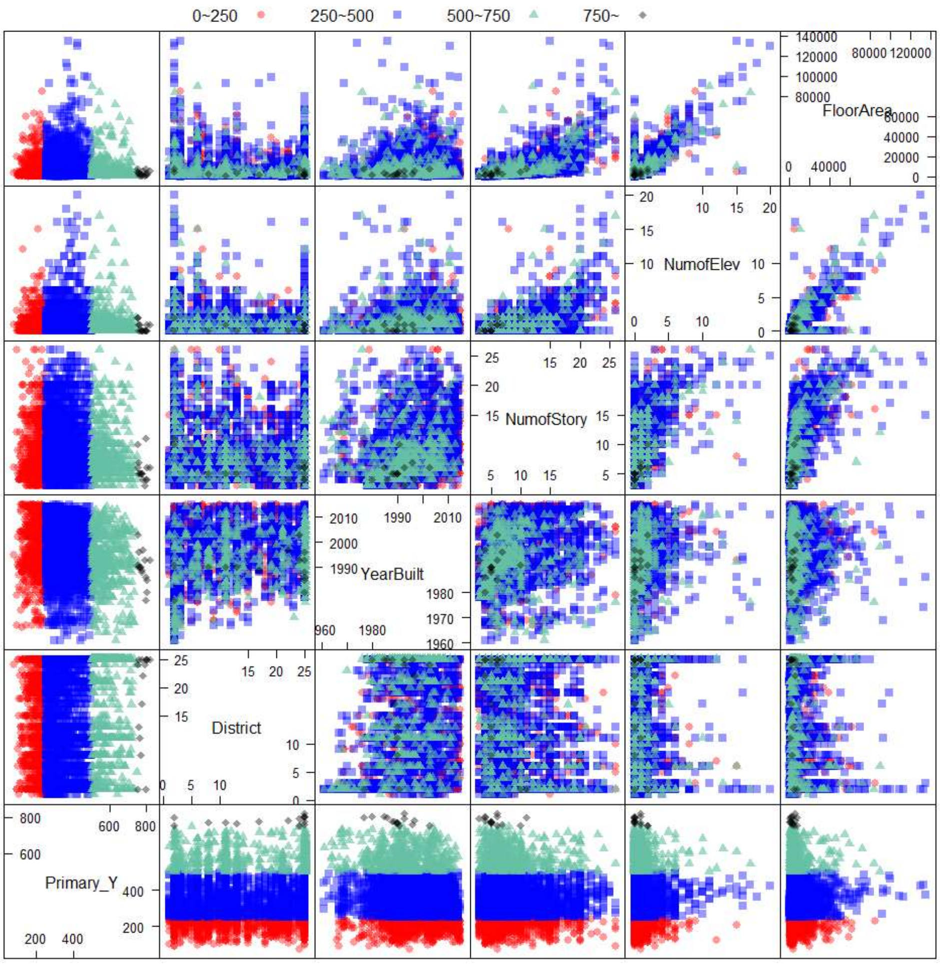


Figure 3. *Scatterplot matrix for yearly primary energy consumption and building information (Primary_Y: yearly primary energy consumption per unit area in 2015 [kWh/m²], District: index of 25 districts, YearBuilt: year built, NumofStory: the number of story, NumofElev: the number of elevator, FloorArea: total floor area [m²])*

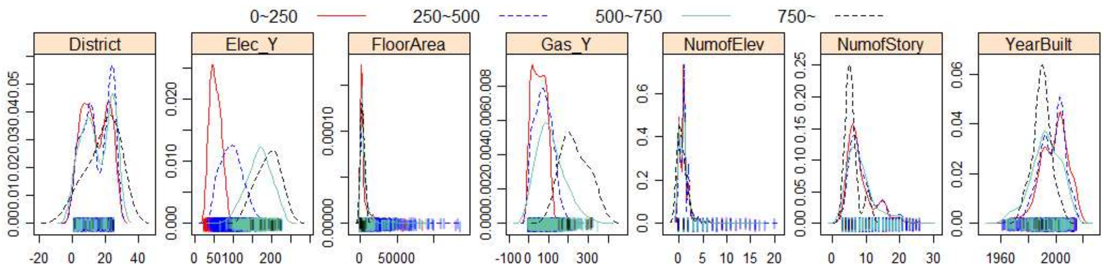


Figure 4. *Density plots of primary energy consumption with respect to the building features (Elec_Y: yearly electric energy consumption per unit area [kWh/m²], Gas_Y: yearly gas energy consumption per unit area [kWh/m²])*

CONCLUSION

This paper presented a big data analysis for national database with the visual data exploration. For this study, the energy database for office building located in Seoul was established based on the public data in South Korea. The visualized database gives the insights and better understanding with respect to the states of energy consumption in the existing buildings and the correlation between the energy consumption and building information.

The key finding in this study is that the relationship between the actual building energy consumption and building age is not strong. In other words, target buildings to be energy-retrofitted should not be determined based on the building age.

The data mining analysis enhanced with visual exploration converts implicit information out of the big-data to intuitive understanding with respect to the correlation between the energy consumption and the building information. The authors found that the correlation between energy use and the building information is not strong.

ACKNOWLEDGEMENTS

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning(KETEP) and the Ministry of Trade, Industry & Energy(MOTIE) of the Republic of Korea (No. 20152020105550).

REFERENCES

Cabena P, Hadjinaian PO, Stadler R, Verhees J, Zanasi A. 1997. Discovering data mining: From concept to implementation. Upper Saddle River, NJ: Prentice Hall.

de Wilde P. 2014. The gap between predicted and measured energy performance of buildings: A framework for investigation, Automation in Construction. 41, pp 40-49.

Keim DA. 2002. Information visualization and visual data mining, IEEE Transactions on Visualization and Computer Graphics. 7(1), 100-107.

Liu WL, Zhang JY, Bluemling B, Mol APJ. 2015. Public participation in energy saving retrofitting of residential buildings in China, Applied Energy. 147, pp 287-296.

Menezes AC, Cripps A, Bouchlaghem D, Buswell R. 2012. Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap, Applied Energy. 97, pp 355-364.

Wu Z, Wang B, and Xiac X. 2016. Large-scale building energy efficiency retrofit: concept, model and control, Energy. 109(15), pp 456-465.