Proceedings of the 2020 Winter Simulation Conference
K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, eds.

CLASSIFICATION OF BUILDING ENERGY PATTERN BASED ON RANSAC AND K-SHAPE ALGORITHM

Dong Hyuk Yi
Young Seo Yoo
Cheol-Soo Park

Department of Architecture and Architectural Engineering
College of Engineering, Seoul National University
1 Gwanak-ro, Gwanak-gu
Seoul, 08826, SOUTH KOREA

ABSTRACT

For efficient energy management of buildings, it is important to classify building energy pattern where dominant design variables could be identified. In this regard, the authors present a classification method of building energy pattern using RANSAC and K-shape algorithm. A reference office building was chosen and populated using the Sobol sequence method, resulting in 1,000 samples. Energy use data of 1,000 sample buildings were obtained from EnergyPlus, a dynamic building simulation tool developed by US DOE. The relationship between outdoor temperature and building energy use was selected as building energy pattern. It is shown that 1,000 buildings can be classified into two clusters (heating dominant vs. cooling dominant) and corresponding four architectural design variables.

1 INTRODUCTION

For efficient energy management of buildings, it is important to evaluate energy use characteristics of existing buildings. In this regard, the authors introduce a classification process of the building energy pattern based on RANSAC (Random sample consensus, Torr and Zisserman 2000) and K-shape algorithm (Paparrizos and Gravano 2015) (Figure 1). In this study, the relationship between daily outdoor temperature and energy use was analyzed as ‘building energy pattern’. RANSAC is a robust regression method that can select optimal data points by ‘survival of the fittest’, leading to minimizing disturbance by outliers. RANSAC can be used to extract individual building’s energy pattern from raw energy data. K-shape algorithm is a time-series clustering method that uses ‘shape-based distance’, or a cross-correlation between two sequences divided by the geometric mean of autocorrelations of individual’s sequences to minimize the scaling & shifting effects in measuring a shape similarity. This method is used in our study to cluster buildings by energy pattern.

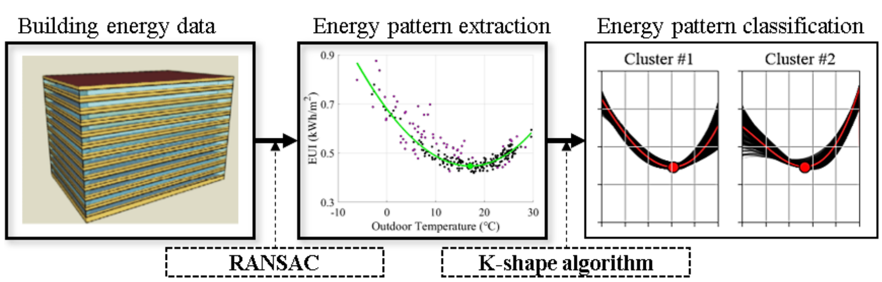

Figure 1: Classification of building energy pattern.

2 CASE STUDY

The EnergyPlus model for a large office type located in Atlanta developed by the US DOE was selected as the reference model. As design variables, Window's U-value, Solar Heat Gain Coefficient (SHGC), lighting power density (LPD), and air infiltration rate (ACH) were considered and the min-max ranges of the aforementioned four design variables are obtained from Yi et al. (2019). 1,000 sampling cases were populated using the Sobol sequence method where daily energy use data (electricity + gas) were obtained.

For RANSAC, the 3rd order polynomial function (the green solid line in the middle box of Figure 1) was applied as the basis function to account for the asymmetry of energy pattern according to the seasons (heating, cooling). To determine the number of clusters in the K-shape algorithm, silhouette scores between 2-11 clusters were measured, and the optimal number of clusters is 2. The red solid lines in Figure 2 show the representative pattern of each cluster. The characteristics of each cluster are summarized as follows:

- Cluster #1: heating energy use is dominant (U-value ↑, SHGC ↓, LPD ↓, ACH ↑)
- Cluster #2: cooling energy use is dominant (U-value ↓, SHGC ↑, LPD ↑, ACH ↓)

In addition, a logistic regression (linear classification) between design variables and energy pattern labels was performed to measure the accuracy of the proposed process. The results were as follows: (1) overall classification accuracy was 96.1%, (2) matching accuracy for cluster #1 was 97.6%, (3) matching accuracy for cluster #2 was 92.3%.

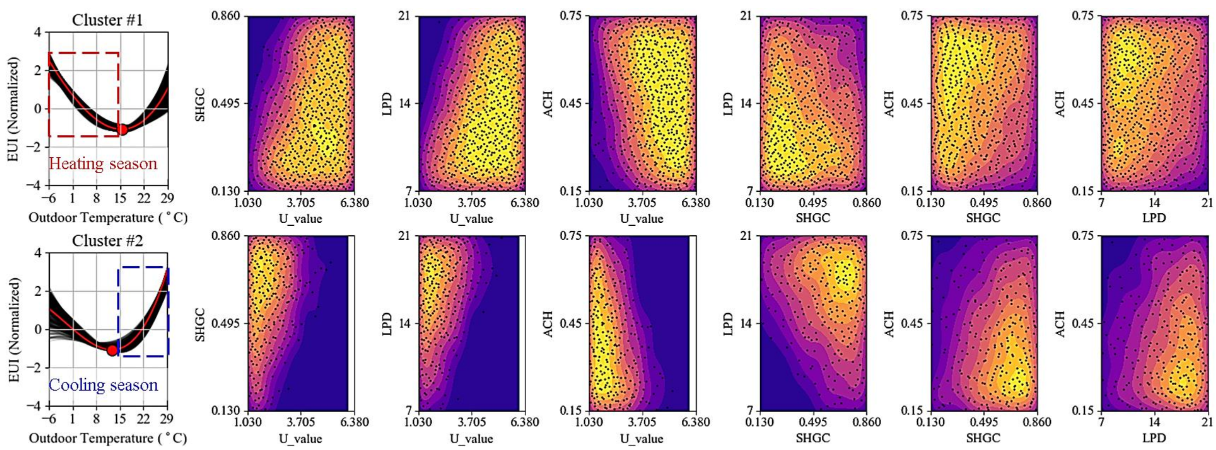

Figure 2: results of building classification (upper: cluster #1, lower: cluster #2, contour plots: densities).

ACKNOWLEDGMENTS

This work was supported by Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korea government (MOTIE) (20202020800030, Development of Smart Hybrid Envelope Systems for Zero Energy Buildings through Holistic Performance Test and Evaluation Methods and Fields Verifications)

REFERENCES

Paparrizos, J., and Gravano, L. 2015. "k-shape: Efficient and accurate clustering of time series". In *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, May 31th-June 4th, Melbourne, Australia, 1855-1870.

Torr, P. H., and Zisserman, A. 2000. "MLESAC: A new robust estimator with application to estimating image geometry. Computer vision and image understanding", 78(1), 138-156.

Yi, D. H., Kim, D. W., and Park, C. S. 2019. "Parameter identifiability in Bayesian inference for building energy models". Energy and Buildings, 198(1), 318-328.