

2025년 추계학술발표대회 : 일반부문

기계설비 대리변수를 활용한 학교 세부용도별 에너지 소비 패턴 이해

Understanding End-Use Energy Patterns in Schools

- Disaggregated EUI Analysis with Surrogate Features and Explainable AI -

○김 한 주*

Hanjoo Kim

허 연 숙**

Yeonsook Heo

Abstract

Energy consumption from buildings accounts for a significant portion of total energy consumption in developed countries. As end-use energy consumption plays a crucial role in prioritizing actions for energy reduction, it has gained increasing attention. This study focuses on 609 school buildings in metropolitan areas of Korea. Simplified energy disaggregation method was used to disaggregate end-uses and various potential influential factors were collected from public data. Random Forest regression models are developed, showing the predictability(R²) of 0.46, 0.51, 0.44 for total, base, and heating EUI. In base EUI, the top three important factors are gas consumption ratio, number of students and teachers per area, and register date. In heating EUI, the top three important factors were gas consumption ratio, register date, and district cooling and heating ratio. The trends between these key variables and the EUI predictions are identified, providing valuable insights for prioritizing energy reduction actions.

키워드 : 건물에너지 소비추정, 에너지 용도분리, 교육시설, 기계학습, SHAP 분석

Keywords : Energy consumption prediction, Energy disaggregation, School buildings, Machine learning, SHapley Additive exPlanations

1. 연구의 배경 및 목적

국제 에너지 기구(IEA)에 따르면, 건물에서 발생하는 에너지 소비는 선진국에서 CO₂ 배출량의 약 3분의 1과 총 에너지 소비의 40%를 차지한다. 이에 대응하여 서울시는 건물 부문의 제로에너지 전환을 촉진하고자 2050 넷제로 전략을 발표하였다. 이러한 맥락에서 건물 에너지 소비를 평가하고 에너지 소비 주요 인자를 식별하는 것은 에너지 절감 목표를 달성하기 위한 효과적인 전략을 수립에서 강조되고 있다.

이에 따라, 다양한 연구들에서 건물 에너지 소비 주요 인자 파악을 위한 데이터 기반 접근법이 수행되어 왔다(Hong, 2013; Cui, 2024). 건물 에너지 소비에는 건물 운영, 열적 특성, 건물 형상, 기계설비, 기후 등 여러 요인에 의해 영향을 받을 수 있음이 알려져 있으나, 일부 건물 정보는 공공 데이터로 수집하는 데에 어려움이 있어 데이터 기반 방법론에서 반영되지 못하는 실정이다. 특히, HVAC와 같은 기계설비 정보는 전체 에너지 소비 중 가스 소비

비율, 또는 1차에너지와 2차에너지의 비율을 통해 간접적으로 반영되었으며 건물에너지 소비에 유의미한 영향을 보임이 확인되었으나 해당 대리변수들과 기계설비 변수 사이의 관계는 제시되지 않았다(Cui, 2024; Won, 2019).

건물의 총에너지 소비강도는(EUI)는 건물 에너지 효율을 평가하기 위해 널리 사용되어 왔다(Trigaux, 2021). 그러나 전체 건물 EUI는 기저, 냉방, 난방에너지 소비를 구별하지 못하는 지표이므로, 총에너지 소비 중 해당 세부용도가 차지하는 비중은 총에너지 소비강도로 보여질 수 없다는 한계를 가진다. 건물에너지 소비를 세부용도(end-use) 수준에서 정량화하고 주요 인자를 식별하는 것은 에너지가 비효율적으로 사용되는 원인을 효과적으로 짚어내는 데에 필요하다, 건물의 세부용도별 에너지소비 정보를 수집하기 위해 센서를 설치하는 데에는 비용적인 문제점이 존재한다(Zhai, 2020). 이에 대한 대안으로 총에너지 소비로부터 세부용도별 에너지소비를 추정하는 다양한 에너지 용도분리 방법론이 개발되어 왔지만(Wortman, 1985; Kissock, 2002), 에너지 용도분리를 통해 얻어진 세부용도별 주요 인자를 식별하는 연구는 부족한 실정이다.

따라서, 본 연구의 목적은 용도별 EUI에 영향을 미치는 핵심 요인을 식별하고, 각 용도와 그 핵심 요인 간의 관계를 이해하는 것이다. 이를 위해 건물의 총에너지소비를 기저, 냉방, 난방 EUI로 분리하였고, 공공데이터로 얻어지는 다양한 건물 정보 및 기계설비 대리변수를 활용하여 머신러닝 기반 랜덤 포레스트 모델을 구축하였다.

* 고려대 건축사회환경공학과 대학원 석사과정

** 고려대 건축사회환경공학과 교수, 공학박사

(Corresponding author : Department of Civil Engineering, Korea University, yeonsookheo@korea.ac.kr)

본 연구는 국토교통부/국토교통과학기술진흥원의 지원으로 수행되었음(과제번호 RS-2023-00244769).

2. 방법론 및 입력자료

그림 1은 교육시설의 용도분리 에너지 소비에 영향을 미치는 핵심 요인을 분석하기 위한 프레임워크를 보여준다. 본 프레임워크는 (1) SED 방법을 이용해 교육시설의 월별 EUI를 용도분리하고, (2) 교육부(MOE)와 행정안전부(MOIS)가 제공하는 공개 데이터셋에서 용도분리 EUI에 잠재적으로 영향을 미칠 수 있는 요인을 수집·전처리하며, (3) 랜덤 포레스트 회귀 모델을 활용해 용도분리 EUI 예측 모델을 구축·검증하고, (4) SHAP 방법을 통해 핵심 요인을 분석하고 용도분리 EUI 예측에 미치는 영향을 해석하는 네 단계로 구성된다.

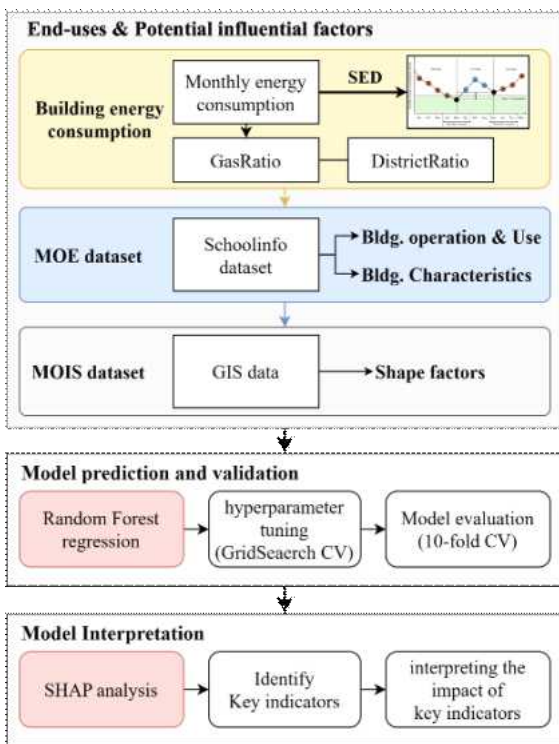


그림 1. 연구 방법론 도식

2.1 교육시설 용도분리 에너지소비량

본 연구는 세부용도별 에너지 소비 추정을 위해 서울 및 경기 지역 교육시설 에너지소비 데이터를 활용하였다. 2018년 서울과 경기 지역에 위치한 총 609개 교육시설(초등학교 375개, 중학교 157개, 고등학교 77개)의 에너지 소비 데이터는 한국부동산원으로부터 수집되었으며, 건물의 에너지소비 강도(EUI)는 이를 국토교통부(MOLIT)가 제공하는 용적률산정 연면적으로 나누어 계산되었다. 그림 2는 학교급별 총에너지 EUI를 비교한 그래프이다 비교하였다. 학교급 별 EUI 분포에서 유의미한 차이가 관찰되지 않았으며, 따라서 본 연구에서는 학교급의 구분 없이 전체 교육시설을 대상으로 분석을 진행하였다.

총에너지 소비강도(EUI)를 기저, 냉방, 난방으로 구분된 세부용도별 에너지소비로 분리하기 위해 Kim (2019)의 단순화 에너지 용도분리(Simplified energy disaggregation, SED)을 수정하여 활용하였다. SED 방법은 냉·난방 부하가 없을 것으로 기대되는 봄, 가을의 중간기의 건물에너지

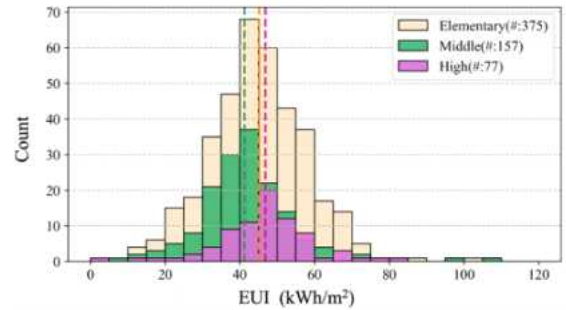


그림 2. 서울 및 경기 소재 초,중,고등학교의 EUI 비교

소비를 고정된 기저소비량으로 가정한다. 기저 소비량은 냉난방 부하가 없거나 적을 것으로 기대되는 봄(3-5월)과 가을(9-11월)의 기간에서 각각 가장 적은 월 에너지소비를 평균하여 계산된다. 냉방 및 난방 소비량은 각각 냉방기간인 여름과 난방기간인 겨울에서 기저소비를 초과하는 잉여 에너지소비량으로 계산된다. 교육시설의 경우 여름방학과 겨울방학 기간동안에는 축소 운영되므로 고정된 월별 기저소비가 과대추정될 가능성이 있어, 본 연구에서는 학사기간(AP) 동안의 기저, 냉방, 난방소비만을 고려하였다. 학사기간 동안의 전체 EUI는 방학을 포함한 연간 전체 대비 63% 수준으로 나타났으며, 학사기간 동안 기저, 냉방, 난방 EUI는 전체 EUI 대비 각각 70%, 27%, 3%수준으로 확인되었다. 교육시설의 냉방 EUI 비율이 낮은 수준으로 나타나므로 본 연구에서는 기저및 난방 EUI만을 분석 대상으로 선정하였다.

2.2 세부용도별 EUI 잠재적 영향 변수

본 연구에서 고려한 잠재적 영향 변수들은 국내의 공공 데이터를 기반으로 하며 표 1에 요약되어 있다. 잠재적 영향 요인은 (1) 건물운영, (2) 건물 열적특성, (3) 건물형상 (4) 그리고 기계설비 시스템 관련 특성의 네 분류를 포함하도록 고려되었다.

건물 운영 관련 변수는 교육부 ‘학교알리미’ (www.schoolinfo.go.kr)를 통해 수집되었으며 인원 수, 기숙사 학생 수, 특수 용도 공간 사용 여부가 포함된다. 인원 표1. 기저 및 난방 추정을 위한 예측 변수 10종

카테고리명	변수명	설명
건물운영	인원수 (m^3)	용적률산정면적 정규화
	기숙사 재실인원 (m^3)	용적률산정면적 정규화
	도서관 수	-
	수영장 수	-
건물 열적특성	강당 수	-
	설립년도(YYYY)	-
건물형상	컴팩트성 비율	동일면적 원 대비 건물폴리곤 둘레
	건물방향 ($^{\circ}$)	건축면적으로 가중평균 된 건물방향
기계설비	가스사용량 비율	기저/난방 EUI 중
	지역냉난방 비율	가스/지역냉난방 사용량 비율

수 및 기숙사 제실인원은 건물의 규모를 반영할 수 있어 EUI와 마찬가지로 용적률·단열면적을 활용해 정규화되었다. 건물 특성 변수에는 학교를 구성하는 건물들의 평균 설립연도가 고려되었다. 설립연도는 시기별 건축관련 법규의 변화와 및 벽체 노후화에 따른 침기·단열 효율 변화를 반영하여 건물의 열적 성능을 간접적으로 반영할 수 있을 것으로 기대된다.

건물 형상정보는 행정안전부의 GIS 자료(business.juso.go.kr)를 활용하여 건물 방향과 건물의 컴팩트성(compactness)의 두 측면에서 고려되었다. GIS 상 하나의 학교는 여러 개의 건물 폴리곤으로 표현될 수 있으며, 건물 방향의 경우 각 폴리곤에 대해 ArcGIS Pro 3.2에서 MBG(Minimum Bounding Geometry) 기법을 사용해 주향을 산정한 뒤(Liao, 2021) 건물의 건축면적을 활용해 가중평균 되었다. 컴팩트성 비율은 건물의 외피 노출이 많을수록 벽체를 통한 열교환이 증가하여 냉방 및 난방 에너지 소비가 증가할 수 있다는 점에서 고려되는 변수다. 컴팩트성 비율(compactness ratio)은 건물 폴리곤의 둘레와 동일 면적의 원의 둘레의 비로 계산되었다(Hong, 2013).

기계설비 변수는 각 세부용도 별 에너지소비 중 가스에너지 소비가 차지하는 비율로 계산되었다. 기저소비량에서의 가스소비량 비율은 급식에 의한 조리시설 규모를 상대적으로 반영하는 지표로, 난방 소비량에서의 가스소비량 비율은 전기 난방기기 대비 가스 난방기기의 상대적 비중을 나타낼 수 있을 것으로 기대되었다.

2.3 랜덤포레스트 모델 및 해석법

본 연구에서는 건물의 세부용도별 EUI를 추정하기 위해 랜덤 포레스트(random forest, RF) 회귀 모델을 적용하였다. RF 회귀 모델은 예측 변수들 간의 비선형성 및 상호작용 효과를 반영할 수 있다는 장점을 가져 건물에너지 소비 추정에 널리 활용되고 있다. RF 회귀는 다수의 결정트리(decision tree)를 배깅(bagging) 기반 앙상블로 결합한 기계학습 회귀 기법이다(Segal, 2004). 각 결정트리는 무작위로 추출된 표본과 변수 세트를 통해 독립적으로 학습되며, 개별 학습 샘플은 연속적인 노드의 구조(decision path)를 따라 하나의 리프 노드에서 예측값을 생성한다. RF 회귀 모델의 예측은 모든 트리 예측의 평균으로 산출되어 단일 트리에서 발생할 수 있는 과적합 위험을 완화하면서 복잡한 관계를 반영할 수 있다. 적합한 모델의 최적화를 위해 격자 탐색 교차검증(Grid search CV)을 수행하여 최적 하이퍼파라미터 조합을 탐색하였고, 10-fold 교차검증을 통해 모델 예측 성능을 평가하였다.

개발된 RF 회귀 모델에서 각 영향 인자와 세부용도별 EUI의 관계를 해석하기 위해 SHAP(Shapley Additive exPlanations) 분석을 적용하였다. SHAP은 게임이론인 Shapley 개념을 활용하여 각 변수의 기여도를 분리해 추정하고, 전체 학습 샘플에서의 누적 효과(총효과)를 정량화하는 모델 해석 기법이다(Lundberg, 2017). SHAP 분석은 변수별 기여도에 대한 가법성을 가정하므로, 특정 표본에 대한 모든 변수 총 효과의 합은 그 표본의 예측값과 같다.

3. 결론

3.1 랜덤포레스트 모델 평가

그림 4은 609개 학교 건물의 총, 기저, 난방 EUI 소비에 대한 RF 회귀 모델의 10-fold 교차검증 결과를 나타낸다. Grid search 교차검증 적용 이후, 예측력은 기저 EUI 추정 모델이 가장 높았고($R^2=0.51$), 그 다음이 총 EUI 예측 모델($R^2=0.46$), 그리고 난방 EUI 예측 모델($R^2=0.33$) 순이었다. 난방 EUI 예측 모델의 낮은 예측력은 직접적인 건물 열적 특성, 미기후, 또는 HVAC 종류 및 효율 관련 변수가 모델에 반영되지 않았기 때문일 수 있다.

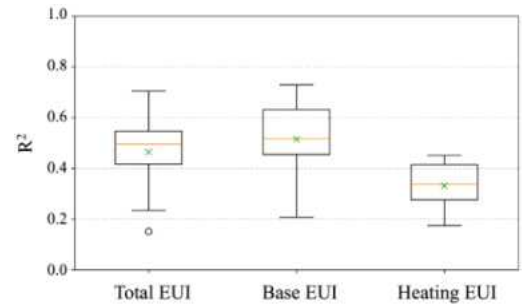
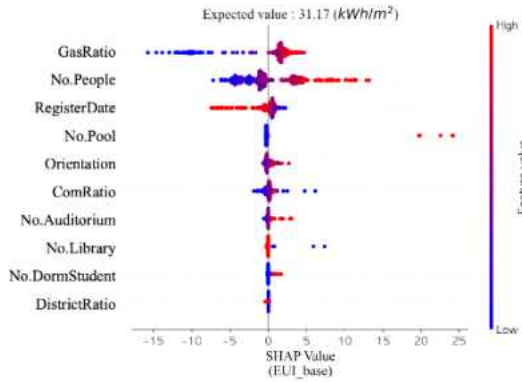


그림 3. 전체, 기저, 난방 모델의 예측 정확도

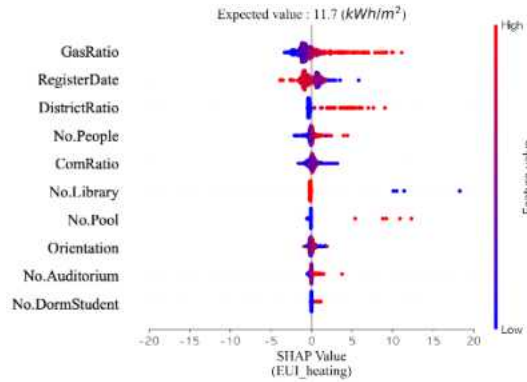
3.2 기저 및 난방 EUI 특성 분석

그림 5는 SHAP 방법론을 사용하여 분석한 기저 EUI와 난방 EUI 예측 모델의 주요 예측 변수를 특성 중요도 순서로 제시한 것이다. 그림에서 0에 해당하는 값은 해당 샘플의 예측값이 전체 샘플의 평균 기대값과 동일함을 의미한다. 각 점은 개별 학교 건물 샘플의 SHAP 값을 나타내며, 점의 x축 위치는 해당 변수가 예측에 미친 총효과, 점의 색상은 해당 변수 값의 상대적 크기를 의미한다. 예측 변수들은 전체 샘플에서 총효과의 절댓값의 평균으로 계산된 변수 중요도 순서로 정렬되므로, 상대적으로 위쪽에 위치한 변수일수록 더 높은 중요도를 가진다. 그림 5(a)에 따르면 평균 예측값이 31.17kWh/m² 일 때, 기저 EUI 예측 모델에서 가장 중요한 요인은 가스사용량 비율이었고, 다음으로 인원수, 설립연도, 수영장 수 순서로 중요하게 나타났다. 상위 네 개 요인의 중요도 합은 전체 중요도의 약 85%를 차지했으며, 그 이외의 변수들은 기저 EUI 예측에서 유의미하지 않은 것으로 나타났다. 특히, 가장 중요한 변수인 가스사용량 비율은 기저 EUI와 양의 경향을 보여, 총 에너지 소비 중 가스 소비 비율이 최소값(0%)에서 최대값(약 85%)으로 증가할수록 기저 EUI도 증가함을 확인할 수 있다. 이는 기저 부하에서의 가스 소비 비율이 급식 조리 활동에 의한 가스 사용 비중을 반영할 수 있음을 시사한다. 학습에 활용된 학교 건물들은 모두 급식 조리를 실시하는 건물로 확인되었는데, LNG 사용량은 공공데이터로 집계되지 않아 가스에너지원 비율이 0%인 경우는 LPG가 아닌 LNG를 사용하는 경우 누락되는 에너지 소비를 나타내는 것으로 예상된다.

난방 EUI 예측 모델에서는 가스사용량 비율, 설립연도, 지역난방 비율, 인원 수의 순서로 중요 변수로 나타났으며, 해당 변수들의 중요도 합은 전체의 86%로 나타났다.



(a) SHAP total effect of base EUI prediction



(b) SHAP total effect of heating EUI prediction

그림 4 교육시설 (a) 기저 및 (b) 난방 EUI의 변수 별 영향 분석

그림 5(b)에 따르면 평균 예측값이 11.7 kWh/m^2 일 때, 기저 EUI 모델과 유사하게 가스사용량 비율이 난방 EUI와 양의 경향을 나타냄을 확인할 수 있다. 난방에서의 가스 소비 비율은 전체 난방 시스템 중 가스 난방 설비 비중을 반영하며 학교시설이 사용하는 주요 냉난방 기기가 전기 히트펌프(EHP), 가스히트펌프(GHP), 가스보일러라는 점을 고려했을 때, 학교 건물에서 가스 난방 시스템의 비중이 높을수록 에너지소비비가 비효율적일 수 있음을 시사한다.

본 연구는 국내 학교 건물의 기저 EUI와 난방 EUI를 예측하기 위한 모델을 개발하고, 기저 EUI 및 난방 EUI와 주요 변수들 간의 관계를 해석하였다. 가스사용량 비율은 조리 및 가스난방 시스템의 비중을 설명하는 대리 변수로서 기저 EUI와 난방 EUI 모두에서 가장 중요한 요인으로 나타났으며 두 지표 모두와 양의 관계를 보였다. 인원수, 설립년도 등의 변수 또한 주요한 변수로 확인되었다. 이러한 결과는 동일한 변수라도 건물 세부용에 따라서 다른 역할과 영향을 가질 수 있음을 보여주며, 세부용도 별 에너지소비 추정을 활용한 건물에너지 소비절감 전략의 설립에 유용한 정보를 제공할 수 있음을 시사한다.

참고문헌

1. Cui, X., et al., Energy consumption prediction and household feature analysis for different residential building types using machine learning and SHAP: Toward energy-efficient buildings. *Energy and Buildings*, 2024. 309: p. 113997.
2. Hong, S.-M., et al., A comparative study of benchmarking approaches for non-domestic buildings: Part 1 - Top-down approach. *International Journal of Sustainable Built Environment*, 2013. 2(2): p. 119-130.
3. International Energy Agency, Total Energy Use in Buildings Analysis and Evaluation Methods, International Energy Agency, 2013.
4. Kim, D.W., Y.M. Kim, and S.E. Lee, Development of an energy benchmarking database based on cost-effective energy performance indicators: Case

- study on public buildings in South Korea. *Energy and Buildings*, 2019. 191: p. 104-116.
5. Kissock, J.K., J.S. Haberl, and D.E. Claridge, Development of a toolkit for calculating linear, change-point linear and multiple-linear inverse building energy analysis models, ashrae research project 1050-rp, final report. 2002, Energy Systems Laboratory, Texas A&M University.
6. Liao, W., Hong, T., & Heo, Y. (2021). The effect of spatial heterogeneity in urban morphology on surface urban heat islands. *Energy and Buildings*, 244, 111027.
7. Lundberg, S.M. and S.-I. Lee, A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 2017. 30.
8. Ryu, S. H., Ryu, J. H., & Hong, W. H. (2016). Field Measurement and Degradation Rate Analysis of Thermal Insulation Performance by Elapsed Year of Public Building Outer Walls in Korea.
9. Segal, M.R., Machine learning benchmarks and random forest regression. 2004.
10. Trigaux, D., Allacker, K., & Debacker, W. (2021). Environmental benchmarks for buildings: a critical literature review. *The international journal of life cycle assessment*, 26, 1-21.
11. Won, C., S. No, and Q. Alhadidi, Factors affecting energy performance of large-scale office buildings: Analysis of benchmarking data from New York City and Chicago. *Energies*, 2019. 12(24): p. 4783.
12. Wortman, D. and C.B. Christensen, Variable-base degree-day correction factors for energy savings calculations. 1985, Solar Energy Research Inst., Golden, CO (USA).
13. Zhai, Z.; Salazar, A. Assessing the implications of submetering with energy analytics to building energy savings. *Energy Built Environ*. 2020.