BS 2019 2-4 SEPT ROME
16th IBPSA INTERNATIONAL CONFERENCE AND EXHIBITION

INTERNATIONAL BUILDING PERFORMANCE SIMULATION ASSOCIATION

A Robust Unsupervised Framework for High-Resolution Building Energy Consumption Profiling

Speaker:

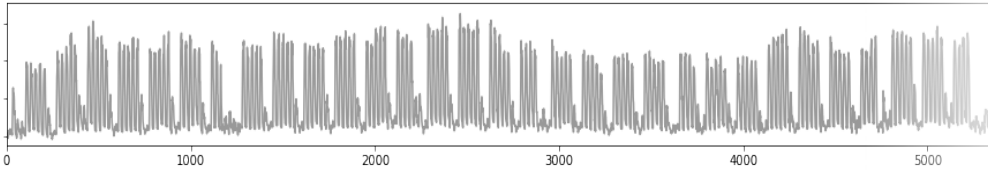
Sicheng Zhan

Authors:

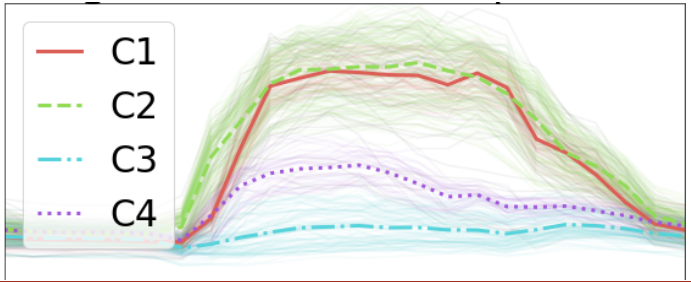
Sicheng Zhan, National University of Singapore
Adrian Chong, National University of Singapore

About building energy consumption profiling

- To extract typical usage patterns from smart meter data
- Provide insights for multiple application
  - Abnormal operation detection
  - Customer classification
  - Schedule inference for BEM
  - Partition period for prediction
- Unsupervised clustering
  - Buildings vary; no ground truth

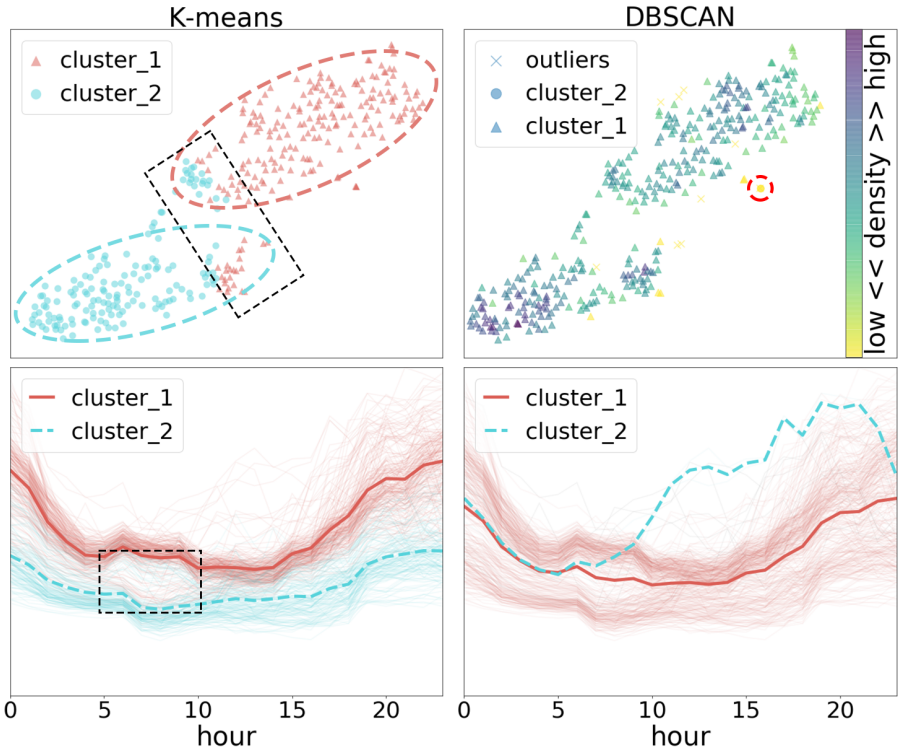


Profiling



A closer look on clustering methods

- K-means
  - Requires a presumption on the cluster number
  - Assumes clusters with spherical variance and similar size
- DBSCAN
  - Requires harder parameter tuning (Epsilon and MinPt)
  - Accepts one density threshold

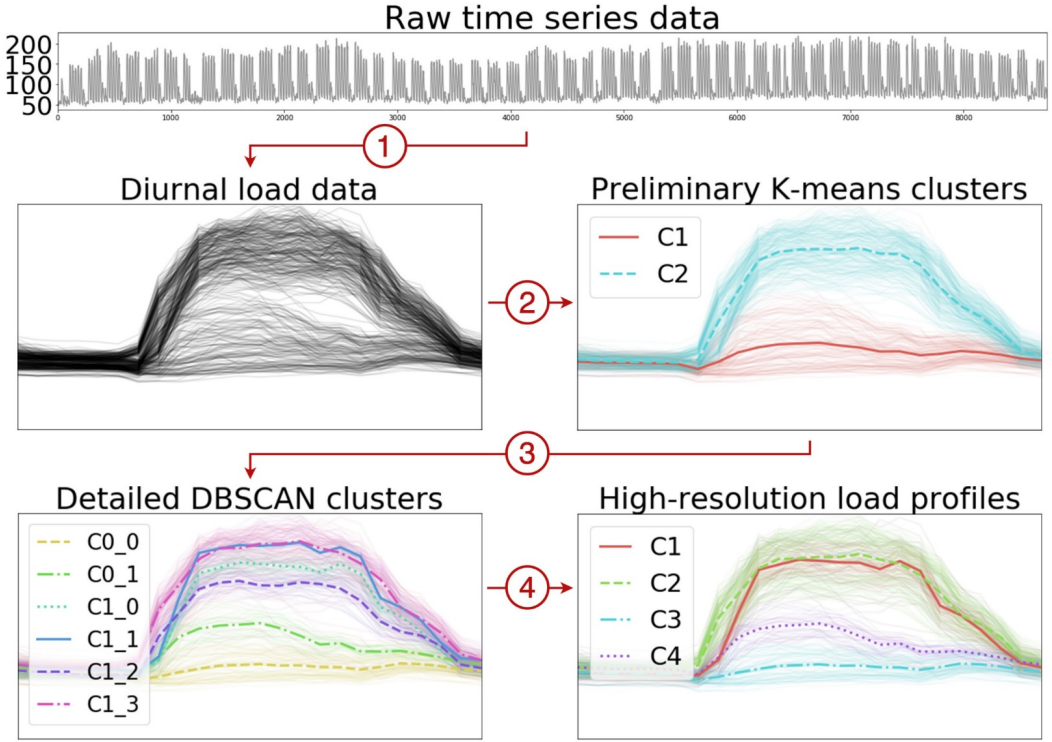


Towards a more robust framework

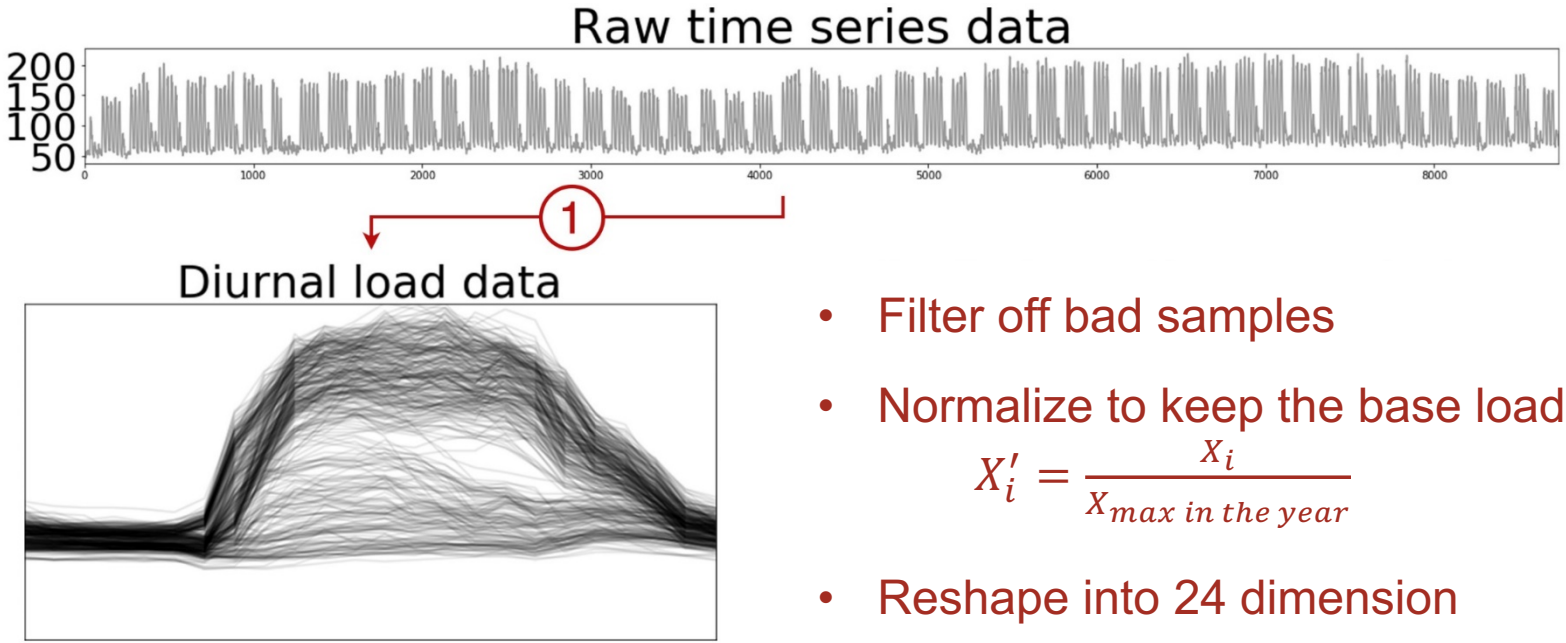
- The profiling results are not always reliable for further application
- The need of human interference makes existing methods not really unsupervised
- An desired framework should be:
  - Able to identify the accurate profiles without human interference
  - Robustly applicable to different buildings

The proposed framework: overview

- Pre-processing
- Preliminary K-means
- Detailed DBSCAN
- Post-processing



The proposed framework: Pre-processing

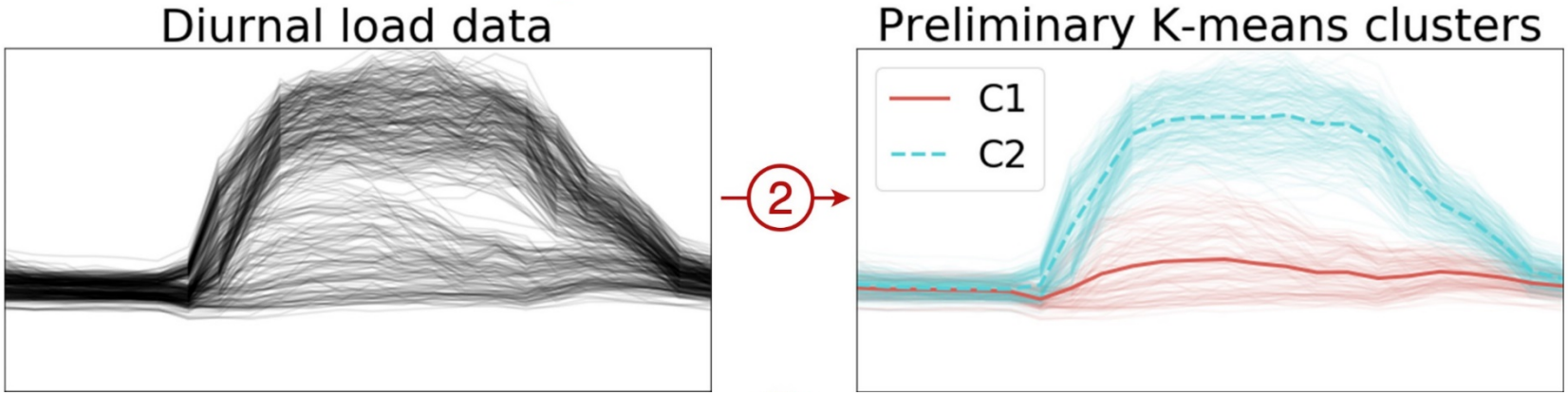


- Filter off bad samples
- Normalize to keep the base load

$$X_i' = \frac{X_i}{X_{max\ in\ the\ year}}$$

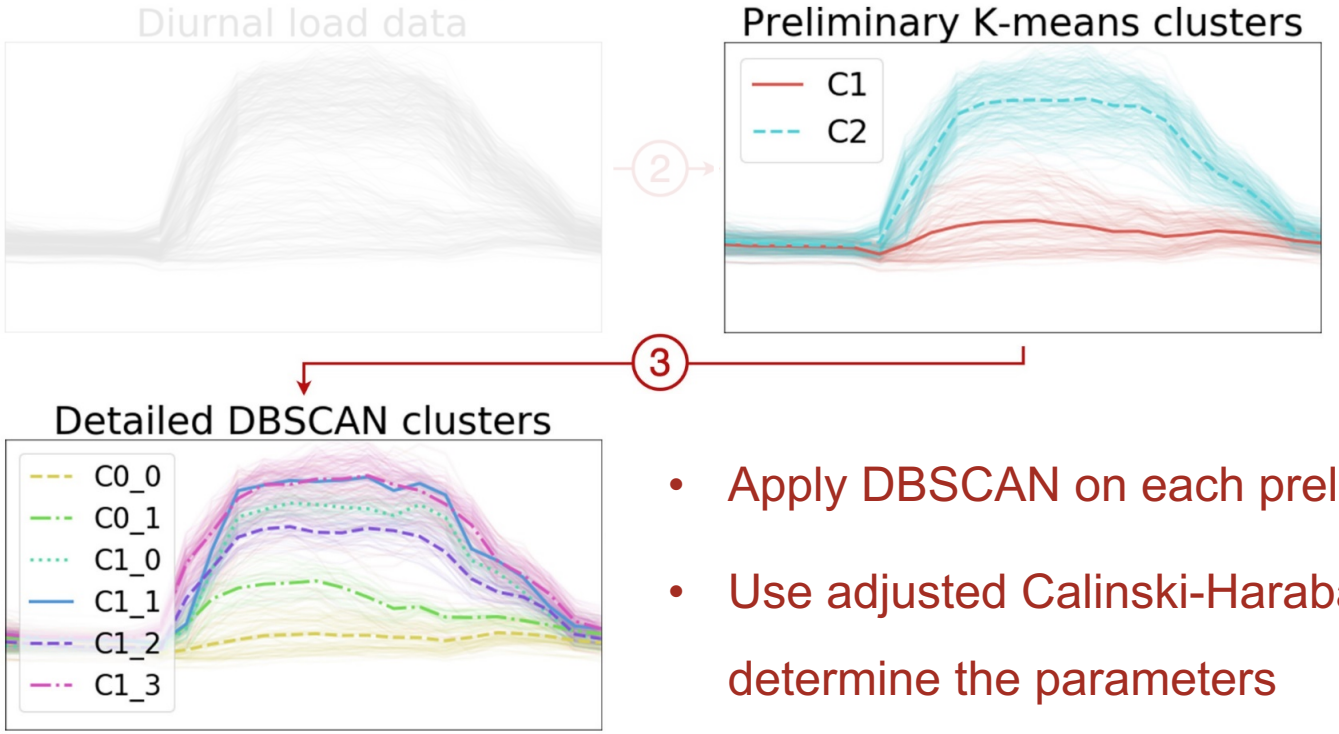
- Reshape into 24 dimension

The proposed framework: K-means



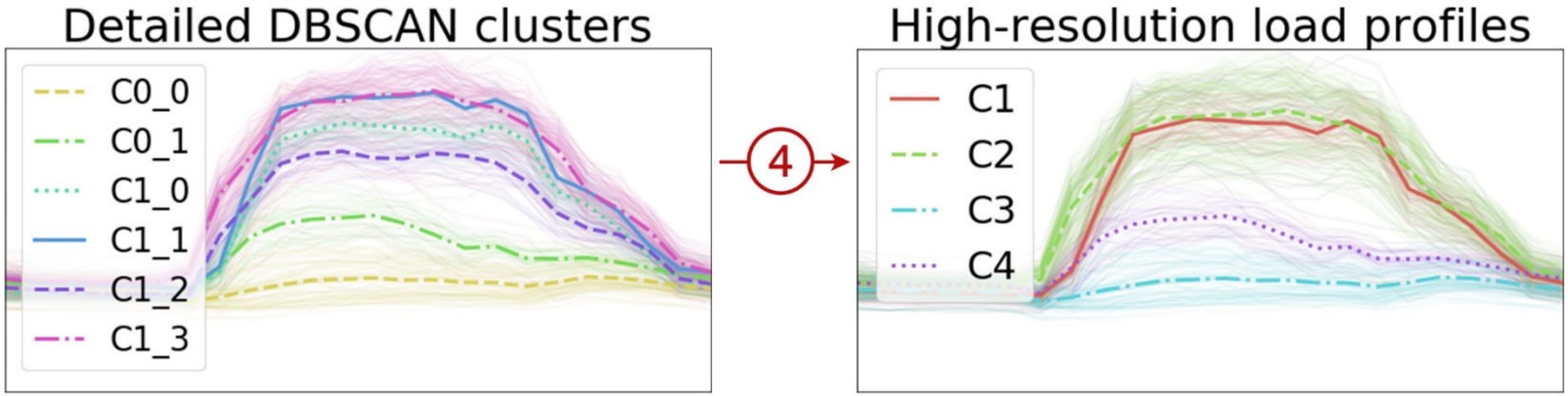
- Apply K-means for preliminary clustering
- Use Calinski-Harabasz index to determine the cluster number K

The proposed framework: DBSCAN



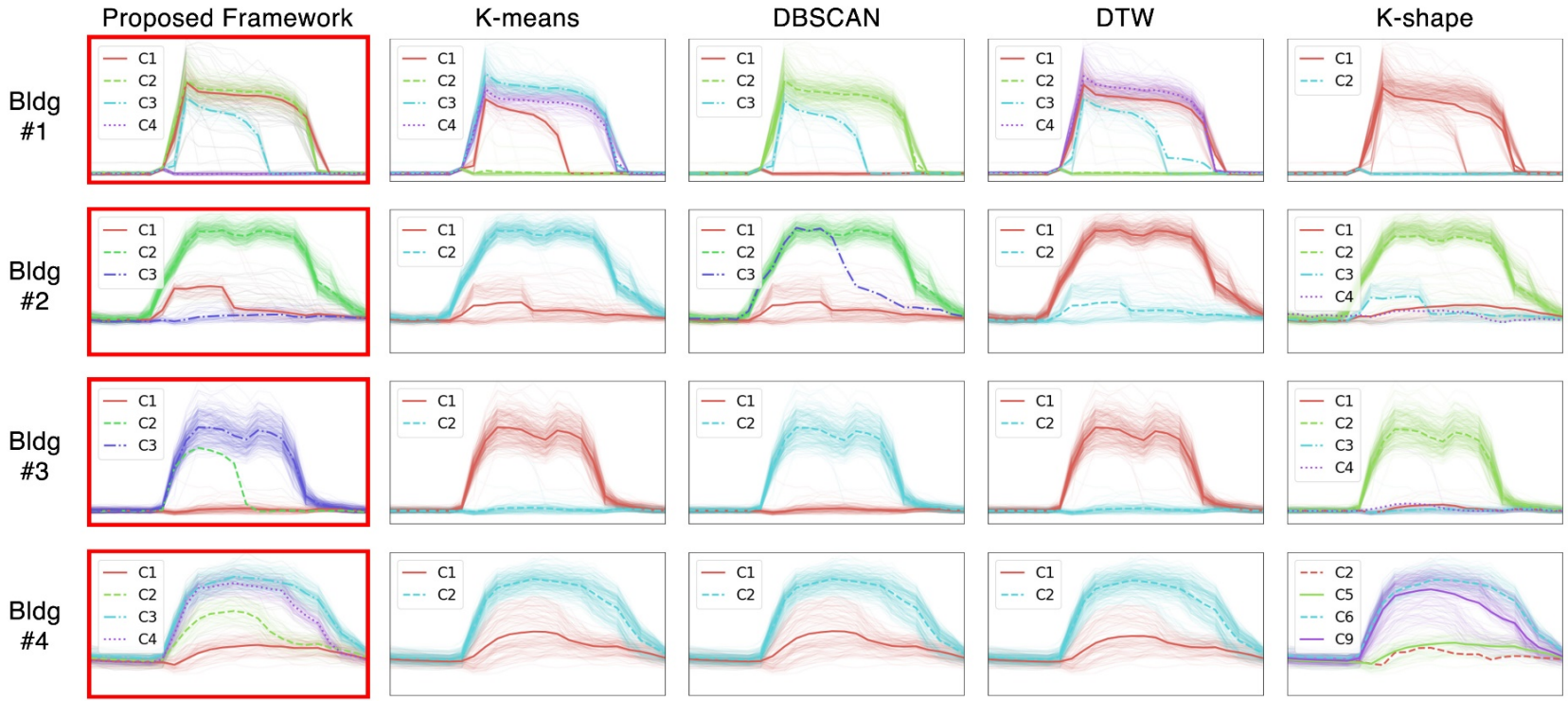
- Apply DBSCAN on each preliminary cluster
- Use adjusted Calinski-Harabasz index to determine the parameters

The proposed framework: Post-processing

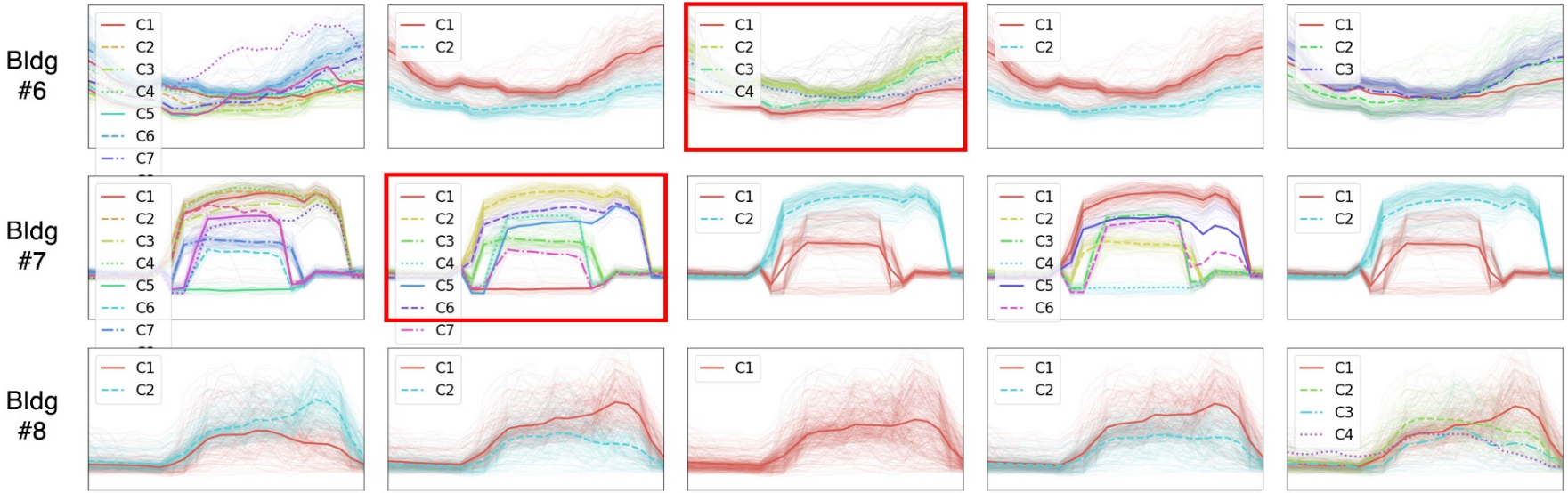


- Merge similar clusters based on Pearson Correlation Coefficient
- Use adjusted Calinski-Harabasz index to determine the threshold

Clustering results comparison (better: 37/50)

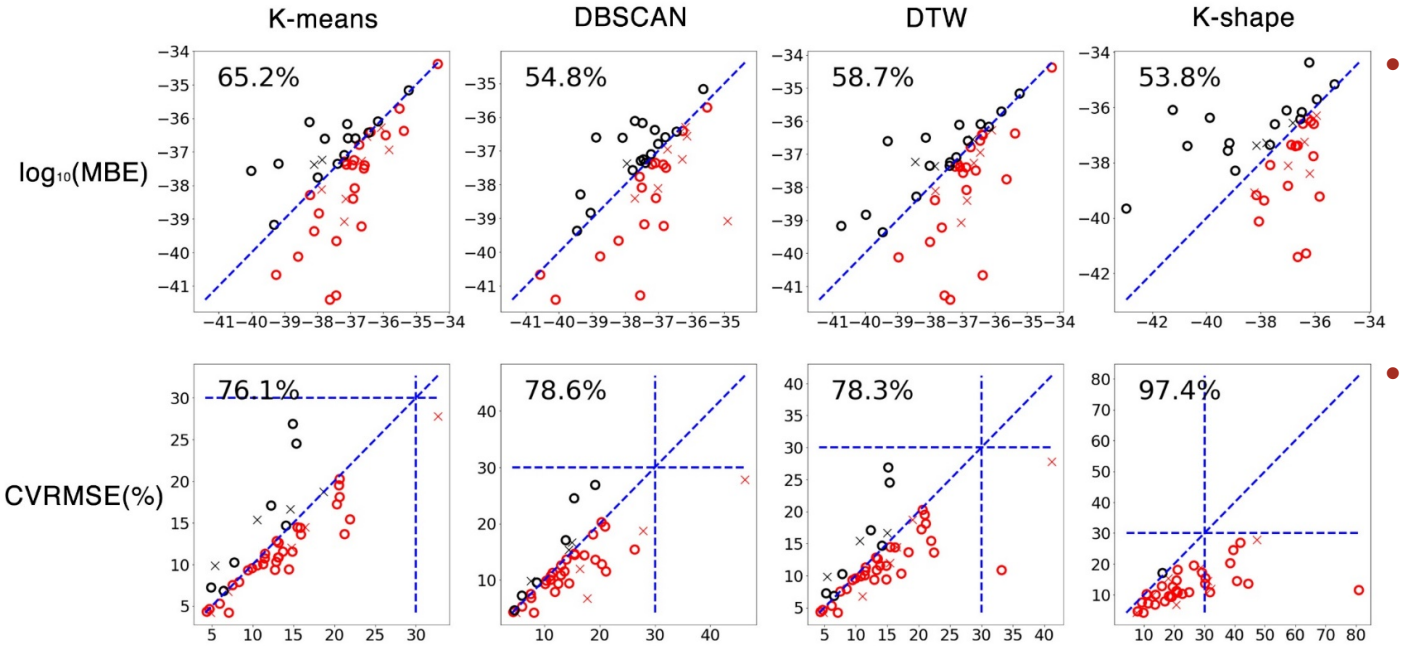


Clustering results comparison (worse: 13/50)



Results comparison through simulation

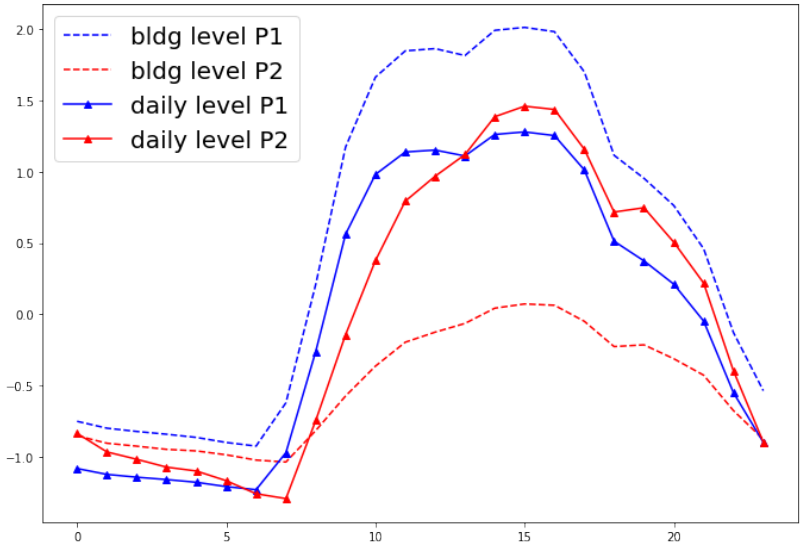
Comparing the simulation results against baseline methods



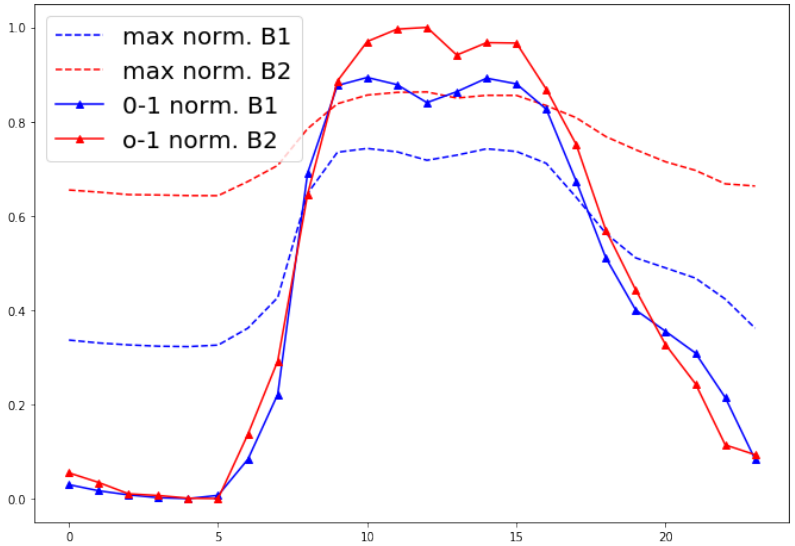
- MBE were all very small because positive and negative errors cancelled
- The improvement in CVRMSE was not large since the changes were mainly on minor days

Observations: normalization matters

Dashed: proper V.S. Solid: misleading



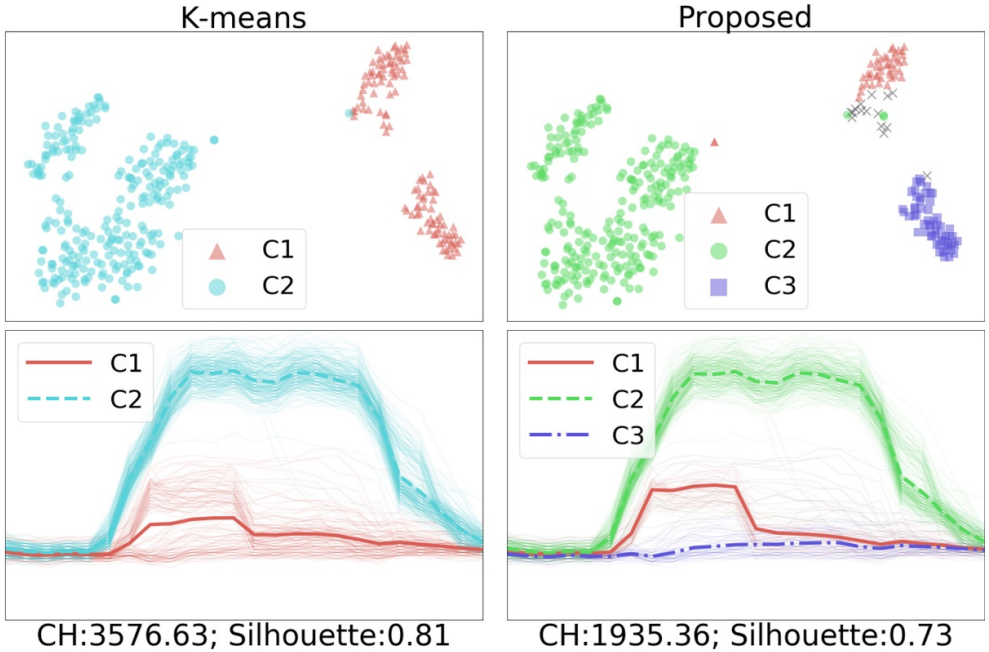
normalization should be done at **the building level**



max normalization should be applied to keep the variance

Observations: indexes fail

- BIC, CH index, Silhouette index, Dunn index and etc.
  - Penalize clusters closer to each other to avoid overfitting
  - Not always reflect the quality of the clusters and result in non-optimal results
  - To promote the scalability with a more robust index



BS 2019 2-4 SEPT ROME 16th IBPSA INTERNATIONAL CONFERENCE AND EXHIBITION

INTERNATIONAL BUILDING PERFORMANCE SIMULATION ASSOCIATION

A Robust Unsupervised Framework for High-Resolution Building Energy Consumption Profiling

Questions and Comments

Speaker:

Sicheng Zhan

Contacts:

szhan@u.nus.edu