BS 2021
1-3 SEPT
BRUGES

INTERNATIONAL
BUILDING
PERFORMANCE
SIMULATION
ASSOCIATION

AI-Based Building Type Assignment

Speaker:

Brett Bass
Oak Ridge National Laboratory

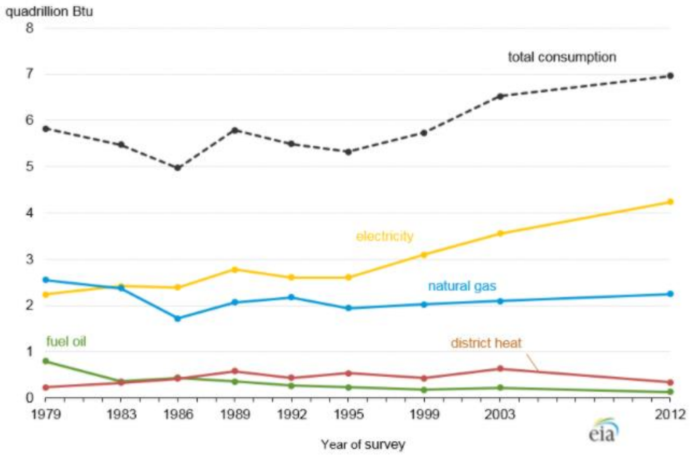
Authors:

Joshua New, Oak Ridge National Laboratory
Evan Ezell, University of Tennessee, Knoxville
Eric Garrison, University of Tennessee, Knoxville
Piljae Im, Oak Ridge National Laboratory
William Copeland, Electric Power Board, Chattanooga

Motivation

- Energy usage increasing in United States
- Buildings use about 40% of energy in the US
- Modeling buildings is important to efficiently design and optimize buildings and building related systems

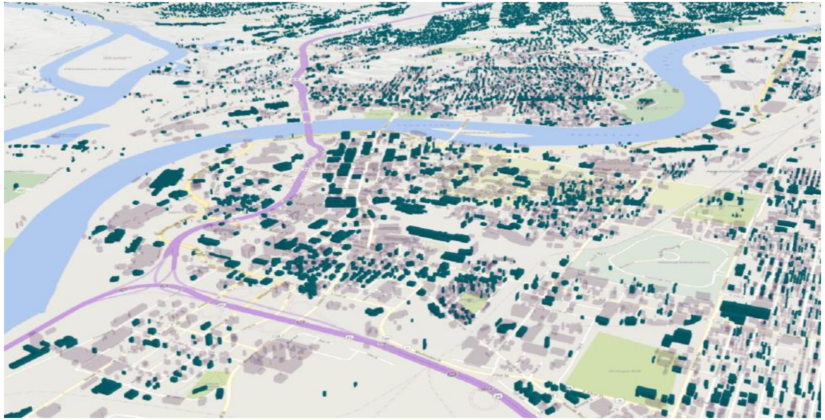
Figure 1. Total electricity usage has increased more than the other energy sources since 2003



Source: U.S. Energy Information Administration, Commercial Buildings Energy Consumption Survey.

Expanding the Scale of Building Energy Modeling (AutoBEM)

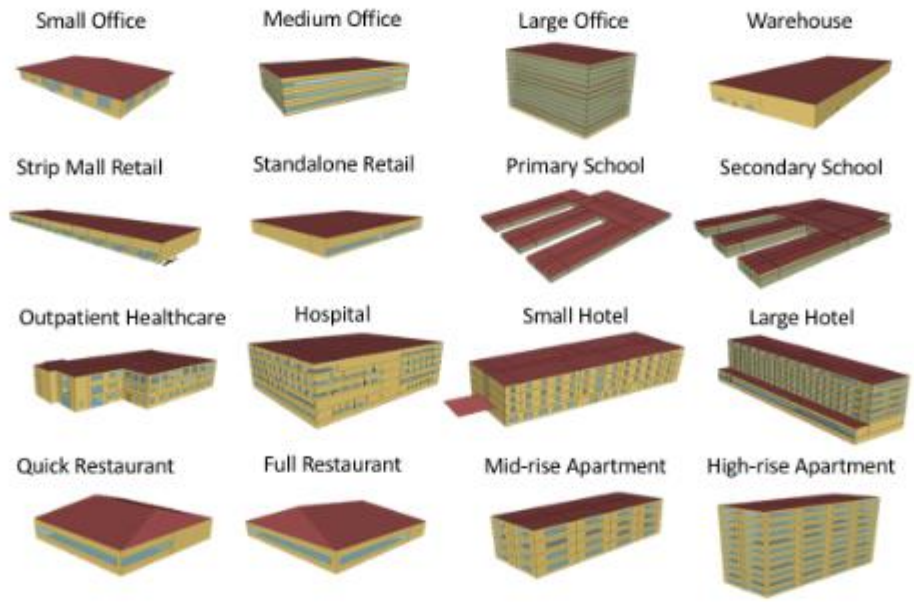
- Modeling large numbers of buildings at a time maximizes impact
- AutoBEM software developed to model large numbers of buildings individually
- Partership formed with Electric Power Board (Chattanooga)
  - Shared 15-min smart meter electricity data
- Properties describing each building are necessary to develop building energy models



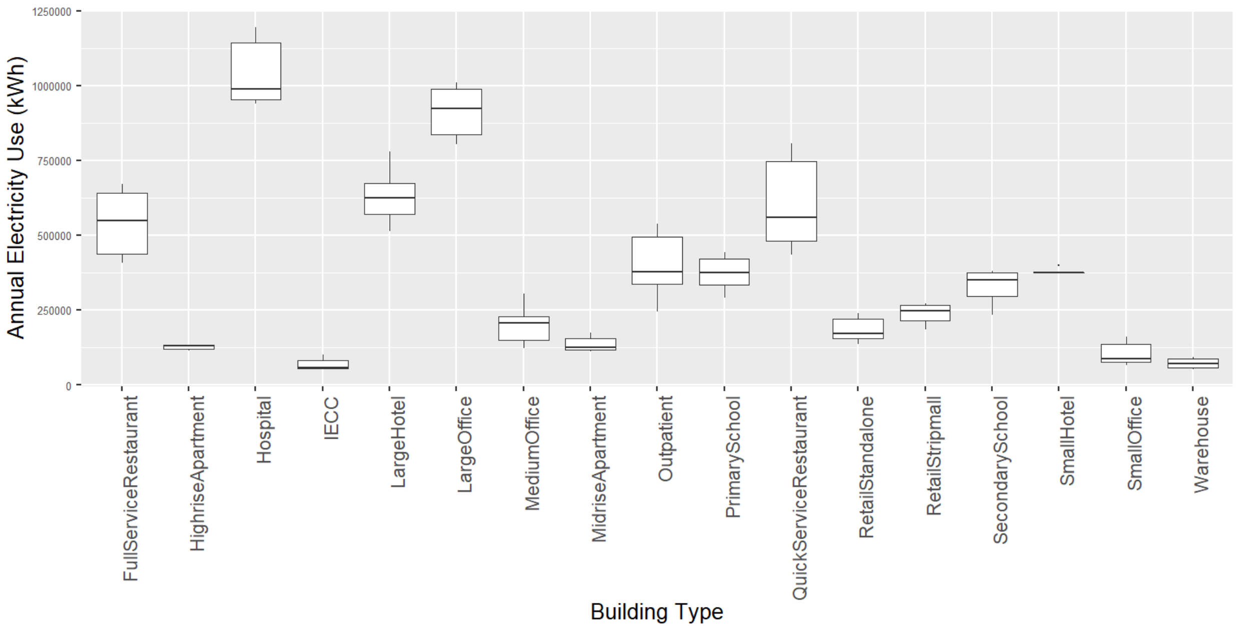
Bass, Brett and New, Joshua, and Copeland, William (2020)

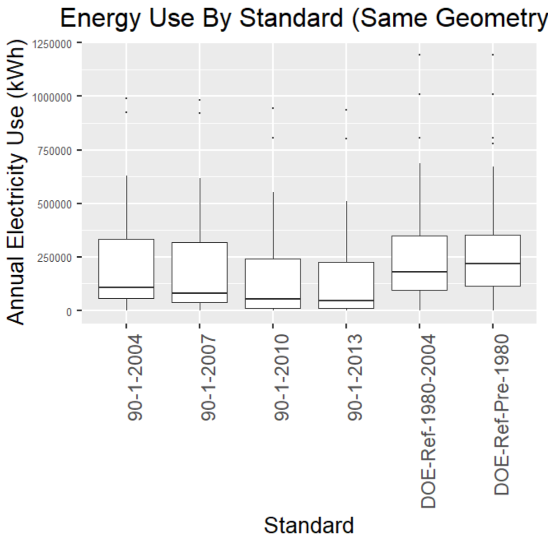
Expanding the Scale of Building Energy Modeling (AutoBEM)

- Physical building properties can be gathered using computer vision or LiDAR
- Aggregating building properties about the buildings energy use patterns can be difficult at a large/non-intrusive scale
- Properties from DOE prototype buildings can be assigned to these individual building geometries

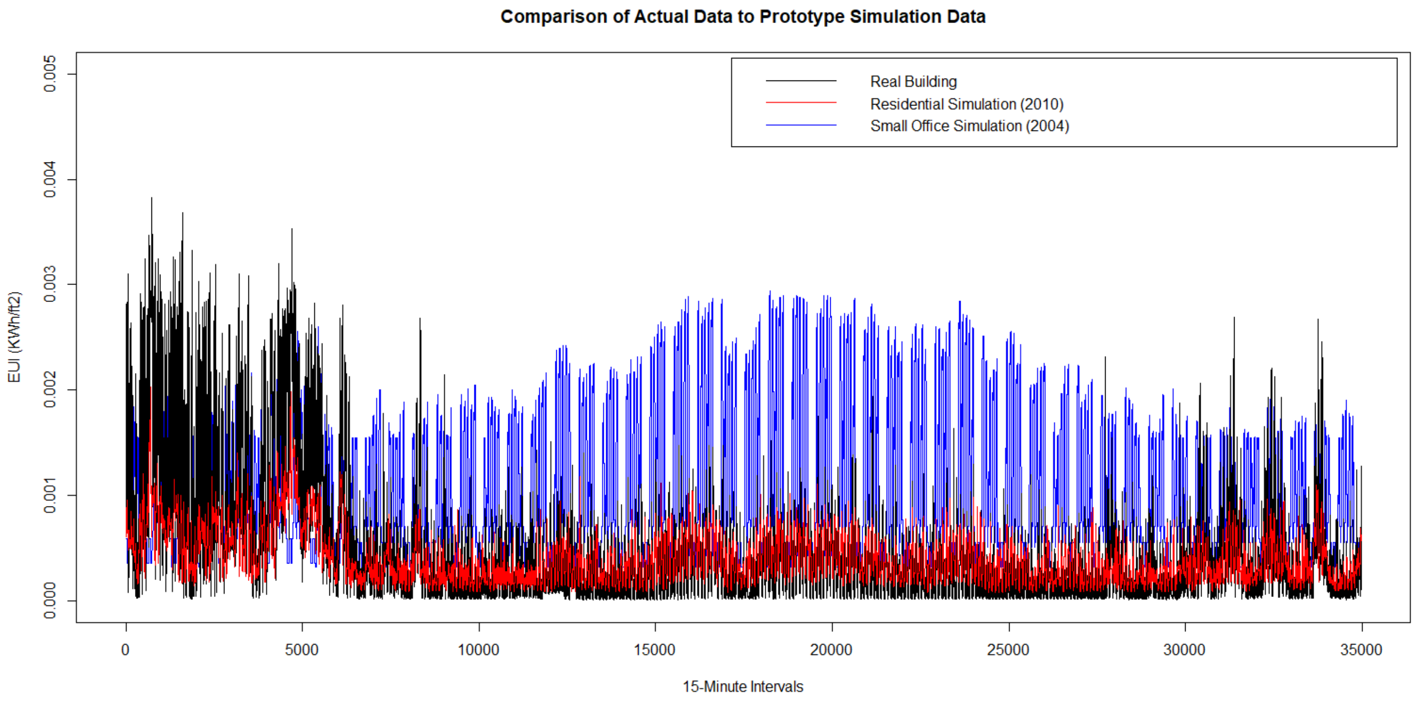


Prototype Building Impact on Energy Use





Prototype Building Comparison



Prototype Building Comparison

- Each real building Energy Use Intensity (EUI) is compared to each of the prototype building/vintage combinations
- Three methods considered for comparing the time series
  - Euclidean Distance
  - Dynamic Time Warping
  - Machine Learning (Random Forest)

Building Type	Standard
Small Office	DOE-Ref-Pre-1980
Medium Office	DOE-Ref-1980-2004
Large Office	90.1-2004
Standalone Retail	90.1-2007
Retail Stripmall	90.1-2010
Primary School	90.1-2013
Secondary School	
Outpatient	
Hospital	
Small Hotel	
Large Hotel	
Warehouse	
Quick-service Restaurant	
Full-service Restaurant	
Mid-rise Apartment	
High-rise Apartment	
Residential	

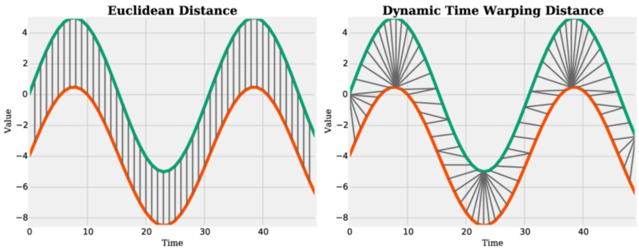
Euclidean Distance

- Straightforward method
- No imputation required
- Low computational cost

$$EucDist = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

Dynamic Time Warping

- Imputation required
- High computational cost (workaround - FastDTW)



Schfer, P. (2015).

Machine Learning

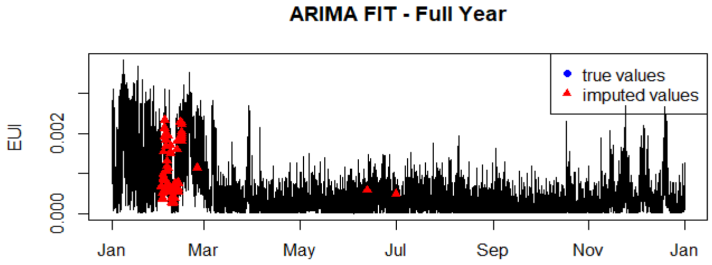
- Imputation required
- Low computational cost
- Extract feature vector from prototype building/vintage simulations

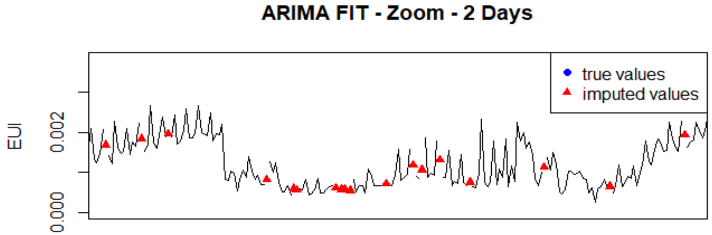
Time Window	Statistics
Monthly	Maximum
Yearly	Mean
Weekends	Median
	Standard Deviation
	Minimum

Method	Median Acc	Mean Acc	Median κ	Mean κ
KNN	78.4%	80.1%	77.1%	78.8%
RF	84.3%	82.2%	83.3%	81.1%
xgbTree	80.3%	81.0%	79.0%	79.7%

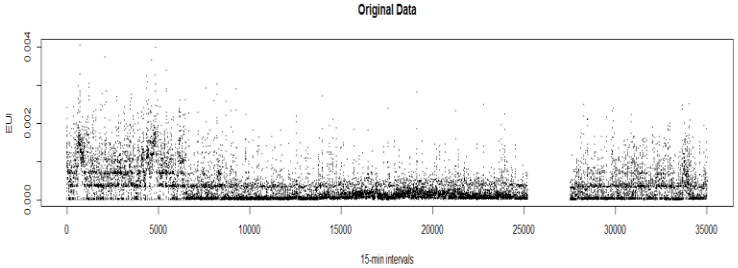
Missing Data Type	Imputation Strategy
Small Gaps (< 1 week)	ARIMA
Large Gaps (> 1 week)	Univariate DTW
> 75% missing	Remove*

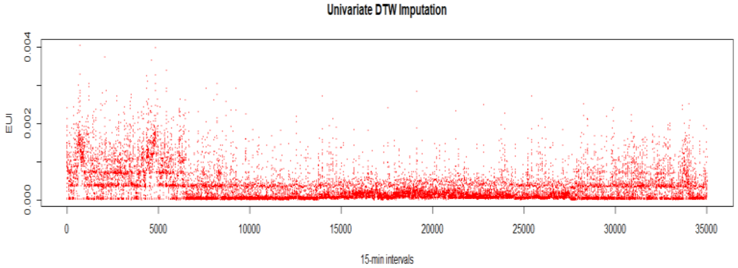
Small Gaps





Large Gaps





Results

- Methods were evaluated on 100 Chattanooga buildings
- Methods were evaluated quantitatively and qualitatively
- Building energy use pattern may not always match function
  - Bitcoin miner in house
- Confidence in each classification can be determined

Quantitative Results

- Compare 15-min simulated energy use from each of the 3 methods to actual energy use
- ASHRAE guideline 14 considers <15% monthly, <30% hourly "investment grade"
  - This is sub-hourly (15-minute data)

$$CVRMSE = \frac{1}{\bar{Y}}\sqrt{\frac{\Sigma_{i=1}^{n}(Y_i - \hat{Y})^2}{N}}$$

Method	Min	Median	Mean	Max
RF	.7%	38.6%	44.1%	206%
Euc	.5%	38.5%	48.6%	545%
DTW	.5%	38.7%	49.1%	560%

Qualitative Results

- Manually labeled 100 sample building to closest prototype
- Direct
  - Comparison of real building type to methodology classification
- General
  - Generalizes building types for common comparison
    - Mid-rise, High-rise Apartments are both only Apartments
- Commercial
  - Is the building a house or a commercial building

Method	Direct	General	Commercial
RF	62%	63%	78%
Euc	80%	80%	81%
DTW	71%	71%	77%

Method	Commercial Sensitivity	Commercial Specificity
RF	78.9%	78.3%
Euc	0.05%	100%
DTW	36.8%	87.8%

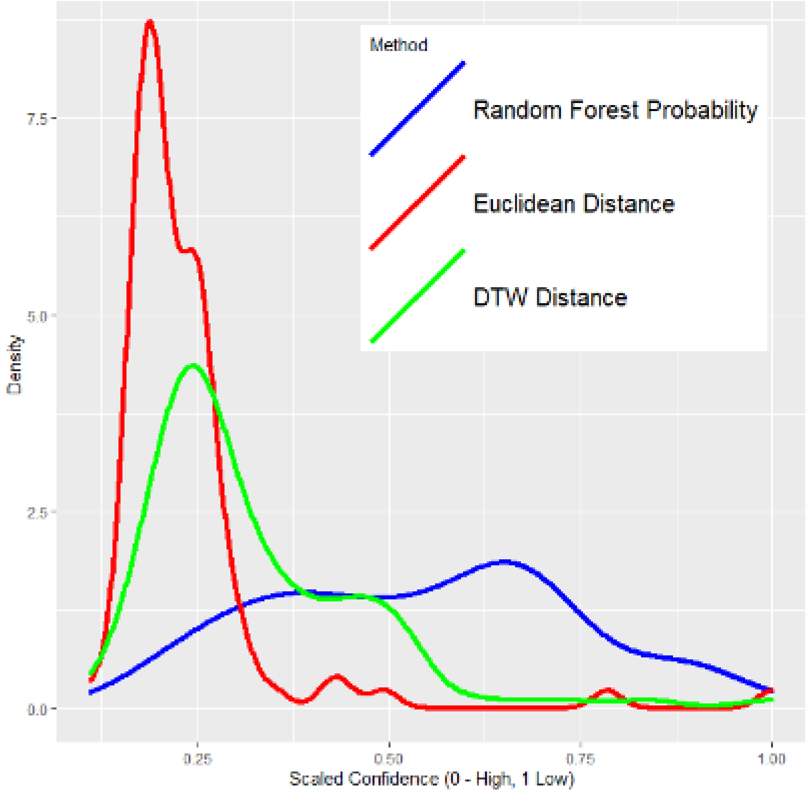
Confidence

Method	Min	Median	Mean	Max	Acc
RF	18.7%	45.8%	51.8%	138%	71%
Euc	5.5%	35.2%	37.5%	78.4%	97%
DTW	20.2%	35.2%	41.3%	206%	87%

Filter above mean confidence (average of 45 buildings across methods)

Method	Min	Median	Mean	Max	Acc
RF	18.7%	48.1%	57.2%	138%	59%
Euc	5.5%	29.9%	31.6%	66.8%	94%
DTW	20.2%	29.9%	32.4%	55.4%	100%

Filter above first quartile confidence (average of 20 buildings across methods)



Conclusions

- Missing and problematic data handling methods explored, and imputation methodology created
- 3 Building Type methodologies compared
- Euclidean distance fastest, simplest, and best classification accuracy
- Random Forest performed best for commercial accuracy
- Runtime of DTW is a limiting factor for large numbers of buildings
- The distance metrics for Euclidean distance and DTW are effective tools for measuring confidence

Future Work

- Expand scope of analysis to all EPB Chattanooga service area
  - Using labeled building types
- Explore different temporal resolutions
  - This this be done at an hourly scale?
- Explore other methods of comparison
  - New time series similarity methodologies

BS 2021
1-3 SEPT
BRUGES

AI-Based Building Type Assignment

Questions and Comments

Speaker:

Brett Bass
Oak Ridge National Laboratory

Contacts:

Joshua New, Oak Ridge National Laboratory
Evan Ezell, University of Tennessee, Knoxville
Eric Garrison, University of Tennessee, Knoxville
Piljae Im, Oak Ridge National Laboratory
William Copeland, Electric Power Board, Chattanooga