

Presenter: Chengxuan Li
PhD Student, Environmental Systems Lab, Cornell University, USA



DERIVING HIGH-FIDELITY RESIDENTIAL BUILDING ARCHETYPES AND TYPICAL USAGE PATTERNS FROM NATIONAL ENERGY USE SURVEYS

TO ENHANCE "INITIAL GUESSES" FOR URBAN BUILDING ENERGY MODEL (UBEM) INPUTS



Scan for paper preview

Chengxuan Li^{1,*}, Timur Dogan¹

¹Environmental Systems Lab, Cornell University, USA

*Corresponding Author, CL2749@cornell.edu

Oral Presentation

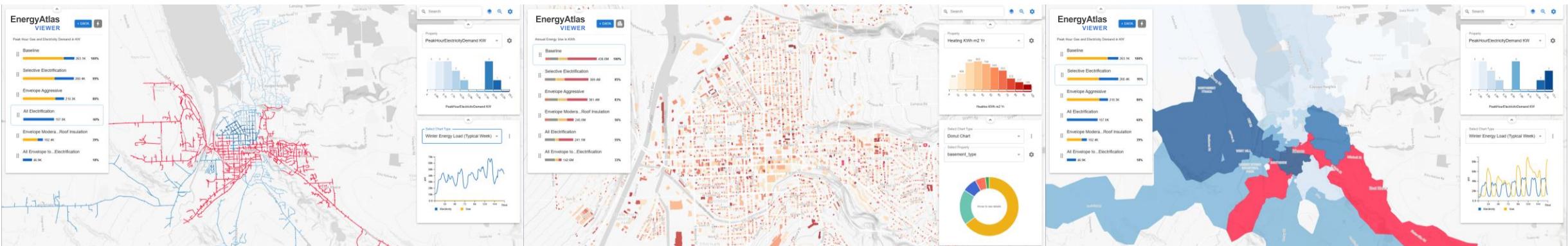


Scan for presentation file



ARCHETYPE-BASED URBAN BUILDING ENERGY MODELING (UBEM)

- UBEM supports energy planning and retrofit prioritization at scale
- Lack of building-level data
 - Simulation fidelity is limited
 - Targeted retrofits/incentives require building-level precision



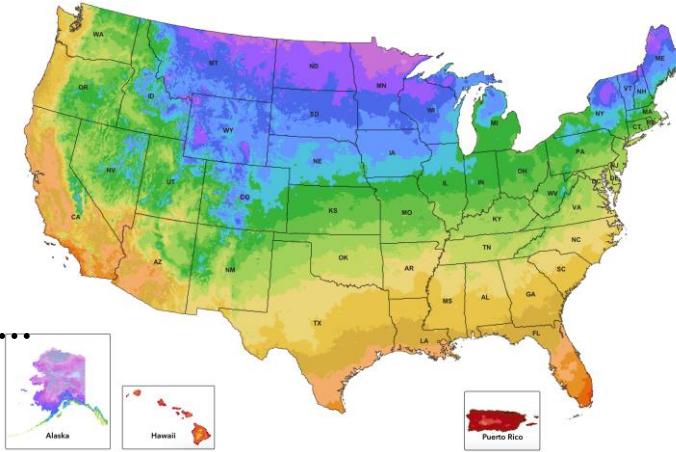
ARCHETYPE-BASED URBAN BUILDING ENERGY MODELING (UBEM)

- UBEM supports energy planning and retrofit prioritization at scale
- Lack of building-level data
 - Simulation fidelity is limited
 - Targeted retrofits/incentives require building-level precision
- Archetypes use stock-level assumptions to fill building-level gaps
 - Traditional Prototype Buildings: only singular prototype for residential
- We need a high-fidelity, modular, data-driven archetype framework



DEPICTING THE FIDELITY OF THE U.S. RESIDENTIAL STOCK

- Residential Energy Consumption Survey (RECS)
 - 5 typologies, 50 states, 17 IECC Climates, 18,496 records ...



Mobile Homes



Single Family Detached



Single Family Attached



Apartments 2-4 Units



Apartments 5+ Units



RECS ATTRIBUTES

- Separation of building attributes

- Construction
- Systems
- Energy loads

Construction

- Combinatory assignment
- Consistency with state-of-the-art modeling frameworks
(ClimateStudio, E+)

Energy Loads

Residential Energy Consumption Survey (RECS)

Basic building information

Housing typology
Floor area

Year of construction
Number of floors

Urban/rural designation
Number of occupants

Construction sub-archetype

Building and envelope properties

- Is housing over a basement?
- Is housing over a crawlspace?
- Is housing over a concrete slab?
- Is housing over other kinds of foundation?
- Is there a finished basement?
- Is there a finished attic?
- Major outside wall material
- Major roofing material
- Level of insulation (respondent-reported)
- Type of glass in most windows
- Window frame material
- Window area per m²
- Door area per m²

Space heating-related EUIs

- Space heating electricity EUI
- Space heating natural gas EUI
- Space heating other fuels EUI

Energy loads sub-archetype

Space use and energy load EUIs

- Equipment power density (electricity, gas) [W/m²]
- Lighting power density [W/m²]
- Hot water power density (electricity, gas) [W/m²]

Systems

Systems sub-archetype

Heating systems properties

- Space heating equipment type
- Space heating fuel type
- Is basement heated?
- Is garage heated?
- Is attic heated?
- Is heating equipment geothermal/ground source?
- Age of heating equipment

Cooling systems properties

- Is there air conditioning equipment?
- Is air conditioning centrally supplied?
- Air conditioning equipment type
- Is basement conditioned?
- Is attic conditioned?
- Is garage conditioned?
- Age of air conditioning equipment

Thermal-related EUIs

- Space heating electricity EUI
- Space heating natural gas EUI
- Space heating other fuels EUI
- Air conditioning electricity EUI
- Latent loads EUI

+ ...





NATIONAL SURVEY TO BUILDING-LEVEL INFO

National/state building stock

Archetype UBEMs

Local
buildings/UBEMs

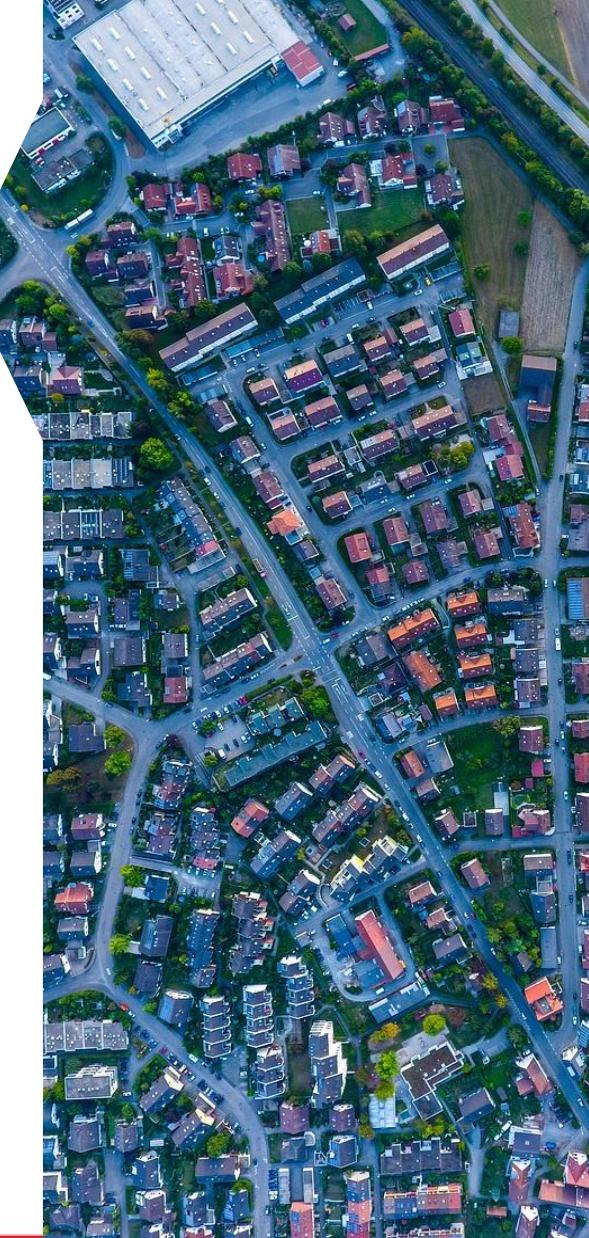


Image credits:
<https://pixabay.com/photos/aerial-view-houses-village-road-2598834/>
<https://www.howardhanna.com/Property/Detail/510-S-Plain-Street-Ithaca-City-NY-14850/IthacaNYSold/R1600388>





ARCHETYPES FROM RECS CLUSTERING

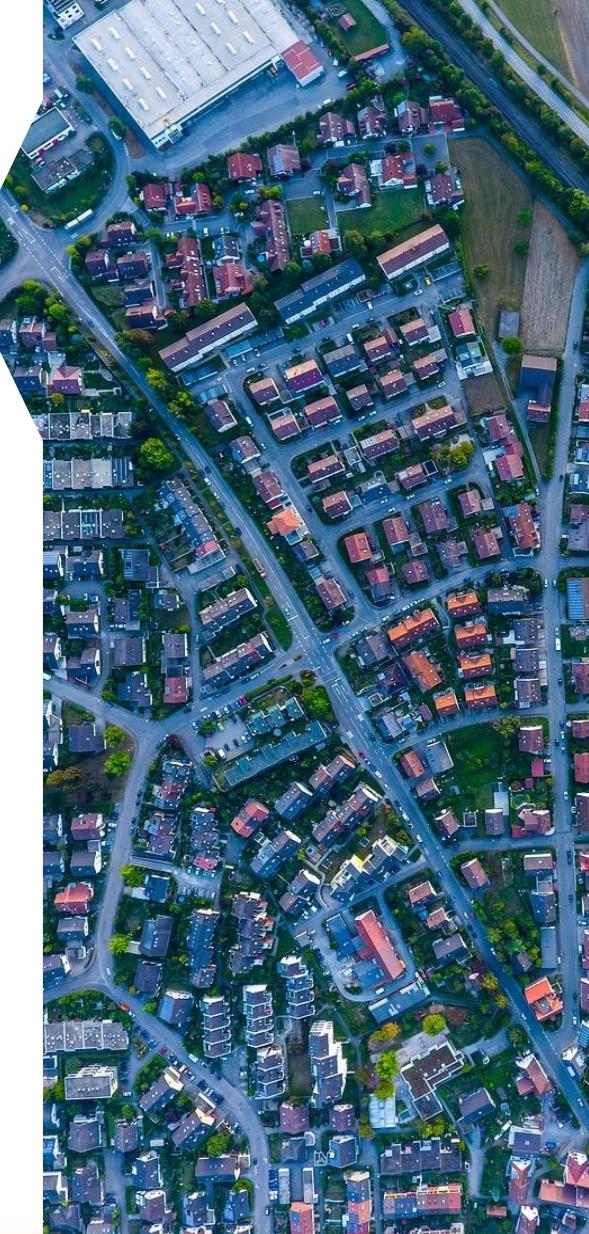
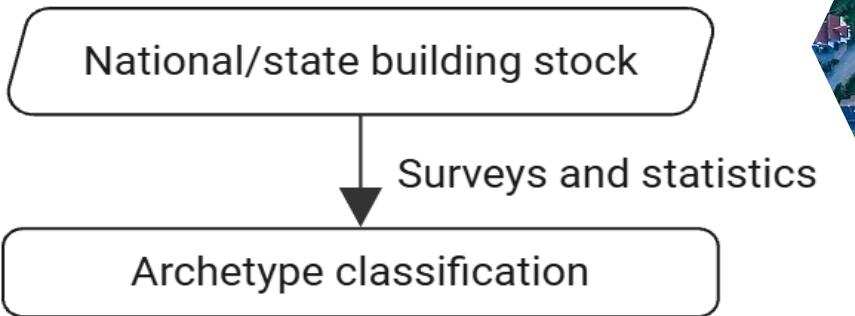
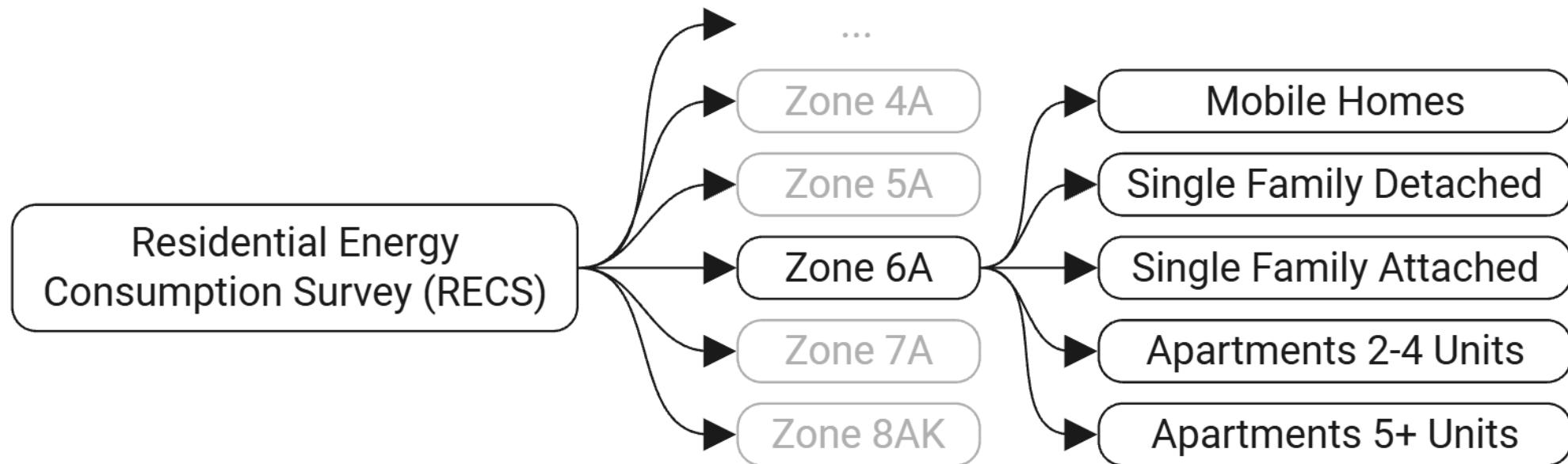


Image credits:
<https://pixabay.com/photos/aerial-view-houses-village-road-2598834/>
<https://www.howardhanna.com/Property/Detail/510-S-Plain-Street-Ithaca-City-NY-14850/IthacaNYSold/R1600388>



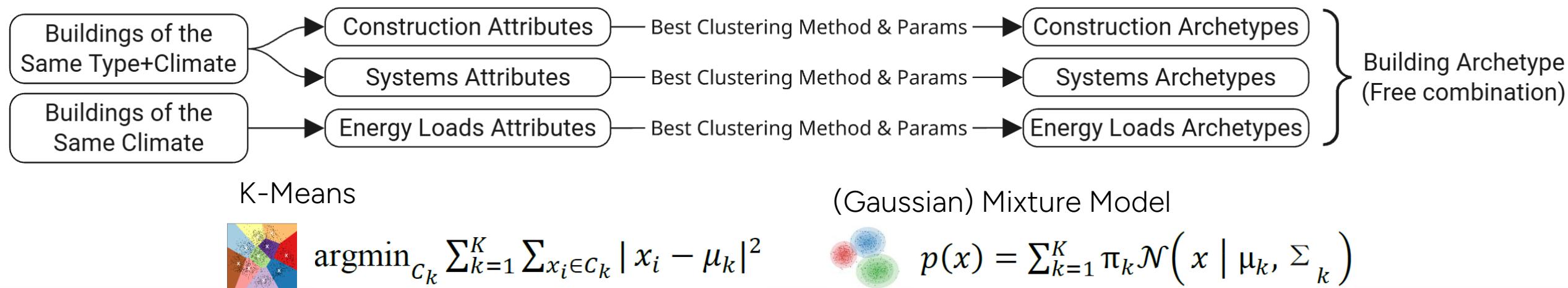
ARCHETYPES FROM RECS CLUSTERING

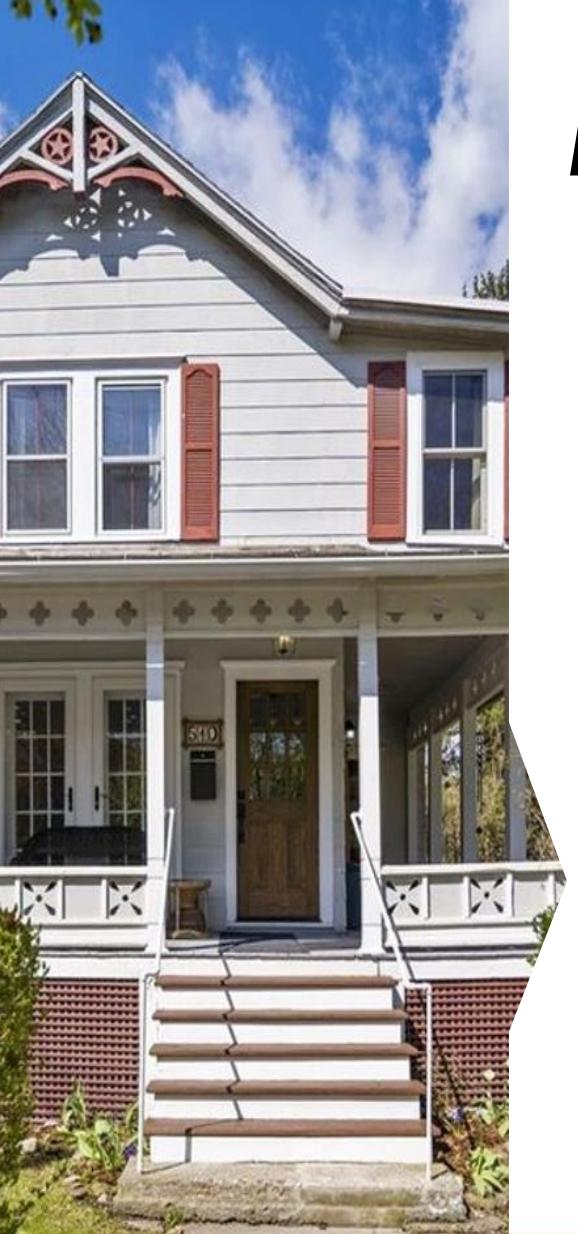
- Stratification of RECS
 - Firstly: group by IECC climate zone
 - Secondly: group by housing typology “TYPEHUQ”



ARCHETYPES FROM RECS CLUSTERING

- Clustering of the identified subsets of buildings
- Clustering by attribute groups:
 - Construction, Systems, Energy Loads (Sub-Archetypes)
 - Building Archetype is a combination of C+S+E
- Clustering methods: K-Means, GMM





ARCHETYPES FROM RECS CLUSTERING

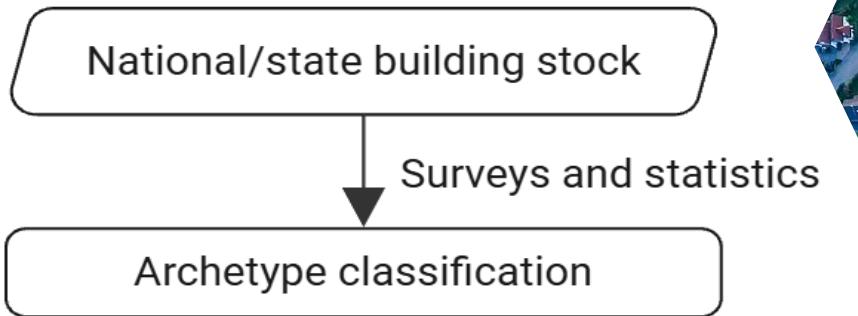
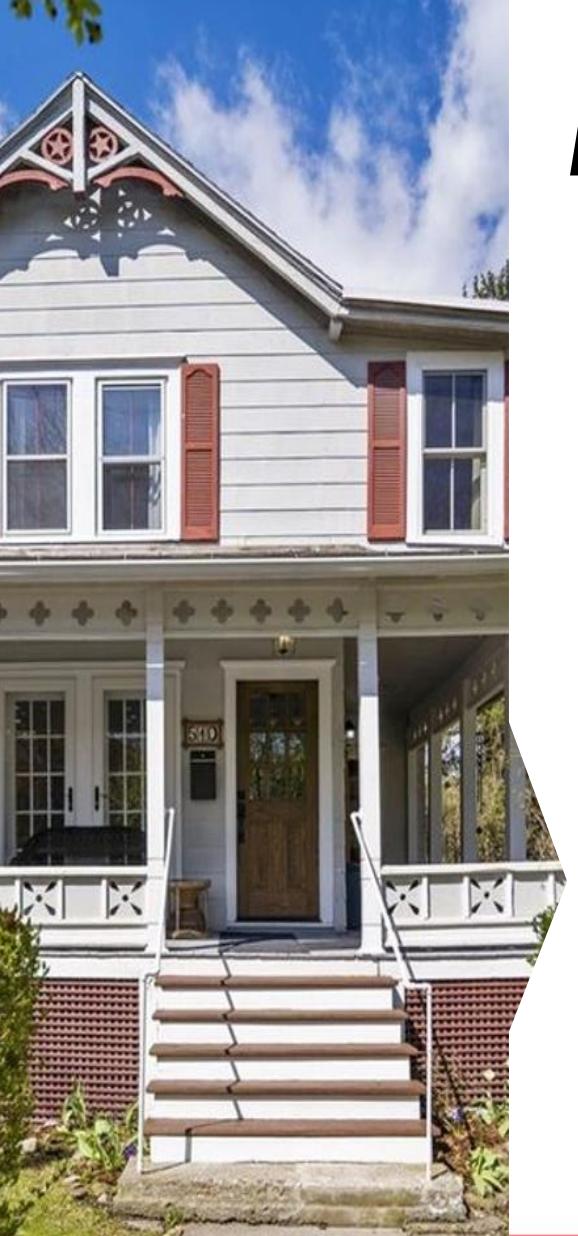
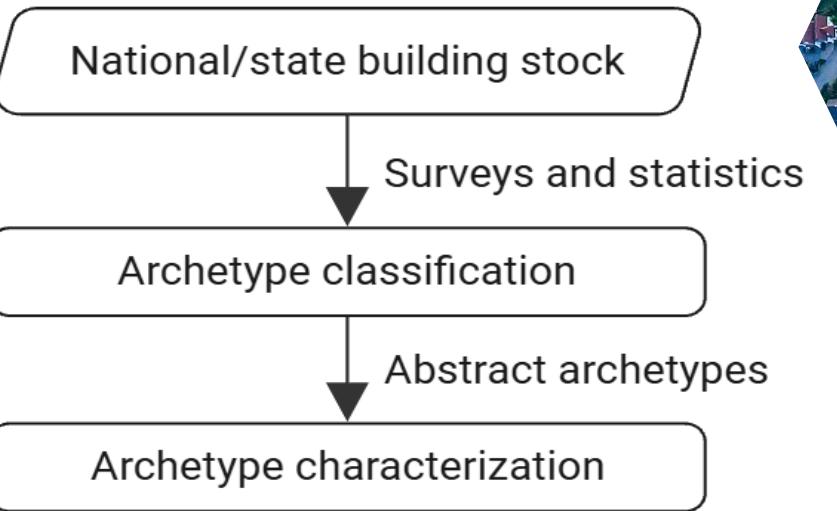


Image credits:
<https://pixabay.com/photos/aerial-view-houses-village-road-2598834/>
<https://www.howardhanna.com/Property/Detail/510-S-Plain-Street-Ithaca-City-NY-14850/IthacaNYSold/R1600388>

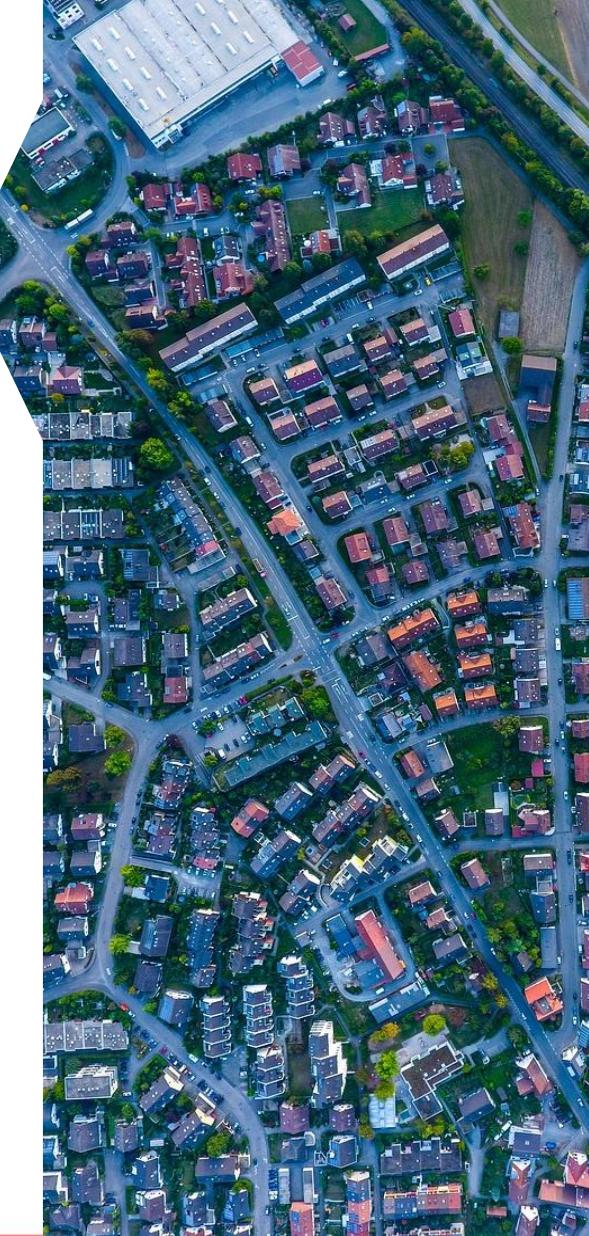


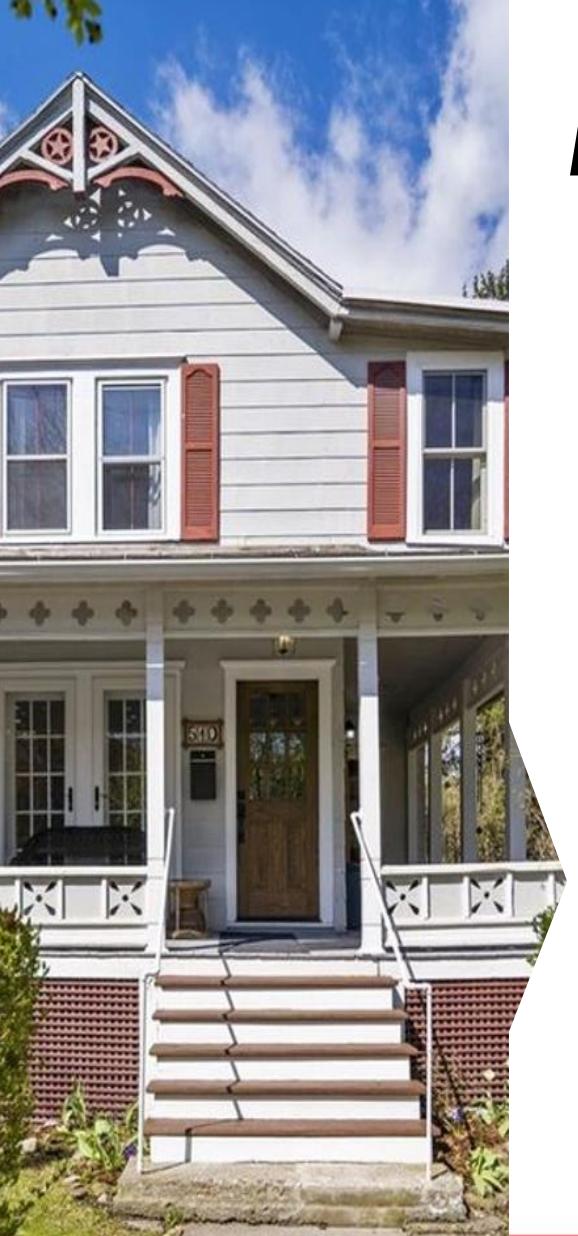


ARCHETYPE CHARACTERIZATION



- Characterization methods
 - Frequency (categorical attributes)
 - Mean value (continuous attributes)
 - Mean plus 5-th, 95-th percentiles (energy loads attributes)





ARCHETYPE ASSIGNMENT

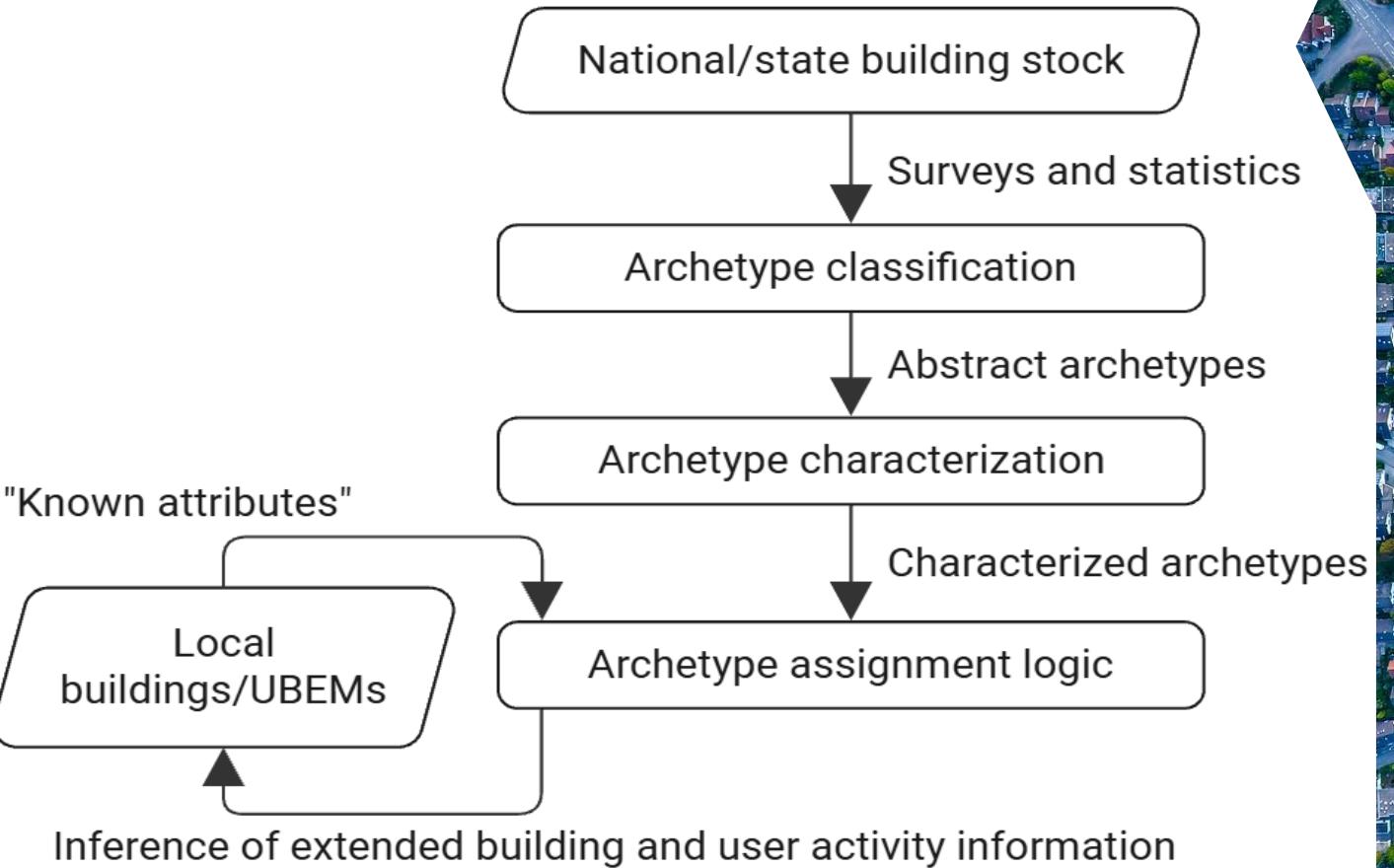
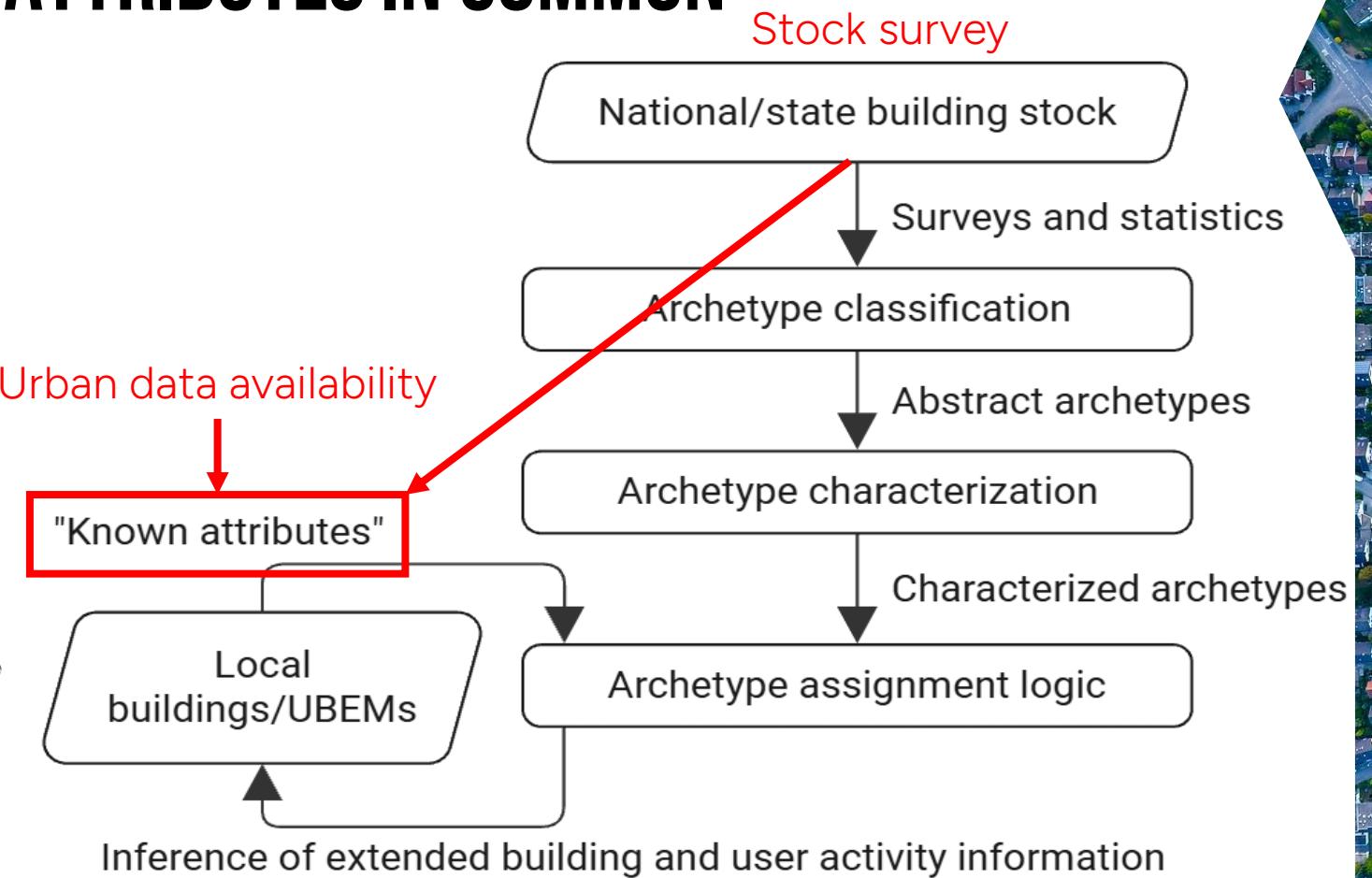


Image credits:
<https://pixabay.com/photos/aerial-view-houses-village-road-2598834/>
<https://www.howardhanna.com/Property/Detail/510-S-Plain-Street-Ithaca-City-NY-14850/IthacaNYSold/R1600388>



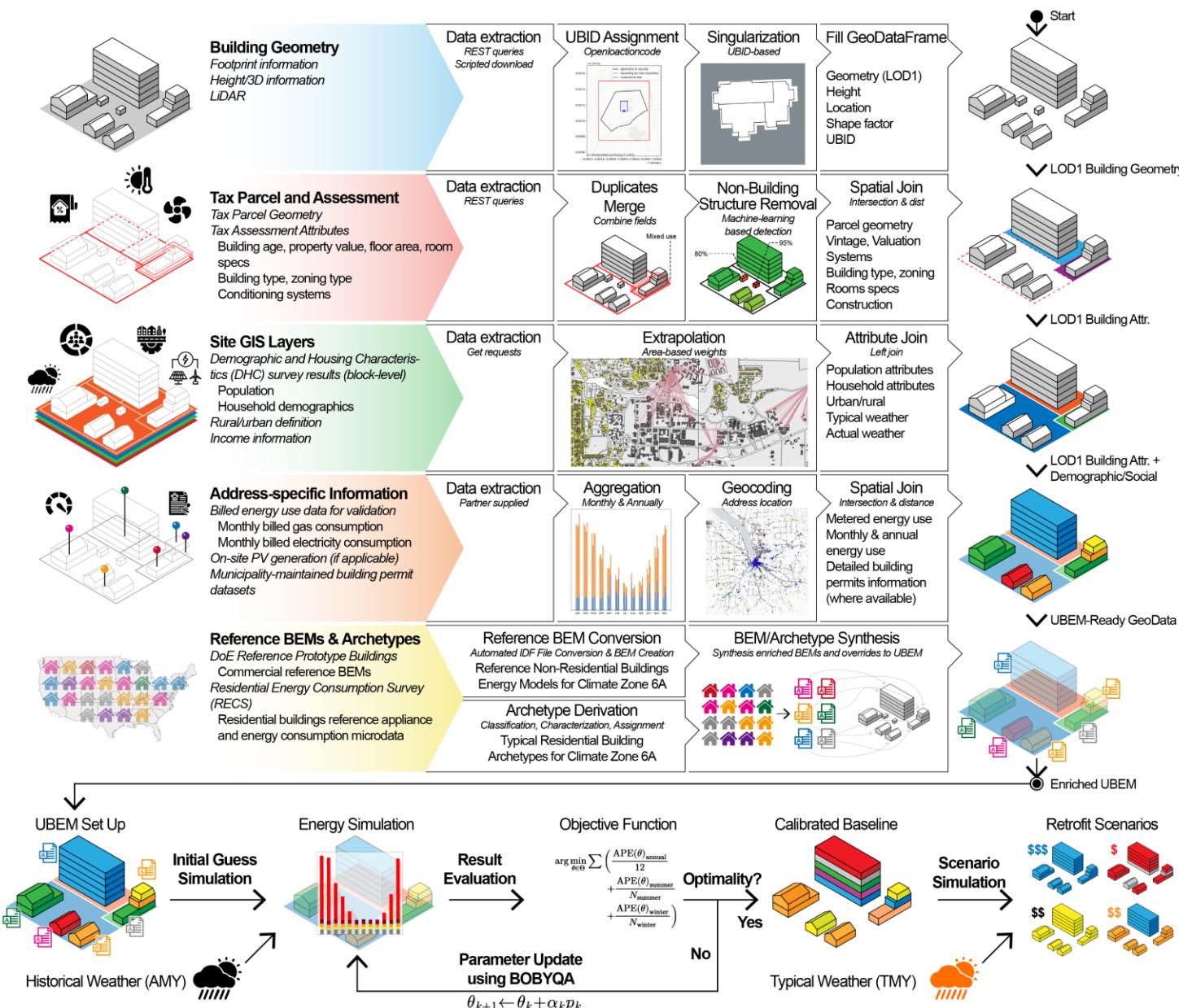


ATTRIBUTES IN COMMON



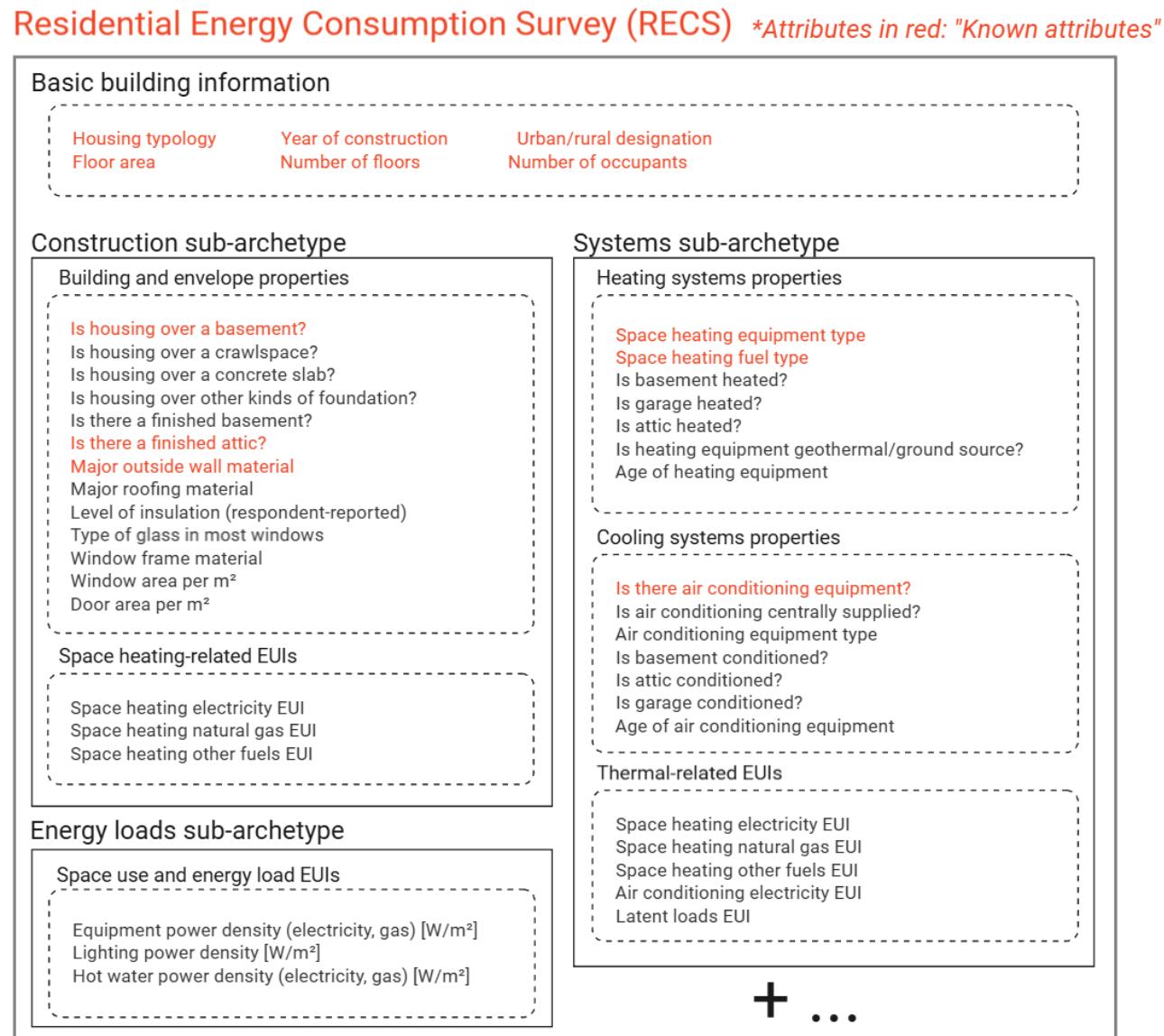
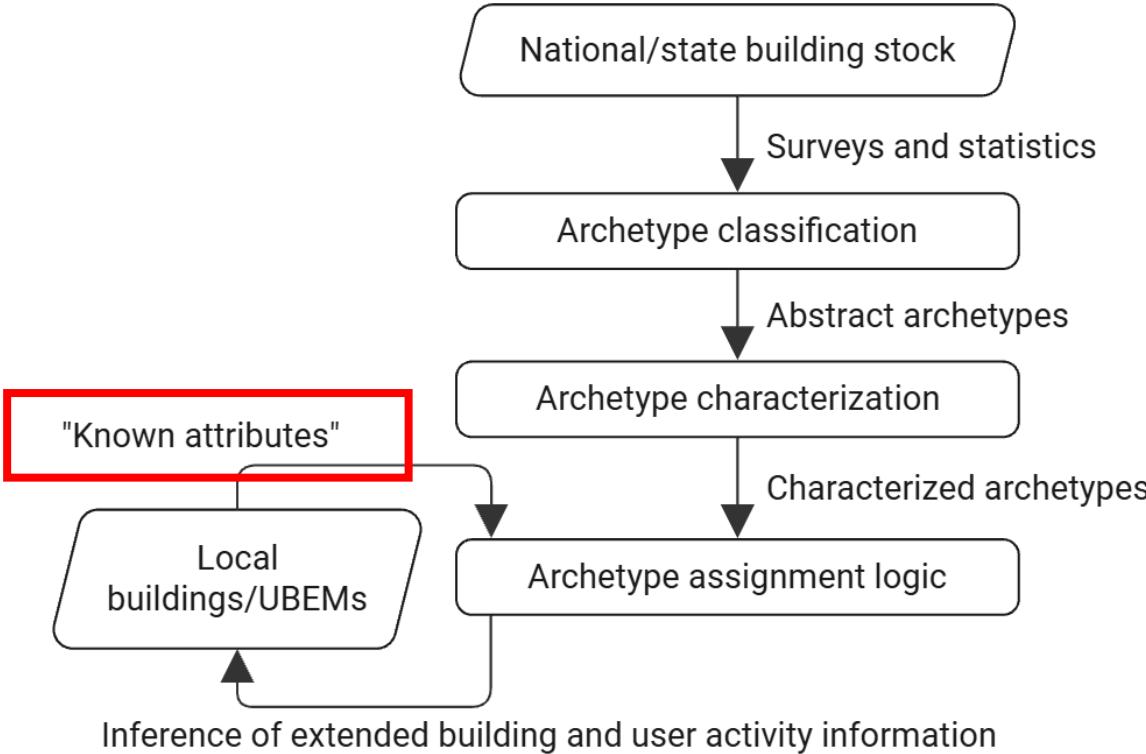
URBAN DATA EXTRACTION, PRE-PROCESSING, AND SYNTHESIS

- Identification of typically available urban data sources
 - County REST
 - Census DHC
 - Tax Assessor Data
 - USGS
- Identification of “Known Attributes”
 - Attributes consistent with/convertible from RECS



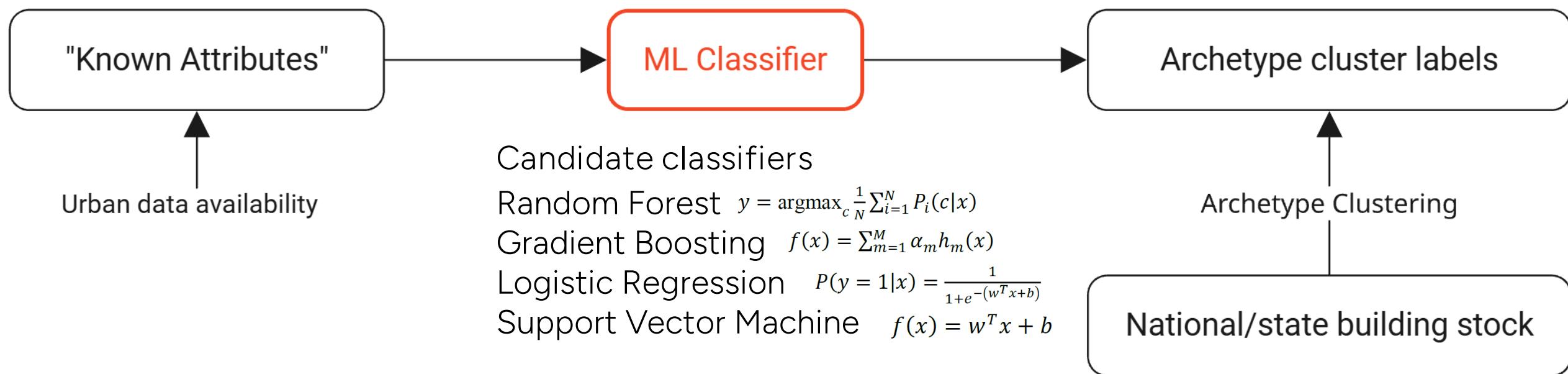
“KNOWN ATTRIBUTES”

Archetype workflow for building-level attribute inference
(infer unknown building attributes from known data)



ARCHETYPE ASSIGNMENT MODEL TRAINING & PREDICTION

- Archetype assignment based on “Known Attributes” in urban data
- Typical ML classifiers reviewed



RESULTS

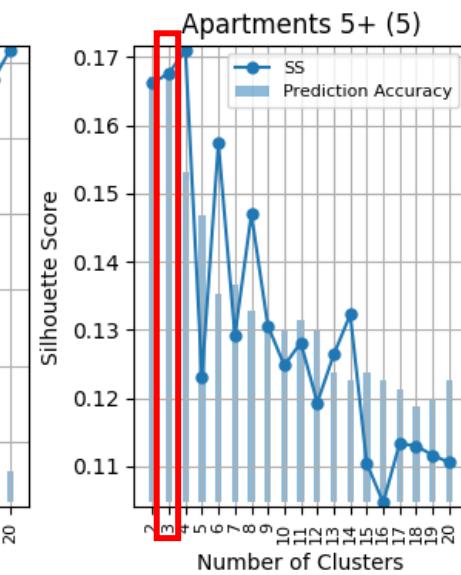
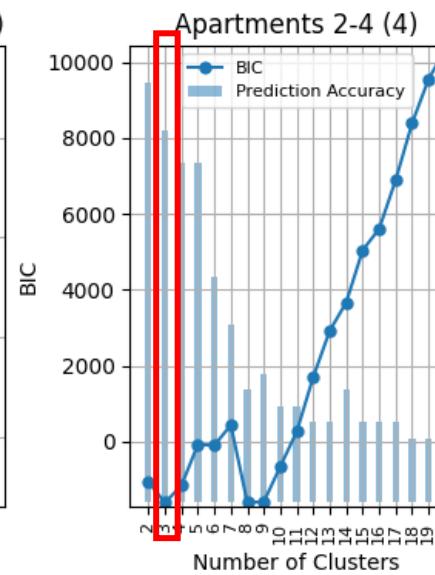
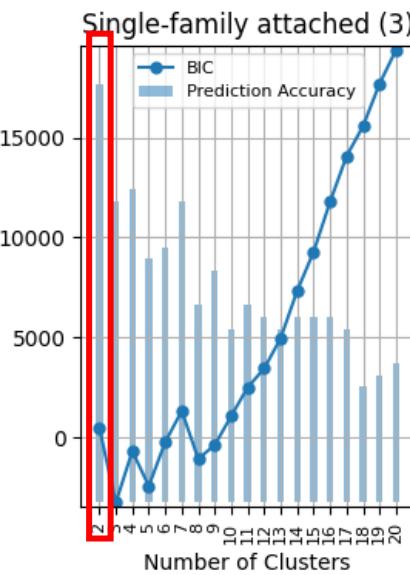
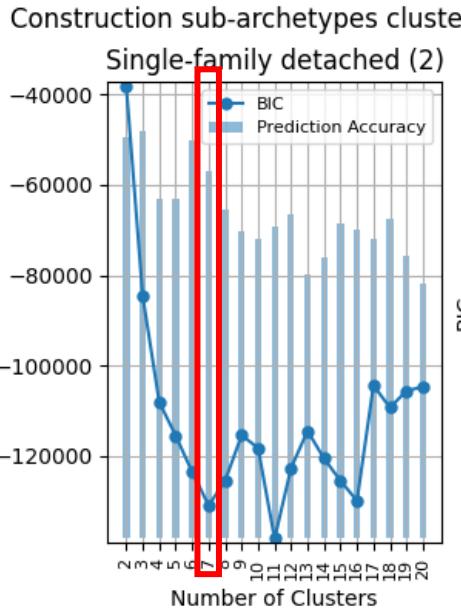
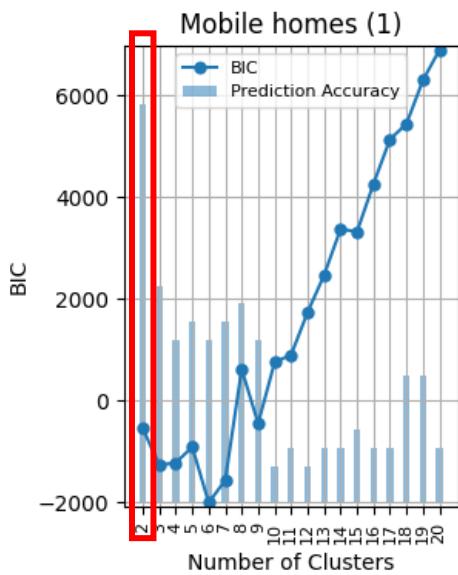
- 41 (sub-)archetypes: 17 construction, 21 systems, 3 energy loads
- 270 possible combinations

Sub-Archetype Category	Housing Typology	# of Sub-Archetypes	Classification Model	Accuracy
Construction	1 (Mobile Home)	2	Random Forest	0.92
	2 (Single-Family Detached)	7	Support Vector Machine	0.81
	3 (Single-Family Attached)	2	Logistic Regression	0.97
	4 (Apartments 2-4 Units)	3	Logistic Regression	0.93
	5 (Apartments 5+ Units)	3	Gradient Boosting	0.96
Systems	1 (Mobile Home)	2	Random Forest	0.96
	2 (Single-Family Detached)	8	Logistic Regression	0.92
	3 (Single-Family Attached)	3	Logistic Regression	0.92
	4 (Apartments 2-4 Units)	5	Support Vector Machine	0.86
	5 (Apartments 5+ Units)	3	Gradient Boosting	1.00
Energy Loads	All	3	Support Vector Machine	0.73



RESULTS: CONSTRUCTION ARCHETYPES

- Determining optimal number of clusters:
 - Combination of “elbow method” and assignment accuracy
 - High cluster counts may divide dataset well, but risk miss-assignment



RESULTS: CONSTRUCTION ARCHETYPES

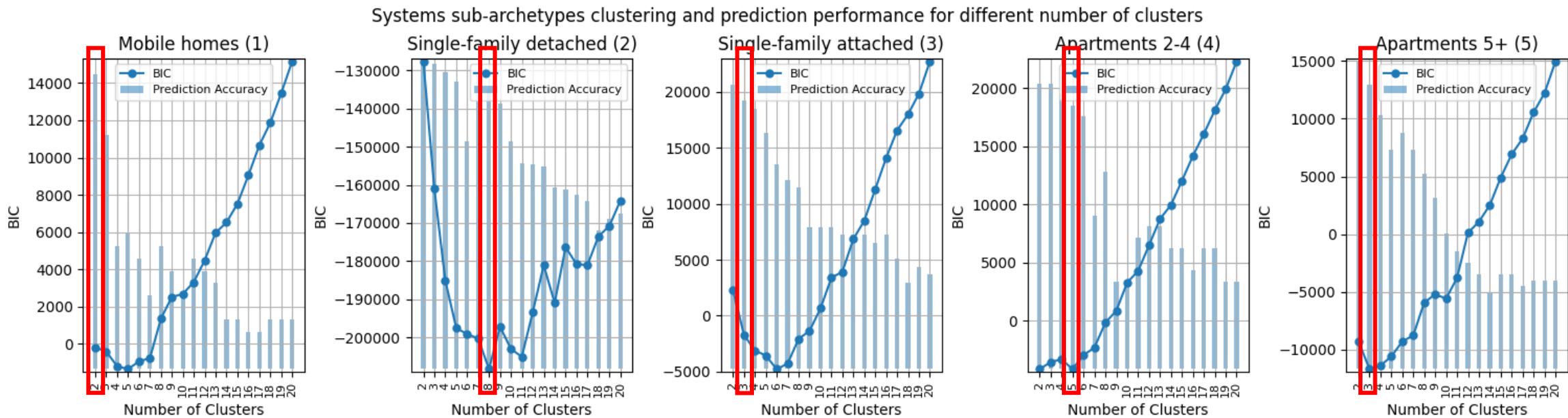
- Construction sub-archetypes reflect façade, insulation, window types and WWR
- Improved inference of envelope performance

Type	Specifications
1	Standard Vinyl Window Mobile Home
	Standard Metal Window Mobile Home
2	Siding Façade Home
	Siding Façade, Finished Basement, Big WWR
	Wood Façade Home
	Siding Façade, Finished Basement, Mod. WWR
	Siding Façade Home with Metal Roof
	Poorly Insulated Shingles Façade Home
3	Well-Insulated Siding Façade Home
	Attached Home with Moderate WWR
4	Attached Home with Big WWR
	Apartment with Big WWR
	Apartment with Moderate WWR
5	Wood Façade Apartment
	Apartment with Metal Windows (Electric Heating)
	Apartment with Metal Windows (Gas Heating)
	Apartment with Wood Windows



RESULTS: SYSTEMS ARCHETYPES

- Determining optimal number of clusters:
 - Combination of “elbow method” and assignment accuracy
 - High cluster counts may divide dataset well, but risk miss-assignment



RESULTS: SYSTEMS ARCHETYPES

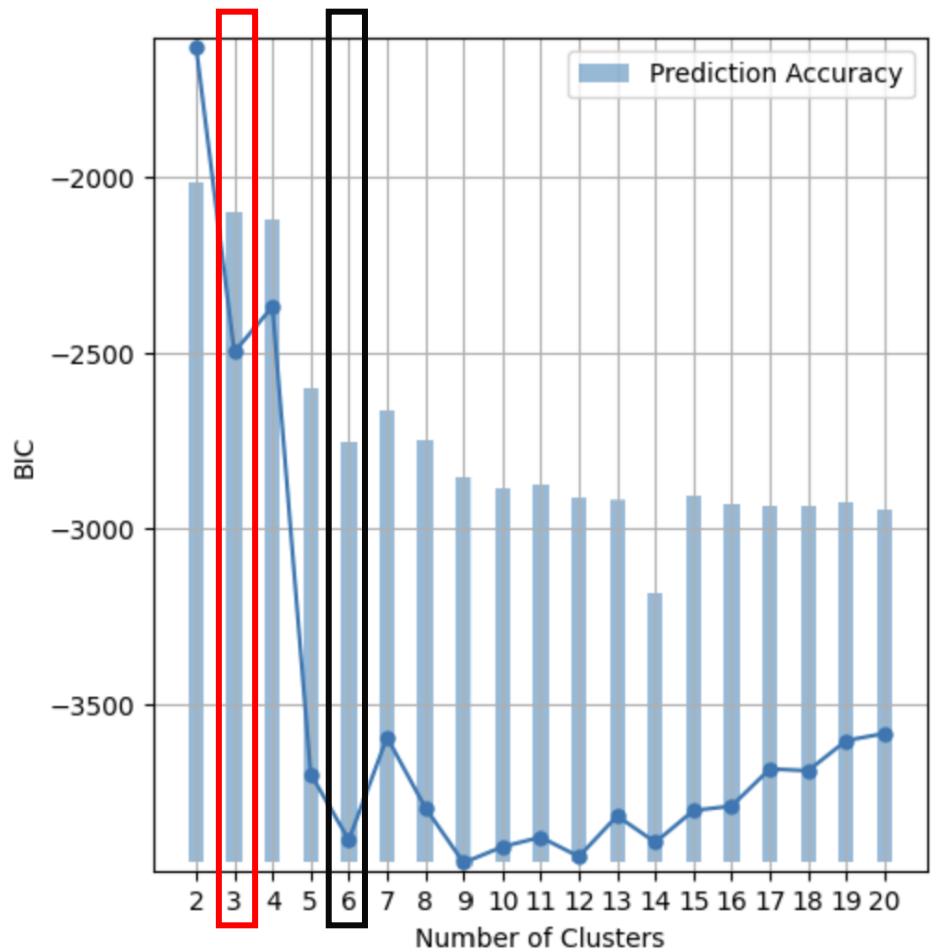
- Heating and cooling system types are primary differentiators
- Archetypes capture equipment efficiency and age
- Improves assumptions for retrofit targeting

Type	Specifications
1	Gas Furnace, Central AC Mobile Home
	Oil Furnace, Wall AC Mobile Home
2	Propane Furnace, Central AC Home
	Typical-Efficiency Gas Furnace, Central AC Home
	Oil Furnace, Wall AC Home with Gas
	Low-Efficiency Gas Furnace, Wall AC Home
	Heat Pump, Central AC Home
	Oil Furnace, Wall AC Home with No Gas
	Electric Heating, Wall AC Home
3	High-Efficiency Gas Furnace, Central AC Home
	Old Gas Furnace, Central AC Attached Home
	Typical Gas Furnace, Central AC Attached Home
4	Propane Furnace, Wall AC Attached Home
	Oil Furnace, Wall AC Apartment
	Gas Steam Heater, Wall AC Apartment
	Gas Furnace, Central AC Apartment
	Electric Heating, Wall AC Apartment
5	Oil Steam Heater, Wall AC Apartment
	Gas Furnace, Central AC Multi-Apartment
	Propane Steam Heater, Wall AC Multi-Apartment
	Electric Heating, Wall AC Multi-Apartment



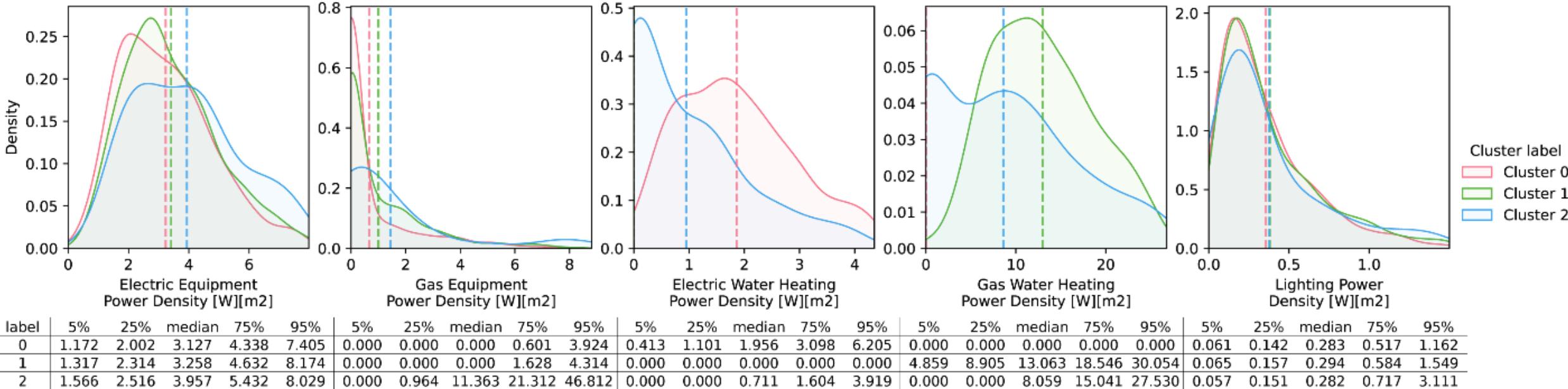
RESULTS: ENERGY LOADS ARCHETYPES

- User behavior remains hard to capture deterministically
- Feasible-range based characterization to supplement the inability to go for higher cluster counts



RESULTS: ENERGY LOADS ARCHETYPES

- Feasible range-based characterization
 - (Electric) equipment power density; (Gas) equipment power density
 - (Electric) DHW power density; (Gas) DHW power density
 - Lighting (constant load) power density



This design of attributes facilitates:

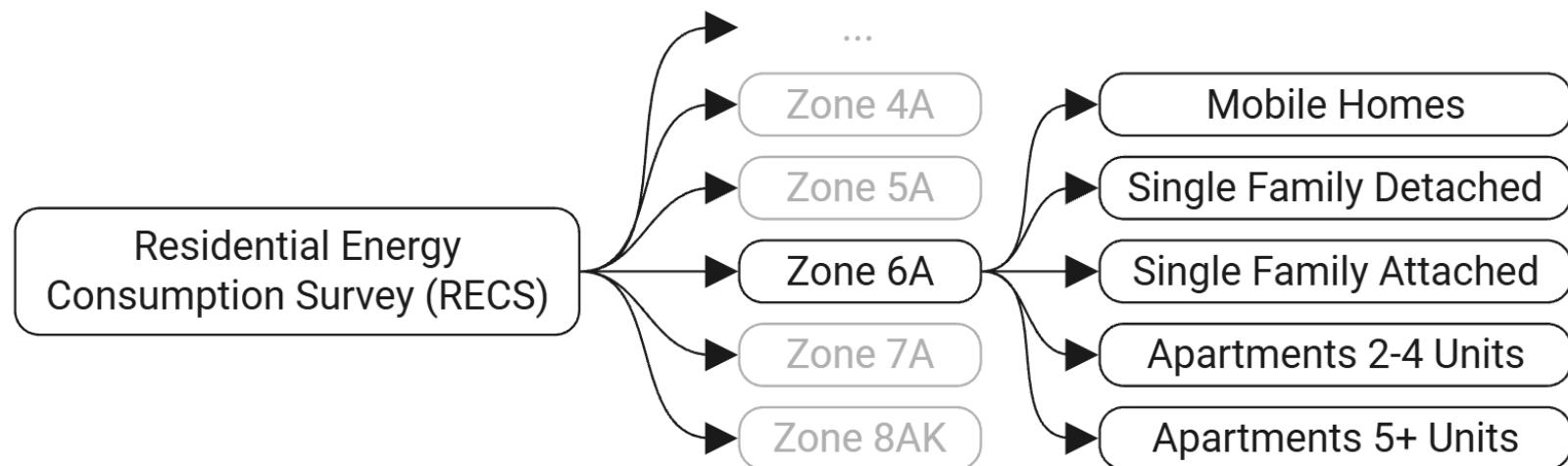
Compatibility with common modeling frameworks: ClimateStudio, E+ etc.

Separation between **fuel** types for retrofit targeting



STRATIFIED GROUPING/CLUSTERING MAKES ARCHETYPES EXPLAINABLE

- Stratification improves clustering interpretability
- Avoids mixing fundamentally different building types
- Each archetype remains consistent and readable

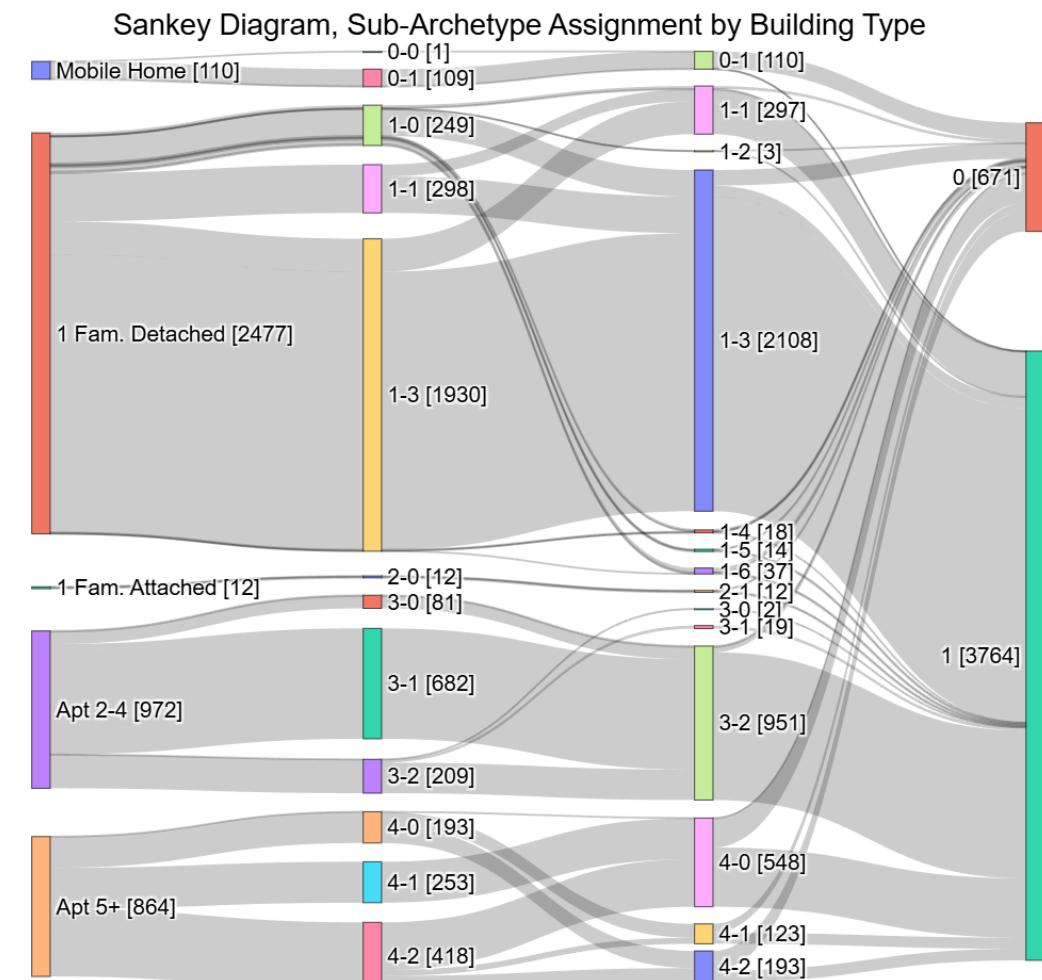


MODULAR ARCHETYPES ARE REUSABLE ENCODINGS OF THE BUILDING STOCK

- Modular decomposition allows flexibility and scalability
- Encoded as reusable JSON objects for common (U)BEM tools

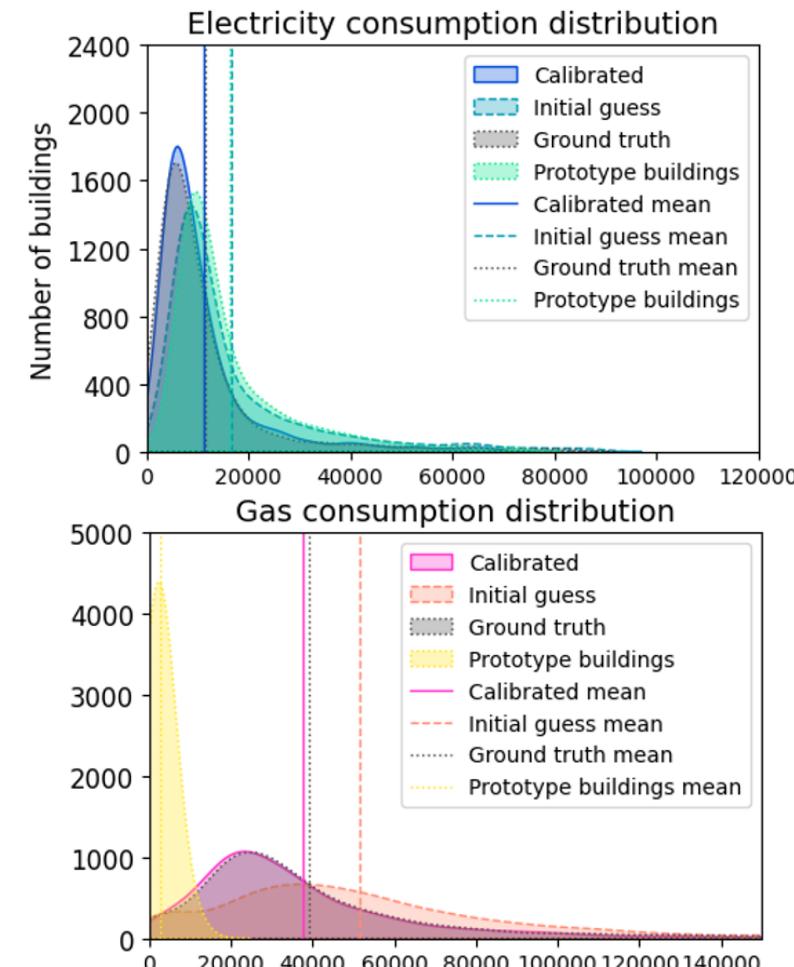
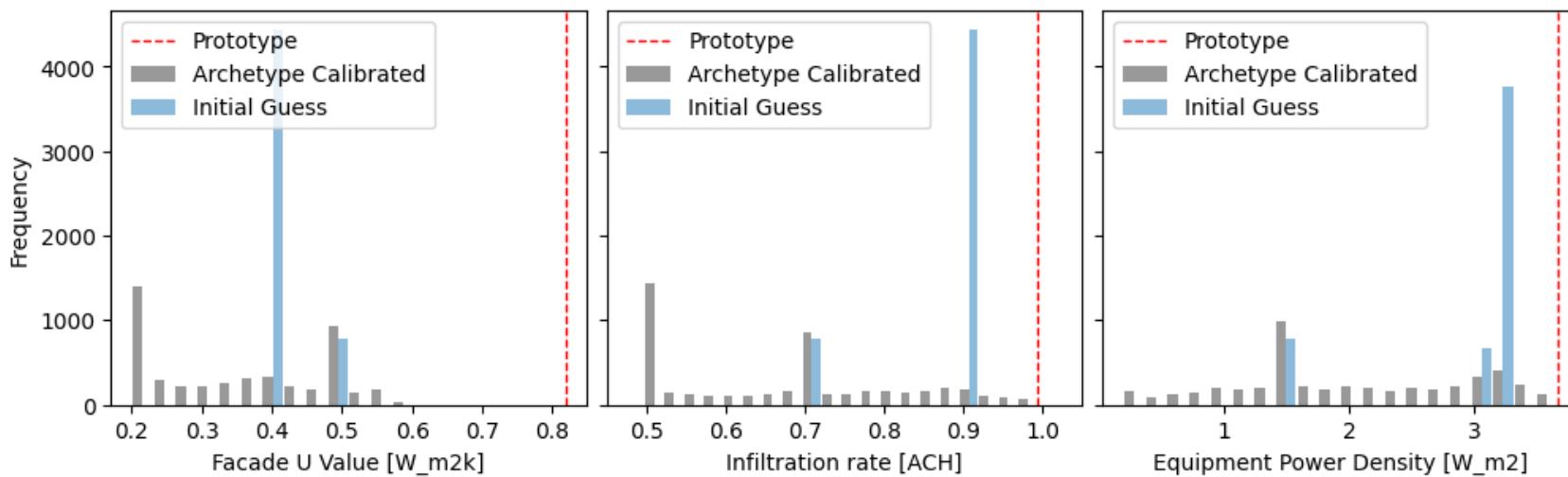
Construction Archetypes
Systems Archetypes
Energy Loads Archetypes

Building Archetype
(Free combination)



ARCHETYPES IMPROVE UBEM INPUT PARAMETER “INITIAL GUESSES”

- Archetype-informed inputs outperform DOE prototypes, and offer “local stock insights”
- Reduces error in early-stage modeling and scenario analysis



FUTURE DIRECTIONS

- Methodological
 - Expert knowledge-informed feature weighting
 - Stratification across different attributes
 - Incorporate smart-meter based schedules
- Application
 - Expand to entire US (straightforward) & worldwide (challenging)
 - Publish of archetype JSON definitions (expected this year)
 - Characterize with empirical distributions & allow Bayesian calibration
 - Develop commercial archetypes from CBECS; and mixed-use buildings



SUMMARY, Q&A

- A modular, data-driven archetype framework
- Combinatory/flexible construction, systems, energy loads sub-archetypes from RECS
- Stratified groups+clustering: more explainable
- ML assignment based on “Known Attributes”
- Deterministic/feasible-range characterization
- Archetype improves “initial guesses”
- Available as JSON, scalable to entire U.S.
- Watchout for Archetype release later this year at
<https://github.com/Chengxuan-Li/USBuildingArchetypes>

Take a look
at the paper!



Scan for paper preview
www.lcx.works/files/recA.pdf

Save a copy of
this presentation!



Scan for presentation file
www.lcx.works/files/recS.pdf

Get a copy of this
presentation



Slide 28

BACKUP SLIDES FOR Q&A

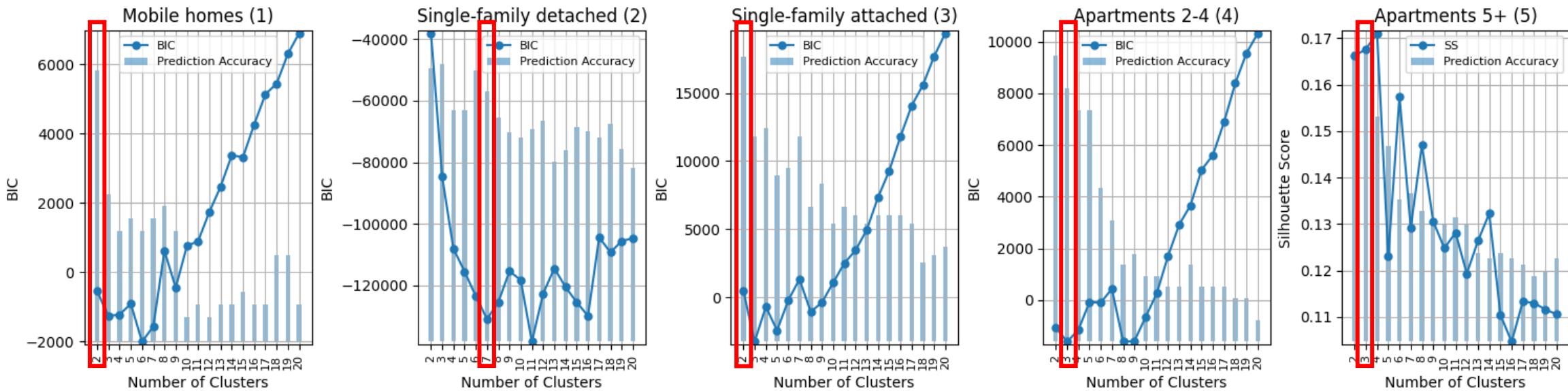


CHOICE OF CLUSTER COUNT, USING “CONSTRUCTION” AS AN EXAMPLE

Total sample counts are imbalanced across types

Type	Mobile Home	Single Family Detached	Single Family Attached	Apartments 2-4	Apartments 5+
# of Archetypes	2	7	2	3	3
# of RECS Records	83	1178	129	91	173
# Records Rep. Per	41.50	168.29	64.50	30.33	57.67
# of Total Homes	320802.60	5116205.51	332276.96	526323.04	1103535.78
# Homes Rep. Per	160401.30	730886.50	166138.48	175441.01	367845.26

Construction sub-archetypes clustering and prediction performance for different number of clusters



CONSTRUCTION ARCHETYPES: RESULT DETAILS

Name	Space heating equipment	Space heating fuel	AC type	Basement AC	Age of heating equipment	AC Age	Heating electricity EUI [kWh/m ²]	Heating natural gas EUI [btu/m ²]	Heating other fuels EUI [btu/m ²]	Cooling EUI [kWh/m ²]	Latent loads EUI [kWh/m ²]
Gas Furnace, Central AC Mobile Home ()	Central furnace	Natural gas	Central AC	No	20+	5-9	12.94	513.68	70.856	1.614	0.787
Oil Furnace, Wall AC Mobile Home	Central furnace	Fuel oil	Window/wall AC	No	20+	2-4	5.998	6.962	573.038	0.855	0.81
Propane Furnace, Central AC Home	Central furnace	Propane	Central AC	Yes	5-9	5-9	17.927	62.041	327.089	2.147	3.1
Typical-Efficiency Gas Furnace, Central AC Home	Central furnace	Natural gas	Central AC	Yes	5-9	5-9	6.469	402.799	4.105	2.136	3.271
Oil Furnace, Wall AC Home with Gas	Central furnace	Fuel oil	Window/wall AC	No	20+	NA	5.151	2.367	645.929	0.903	2.897
Low-Efficiency Gas Furnace, Wall AC Home	Central furnace	Natural gas	Window/wall AC	No	20+	NA	5.23	578.951	8.646	0.655	2.341
Heat Pump, Central AC Home	Central heat pump	Electricity	Central AC	Yes	5-9	10-14	15.922	6.317	137.183	2.023	0.961
Oil Furnace, Wall AC Home with No Gas	Central furnace	Fuel oil	Window/wall AC	No	20+	NA	5.382	0	614.55	0.543	1.613
Electric Heating, Wall AC Home	Built-in electric units	Electricity	Window/wall AC	No	20+	5-9	74.28	15.861	96.146	0.81	3.274
High-Efficiency Gas Furnace, Central AC Home	Central furnace	Natural gas	Central AC	No	<2	5-9	8.038	337.489	241.771	1.821	21.887
Low-Efficiency Gas Furnace, Central AC Attached Home	Central furnace	Natural gas	Central AC	No	10-14	<2	5.502	390.756	50.365	1.35	18.571
Typical-Efficiency Gas Furnace, Central AC Attached Home	Central furnace	Natural gas	Central AC	No	5-9	5-9	11.963	298.159	0	1.192	1.071
Propane Furnace, Wall AC Attached Home	Central furnace	Propane	Window/wall AC	No	20+	NA	4.866	0	430.784	0.372	1.03
Oil Furnace, Wall AC Apartment	Central furnace	Fuel oil	Window/wall AC	No	20+	NA	5.24	0	326.196	0.105	1.804
Gas Steam Heater, Wall AC Apartment	Steam/hot water	Natural gas	Window/wall AC	No	10-14	5-9	2.346	550.715	0	0.181	1.589
Gas Furnace, Central AC Apartment	Central furnace	Natural gas	Central AC	No	2-4	2-4	5.424	140.45	3.541	0.629	0.185
Electric Heating, Wall AC Apartment	Built-in electric units	Electricity	Window/wall AC	No	20+	NA	43.443	0	0	0.593	2.992
Oil Steam Heater, Wall AC Apartment	Steam/hot water	Fuel oil	Window/wall AC	No	15-19	5-9	6.8	79.255	198.556	3.578	0.053
Gas Furnace, Central AC Multi-Apartment	Central furnace	Natural gas	Central AC	No	20+	5-9	2.098	259.575	0.08	0.928	0.952
Propane Steam Heater, Wall AC Multi-Apartment	Steam/hot water	Propane	Window/wall AC	No	20+	5-9	1.408	0	221.281	0.956	0.793
Electric Heating, Wall AC Multi-Apartment	Built-in electric units	Electricity	Window/wall AC	No	20+	10-14	21.697	0.199	0	0.776	0.363



SYSTEMS ARCHETYPES: RESULT DETAILS

	#	Name	Basement	Finished basement	Attic space	Major outside wall material	Major roof material	Insulation level, respondent-reported	Type of glass in most windows	Window frame material	Window area per sqm	Door area per sqm	Heating electricity EUI [kWh/m ²]	Heating natural gas EUI [btu/m ²]	Heating other fuels EUI [btu/m ²]
1 (Mobile Home)	0	Standard Vinyl Window Mobile Home	No	No	No	Siding	Shingles	Adequately	Double	Vinyl	0.106	4.20E-03	13.573	496.557	54.324
	1	Standard Metal Window Mobile Home	No	No	No	Siding	Shingles	Adequately	Double	Metal	0.123	3.57E-03	8.519	7.093	591.871
2 (Single Family Detached)	0	Siding Façade Home	Yes	No	Yes	Siding	Shingles	Adequately	Double	Wood	0.103	5.03E-03	17.14	0	538.887
	1	Siding Façade Home with Finished Basement and Big WWR	Yes	Yes	Yes	Siding	Shingles	Adequately	Double	Wood	0.114	2.38E-03	11.702	493.27	4.06
	2	Wood Façade Home	Yes	No	Yes	Wood	Shingles	Adequately	Double	Wood	0.151	1.46E-03	7.425	38.555	910.399
	3	Siding Façade Home with Finished Basement and Moderate WWR	Yes	Yes	Yes	Siding	Shingles	Adequately	Double	Wood	0.089	4.75E-03	13.565	378.122	13.804
	4	Siding Façade Home with Metal Roof	Yes	No	Yes	Siding	Metal	Adequately	Double	Wood	0.104	4.42E-03	31.496	306.163	76.789
	5	Poorly-Insulated Shingles Façade Home	Yes	No	Yes	Shingles	Shingles	Adequately	Single	Wood	0.127	2.30E-03	14.849	699.074	5.065
	6	Well-Insulated Siding Façade Home	Yes	No	Yes	Siding	Shingles	Well	Double	Composite	0.094	5.87E-03	10.887	1.92	445.379
3 (Single Family Attached)	0	Standard Attached Home with Moderate WWR	Yes	No	No	Siding	Shingles	Adequately	Double	Wood	0.064	4.59E-03	14.664	303.351	1.078
	1	Standard Attached Home with Big WWR	Yes	No	No	Siding	Shingles	Adequately	Double	Wood	0.1	7.59E-03	7.192	0	439.279
4 (Apartment 2-4 Units)	0	Standard Apartment with Big WWR	No	No	No	Siding	Shingles	Adequately	Double	Wood	0.097	1.60E-03	7.629	0	275.373
	1	Standard Apartment with Moderate WWR	No	No	No	Siding	Shingles	Adequately	Double	Wood	0.063	3.42E-03	38.868	25.576	0
	2	Wood Façade Apartment	No	No	No	Wood	Shingles	Adequately	Double	Wood	0.084	3.17E-03	4.117	454.146	0
5 (Apartment 5+ Units)	0	Standard Multi-Apartment with Metal Windows (Electricity Heating)	No	No	No	Siding	NA	Adequately	Double	Metal	0.053	4.77E-03	18.144	3.985	39.339
	1	Standard Multi-Apartment with Metal Windows (Gas Heating)	No	No	No	Siding	NA	Adequately	Double	Metal	0.047	6.10E-03	3.391	289.319	0.097
	2	Standard Multi-Apartment with Wood Windows	No	No	No	Siding	NA	Adequately	Double	Wood	0.128	0.026	10.727	127.941	0



FEATURE IMPORTANCE

- Current assumption: independent + equal importance
- Justifications
 - Attributes clustered are designed to be as collinear as possible
 - Interpretability concerns
 - No expert knowledge required (no bias) – good for automation & scaling
- Potential improvements
 - Automatic feature weighting
 - PCA embedding



CLUSTERING WITH SAMPLE WEIGHT

- RECS has a “ZWEIGHT” column denoting sample weight
 - i.e. how many “homes” or “residential units” a record statistically represents in the U.S. residential housing stock
- Weighted K-Means: easy implementation
- Weighted GMM: needs extra work to implemented (already done, not within the scope of this paper)
 - More important for climate zones with fewer records, e.g. Alaska/8AK



ARCHETYPE REUSABILITY

- How far is this archetype method away from a ready-to-use JSON archetype for common BEM tools, such as e+ or ClimateStudio?
- Current Features:
 - Consistent schematic definitions of “power density”
 - Separation of Systems and Energy Loads (Space Loads)
- Areas for Improvement:
 - No archetypical schedules (RECS does not have them) – could be compensated with smart meter derived archetypes

