Same Data, Multiple Scales

Aggregates

Web Viewer

Individual Buildings

Real data will be spikier
Pre-aggregated Load Profiles

Pre-aggregated EULPs by building type for:
- U.S. States (contiguous)
- ASHRAE Climate Zones
- DOE Building America Climate Zones
- Electric System ISOs
- U.S. Census Public Use Microdata Area*
- U.S. Counties

Format:
- CSV files (for Excel, etc. ease of use)

Additional Data:
- Count of models included per aggregation
- List of model IDs per aggregation
- Model characteristics by ID
- Timeseries mean, stdev, and range

*PUMA is an area with ~200k people; ~2,400 in U.S.
VizStock Web Interface

Aggregates

Web Viewer

Individual Buildings

- View End Use Load Profiles
- View distributions of building characteristics
- Filter by building characteristic
- Filter by geography
- Select time window
- Download CSV of results
Individual Building End Use Load Profiles

- ~450,000 residential
- ~350,000 commercial
- Full dataset will be 10’s of terabytes
- Plan to include high-level instructions for loading this dataset using one cloud-based big-data analysis tool

Format:
- Folders with a series of Apache parquet* files
  - Likely 1 file per building, with IDs in names
- In Amazon S3 bucket or similar

Additional Data:
- Model characteristics by ID
- Model in OpenStudio (.osm) format

*https://parquet.apache.org/
Typical Meteorological Year (TMY3)
- Widely accepted/expected by utilities, regulators, etc.
- Weather is not coordinated across regions

<table>
<thead>
<tr>
<th>Month</th>
<th>Weather Data from Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>Denver, CO 1995</td>
</tr>
<tr>
<td>February</td>
<td>Boulder, CO 1987</td>
</tr>
<tr>
<td>March</td>
<td>Denver, CO 1994</td>
</tr>
<tr>
<td>April</td>
<td>Boulder, CO 1990</td>
</tr>
</tbody>
</table>

Actual Meteorological Year (AMY)
- Using 2018 NOAA data

Format:
- CSV timeseries data for each location used
  - Dry bulb temperature
  - Relative humidity
  - Solar direct normal irradiation
  - Solar diffuse horizontal irradiation
  - Wind speed
  - Building characteristics
- Location used for each Model

2 locations 40 miles apart use data from different years for the same month
Time Stamps & Time Zones

Time Zones:
• Data will be provided in UTC

Time Stamps:
• Wrap data from first few hours of year back to the end
• Creates a single, aligned 1 year worth of data
## Residential Building Types & End Uses

<table>
<thead>
<tr>
<th>Residential Building Types</th>
<th>Residential End Uses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-Family Detached</td>
<td>Heating</td>
</tr>
<tr>
<td>Multifamily (low-rise) Single-Family Attached</td>
<td>Cooling</td>
</tr>
<tr>
<td>Multifamily (low-rise) 2–4 Units</td>
<td>Furnace/AC fan</td>
</tr>
<tr>
<td>Multifamily (low-rise) 5+ Units</td>
<td>Boiler pumps</td>
</tr>
<tr>
<td></td>
<td>Vent. fans</td>
</tr>
<tr>
<td></td>
<td>Water heating</td>
</tr>
<tr>
<td></td>
<td>Interior Lights</td>
</tr>
<tr>
<td></td>
<td>Exterior Lights</td>
</tr>
<tr>
<td></td>
<td>Misc. plug loads</td>
</tr>
<tr>
<td></td>
<td>Refrigerator</td>
</tr>
<tr>
<td></td>
<td>Clothes washer</td>
</tr>
<tr>
<td></td>
<td>Clothes dryer</td>
</tr>
<tr>
<td></td>
<td>Dishwasher</td>
</tr>
<tr>
<td></td>
<td>Cooking Range</td>
</tr>
</tbody>
</table>
## Commercial Building Types & End Uses

<table>
<thead>
<tr>
<th>Commercial Building Types</th>
<th>Commercial End Uses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Office</td>
<td>Heating</td>
</tr>
<tr>
<td>Medium Office</td>
<td>Cooling</td>
</tr>
<tr>
<td>Large Office</td>
<td>Interior Lighting</td>
</tr>
<tr>
<td>Stand-alone Retail</td>
<td>Exterior Lighting</td>
</tr>
<tr>
<td>Strip Mall</td>
<td>Interior Equipment</td>
</tr>
<tr>
<td>Primary School</td>
<td>Exterior Equipment</td>
</tr>
<tr>
<td>Secondary School</td>
<td>Fans</td>
</tr>
<tr>
<td>Outpatient Healthcare</td>
<td>Pumps</td>
</tr>
<tr>
<td>Hospital</td>
<td>Heat Rejection</td>
</tr>
<tr>
<td>Small Hotel</td>
<td>Humidification</td>
</tr>
<tr>
<td>Large Hotel</td>
<td>Heat Recovery</td>
</tr>
<tr>
<td>Warehouse (non-refrigerated)</td>
<td>Water Systems</td>
</tr>
<tr>
<td>Quick Service Restaurant</td>
<td>Refrigeration</td>
</tr>
<tr>
<td>Full Service Restaurant</td>
<td></td>
</tr>
<tr>
<td>Mid-rise Apartment</td>
<td></td>
</tr>
<tr>
<td>High-rise Apartment</td>
<td></td>
</tr>
</tbody>
</table>
# Residential Building Characteristics

<table>
<thead>
<tr>
<th>Residential Model Characteristics</th>
<th>(continued)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ahhs_region</td>
<td>heating_fuel</td>
</tr>
<tr>
<td>applicable</td>
<td>heating_setpoint</td>
</tr>
<tr>
<td>ashrea_icec_climate_zone_2004</td>
<td>heating_setpoint_has_offset</td>
</tr>
<tr>
<td>bathroom_spot_vent_hour</td>
<td>heating_setpoint_offset_magnitude</td>
</tr>
<tr>
<td>bedrooms</td>
<td>heating_setpoint_offset_period</td>
</tr>
<tr>
<td>building_america_climate_zone</td>
<td>holiday_lighting</td>
</tr>
<tr>
<td>ceiling_fan</td>
<td>hot_water_distribution</td>
</tr>
<tr>
<td>census_division</td>
<td>hot_water_fixtures</td>
</tr>
<tr>
<td>census_region</td>
<td>hvac_system_cooling</td>
</tr>
<tr>
<td>climate_zone_ba</td>
<td>hvac_system_cooling_type</td>
</tr>
<tr>
<td>climate_zone_icecc</td>
<td>hvac_system_heat_pump</td>
</tr>
<tr>
<td>clothes_dryer</td>
<td>hvac_system_heating_electricity</td>
</tr>
<tr>
<td>clothes_washer</td>
<td>hvac_system_heating_fuel_oil</td>
</tr>
<tr>
<td>clothes_washer_presence</td>
<td>hvac_system_heating_natural_gas</td>
</tr>
<tr>
<td>cooking_range</td>
<td>hvac_system_heating=None</td>
</tr>
<tr>
<td>cooking_range_schedule</td>
<td>hvac_system_heating_thermohydrogen</td>
</tr>
<tr>
<td>cooling_setpoint</td>
<td>hvac_system_heating_propane</td>
</tr>
<tr>
<td>cooling_setpoint_has_offset</td>
<td>hvac_system_is_heat_pump</td>
</tr>
<tr>
<td>cooling_setpoint_offset_magnitude</td>
<td>hvac_system_is_shared</td>
</tr>
<tr>
<td>cooling_setpoint_offset_period</td>
<td>hvac_system_shared_electricity</td>
</tr>
<tr>
<td>corridor</td>
<td>hvac_system_shared_fuel_oil</td>
</tr>
<tr>
<td>county</td>
<td>hvac_system_shared_natural_gas</td>
</tr>
<tr>
<td>days_shifted</td>
<td>hvac_system_shared=None</td>
</tr>
<tr>
<td>dehumidifier</td>
<td>hvac_system_shared_other_fuel</td>
</tr>
<tr>
<td>dishwasher</td>
<td>hvac_system_shared_propane</td>
</tr>
<tr>
<td>door_area</td>
<td>infiltration</td>
</tr>
<tr>
<td>doors</td>
<td>insulation_crawlspace</td>
</tr>
<tr>
<td>ducts</td>
<td>insulation_finished_basement</td>
</tr>
<tr>
<td>eaves</td>
<td>insulation_finished_roof</td>
</tr>
<tr>
<td>electric_vehicle</td>
<td>insulation_interzonal_floor</td>
</tr>
<tr>
<td>geometry_building_number_units_hl</td>
<td>insulation_pier_beam</td>
</tr>
<tr>
<td>geometry_building_number_units_mf</td>
<td>insulation_slab</td>
</tr>
<tr>
<td>geometry_building_number_units_sfa</td>
<td>insulation_unfinished_attic</td>
</tr>
<tr>
<td>geometry_building_type_acs</td>
<td>insulation_unfinished_basement</td>
</tr>
<tr>
<td>geometry_building_type_ress</td>
<td>insulation_watt</td>
</tr>
<tr>
<td>geometry_floor_area</td>
<td>interior_shading</td>
</tr>
<tr>
<td>geometry_floor_area_bin</td>
<td>iso_rto_region</td>
</tr>
<tr>
<td>geometry_foundation_type</td>
<td>lighting</td>
</tr>
<tr>
<td>geometry_garage</td>
<td>lighting_interior_use</td>
</tr>
<tr>
<td>geometry_perimeter_footprint_ratio</td>
<td>lighting_other_use</td>
</tr>
<tr>
<td>geometry_stories</td>
<td>location</td>
</tr>
<tr>
<td>geometry_wall_type</td>
<td>location_city</td>
</tr>
</tbody>
</table>

NREL | 11
### Commercial Building Characteristics

<table>
<thead>
<tr>
<th>Commercial Model Characteristics</th>
<th>(continued)</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>building_type</code></td>
<td><code>cooling_source_fuel</code></td>
</tr>
<tr>
<td><code>climate_zone</code></td>
<td><code>heating_source_fuel</code></td>
</tr>
<tr>
<td><code>weather_file_name</code></td>
<td><code>hvac_delivery_type</code></td>
</tr>
<tr>
<td><code>rentable_area</code></td>
<td><code>service_water_heating_source_fuel</code></td>
</tr>
<tr>
<td><code>number_stories</code></td>
<td><code>kitchen_makeup</code></td>
</tr>
<tr>
<td><code>aspect_ratio</code></td>
<td><code>exterior_lighting_zone</code></td>
</tr>
<tr>
<td><code>total_bldg_floor_area</code></td>
<td><code>onsite_parking_fraction</code></td>
</tr>
<tr>
<td><code>bottom_story_groundExposed_floor</code></td>
<td><code>energy_code_when_built</code></td>
</tr>
<tr>
<td><code>building_height_relative_to_neighbors</code></td>
<td><code>energy_code_when_envelope_last_updated</code></td>
</tr>
<tr>
<td><code>building_rotation</code></td>
<td><code>energy_code_when_exterior_lighting_last_updated</code></td>
</tr>
<tr>
<td><code>floor_to_floor_height</code></td>
<td><code>energy_code_when_hvac_last_updated</code></td>
</tr>
<tr>
<td><code>party_wall_stories_west</code></td>
<td><code>party_wall_fraction</code></td>
</tr>
<tr>
<td><code>single_floor_area</code></td>
<td><code>party_wall_stories_east</code></td>
</tr>
<tr>
<td><code>story_multiplier</code></td>
<td><code>party_wall_stories_north</code></td>
</tr>
<tr>
<td><code>top_story_exterior_exposed_roof</code></td>
<td><code>party_wall_stories_south</code></td>
</tr>
<tr>
<td><code>window_to_wall_ratio</code></td>
<td><code>energy_code_when_interior_equipment_last_updated</code></td>
</tr>
<tr>
<td><code>hvac_system_type</code></td>
<td><code>energy_code_when_interior_lighting_last_updated</code></td>
</tr>
<tr>
<td><code>energy_code_when_service_water_heating_last_updated</code></td>
<td></td>
</tr>
<tr>
<td><code>weekday_operation_start_time</code></td>
<td></td>
</tr>
<tr>
<td><code>weekday_operation_duration</code></td>
<td></td>
</tr>
<tr>
<td><code>weekend_operation_start_time</code></td>
<td></td>
</tr>
<tr>
<td><code>weekend_operation_duration</code></td>
<td></td>
</tr>
</tbody>
</table>
Commercial Region 1 Calibration

Andrew Parker
Matthew Dahlhausen, Ph.D.
September 22, 2020
Calibration Strategy
Model Architecture

Building stock characteristics database

Physics-based computer modeling

National
Climate/Region
State
County

Modeling Algorithms
Schedules
Human Behavior

Performance Curves
Component Properties
Weather Data
Calibration Process for One Region

Focus on reducing error for one region at a time

Keep an eye on impacts to other regions

Before Calibration

After Calibration

Error

Region 1 Calibration | Region 2 Calibration | Region 3 Calibration | Region 4 Calibration | Region 5 Calibration

Region 1
Calibration Process Over Time

Focus on reducing error for one region at a time

Keep an eye on impacts to earlier regions

Error

Region 1 Calibration  Region 2 Calibration  Region 3 Calibration  Region 4 Calibration  Region 5 Calibration

Region 1  Region 2
Calibration Process Over Time

Calibration efforts for earlier regions create better starting point for later regions

Error
Calibration Process Over Time

Improvements from later regions will improve results for regions focused on earlier.
Region 1 Focus: Major Schedules

Building stock characteristics database

Physics-based computer modeling

Lighting schedules
Plug load schedules
Thermostat setpoints

National
Climate/Region
State
County
Region 1

- Fort Collins, Colorado (pop. ~160k)
- Municipal Utility
- Commercial Stock Summary

<table>
<thead>
<tr>
<th>Building Type</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>small_office</td>
<td>369</td>
</tr>
<tr>
<td>strip_mall</td>
<td>223</td>
</tr>
<tr>
<td>retail</td>
<td>181</td>
</tr>
<tr>
<td>warehouse</td>
<td>153</td>
</tr>
<tr>
<td>outpatient</td>
<td>96</td>
</tr>
<tr>
<td>full_service_restaurant</td>
<td>72</td>
</tr>
<tr>
<td>quick_service_restaurant</td>
<td>29</td>
</tr>
<tr>
<td>medium_office</td>
<td>28</td>
</tr>
<tr>
<td>primary_school</td>
<td>21</td>
</tr>
<tr>
<td>large_hotel</td>
<td>8</td>
</tr>
<tr>
<td>small_hotel</td>
<td>7</td>
</tr>
<tr>
<td>large_office</td>
<td>4</td>
</tr>
</tbody>
</table>

- Primarily used AMI data from 2016
ComStock Sensitivity Analysis

Analyzed impact on ~15 QOIs ... focused on inputs highly rated for many QOIs
List of updates

New capabilities
• None

Baseload updates
• Interior lighting schedules
• Plug load schedules

HVAC updates
• Thermostat setpoints
ComStock Calibration Before EULP

ComStock & CBECS Energy Estimate Comparisons by Census Division

TWh

New England  Mid-Atlantic  East North Central  West North Central  South Atlantic  East South Central  West South Central  Mountain  Pacific

CBECS Target

ComStock (CS) Results
Challenges of Commercial AMI data
# Sanity Checking AMI Data vs CBECS

<table>
<thead>
<tr>
<th>building_type</th>
<th>CBECS Mountain Region (kWh/sf)</th>
<th>Fort Collins AMI (kWh/sf)</th>
</tr>
</thead>
<tbody>
<tr>
<td>full_service_restaurant</td>
<td>34*</td>
<td>67</td>
</tr>
<tr>
<td>hospital</td>
<td>27</td>
<td>NA</td>
</tr>
<tr>
<td>large_hotel</td>
<td>22</td>
<td>12</td>
</tr>
<tr>
<td>large_office</td>
<td>14*</td>
<td>16</td>
</tr>
<tr>
<td>medium_office</td>
<td>14*</td>
<td>14</td>
</tr>
<tr>
<td>outpatient</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>primary_school</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>quick_service_restaurant</td>
<td>34*</td>
<td>123</td>
</tr>
<tr>
<td>retail</td>
<td>24</td>
<td>40</td>
</tr>
<tr>
<td>secondary_school</td>
<td>10</td>
<td>NA</td>
</tr>
<tr>
<td>small_hotel</td>
<td>22</td>
<td>9</td>
</tr>
<tr>
<td>small_office</td>
<td>14*</td>
<td>31</td>
</tr>
<tr>
<td>strip_mall</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>warehouse</td>
<td>4</td>
<td>22</td>
</tr>
</tbody>
</table>

Mean EUI of AMI data is much higher/lower than CBECS for some building types... why?

* CBECS doesn’t break out quick vs. full svc. Restaurant or office size category
## Distribution of AMI Data

<table>
<thead>
<tr>
<th></th>
<th>Annual EUI (kWh/sqft)</th>
<th>Percentiles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>P05</td>
</tr>
<tr>
<td>full_service_restaurant</td>
<td>76.8</td>
<td>11.7</td>
</tr>
<tr>
<td>large_hotel</td>
<td>20.2</td>
<td>9.0</td>
</tr>
<tr>
<td>large_office</td>
<td>18.3</td>
<td>7.2</td>
</tr>
<tr>
<td>medium_office</td>
<td>17.0</td>
<td>4.8</td>
</tr>
<tr>
<td>midrise_apartment</td>
<td>29.3</td>
<td>3.4</td>
</tr>
<tr>
<td>outpatient</td>
<td>15.1</td>
<td>3.3</td>
</tr>
<tr>
<td>primary_school</td>
<td>9.1</td>
<td>1.8</td>
</tr>
<tr>
<td>quick_service_restaurant</td>
<td>174.4</td>
<td>48.1</td>
</tr>
<tr>
<td>retail</td>
<td>40.8</td>
<td>2.6</td>
</tr>
<tr>
<td>small_hotel</td>
<td>9.8</td>
<td>1.3</td>
</tr>
<tr>
<td>small_office</td>
<td>30.0</td>
<td>3.8</td>
</tr>
<tr>
<td>strip_mall</td>
<td>28.0</td>
<td>5.0</td>
</tr>
<tr>
<td>warehouse</td>
<td>62.2</td>
<td>1.3</td>
</tr>
</tbody>
</table>

Mean EUI almost as high as 95th percentile!
Outliers Skewing Mean AMI Data

Are these really the building type we think they are?
V1 outlier removal methods:
Outliers change AMI significantly: +/- 30%

AMI w/ no outliers removed
AMI w/ high & low outliers removed
AMI w/ high outliers removed
Investigating V1 Outliers with Google Maps

- strip_mall (23 outliers)
  - 9 are convenience store/gas stations
  - 11 are restaurants (or primarily restaurants)
- warehouse (18 outliers)
  - 13 are manufacturing
  - 2 are autobody shops
- small_office (13 outliers)
  - 2 are manufacturing
  - 1 is a nursery/greenhouse
  - 1 is a multifamily condo w/ maybe office space on first floor?
  - The rest are just normal-looking offices
- retail (5 outliers)
  - 3 are nursery/greenhouses
- outpatient (4 outliers)
  - All appear to legitimately be outpatient... perhaps some specialties use much more energy?
- quick_service_restaurant (1 outlier)
  - Drive-through where service is not tied to floor area

Conclusion: Most “outliers” were actually misclassified buildings, not truly “outliers” of the target building type
Identifying Misclassified AMI Systematically

Refine the misclassification/outlier detection process.

Goals:
• Keep as much AMI as possible
• Can’t rely on manually classifying every AMI building

Approach:
1. Run each method on AMI
2. Find method that identifies largest number of buildings
3. Manually classify all these outliers using google maps
4. Compare the results of the various methods
### Identifying Misclassified AMI - Results

Methods tested:

1. +/- 1.5x IQR: original approach, only removes super high EUIs
2. +/- 1.5x natural log(IQR): incorrectly removes low but realistic EUIs
3. 25th-75th percentile: too conservative, removes half of buildings every time
4. +/- 10x median: misses low EUIs
5. +/- 5x median: slightly better
6. +/- 3x median: best balance we’ve found

<table>
<thead>
<tr>
<th>Building type</th>
<th>Total AMI count</th>
<th>+/- 1.5x IQR</th>
<th>+/- 1.5x ln(IQR)</th>
<th>25th-75th percentile</th>
<th>10x median</th>
<th>5x median</th>
<th>3x median</th>
</tr>
</thead>
<tbody>
<tr>
<td>small_office</td>
<td>369</td>
<td>28</td>
<td>18</td>
<td>184</td>
<td>11</td>
<td>26</td>
<td>58</td>
</tr>
<tr>
<td>strip_mall</td>
<td>223</td>
<td>24</td>
<td>10</td>
<td>110</td>
<td>14</td>
<td>37</td>
<td>65</td>
</tr>
<tr>
<td>retail</td>
<td>181</td>
<td>16</td>
<td>6</td>
<td>90</td>
<td>10</td>
<td>23</td>
<td>57</td>
</tr>
<tr>
<td>warehouse</td>
<td>153</td>
<td>18</td>
<td>18</td>
<td>76</td>
<td>16</td>
<td>25</td>
<td>45</td>
</tr>
<tr>
<td>outpatient</td>
<td>96</td>
<td>4</td>
<td>14</td>
<td>48</td>
<td>9</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>full_service_restaurant</td>
<td>72</td>
<td>3</td>
<td>2</td>
<td>36</td>
<td>1</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>quick_service_restaurant</td>
<td>29</td>
<td>2</td>
<td>1</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>medium_office</td>
<td>28</td>
<td>1</td>
<td>2</td>
<td>14</td>
<td>2</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>primary_school</td>
<td>21</td>
<td>0</td>
<td>2</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>large_hotel</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>small_hotel</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>large_office</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Final Method Selection: +/- 3x Median

Still uses most of the AMI

<table>
<thead>
<tr>
<th>Building type</th>
<th>Total AMI Count</th>
<th>True Positive</th>
<th>False Positive</th>
<th>False Positive Rate</th>
<th>Valid AMI Preserved (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>full_service_restaurant</td>
<td>72</td>
<td>5</td>
<td>5</td>
<td>50%</td>
<td>93%</td>
</tr>
<tr>
<td>large_hotel</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100%</td>
</tr>
<tr>
<td>large_office</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100%</td>
</tr>
<tr>
<td>medium_office</td>
<td>28</td>
<td>3</td>
<td>2</td>
<td>40%</td>
<td>92%</td>
</tr>
<tr>
<td>outpatient</td>
<td>96</td>
<td>9</td>
<td>7</td>
<td>44%</td>
<td>92%</td>
</tr>
<tr>
<td>primary_school</td>
<td>21</td>
<td>2</td>
<td>0</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>quick_service_restaurant</td>
<td>29</td>
<td>0</td>
<td>1</td>
<td>100%</td>
<td>97%</td>
</tr>
<tr>
<td>retail</td>
<td>181</td>
<td>46</td>
<td>11</td>
<td>19%</td>
<td>92%</td>
</tr>
<tr>
<td>small_hotel</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>50%</td>
<td>83%</td>
</tr>
<tr>
<td>small_office</td>
<td>369</td>
<td>45</td>
<td>13</td>
<td>22%</td>
<td>96%</td>
</tr>
<tr>
<td>strip_mall</td>
<td>223</td>
<td>58</td>
<td>7</td>
<td>11%</td>
<td>96%</td>
</tr>
<tr>
<td>warehouse</td>
<td>153</td>
<td>38</td>
<td>7</td>
<td>16%</td>
<td>94%</td>
</tr>
<tr>
<td>Grand Total</td>
<td>1191</td>
<td>207</td>
<td>54</td>
<td>21%</td>
<td>95%</td>
</tr>
</tbody>
</table>

True Positive: correctly identified a misclassified building
False Positive: identified a building as misclassified, but the original classification was correct
Misclassified AMI: +/- 3x Median

<table>
<thead>
<tr>
<th>Building type</th>
<th>false positive</th>
<th>quick_service_restaurant</th>
<th>empty or abandoned</th>
<th>retail or strip malls that are mostly or entirely restaurants</th>
<th>full_service_restaurant</th>
<th>medium_office</th>
<th>primary_school</th>
<th>quick_service_restaurant</th>
<th>retail</th>
<th>outpatient</th>
<th>small_hotel</th>
<th>small_office</th>
<th>strip_mall</th>
<th>warehouse</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Retail or strip malls that are mostly or entirely restaurants</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>B. Empty or abandoned buildings (don’t use much energy)</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>C. Residential homes that are converted into dentist/doctor/counseling/etc. offices</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>D. More restaurants</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>E. Strip malls or retail that are convenience stores, which have refrigeration</td>
<td>11</td>
<td>10</td>
<td>8</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>F. Warehouses that are light manufacturing of some sort</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
CoStar (our classification starting point) classifies based on real-estate needs
- Some are clear: offices, outpatient, standalone retail
- Some are ambiguous: strip malls, warehouses

For region 1, addresses were available per-building
- Were able to investigate specific outliers
- Must use this opportunity to refine classification & outlier detection

Was important for Region 1
Will be a major focus for Region 2
Matching AMI Data to Buildings

- Some datasets have one service address per-AMI meter
- Ft. Collins...not so simple

- Many are 1 AMI meter, 1 parcel, 1 account, 1 building
- Some are not
- Getting area right is critical for EUI calculation
Processing Commercial end use data
Commercial End Use Data

• Res end use data – typically collected with energy research in mind
• Com end use data – typically collected for another reason & purchased for EULP
  • Data format varies – a few days to reformat/clean up each dataset
  • Labeling/data not always certain
    • A meter is labeled “lighting,” but does that circuit have any other loads on it?
    • How much of the building area is served by a given circuit?

Takes time & effort to use each dataset
Baseload Updates
## Update: Lighting Schedules

<table>
<thead>
<tr>
<th>Task</th>
<th>Affected Building Type</th>
<th>Considerations</th>
</tr>
</thead>
</table>
| Update normalized lighting schedules | retail, full service restaurant, warehouse, office         | • Calculated average daily (in hourly interval) profiles for each building type and for each day of week.  
• Initial understanding of variability of profiles was also explored. |
|                               | primary school and secondary school                      | • Calculated average daily (in hourly interval) profiles for each building type and for each day of week.  
• Differentiated average profiles between academic and summer break periods. |
| Update lighting power density  | retail, full service restaurant, warehouse, office         | • Normalized profile can underestimate usage since it is based on measured peak power.  
• Quantified approximate difference between design peak power and measured peak power (≈5%).  
• Lighting power density adjusted. |
Details of using procured end use data

<table>
<thead>
<tr>
<th>Data source 1</th>
<th>Metadata</th>
<th>Timeseries Data</th>
<th>Two different sources in two different formats</th>
<th>Total of 7355 unique submeter labels</th>
<th>Multiple emails and phone calls</th>
<th>Grouping building types and enduse types separately</th>
<th>~1000 buildings for each enduse type</th>
<th>Ready for post-process</th>
</tr>
</thead>
<tbody>
<tr>
<td>~300 GB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>~3000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>buildings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1min interval</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data source 2</th>
<th>Metadata</th>
<th>Timeseries Data</th>
<th>Two different sources in two different formats</th>
<th>Total of 7355 unique submeter labels</th>
<th>Multiple emails and phone calls</th>
<th>Grouping building types and enduse types separately</th>
<th>~1000 buildings for each enduse type</th>
<th>Ready for post-process</th>
</tr>
</thead>
<tbody>
<tr>
<td>~50 GB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>~70 buildings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15mins interval</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Pre-process

Post-process

<table>
<thead>
<tr>
<th>Building Type</th>
<th>Building Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mercantile Retail</td>
<td>787</td>
</tr>
<tr>
<td>FoodService Restaurant</td>
<td>135</td>
</tr>
<tr>
<td>Education Restaurant</td>
<td>27</td>
</tr>
<tr>
<td>Warehouse</td>
<td>26</td>
</tr>
<tr>
<td>Office</td>
<td>25</td>
</tr>
<tr>
<td>FoodSales Grocery</td>
<td>23</td>
</tr>
<tr>
<td>Other</td>
<td>17</td>
</tr>
<tr>
<td>Service</td>
<td>11</td>
</tr>
<tr>
<td>Warehouse Refrigerated</td>
<td>5</td>
</tr>
<tr>
<td>Data Center</td>
<td>2</td>
</tr>
<tr>
<td>Lodging Hotel</td>
<td>1</td>
</tr>
<tr>
<td>Office Medical</td>
<td>1</td>
</tr>
</tbody>
</table>

Understanding formatting
Review/Classify submeter labels
Review with data owner
Create common metadata
Filter down to certain enduse type
Create common data frame

Normalized Lighting Power [W]
Impact: Lighting Schedules

Before

Lights go on later in the morning on weekdays

After

Lights stay on later into the evening on weekdays and weekends
## Update: Plug Load Schedules

<table>
<thead>
<tr>
<th>Task</th>
<th>Affected Building Type</th>
<th>Considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Update normalized plug load schedules</td>
<td>retail, full service restaurant, grocery, warehouse, office</td>
<td>• Calculated average daily (in hourly interval) profiles for each building type and for each day of week.</td>
</tr>
<tr>
<td></td>
<td>primary school, secondary school</td>
<td>• Calculated average daily (in hourly interval) profiles for each building type and for each day of week.</td>
</tr>
<tr>
<td></td>
<td>retail, full service restaurant, grocery, warehouse, office, primary school, secondary school</td>
<td>• Differentiated average profiles between academic and summer break periods.</td>
</tr>
<tr>
<td>Update equipment power density</td>
<td>retail, full service restaurant, grocery, warehouse, office, primary school, secondary school</td>
<td>• Normalized profile can underestimate usage since it is based on measured peak power.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Equipment power density adjusted.</td>
</tr>
</tbody>
</table>
Details of using procured end use data

<table>
<thead>
<tr>
<th>Building Type</th>
<th>Building Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>FoodService_Restaurant</td>
<td>697</td>
</tr>
<tr>
<td>Mercantile_Retail</td>
<td>59</td>
</tr>
<tr>
<td>Warehouse</td>
<td>31</td>
</tr>
<tr>
<td>FoodSales_Grocery</td>
<td>30</td>
</tr>
<tr>
<td>Education_School</td>
<td>28</td>
</tr>
<tr>
<td>Office</td>
<td>27</td>
</tr>
<tr>
<td>Lodging_Hotel</td>
<td>1</td>
</tr>
<tr>
<td>Office_Medical</td>
<td>1</td>
</tr>
<tr>
<td>Service</td>
<td>0</td>
</tr>
<tr>
<td>DataCenter</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
</tr>
</tbody>
</table>

Building types with greater than 20 sample buildings were included.

Manually identified meter data that is not relevant to plug loads. About 90 keywords identified as irrelevant. Examples:

- “Lighting”
- “Air Conditioner”
- “Unknown”

Calculated median building plug load schedule for each day of the week and each building type.

Visually compared load profiles between the days of the week to determine whether days could be combined. For example, warehouses have similar weekday profiles, but weekends should have a unique schedule.

Figure 1. Warehouse profiles
Impact: Plug Load Schedules

Plug loads bring building energy use closer to the base load.
HVAC Updates
# Update: Thermostat Setpoints

<table>
<thead>
<tr>
<th>Task</th>
<th>Affected Building Type</th>
<th>Methods</th>
</tr>
</thead>
</table>
| Update normalized thermostat heating and cooling setpoint schedules | retail, strip mall, quick-service restaurant, full-service restaurant, grocery, office | • Calculated average daily (in hourly interval) profiles for each building type and each day of the week.  
• Initial understanding of variability of profiles was also explored for future calibration efforts. |
Details of using procured end use data

1. Identified heating and cooling thermostat setpoint meters in datasets.

2. Calculated average heating and cooling profiles for each building for each day of the week.

3. Profiles were aggregated to produce representative heating and cooling setpoint profiles

### Building Types

<table>
<thead>
<tr>
<th>Building Type</th>
<th>Building Count With Thermostat Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>FoodService_Restaurant</td>
<td>1817</td>
</tr>
<tr>
<td>Mercantile_Retail</td>
<td>1692</td>
</tr>
<tr>
<td>FoodSales_Grocery</td>
<td>164</td>
</tr>
<tr>
<td>Office</td>
<td>31</td>
</tr>
<tr>
<td>Education_School</td>
<td>16</td>
</tr>
<tr>
<td>Warehouse</td>
<td>4</td>
</tr>
<tr>
<td>Lodging_Hotel</td>
<td>4</td>
</tr>
<tr>
<td>Hospital</td>
<td>2</td>
</tr>
<tr>
<td>Outpatient</td>
<td>2</td>
</tr>
</tbody>
</table>

Building types with greater than 20 sample buildings were included.

Example of cooling thermostat data producing average Monday profiles.
Impact: Thermostat Setpoints

Before

Cooling & Fan
Energy Increased

After

Overshot
Weekends

Summer_Weekday

Summer_Weekend

Strip mall, Day Type Comparison by Enduse

Summer_Weekday

Summer_Weekend

Hour of day (0-23)

Hour of day (0-23)

Electric Load (kWh/h)

Electric Load (kWh/h2)

refrigeration
heating
cooling
pumps
fans
heat_recovery
hot_water
interior_equipment
interior_lighting
exterior_lighting
ami_2016_3x_median_filter +5%
ami_2016_3x_median_filter
ami_2016_3x_median_filter -5%
Commercial stock
end-use summary
Total building stock load by day type

Missing summer cooling?
Small office by day type

Graph showing electric load (kWh/ft²) for different days and hours, with labels for various loads such as refrigeration, heating, cooling, fans, pumps, heat recovery, hot water, interior equipment, interior lighting, exterior lighting. Annotations indicate missing daytime cooling and heating loads.
Strip mall by day type

Seasonal difference in base loads?
Retail by day type

Missing nighttime load, some cooling, mostly base
Warehouse by day type

Bias in dataset used to pull lighting and plug load schedules?
Base run does not show this bias
Tracking Quantities of Interest
Region 1 Focus: Annual Error

Relative error: annual electricity use per ft²

Low annual usage overall; missing nighttime load and summer cooling in some building types
Region 1 Focus: Total Error Metrics

Timing of peak heating:
- Summer peak timing is relatively accurate.
- Little high on winter days.

Graphs showing:
- Average of All Days
- Top 10 Days
- Peak Timing

AMI no outlier filtering vs. AMI 3x median filtering.
Areas for Improvement
Impact of Outliers

Outliers & classification: more impactful than any model changes we’ve done thus far!

V1 outlier identification methods
Next Steps: AMI Classification & Outliers

Investigate approach using a much bigger dataset

- Have dataset of ~500k monthly meters from Xcel Energy
- ~200k have CoStar matches (for metadata)

1. Take a random sample and manually classify
2. Calculate rate of true/false positives & true/false negatives
3. Look for systematic misclassifications or reasons for outliers
4. Improve outlier/misclassification detection (if possible)
Next Steps: Model Improvements

• Update energy code adoption by state/year
• Implement RTU efficiency & performance changes
  • RTU data analyzed by excellent intern this summer!

Ensure point of comparison is correct first

• Need to ensure correct classification, perhaps change schedules space use types for things like warehouses, which act more like offices. None of the model updates considered warehouses.
• Consider exterior lighting.
• Add schedule, enduse lighting/equipment variability.
• Restaurant kitchen equipment, particularly hot water use.
Conclusions

• Ran 6 iterations of ComStock incorporating 3 discrete changes
• Total and individual building type load shapes mostly look good, ignoring magnitude
• Determined that AMI classification & screening needs improvement
  • Will impact AMI comparisons moving forward
  • Have a plan to improve this
• Will be moving on to Region 2, but continue tracking Region 1 metrics
Calibration Strategy
Model Architecture

ResStock

Housing stock characteristics database

National
Climate/Region
State
Public Use Microdata Area (PUMA)

Physics-based computer modeling

Modeling Algorithms
Schedules
Human Behavior
Performance Curves
Component Properties
Weather Data
Region 2 Focus: Nationally-Relevant Updates

Housing stock characteristics database

Miscellaneous electric loads (MELs) regression equations

Physics-based computer modeling

Appliance baseload schedules

Cooling/heating Setpoints

Heating fuel type

National

Climate/Region

State

Public Use Microdata Area (PUMA)

Modeling Algorithms

Schedules

Human Behavior

Performance Curves

Component Properties

Weather Data
Calibration Process for One Region

Focus on reducing error for one region at a time

Keep an eye on impacts to other regions

Before Calibration

After Calibration

Error

Region 1 Calibration | Region 2 Calibration | Region 3 Calibration | Region 4 Calibration | Region 5 Calibration

Region 1
Calibration Process Over Time

Focus on reducing error for one region at a time

Keep an eye on impacts to earlier regions
Calibration Process Over Time

Calibration efforts for earlier regions create better starting point for later regions.
Calibration Process Over Time

Improvements from later regions will improve results for regions focused on earlier.
Region 2 Calibration Strategy

- Heating fuel type by PUMA
- Cooling/Heating Setpoints

- Baseload schedules
- MELS annual load by building type

- Switch HVAC housing characteristics
- Vacant units

- Weighting factors for mapping to EIA form 861 service territory

- RECS end-use comparisons
- 2018 Update EIA scatterplots
Residential Calibration Dimensions

- AMI data from future region 5
- AMI data from future region 4
- AMI data from Seattle City Light, WA
- AMI data from Fort Collins municipal service territory

Calibration effort

- EIA Form 861
- EIA natural gas data
- RECS end-use scatterplots
- Utility load research data (LRD)
- Submeter end-use data

- Annual electric sales of all utilities in U.S.
- Annual and monthly natural gas consumption by state
- Annual end-use loads of occupied dwelling units
  - Building type
  - Climate zone
  - Fuel (electricity, natural gas, propane, fuel oil)
- Load duration curves and seasonal load shapes of >20 utilities around U.S.
- Sub-metered end-use load data power levels and load shapes

Advanced metering infrastructure (AMI) data from ComEd service territory
Region 2 – Fort Collins, CO

- Fort Collins, Colorado (pop. ~160k)
- Municipal Utility
- Primarily used AMI data from 2018

<table>
<thead>
<tr>
<th>Residential Building Stock Summary</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Building Type Distribution</strong></td>
<td></td>
</tr>
<tr>
<td>Single-Family Detached</td>
<td>66.0%</td>
</tr>
<tr>
<td>Multi-Family 5+ Units</td>
<td>19.7%</td>
</tr>
<tr>
<td>Single-Family Attached</td>
<td>7.7%</td>
</tr>
<tr>
<td>Multi-Family 2-4 Units</td>
<td>6.6%</td>
</tr>
<tr>
<td><strong>Heating Fuel Distribution</strong></td>
<td></td>
</tr>
<tr>
<td>Natural Gas</td>
<td>69.7%</td>
</tr>
<tr>
<td>Electricity</td>
<td>23.5%</td>
</tr>
<tr>
<td>Propane</td>
<td>4.0%</td>
</tr>
<tr>
<td>Other</td>
<td>2.6%</td>
</tr>
</tbody>
</table>
List of updates

New validation comparisons
• Updated EIA form 861 annual residential sales scatterplot to 2018
• Updated EIA natural gas data comparison to 2018
• Updated EIA natural gas data to standard temperature and pressure (STP) by TAG suggestion
• RECS 2009 and RECS 2015 annual end-use scatterplots by fuel type, building type, and climate

New capabilities
• Restructured HVAC housing characteristics to allow for regional and more granular datasets
• Introduced vacant units

Baseload updates
• Appliance schedules informed by stochastic occupancy model
• Annual miscellaneous electric loads (MELs) broken out by building type

HVAC updates
• Setpoints now described by IECC climate zones and RECS building type
• Heating fuel type described by Public Use Microdata Area (PUMA)
• More diversity added to cooling and heating setback schedules
• Investigated effect of number of days for natural ventilation
• Investigated impact of new HVAC housing characteristics structure
New validation comparisons
RECS end-use scatterplots

- Scatterplots for annual fuel consumption
- Dimensions:
  - RECS 2009 and RECS 2015
  - Building America climate zone
  - Building Type
  - Fuel (electricity, natural gas, propane, fuel oil)

Most likely over predicting cooling

** End-use totals in RECS are modeled, but comparisons are useful for determining large errors and potential end-uses to improve.
2018 EIA data comparisons (updated from 2012)

- Adjusted to standard temperature and pressure (STP), TAG suggestions

Relatively accurate

Low on gas consumption
Scaling Factors Update
Update: EIA service territory and customer mapping

Before Region 2 calibration
• Data sources
  • Service territory shape files (2012)
  • NSRDB grid cells
  • Census tracts
  • Dwelling unit counts

After Region 2 calibration
• Data sources
  • EIA Form 861
    • Customer counts
    • Service territory (counties)
  • American Community Survey (ACS)
    • Dwelling unit counts by county

- Only have mapping for 2012
- All units get allocated causing errors
- Utility service territory to NSRDB grid cell
- Dwelling units to NSRDB grid cell
- Allocate units based on area service territory weight
- Does not preserve number of customers

- Can be updated for any year
- Preserves number of customers
- Join EIA form 861 service territory to ACS unit counts
- Calculate weighting factors from ACS
- Allocate customers based on weights
Impact: EIA service territory and customer mapping

Before mapping update

After mapping update

ResStock: Annual Retail Sales (TWh)

EIA Form 861: Annual Retail Sales (TWh)

2012 Residential Annual Retail Sales by Utility

Error line: 00%
Error line: 10%
Error line: 20%
Utility

Florida Power & Light Co
Pacific Gas & Electric Co
Commonwealth Edison Co
Virginia Electric & Power Co
Enterov Texas Inc
Consolidated Edison Co-NY Inc

R-squared: 0.95
Slope: 0.97
Intercept: 0.00

ResStock: Annual Retail Sales (TWh)

EIA Form 861: Annual Retail Sales (TWh)

2012 Residential Annual Retail Sales by Utility

Error line: 00%
Error line: 10%
Error line: 20%
Utility

Pacific Gas & Electric Co
Commonwealth Edison Co
Duke Energy Carolinas, LLC
Consolidated Edison Co-NY Inc
Public Service Elec & Gas Co

R-squared: 0.99
Slope: 1.18
Intercept: 0.00

Resolves large scaling issues
Increases $R^2$
Over predicting
Added Capabilities
Update: Introduce Vacant Units

- American Community Survey (ACS) Public Use Microdata Sample (PUMS) 2016
  - Vacant dwelling units by building type at PUMA-region resolution (N=\~2,400)
  - \~14,000,000 vacant units in U.S.
  - Higher fraction of vacant multi-family units in city centers
  - *Modeled reduced baseloads, but buildings still heated and cooled*
Update: Introduce Vacant Units

- American Community Survey (ACS) Public Use Microdata Sample (PUMS) 2016
- Vacant dwelling units by building type at PUMA-region resolution (N=~2,400)
- ~14,000,000 vacant units in U.S.
- Higher fraction of vacant multi-family units in city center
- Modeled reduced baseloads, but buildings still heated and cooled

Multi Family Units

Single Family Units
Impact: Introduce Vacant Units

Cohort: Single-Family Detached, Fort Collins

The PUMA containing Fort Collins has 7.3% vacant units
Impact: Introduce Vacant Units

Without Vacant Units

2018 Residential Annual Retail Sales by Utility

With Vacant Units

2018 Residential Annual Retail Sales by Utility

Increased accuracy in slope
Update: Improve HVAC housing characteristics

- HVAC distributions derived from RECS:
  - Heating and cooling equipment type
  - Equipment efficiency levels (based on age of equipment from RECS and efficiency vs. age distributions derived from AHRI data and other sources)
- Automated update process and improved structure of dependencies,
  - Simplified from 37,071 rows in 17 tables to 2,416 rows in 7 tables
  - Will make future updates easier
  - Makes outputs more transparent and easier to understand
  - Simplifies handling of heat pumps and HVAC systems serving multiple units
  - Separates assignment of equipment type (furnace, boiler, heat pump, etc.) from efficiency level
    - Enables us to use American Housing Survey (N=114,860) instead of RECS (N=12,083) for HVAC equipment type, improving geographic resolution
    - Efficiency distributions still depend on equipment type
Update: Improve HVAC housing characteristics

Improved distribution of **Room AC efficiency** using historical ENERGY STAR saturation data, ENERGY STAR minimum efficiency values over time, and federal minimum efficiency values over time.

<table>
<thead>
<tr>
<th>Year</th>
<th>Product</th>
<th>Metric</th>
<th>Shipment Weighted Value</th>
<th>Federal Minimum Value</th>
<th>ENERGY STAR Minimum</th>
<th>Percent ENERGY STAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1970</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>5.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>9.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>9.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>9.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>9.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>9.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>9.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>9.63</td>
<td>9.79</td>
<td>10.88</td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>9.75</td>
<td>9.79</td>
<td>10.88</td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>9.75</td>
<td>9.79</td>
<td>10.88</td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>9.71</td>
<td>9.79</td>
<td>10.88</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>9.95</td>
<td>9.79</td>
<td>10.88</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>10.02</td>
<td>9.79</td>
<td>10.88</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>9.81</td>
<td>9.79</td>
<td>10.88</td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>9.93</td>
<td>9.79</td>
<td>10.88</td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>9.93</td>
<td>9.79</td>
<td>10.88</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>9.79</td>
<td>9.79</td>
<td>10.88</td>
<td>33%</td>
</tr>
<tr>
<td>2011</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>9.79</td>
<td>9.79</td>
<td>10.88</td>
<td>62%</td>
</tr>
<tr>
<td>2012</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>9.79</td>
<td>9.79</td>
<td>10.88</td>
<td>58%</td>
</tr>
<tr>
<td>2013</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>9.79</td>
<td>9.79</td>
<td>10.88</td>
<td>72%</td>
</tr>
<tr>
<td>2014</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>11</td>
<td>11</td>
<td>10.88</td>
<td>50%</td>
</tr>
<tr>
<td>2015</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>11</td>
<td>11</td>
<td>10.88</td>
<td>54%</td>
</tr>
<tr>
<td>2016</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>11</td>
<td>12</td>
<td>10.88</td>
<td>38%</td>
</tr>
<tr>
<td>2017</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>11</td>
<td>12</td>
<td>10.88</td>
<td>34%</td>
</tr>
<tr>
<td>2018</td>
<td>Room AC</td>
<td>EER/CEER</td>
<td>11</td>
<td>12</td>
<td>10.88</td>
<td>42%</td>
</tr>
</tbody>
</table>

Table sources/notes:
- Shipment weighted values for Room AC comes from AHRI data from Home Energy Score documentation (1970–2008; 2009 is a copy of 2008 values)
- Federal minimum values from a descriptive paragraph at [https://appliance-standards.org](https://appliance-standards.org)
- ENERGY STAR, 2014-2015 values are a simplification based on a document at energystar.gov
- ENERGY STAR, 2016-2020 values are a simplification based on a document at energystar.gov
- ENERGY STAR EER value simplification based on data from data.energystar.gov (majority of typical units are EER=12.0)
- Percent ENERGY STAR comes from data found on energystar.gov (no data for 2009 and prior; 2019 and 2020 copied from 2018)
Update: Improve HVAC housing characteristics

- Updated reference year for HVAC and refrigerator ages (and thus efficiencies) from 2009 to 2018.

For example:

- In 2009, a 15-year-old furnace was made in 1994:

- In 2018, a 15-year-old furnace was made in 2003:
Impact: Improve HVAC housing characteristics

Before HVAC improvement

After HVAC improvement

2018 Residential Annual Retail Sales by Utility

- Error line: 00%
- Error line: 10%
- Error line: 20%
- Utility

ResStock: Annual Retail Sales (TWh)

EIA Form 861: Annual Retail Sales (TWh)

R-squared: 0.99
Slope: 1.12
Intercept: 0.01

R-squared: 0.98
Slope: 1.06
Intercept: 0.02

Decrease in slope
Impact: Improve HVAC housing characteristics

Fort Collins, CO

ComEd, IL

Small increase in cooling

Significant decrease in cooling

Winter largely not affected

Electric Load (kWh/unit)

Hour of day (0-23)

After HVAC restructure
Before HVAC restructure
AMI or LRD +5%
AMI or LRD average
AMI or LRD -5%
Baseload Updates
Update: Integrated stochastic occupancy model

→ Stochastic occupant-driven load model is now used for every ResStock run

Example of changes for one home:
**Impact: Integrated stochastic occupancy model**

**Cohort: Total Residential Stock Fort Collins**

Jagged loads signal lack of stochasticity

Smoother load with same number of samples

Overall load shape improved
Update: MELs equations by building type

Previous approach
- Single regression equation for all building types (RECS 2015)

\[ MELS = 908.91 + 277.75n_{\text{occupants}} + 0.39 ffa \]

New approach
- Separate regression equation for each building type (RECS 2015)

\[ MELS_{SFD} = 1146.95 + 296.94 n_{\text{occupants}} + 0.30 ffa \]
\[ MELS_{SFA} = 1395.84 + 136.53 n_{\text{occupants}} + 0.16 ffa \]
\[ MELS_{MF} = 875.22 + 184.11 n_{\text{occupants}} + 0.38 ffa \]

\( n_{\text{occupants}} \): Number of occupants
\( ffa \): Finished floor area
SFD: Single-Family Detached
SFA: Single-Family Attached
MF: Multi-Family

** MELS are defined by the following fields in RECS 2015: televisions, microwaves, humidifiers, and other devices not elsewhere classified
Impact: MELs equations by building type

Previous approach
• Single regression equation for all building types

\[
MELS = 908.91 + 277.75 n_{\text{occupants}} + 0.39 ffa
\]

New approach
• Separate regression equation for each building type

\[
\begin{align*}
MELS_{\text{SFD}} &= 1146.95 + 296.94 n_{\text{occupants}} + 0.30 ffa \\
MELS_{\text{SFA}} &= 1395.84 + 136.53 n_{\text{occupants}} + 0.16 ffa \\
MELS_{\text{MF}} &= 875.22 + 184.11 n_{\text{occupants}} + 0.38 ffa
\end{align*}
\]

\(n_{\text{occupants}}\): Number of occupants
\(ffa\): Finished floor area
SFD: Single-Family Detached
SFA: Single-Family Attached
MF: Multi-Family

RECS 2015 percent differences

<table>
<thead>
<tr>
<th>Building Type</th>
<th>Error: 1 equation</th>
<th>Error: 3 equations</th>
</tr>
</thead>
<tbody>
<tr>
<td>MF</td>
<td>21.4%</td>
<td>6.1%</td>
</tr>
<tr>
<td>SFA</td>
<td>13.5%</td>
<td>2.0%</td>
</tr>
<tr>
<td>SFD</td>
<td>-12.4%</td>
<td>-9.2%</td>
</tr>
</tbody>
</table>

** MELS are defined by the following fields in RECS 2015: televisions, microwaves, humidifiers, and other devices not elsewhere classified
HVAC Updates
Update: Heating Fuel described by PUMA

Before: the EULP project
Heating fuel described by groups of states

After: Calibration region 2
Heating fuel described by PUMA
Impact: Heating Fuel described by PUMA

Cohort: Total Residential Stock

Region: Fort Collins

Region: ComEd

Winter_Weekday

Electric Load (kwh/unit)

Hour of day

- Heating fuel by PUMA
- Heating fuel by State
- AMI +5%
- AMI average
- AMI -5%

Change not substantial in Fort Collins

Significant improvement in ComEd
Update: Setpoint & setback distributions and setback time period diversity

- Before Calibration Region 2
  - Setpoints by AIA Climate Zone

- After Calibration Region 2
  - Setpoints by IECC 2004 climate and moisture zone
  - Increased diversity of setback schedules (±5 hours)
Update: Setpoint & setback distributions and setback time period diversity

- After Calibration Region 2
  - Setpoints by IECC 2004 climate and moisture zone
  - Increased diversity of setback schedules (±5 hours)

Example: Average Cooling Setpoint after update (we use distributions)

Example: Average Heating Setpoint after update (we use distributions)
Update: Setpoint & setback distributions and setback time period diversity

On average:

- Dry climates use higher setpoint at night
- Moist climates use cooler setpoint at night

Average difference between nighttime and home AC setpoints, by IECC Climate Zone (source: RECS 2009)
Impact: Setpoint distributions and diversity

Fort Collins, CO
Total Residential Stock

ComEd, IL
Total Residential Stock

Overall, total stock load not affected

Setpoint Diversity resulting in better shape
Impact: Setpoint distributions and diversity

Fort Collins, CO
Selected cohorts

Coiling energy changed for some cohorts

ComEd, IL
Selected cohorts

Some shape improvement to electric heating cohort
Residential stock end-use summary

Fort Collins municipal utility, CO
Seasonal end-use loads by day type

Fort Collins municipal utility, CO
Seasonal end-use loads by day type

Fort Collins municipal utility, CO

![Graph showing seasonal end-use loads by day type for Fort Collins municipal utility, CO. The graphs compare electric load across different times of the day, with indicators for various household activities and load scenarios.]
Seasonal end-use loads by day type

Fort Collins municipal utility, CO

Winter_Weekday

Winter_Weekend

Electric Load (kWh/unit)

Hour of day (0-23)
Residential stock end-use summary

ComEd service territory, IL
Seasonal end-use loads by day type

ComEd service territory, IL

[Graph showing seasonal end-use loads by day type for summer weekday and weekend. The graph illustrates the hourly electric load for different load categories such as heating, cooling, and various appliances like dishwashers, clothes dryers, and refrigerators. The data is presented for a service territory in Illinois.]
Seasonal end-use loads by day type

ComEd service territory, IL
Seasonal end-use loads by day type

ComEd service territory, IL

Winter_Weekday

Winter_Weekend

Electric Load (kWh/unit)

Hour of day (0-23)

LRD +5%

LRD

LRD -5%
Tracking Quantities of Interest
Fort Collins, CO: Annual Error

Relative error: annual electricity use per unit

Reasons
- Combination of a set of end-uses consuming too much electricity
- Cooling energy still too high
- Lighting, water heating, laundry, and/or cooking may also be too high
ComEd, IL: Annual Error

Reasons
- The increase initially at the start of region 2 is due to the stochastic occupant model
- Heating fuel helped correction
- Vacant units helped correction
- HVAC Restructure helped correction

Relative error: annual electricity use per unit

Stochastic occupancy model integration

Only slightly low after corrections
Fort Collins, CO: Total Error Metrics

Average of All Days

Top 10 Days

Peak Timing

Base Load

Average Peak

Top 10 Peaks

Mild summer days issues

Timing of peak heating relatively accurate

High peak load

Fixed winter peak timing
ComEd, IL: Total Error Metrics

Average of All Days

Top 10 Days

Baseline still an issue

Improved winter peak

HVAC restructure

Timing of peak heating relatively accurate

Peak Timing

Base Load

Average Peak

Top 10 Peaks

Region 1

Region 2

Summer

Shoulder

Winter

0 5 10 15 20 25 30

Run Number

0 5 10 15 20 25 30

Run Number

0 5 10 15 20 25 30

Run Number

Average Peak

Top 10 Peaks

Region 1

Region 2

Summer

Shoulder

Winter

0 5 10 15 20 25 30

Run Number

0 5 10 15 20 25 30

Run Number

0 5 10 15 20 25 30

Run Number
Areas for Improvement
Two regions provides additional insight into areas for improvement

Fort Collins, CO

Fort Collins still shows too much cooling, especially in the shoulder season

ComEd, IL
Next Region: Likely Areas for Improvement

Two regions provide additional insight into areas for improvement:

Fort Collins, CO

Fort Collins still shows too much cooling, especially in the shoulder season.

→ Incorporate more seasonal usage of AC.

ComEd, IL

Fort Collins still shows too much cooling, especially in the shoulder season.

→ Incorporate more seasonal usage of AC.
Next Region: Likely Areas for Improvement

Two regions provides additional insight into areas for improvement

Fort Collins, CO

Fort Collins still shows too much cooling, especially in the shoulder season

› Incorporate more seasonal usage of AC

ComEd, IL

ComEd peak magnitude is good, but still too low at night

Lighting drops off too quickly?
Two regions provides additional insight into areas for improvement

**Fort Collins, CO**

- Fort Collins still shows too much cooling, especially in the shoulder season
- incorporate more seasonal usage of AC

**ComEd, IL**

- ComEd peak magnitude is good, but still too low at night
- Use end-use datasets and the American Time Use Survey to investigate regionality of baseload schedules
- Investigate microclimate and moisture capacitance effects

Lighting drops off too quickly?
HVAC structure improvement enables use of more granular HVAC data:

- Transition from RECS to American Housing Survey (AHS)
  - Cooling saturation and type (central vs. window AC)
  - Heating system type (furnace vs. boiler; electric baseboard vs. heat pump)
- Water heating fuel type
- AHS is broken into census divisions and the largest 15 metro areas
  - Help towards improving urban and rural models
  - AHS has more samples than RECS, which enables more accurate slicing by vintage, region, and building type
Conclusions (1)

- Ran 14 iterations of ResStock incorporating 8 discrete changes
  - Saw general improvements in QOI metrics
  - Most of the improvements made will carry over to the entire U.S.
- Structural changes to HVAC housing characteristics
  - Enables use for more granular data
  - Will benefit all regions moving forward
- Integrated the residential stochastic occupant-driven load model
  - Now used for every run
- New/Updated visualizations
  - Updated EIA data comparisons from 2012 to 2018
  - RECS 2009 and 2015 end-use scatterplots
Conclusions (2)

• Summary of changes
  • Reduced baseload by adding dwelling unit vacancy
  • More accurate scaling factors for mapping to utility service territories
  • Heating fuel distributions refined with sub-state resolution
  • Improved climate dependence of cooling & heating setpoints and setbacks
  • Improved diversity of cooling and heating setback period start and end times
  • Separated MELs regression equations by building type

• Priority areas for improvement for next region
  • Regionally variable schedule and power level baseloads
  • Potentially missing thermal mass or nighttime baseloads in ComEd service territory
  • Geographically granular HVAC system saturations and water heating fuel type saturations

• Will be moving on to Regional Dataset 3 (Seattle, WA), but continue tracking metrics for the first two region datasets