

U.S. DEPARTMENT OF  
**ENERGY**

Office of  
ENERGY EFFICIENCY &  
RENEWABLE ENERGY

# End-Use Load Profiles for the U.S. Building Stock

Methodology and Results of Model Calibration, Validation,  
and Uncertainty Quantification

March 2022



(This page intentionally left blank)



## Disclaimer

This work was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or any third party's use or the results of such use of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof or its contractors or subcontractors. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, its contractors or subcontractors.

This research was performed using computational resources sponsored by the Department of Energy's Office of Energy Efficiency and Renewable Energy and located at the National Renewable Energy Laboratory and Argonne National Laboratory.

## Authors

The authors of this report are:

Eric J.H. Wilson, National Renewable Energy Laboratory (NREL)  
Andrew Parker, NREL  
Anthony Fontanini, NREL  
Elaina Present, NREL  
Janet L. Reyna, NREL  
Rajendra Adhikari, NREL  
Carlo Bianchi, NREL  
Christopher CaraDonna, NREL  
Matthew Dahlhausen, NREL  
Janghyun Kim, NREL  
Amy LeBar, NREL  
Lixi Liu, NREL  
Marlena Praprost, NREL  
Liang Zhang, NREL  
Peter DeWitt, NREL  
Noel Merket, NREL  
Andrew Speake, NREL  
Tianzhen Hong, Berkeley Lab  
Han Li, Berkeley Lab  
Natalie Mims Frick, Berkeley Lab  
Zhe Wang, Berkeley Lab  
Aileen Blair, NREL  
Henry Horsey, NREL  
David Roberts, NREL  
Kim Trenbath, NREL  
Oluwatobi Adekanye, former NREL  
Eric Bonnema, NREL  
Rawad El Kontar, NREL  
Jonathan Gonzalez, former NREL  
Scott Horowitz, NREL  
Dalton Jones, former NREL  
Ralph T. Muehleisen, Argonne  
Siby Platthotam, Argonne  
Matthew Reynolds, NREL  
Joseph Robertson, NREL  
Kevin Sayers, NREL  
Qi Li, former Argonne

Contributions of each author are listed in Appendix B.

The technical monitor of this project is:

Monica Neukomm, U.S. Department of Energy (DOE)

## Acknowledgments

The authors would like to acknowledge the valuable guidance and input provided by the end-use load profile technical advisory group over the three years of work documented in this report. The list of technical advisory group members is listed in Appendix A.

We are grateful to the data sharing partners who shared anonymized data with the project team or helped the team find data: Adams 12 Five Star Schools, AES Indiana, the Alliance Center, Ameren Missouri, Bert Brains, Bonneville Power Administration, Center for Energy and Environment, Center for the Built Environment, Cherryland Electric Cooperative, City of Fort Collins Utilities and Colorado State University, City of Tallahassee Utilities, Clarkson University, DNV, Ecotope, Efficiency Maine, Elevate, EPB, Horry Electric Cooperative, Hot Water Research, kW Engineering, Los Angeles Department of Water and Power, the Massachusetts Program Administrators, the National Rural Electric Cooperative Association, New City Energy, New York State Energy Research and Development Authority, Northeast Energy Efficiency Partnerships, NV Energy, PacifiCorp, PEPCO and Exelon, Portland General Electric, Powerhouse Dynamics, Resource Central, Seattle City Light, Southern Company, Vermont Energy Investment Corporation and Green Mountain Power, Xcel Energy, and numerous additional partners who wished to remain publicly anonymous. Northwest Energy Efficiency Alliance, Pecan Street, Florida Solar Energy Center (FSEC), Southern Company, and Massachusetts Energy Efficiency Advisory Council deserve special mention for their leadership and foresight into end-use load research, as do ComEd and ecobee for making anonymized data available to researchers. Future users of the load profiles created by this project owe these partners a debt of gratitude; without their significant effort and contributions, the project would not have been possible.

We thank the staff at the U.S. Energy Information Administration (EIA) who collect and analyze data from Forms 861, 861M, and 176, as well as the RECS and CBECS teams. We are particularly grateful to Greg Lawson of EIA who served on our technical advisory group and provided invaluable contributions related to our use of 861M data.

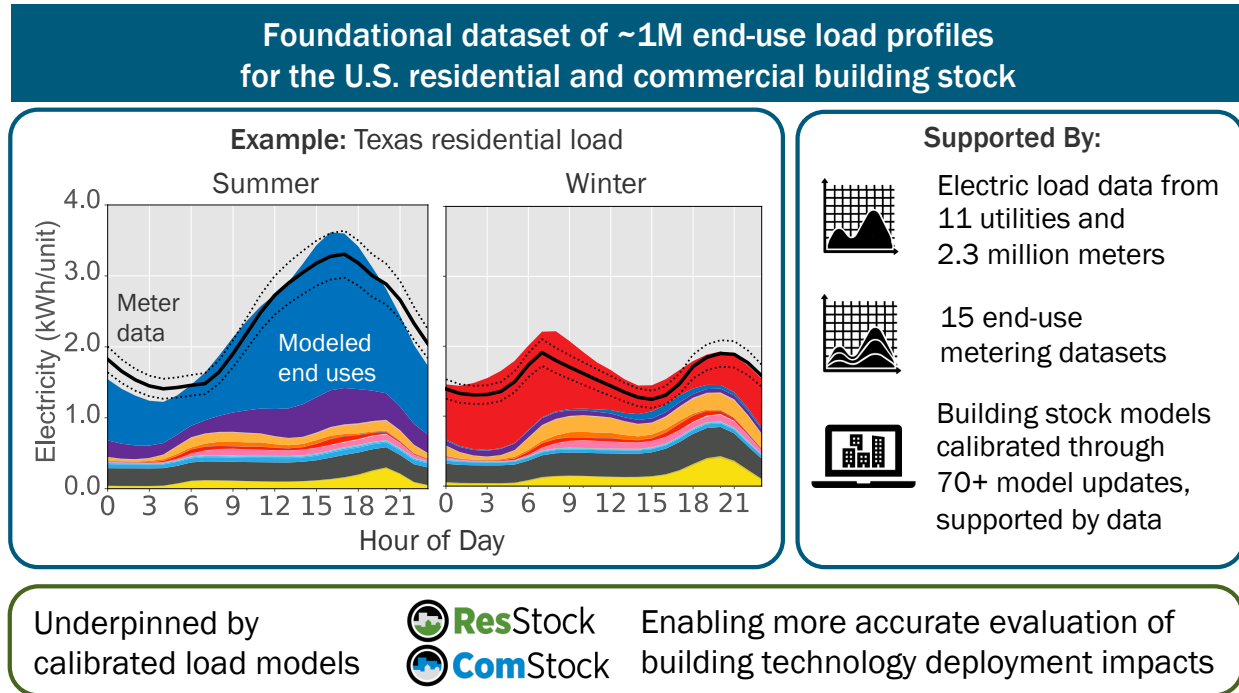
Finally, we would like to thank the project's technical monitor, Monica Neukomm at the U.S. Department of Energy (DOE), for guiding the project team over the past three years and ensuring our success. We also thank David Nemtzw, Erika Gupta, Joan Glickman, Dale Hoffmeyer, Eric Werling, Amy Jiron, Harry Bergmann, and Amir Roth at DOE for their support of our work and past support of ResStock<sup>TM</sup> and ComStock<sup>TM</sup>.

## Acronyms

ACH <sub>50</sub>	air changes per hour at an induced pressure differential of 50 Pa
ACS	American Community Survey
AHS	American Housing Survey
AIA	American Institute of Architects
AMI	advanced metering infrastructure
AMY	annual meteorological year
ANL	Argonne National Laboratory
ATUS	American Time Use Survey
BAS	building automation system
BPA	Bonneville Power Administration
BPD	Building Performance Database
BPR	base-to-peak ratio
CB ECS	Commercial Buildings Energy Consumption Survey
CBSA	core-based statistical area
CDD	cooling degree days
DOE	U.S. Department of Energy
DHW	domestic hot water
EER	energy efficiency ratio
EIA	U.S. Energy Information Administration
EF	energy factor
EPB	formerly known as the Electric Power Board of Chattanooga
EPW	EnergyPlus weather
EUI	energy use intensity
EUL	effective useful life
EULP	end-use load profile
GEB	grid-interactive efficient building
HDD	heating degree days
HEMS	home energy management system
HPC	high-performance computing
HPWH	heat pump water heater
HVAC	heating, ventilating, and air conditioning
IEA	International Energy Agency
IQR	interquartile range
ISD	Integrated Surface Database project
ISO	independent system operator
LA100	Los Angeles 100% Renewable Energy Study
LBNL	Lawrence Berkeley National Laboratory
LRD	load research data
MERRA	NASA's Modern-Era Retrospective Analysis for Research and Applications dataset
MERRA-2	version 2 of NASA's MERRA dataset
MF	multifamily
NEEA	Northwest Energy Efficiency Alliance
NMBE	normalized mean bias error
NOAA	National Oceanic and Atmospheric Administration
NRECA	National Rural Electric Cooperative Association
NREL	National Renewable Energy Laboratory
NSRDB	NREL's National Solar Radiation Database
NYSERDA	New York State Energy Research and Development Authority

ORNL	Oak Ridge National Laboratory
PNNL	Pacific Northwest National Laboratory
PUMA	Public Use Microdata Areas
PUMS	Public Use Microdata Samples
PV	photovoltaic
QOI	quantities of interest
RBSA	Residential Building Stock Assessment
RBSAM	RBSA Metering
RECS	Residential Energy Consumption Survey
RTO	regional transmission organization
RTU	rooftop unit
SFA	single-family attached
SFD	single-family detached
TAG	technical advisory group
TMY	typical meteorological year
TMY3	typical meteorological year, version 3
UQ	uncertainty quantification
VEIC	Vermont Energy Investment Corporation
VVUQ	verification, validation, and uncertainty quantification
WWR	window-to-wall ratio

## Executive Summary



**Figure ES-1. Graphical summary of the *End-Use Load Profiles* project**

### Motivation and Background

The United States is embarking on an ambitious transition to a 100% clean energy economy by 2050, which will require improving the flexibility of electric grids. One way to achieve grid flexibility is to shed or shift demand to align with changing grid needs. To facilitate this, it is critical to understand *how* and *when* energy is used. High-quality end-use load profiles (EULPs) provide this information, and can help cities, states, and utilities understand the time-sensitive value of energy efficiency, demand response, and distributed energy resources.

Publicly available EULPs have traditionally had limited application because of age and incomplete geographic representation (Frick, Eckman, and Goldman 2017; Frick 2019). To help fill this gap, the U.S. Department of Energy (DOE) funded a three-year project—*End-Use Load Profiles for the U.S. Building Stock*—that culminated in the release of a publicly available dataset<sup>1</sup> of simulated EULPs representing residential and commercial buildings across the contiguous United States. The motivation for this work is further detailed in a November 2019 report: *Market Needs, Use Cases, and Data Gaps* (Mims Frick et al. 2019).

This *Methodology and Results* report provides detailed descriptions of how the dataset was developed, intended for an audience of dataset and model users interested in the technical details. These details include descriptions of all of the model improvements made for calibration and the final comparisons to empirical data sources. A companion report, *End-Use Load Profiles for the U.S. Building Stock: Applications and Opportunities*, will be published subsequently and will describe example applications and considerations for using the dataset, intended for an audience of general dataset users.

### Project Team

The project team included researchers from the National Renewable Energy Laboratory (NREL), Lawrence Berkeley National Laboratory (LBNL), and Argonne National Laboratory (ANL). The project was guided by an extensive

<sup>1</sup> As of October 28, 2021, the dataset is available at <https://www.nrel.gov/buildings/end-use-load-profiles.html>

technical advisory group (TAG) of 92 individuals representing 61 organizations, including stakeholders from electric utilities, independent system operators (ISOs) and regional transmission organizations (RTOs), public utility commissions, state and local governments, consulting firms, software companies, academic institutions, nongovernmental organizations representing utilities and regional efficiency groups, and DOE. A full list of TAG members is included in Appendix A. As a project partner, the Electric Power Research Institute assisted the project team with utility data outreach. Northeast Energy Efficiency Partnerships received funding from the New York State Energy Research and Development Authority and the Massachusetts Clean Energy Center to engage with stakeholders in the Northeast, assist with data gathering and outreach, and develop a data inventory and needs assessment for the Northeast (Titus and McChalicher 2021).

## Methodology

Historical and contemporary efforts to develop EULPs for particular regions have used direct submetering of a statistically representative sample of buildings. Applying such an approach to the contiguous United States would have cost an order of magnitude higher than the already significant budget for this project, would have likely been limited in coverage of building types and end uses, and would not have resulted in calibrated models that enable future what-if analyses of scenarios involving energy efficiency, electrification, demand flexibility, and changes in climate or behavior.

Our hybrid approach—using a wide range of empirical data to inform updates to detailed physics-simulation building stock models—produced EULPs covering all major commercial and residential building types and end uses, for all locations of the contiguous United States. More importantly, the calibrated building models enable what-if scenario analyses. Although a pure submetering approach was not feasible for the national DOE EULP effort, our approach would not have been possible without the foresight of organizations that have invested in regional end-use submetering load research. Our work was also made possible by the ratepayer- and taxpayer-funded investments in advanced metering infrastructure over the past decade. We have fully documented our hybrid approach methodology in Section 2 and model calibration updates in Section 3.

To implement our hybrid approach, we adopted a framework for calibration, validation, and uncertainty quantification (UQ) and defined quantities of interest (QOI) used to evaluate calibration progress, validation accuracy, and uncertainty (Section 2.1.1). As documented in Section 2.3, we reviewed the types of calibration relevant to building energy modeling—manual, automated, and output calibration—and the advantages and disadvantages of each. We summarized the few examples of load profile model calibration that existed prior to this project, which all informed our work to some degree, and established a calibration philosophy to guide the calibration for this project.

## Empirical Data Sources

We worked with more than 30 data sharing partners to obtain the empirical data used for calibration and validation. This involved navigating privacy concerns and data transfer issues in order to obtain customer meter data and building metadata from about a dozen utilities (Sections 2.3.6 and 2.3.7). We developed an approach to process the meter data, associate it with building characteristics, and clean it for use in validation (Section 2.3.5). Whereas utility meter data enabled validation of whole-building and whole-sector load, we needed empirical timeseries data broken out by end use to calibrate our modeled end uses. There were a variety of residential end-use datasets available for use in this project (Section 2.3.6). In contrast, lack of commercial end-use data was a major data gap identified at the start of the project. Undertaking an end-use monitoring study would not have yielded data in time to use it for calibration, so we pursued a major outreach effort that resulted in procuring existing commercial end-use data from a range of unconventional sources (Section 2.3.5). In total, we used 33 empirical data sources for residential calibration and validation (Table 2) and 26 sources for commercial calibration and validation (Table 3).

We studied whether residential end-use data were transferable between regions, concluding that, with some exceptions, the *shapes* of most non-weather dependent end uses can be considered transferable between regions, though the magnitude of the end uses will vary depending on factors such as the saturation of end-use equipment in a given location (Section 2.3.9). Our project relies on physics-based simulations to model how end uses that depend on weather (cooling, space heating, and water heating) or location (lighting) vary from region to region. As such, historical weather data were another key aspect of empirical data for this project. We developed new capabilities to construct historical weather data files using ground-based measurements for most variables (temperature, humidity, wind speed/direction, and pressure) and satellite-derived solar radiation data (Section 2.4).



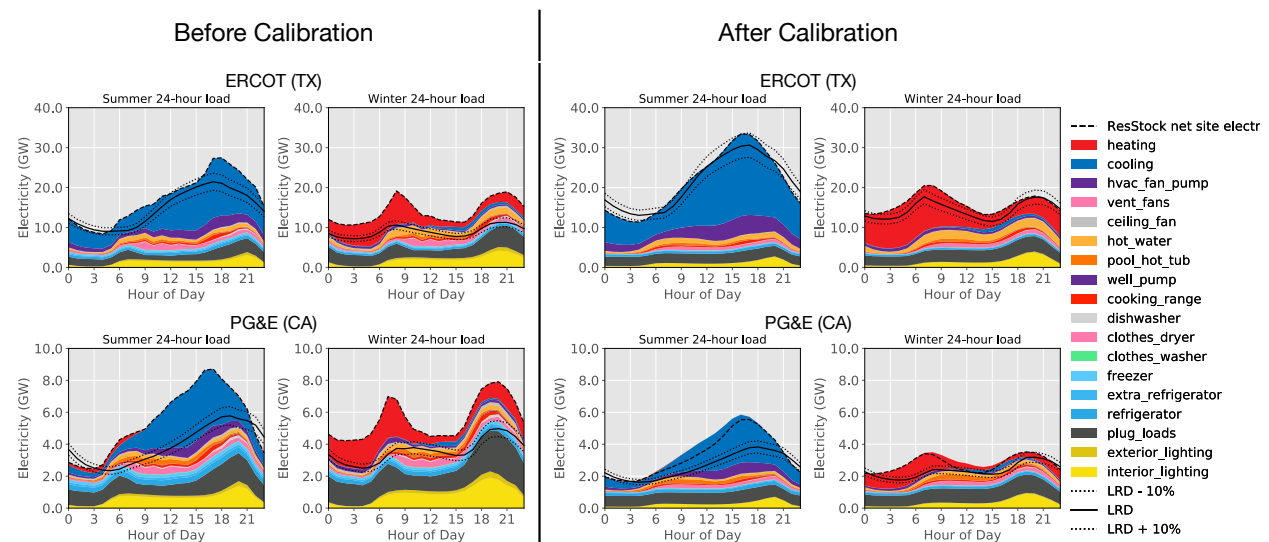
## Calibration

To calibrate the EULP dataset, we made more than 70 improvements to the ResStock™ and ComStock™ models. To align with the phased timing of when we obtained empirical meter datasets, we divided the calibration effort into five residential and four commercial regional phases, each including comparisons to one or more utility meter datasets in a similar region. By design, calibration did *not* involve automated or manual tuning of inputs to minimize error, which often leads to “getting the right answer for the wrong reason.” Instead, the objective of calibration was to make model improvements that reduce model error, but only when supported by data, such as weather, census, real estate, time use, EIA surveys, and submetering data (see Tables 2 and 3 for full lists of ResStock and ComStock data sources). The calibration process relied on both data science and buildings domain expertise.

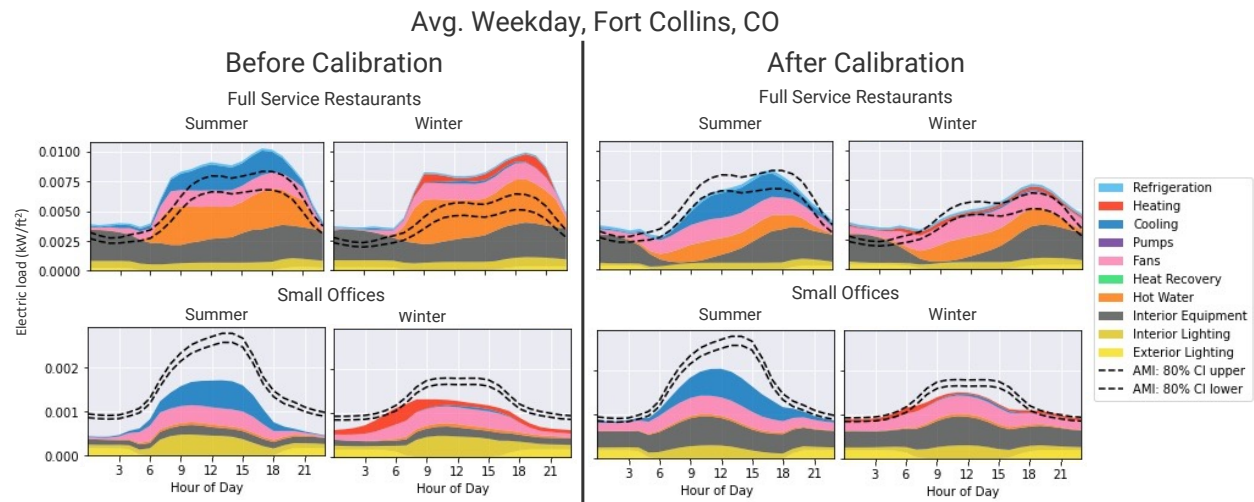
We developed a novel approach to building stock model sensitivity analysis to evaluate which input parameters are more important for the improvements made for calibration (Section 3.1). Model improvements were typically in the form of increased accuracy, diversity, or resolution of input parameter distributions, as documented comprehensively in Section 3. To improve the realism of individual housing unit load profiles, we developed a novel approach to simulating stochastic occupant behavior schedules and integrated it into ResStock, as discussed in Section 3.2.1. As a final step in residential calibration, we developed an output correction model to reduce some of the remaining model error, as presented in Section 3.2.10.

## Validation

A “validated” model does not mean that the outputs perfectly match the available empirical data; it means that model accuracy was evaluated for the quantities of interest and reported so that EULP data users know what level of confidence they should have when putting them to use. EULP data users should review the detailed validation results presented in Section 4, where, for each comparison, we provide discussion of accuracy and possible explanations for discrepancies. One finding from this research is that there can be a large degree of uncertainty in empirical load data, particularly for commercial buildings when disaggregating by building type or calculating energy use intensity, because of the metadata matching process (Section 4.2.2). Readers and data users may also be interested to understand how much accuracy improved as a result of the model input updates made for calibration; see examples in Figures ES-2 and ES-3 and discussion in Section 5.1.2.



**Figure ES-2. Before and after calibration example: ResStock comparisons to load research data (LRD) average daily profiles (note: before and after years are different). See Section 5.1.2 for a full discussion of how accuracy improved as a result of the model input updates made through calibration.**



**Figure ES-3. Before and after calibration example: ComStock comparisons to Fort Collins advanced metering infrastructure (AMI) average daily profiles.** See Section 5.1.2 for a full discussion of how accuracy improved as a result of the model input updates made through calibration.

### Uncertainty Quantification

Uncertainty in results is discussed at a high level in Section 5.1.3. We quantified the uncertainty ranges around model outputs by developing and applying a novel approach to building stock model uncertainty quantification that used trained timeseries surrogate models to evaluate millions of permutations of model inputs in order to propagate input uncertainty ranges through to outputs (Section 4.4). Results of this input uncertainty quantification indicate that the uncertainty in seasonal or top 10 day average peak magnitude is typically in the 3–9% range for ResStock and 4–11% for ComStock, depending on the season, building type, and region. The daily minimum base load magnitude has slightly lower uncertainty, with values in the 1–4% range for ResStock, and 3–6% for ComStock. Annual energy uncertainty is in the 3–6% range for ResStock and 3–8% for ComStock. Peak time uncertainty is typically 45–90 minutes for ResStock and 30–90 minutes for ComStock, depending on the season, building type, and region.

The other major source of model uncertainty is uncertainty from the use of insufficient numbers of ResStock or ComStock samples to estimate quantities of interest in particular locations or segments of the building stock (Section 5.1.3). This is an area of ongoing work, but our preliminary ResStock analysis showed that a sample of 1,000 dwelling units results in sample size uncertainty—that is, the error relative to the result with 1,000,000 samples—of less than 15% for all quantities of interest. With 10,000 samples, the uncertainty drops to less than 10% for all quantities of interest. Therefore, we recommend users ensure that there are at least 1,000 samples in a query of the data or in a downloaded aggregate data file. If there are less than 1,000 samples for a given query or aggregate, we recommend combining nearby geographies (e.g., counties) to increase the sample size.

### Conclusion

The final product of this project is the first version (1.0) of an EULP dataset containing calibrated and validated 15-minute resolution load profiles for all major residential and commercial building types and end uses, for all locations of the contiguous United States. Although our hybrid approach was similar to some previous examples that used tens or hundreds of building energy models, the scale of our application was unprecedented, both in terms of empirical data gathered—2.3 million meters worth of hourly data from 11 utilities—and in terms of the granularity of building stock simulation—900,000 building energy models representing 58 billion ft<sup>2</sup> of commercial buildings and 133 million residential dwelling units.

The EULP dataset is available in three formats<sup>2</sup>—(1) via a web viewer, (2) as downloadable spreadsheet files, and (3) in a detailed format that can be queried with big data tools—to maximize accessibility for different types of users.

<sup>2</sup>All three methods for accessing the EULP dataset can be found on the dataset website: <https://www.nrel.gov/buildings/end-use-load-profiles.html>

The OpenStudio® model input files are also available for building energy modelers to use for their own analyses. Utility planners, consultants, regulators, state energy offices, researchers, and building owners are now able to use these resources, along with tools such as Berkeley Lab's forthcoming Time-Sensitive Value Calculator, to estimate the value of energy efficiency, demand response, and other distributed energy resources. Such analysis can be used to guide utility resource and distribution system planning, research and development prioritization, and state and local energy planning and regulation.

The EULP dataset can be updated over time by incorporating new input data (weather, census, real estate, time use, EIA surveys, and so on) into the models as they become available, giving the load models longevity beyond this calibration effort and beyond EULPs developed through a pure end-use submetering approach. Additionally, the calibrated models can be a foundation to develop *end-use savings shapes* that describe the difference in energy consumption between the current building stock and the building stock with an energy saving, electrification, flexibility, or other measure applied.

The EULP dataset and calibrated building stock models can play a key role by helping us understand, with more accuracy than ever before, how buildings and their occupants interact with the national electricity system, and the role that high-performing, energy-efficient, and demand-flexible buildings can play in an equitable transition to a decarbonized, affordable, and reliable energy system.

## Table of Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Project Overview	1
1.1.1	Motivation and Background	1
1.1.2	Project Approach and Components	1
<b>2</b>	<b>Methodology</b>	<b>4</b>
2.1	Overall Approach to Calibration, Validation, and Uncertainty Quantification	4
2.1.1	Definitions and Framework	4
2.1.2	Quantities of Interest	5
2.1.3	Organization of the Methodology Section	6
2.2	Methodology: ResStock and ComStock	6
2.3	Methodology: Calibration and Validation	13
2.3.1	Background	13
2.3.2	Approaches to Building Energy Model Calibration	13
2.3.3	Load Profile Calibration Examples	14
2.3.4	Our Calibration Philosophy	16
2.3.5	Obtaining and Processing Empirical Data	16
2.3.6	Empirical Data Used for Residential Calibration and Validation	24
2.3.7	Empirical Data Used for Commercial Calibration and Validation	32
2.3.8	Calibration and Validation in Multiple Dimensions	38
2.3.9	End-Use Transferability Study	41
2.3.10	Methods to Understand Variability in Commercial Building Operation	49
2.4	Methodology: Weather Data	53
2.4.1	Background on Historical Weather Data	53
2.4.2	Process to Generate Weather Data Files	54
2.5	Methodology: Uncertainty Quantification	57
2.5.1	Overview of Uncertainty Propagation Approach	57
2.5.2	Selection of Input Parameters and Development of Input Uncertainty Ranges	58
2.5.3	Development of Surrogate Models	58
2.5.4	Application of Surrogate Models to Uncertainty Propagation	59
<b>3</b>	<b>Model Calibration</b>	<b>61</b>
3.1	Sensitivity Analysis	61
3.1.1	Methodology	61
3.1.2	Sensitivity Analysis Results	62
3.1.3	Application to Calibration and Validation	66
3.2	Residential Model Updates	68
3.2.1	Occupant-Driven Schedule Updates	68
3.2.2	Geospatial and Weather Data Updates	77
3.2.3	Geometry Updates	82
3.2.4	Appliance and Plug Load Updates	85
3.2.5	Lighting Updates	89
3.2.6	Water Heating Updates	94
3.2.7	HVAC Updates	95
3.2.8	Thermal Envelope Updates	110
3.2.9	On-Site Solar Photovoltaic Generation	119
3.2.10	Residential Output Correction Model	122
3.3	Commercial Model Updates	131
3.3.1	Interior Lighting Schedules	131
3.3.2	Interior Equipment Schedules	133
3.3.3	Thermostat Setpoints	135

3.3.4	Interior Lighting Schedule Magnitude Variability . . . . .	136
3.3.5	Interior Equipment Schedule Magnitude Variability . . . . .	139
3.3.6	Exterior Lighting Power . . . . .	140
3.3.7	Off-Cycle Controls for Packaged Single-Zone Systems . . . . .	141
3.3.8	Demand Control Ventilation and Energy Recovery . . . . .	142
3.3.9	Nighttime HVAC System Operation . . . . .	142
3.3.10	Window-to-Wall Ratio . . . . .	145
3.3.11	Restaurants in Strip Malls . . . . .	146
3.3.12	Lighting Power Density . . . . .	147
3.3.13	Energy Code Adoption . . . . .	149
3.3.14	Building System Turnover . . . . .	150
3.3.15	Heating and Service Water Heating Fuel Determination . . . . .	154
3.3.16	Thermostat Variability . . . . .	162
3.3.17	Hours of Operation Schedules . . . . .	165
<b>4</b>	<b>Results . . . . .</b>	<b>168</b>
4.1	Final Calibration and Validation Results—Residential . . . . .	168
4.1.1	Residential EIA Data Validation . . . . .	168
4.1.2	Residential AMI Data Validation . . . . .	175
4.1.3	Residential LRD Validation . . . . .	182
4.1.4	Residential End-Use Data Validation . . . . .	185
4.1.5	Residential Quantities of Interest Over Time . . . . .	188
4.2	Final Calibration and Validation Results—Commercial . . . . .	199
4.2.1	Commercial EIA Data Validation . . . . .	199
4.2.2	Commercial AMI Data Validation . . . . .	210
4.2.3	Final Commercial QOIs . . . . .	293
4.3	Description of Published Dataset . . . . .	334
4.4	Uncertainty Quantification Results . . . . .	337
4.4.1	Uncertainty Propagation Results—Commercial . . . . .	337
4.4.2	Uncertainty Propagation Results—Residential . . . . .	339
<b>5</b>	<b>Discussion and Conclusions . . . . .</b>	<b>342</b>
5.1	Interpreting and Using Results . . . . .	342
5.1.1	How Accurate Are the Results? . . . . .	342
5.1.2	How Much Did Accuracy Improve Through Calibration? . . . . .	342
5.1.3	How Much Uncertainty Is There in the Results? . . . . .	346
5.1.4	How Does the EULP Dataset Compare to Previously Available EULPs? . . . . .	347
5.2	Conclusion . . . . .	348
	<b>References . . . . .</b>	<b>350</b>
	<b>Appendix A List of Technical Advisory Group Members and Organizations . . . . .</b>	<b>357</b>
	<b>Appendix B Roles of Authors and Non-Author Contributors . . . . .</b>	<b>358</b>
	<b>Appendix C Supplemental Residential AMI Comparisons . . . . .</b>	<b>359</b>
	<b>Appendix D Supplemental Residential End-Use Dataset Analysis . . . . .</b>	<b>370</b>
	<b>Appendix E Supplemental Figures for Residential Stochastic Occupant-Driven Loads . . . . .</b>	<b>376</b>
	<b>Appendix F Residential EIA Monthly Electricity and Natural Gas Consumption Comparisons by State . . . . .</b>	<b>381</b>

## List of Figures

Figure ES-1. Graphical summary of the End-Use Load Profiles project . . . . .	vii
Figure ES-2. Before and after calibration example: ResStock comparisons to LRD average daily profiles . . . . .	ix
Figure ES-3. Before and after calibration example: ComStock comparisons to Fort Collins AMI average daily profiles . . . . .	x
Figure 1. Major components of the End-Use Load Profiles for the U.S. Building Stock project . . . . .	2
Figure 2. Verification, validation, and uncertainty quantification framework . . . . .	5
Figure 3. Season assignments for three example regions . . . . .	6
Figure 4. Sankey diagram of commercial AMI data outreach . . . . .	18
Figure 5. Sankey diagram of commercial end-use data outreach . . . . .	21
Figure 6. Residential calibration dimensions . . . . .	24
Figure 7. Map of residential calibration datasets (AMI and submetering) . . . . .	27
Figure 8. Map of residential load research data . . . . .	30
Figure 9. Commercial calibration dimensions . . . . .	32
Figure 10. CBECS 2012 and Form EIA-861M 2012 electricity consumption comparison . . . . .	35
Figure 11. CBECS 2012 and Form EIA-176 2012 natural gas consumption comparison . . . . .	35
Figure 12. Map of commercial calibration datasets (AMI) . . . . .	36
Figure 13. Tiered strategy for developing end-use data . . . . .	39
Figure 14. Region-by-region calibration strategy . . . . .	40
Figure 15. End-use load shape transferability—clothes washer . . . . .	42
Figure 16. End-use load shape transferability—clothes dryer . . . . .	42
Figure 17. End-use load shape transferability—cooking range . . . . .	43
Figure 18. End-use load shape transferability—dishwasher . . . . .	43
Figure 19. End-use load shape transferability—plug loads . . . . .	43
Figure 20. End-use load shape transferability—refrigerator . . . . .	44
Figure 21. End-use load shape transferability—water heating . . . . .	44
Figure 22. End-use load shape transferability—lighting . . . . .	44
Figure 23. End-use load shape transferability—pool and spa pumps . . . . .	45
Figure 24. End-use load shape transferability—pool and spa heaters . . . . .	45
Figure 25. Monthly variation in end-use loads . . . . .	47
Figure 26. Monthly appliance end-use multipliers averaged across six end-use datasets . . . . .	48
Figure 27. Two-path approach to characterize load profile variability . . . . .	49
Figure 28. Key parameters to characterize building load shape from the time domain . . . . .	49
Figure 29. Workflow of the frequency-domain load profile characterization . . . . .	50
Figure 30. Example frequency spectrums extracted from daily load profiles . . . . .	51
Figure 31. Example load profiles and the binned daily frequency features . . . . .	52
Figure 32. Two example daily load profiles . . . . .	52

Figure 33.	Illustration of typical meteorological year weather data construction . . . . .	53
Figure 34.	Comparison of MERRA-2 and NOAA ISD weather . . . . .	54
Figure 35.	Map of weather stations for ComStock and ResStock (TMY3) . . . . .	56
Figure 36.	Map of weather stations for ComStock and ResStock (2018) . . . . .	56
Figure 37.	Graphical summary of the surrogate framework . . . . .	59
Figure 38.	Monte Carlo simulation methodology . . . . .	60
Figure 39.	Weighted feature importance of input parameters, aggregated across QOIs, for Fort Collins, CO .	63
Figure 40.	Weighted feature importance of input parameters for the U.S. commercial building stock . . . . .	64
Figure 41.	ResStock summer peak sensitivity . . . . .	65
Figure 42.	ResStock average QOIs for ComEd . . . . .	66
Figure 43.	Comparison of plug load schedules from different submetered data sources . . . . .	68
Figure 44.	Impact of lighting technology saturation and plug load schedule changes . . . . .	69
Figure 45.	Validation of residential stochastic occupant-driven loads . . . . .	70
Figure 46.	Impact of new residential schedule generator in an example home . . . . .	70
Figure 47.	Impact of new residential schedule generator on EULPs for a typical week in 1,000 example homes	71
Figure 48.	Impact of new residential schedule generator on Fort Collins AMI comparison . . . . .	72
Figure 49.	Comparison of occupant schedules between MSA and non-MSA counties . . . . .	73
Figure 50.	Comparison of occupant schedules between Manhattan and upstate New York . . . . .	73
Figure 51.	Average bedtime by county and average ATUS shift by state . . . . .	74
Figure 52.	Impact of occupant schedule shifts on AMI comparisons . . . . .	74
Figure 53.	Unoccupied dwelling units by housing type at PUMA-region resolution . . . . .	75
Figure 54.	Timeseries impact of introduction of unoccupied units . . . . .	76
Figure 55.	Utility-scale annual impact of introduction of unoccupied units . . . . .	76
Figure 56.	Increased ResStock geospatial resolution . . . . .	77
Figure 57.	Updated mapping of dwelling units to utilities . . . . .	78
Figure 58.	Impact of updated mapping of dwelling units to utilities . . . . .	79
Figure 59.	Comparison of MERRA-2 and NOAA ISD weather . . . . .	80
Figure 60.	Impact of weather data change . . . . .	80
Figure 61.	Increased geographic resolution of weather data . . . . .	81
Figure 62.	Impact of increased geographic resolution of weather data on electricity . . . . .	81
Figure 63.	Impact of increased geographic resolution of weather data on natural gas . . . . .	82
Figure 64.	Modeled floor area distributions by AHS 2017 floor area bin . . . . .	82
Figure 65.	Comparison of annual results from multifamily modeling approaches . . . . .	84
Figure 66.	Impact of multifamily modeling speed improvements . . . . .	84
Figure 67.	Impact of expanding multifamily building height options . . . . .	85
Figure 68.	Impact of major appliance saturation and floor area updates . . . . .	86
Figure 69.	Map of household sizes by PUMS region . . . . .	87
Figure 70.	National distributions of occupants per unit from RECS 2015 versus PUMS 2017 . . . . .	88



Figure 71.	Impact of more granular house sizes . . . . .	88
Figure 72.	Comparison of RECS 2015 to ResStock plug loads by census division . . . . .	89
Figure 73.	Comparison of national average lighting efficiency distributions from four sources and ResStock . . . . .	90
Figure 74.	Regional variation in lighting efficiency . . . . .	91
Figure 75.	Impact of lighting, plug load, and appliance changes on Seattle AMI comparison . . . . .	92
Figure 76.	Impact of lighting technology saturation change during Region 4 of calibration . . . . .	93
Figure 77.	Impact of updating lighting algorithm to the ANSI/RESNET/ICC 301-2019 approach . . . . .	94
Figure 78.	Updated water heater distributions . . . . .	94
Figure 79.	Impact of water heater efficiency updates . . . . .	95
Figure 80.	Distribution of furnace efficiencies in 1994 and 2003 . . . . .	97
Figure 81.	Impact of updated HVAC efficiency distributions on annual electricity sales by utility . . . . .	98
Figure 82.	Impact of updated HVAC efficiency distributions on AMI comparisons . . . . .	98
Figure 83.	Saturations of heating fuels by state and PUMA area . . . . .	99
Figure 84.	Impact of moving from state to PUMA area for heating fuel geographic resolution . . . . .	99
Figure 85.	AIA and IECC 2004 climate zone maps . . . . .	100
Figure 86.	Average heating and cooling setpoint values by IECC climate zone . . . . .	100
Figure 87.	Average difference between nighttime and home air-conditioning setpoints by IECC climate zone . . . . .	101
Figure 88.	Impact of changing setpoint schedule diversity ( $\pm 2$ hours) . . . . .	102
Figure 89.	Impact of changing setpoint schedule diversity ( $\pm 5$ hours) and distributions (total stock) . . . . .	103
Figure 90.	Impact of changing setpoint schedule diversity ( $\pm 5$ hours) and distributions (selected cohorts) . . . . .	103
Figure 91.	Impact of air-conditioning saturation climate dependency . . . . .	104
Figure 92.	Population density of Horry County, SC . . . . .	105
Figure 93.	Sensitivity of Horry County load to non-coastal weather . . . . .	106
Figure 94.	Impact of zonal electric heating setpoint update on AMI comparisons . . . . .	107
Figure 95.	Cooling setpoint distributions in RASS 2009 and ResStock . . . . .	108
Figure 96.	Impact of room air-conditioner setpoint update . . . . .	108
Figure 97.	Impact of updated room air-conditioner performance curves . . . . .	109
Figure 98.	Impact of removing central space and water heating from Seattle multifamily AMI comparisons . . . . .	110
Figure 99.	Residential window-to-wall ratio distributions . . . . .	110
Figure 100.	Distribution of ResStock infiltration rates before and after update . . . . .	111
Figure 101.	Updated roof material distributions example . . . . .	112
Figure 102.	Impact of updated roof material distributions . . . . .	112
Figure 103.	Foundation type distributions before and after ResStock update . . . . .	113
Figure 104.	Impact of foundation distribution update . . . . .	113
Figure 105.	Impact of modeling masonry wall constructions on ComEd LRD comparison . . . . .	114
Figure 106.	Wall type saturation by vintage bin and census region . . . . .	115
Figure 107.	Wall exterior finish saturation by vintage bin and census region . . . . .	116
Figure 108.	Impact of wall type and exterior finish updates . . . . .	117

Figure 109. Impact of new window type options and distributions on total electric energy consumption . . . .	119
Figure 110. Fraction of dwelling units with rooftop PV, by county . . . . .	120
Figure 111. Rooftop PV system size by state (left) and system orientation (right) . . . . .	120
Figure 112. Impact of the new PV distributions and resulting loads for SCE and PG&E service territory . . . .	121
Figure 113. Illustration of billing month reporting for February . . . . .	123
Figure 114. Example EIA model fit for Illinois . . . . .	125
Figure 115. Correction model performance example for Florida . . . . .	126
Figure 116. Correction model performance matrix by state . . . . .	128
Figure 117. Correction model performance example for VEIC . . . . .	129
Figure 118. Correction model performance example for Tallahassee . . . . .	129
Figure 119. Correction model performance matrix for various utility regions . . . . .	130
Figure 120. Data processing workflow for commercial end-use data . . . . .	131
Figure 121. Sample space covered by the commercial building end-use data: Interior lighting . . . . .	132
Figure 122. Median schedules derived from the commercial end-use data: Interior lighting . . . . .	132
Figure 123. Impact of the commercial building interior lighting schedule . . . . .	133
Figure 124. Sample space covered by the end-use data: Commercial interior equipment . . . . .	134
Figure 125. Median schedules derived from the commercial end-use data: Interior equipment . . . . .	134
Figure 126. Impact of the commercial building interior equipment schedule . . . . .	135
Figure 127. Commercial end-use data thermostat setpoints . . . . .	136
Figure 128. Strip mall thermostat setpoint impact . . . . .	136
Figure 129. Base-to-peak ratio distribution: Interior lighting . . . . .	137
Figure 130. Example BPR measure schedule adjustment . . . . .	138
Figure 131. Impact of interior lighting BPR variability measure . . . . .	139
Figure 132. Base-to-peak ratio distribution: Interior equipment . . . . .	139
Figure 133. Impact of interior equipment BPR measure . . . . .	140
Figure 134. Impact of exterior lighting power density changes . . . . .	141
Figure 135. Impact of off-cycle control: example on retail buildings . . . . .	142
Figure 136. Nighttime HVAC operating mode from BAS data . . . . .	144
Figure 137. Commercial calibration nighttime HVAC impact . . . . .	145
Figure 138. Window-to-wall ratio breakdowns by vintage and rentable floor area . . . . .	146
Figure 139. Window-to-wall ratio distributions before and after incorporating NFRC data . . . . .	146
Figure 140. Distribution of restaurant fraction in strip malls . . . . .	147
Figure 141. Strip mall electric EUIs with and without restaurant space type . . . . .	147
Figure 142. Impact of commercial lighting power density update . . . . .	149
Figure 143. Reliability analysis for windows in commercial buildings . . . . .	152
Figure 144. Survival curves and derived lifespan probability density functions for HVAC equipment . . . . .	153
Figure 145. Comparison of heating fuel distribution between CBECS and ACS by census division . . . . .	155
Figure 146. Map of the revised saturation of natural gas heating in commercial buildings, by county . . . . .	156

Figure 147. Map of the revised saturation of electric heating in commercial buildings, by county . . . . .	157
Figure 148. Map of the revised saturation of fuel oil heating in commercial buildings, by county . . . . .	158
Figure 149. Map of the revised saturation of propane heating in commercial buildings, by county . . . . .	159
Figure 150. Map of the revised saturation of district heating in commercial buildings, by county . . . . .	160
Figure 151. Map of the revised saturation of unheated commercial buildings, by county . . . . .	161
Figure 152. Thermostat setpoint variability scatterplot . . . . .	162
Figure 153. Thermostat setback fractions in CBECS . . . . .	163
Figure 154. Thermostat setpoint variability end-use impact . . . . .	163
Figure 155. Thermostat setpoint variability EUI distribution impact . . . . .	164
Figure 156. Number of samples by building type and utility in the commercial AMI dataset . . . . .	165
Figure 157. Distribution of small office hours of operations, by day type and season . . . . .	166
Figure 158. Distribution of small office hours of operation, by utility and day type . . . . .	167
Figure 159. ResStock modeled annual retail sales compared to 2018 EIA-861 by utility . . . . .	169
Figure 160. ResStock modeled annual total site electricity and gas compared to 2018 EIA-861 and EIA-176 . . . . .	169
Figure 161. ResStock monthly electricity and gas compared to EIA data for eight states . . . . .	170
Figure 162. Annual end-use electricity totals from ResStock and RECS . . . . .	172
Figure 163. Annual end-use electricity totals from ResStock and RECS by building type . . . . .	173
Figure 164. Annual end-use electricity totals from ResStock and RECS by climate zone . . . . .	174
Figure 165. ResStock modeled stacked end-use loads compared to summer weekday meter data . . . . .	176
Figure 166. ResStock modeled stacked end-use loads compared to summer weekend meter data . . . . .	177
Figure 167. ResStock modeled stacked end-use loads compared to winter weekday meter data . . . . .	178
Figure 168. ResStock modeled stacked end-use loads compared to winter weekend meter data . . . . .	179
Figure 169. ResStock modeled stacked end-use loads compared to shoulder weekday meter data . . . . .	180
Figure 170. ResStock modeled stacked end-use loads compared to shoulder weekend meter data . . . . .	181
Figure 171. ResStock modeled stacked end-use loads compared to 2018 summer LRD . . . . .	182
Figure 172. ResStock modeled stacked end-use loads compared to 2018 winter LRD . . . . .	183
Figure 173. ResStock modeled stacked end-use loads compared to 2018 shoulder LRD . . . . .	184
Figure 174. ResStock modeled daily sums of electricity per unit (kWh/unit/day) compared to 2018 LRD . . . . .	185
Figure 175. Comparison of HEMS and RBSAM submetering data to Seattle AMI data . . . . .	186
Figure 176. Comparison of HEMS and RBSAM submetering data to Seattle AMI data and ResStock end uses . . . . .	187
Figure 177. Residential QOIs over time: annual electricity per dwelling unit . . . . .	192
Figure 178. Residential QOIs over time: timeseries for ComEd . . . . .	193
Figure 179. Residential QOIs over time: timeseries for the City of Fort Collins . . . . .	193
Figure 180. Residential QOIs over time: timeseries for Seattle City Light . . . . .	194
Figure 181. Residential QOIs over time: timeseries for EPB . . . . .	194
Figure 182. Residential QOIs over time: timeseries for Horry Electric Cooperative . . . . .	195
Figure 183. Residential QOIs over time: timeseries for the City of Tallahassee . . . . .	195
Figure 184. Residential QOIs over time: timeseries for VEIC . . . . .	196

Figure 185. Residential QOIs over time: timeseries for Cherryland Electric Co-op . . . . .	196
Figure 186. Magnitude QOI discrepancies for the total residential building stock in kW/unit . . . . .	197
Figure 187. Magnitude QOI discrepancies for the total residential building stock relative to the meter data . . .	198
Figure 188. Timing QOI discrepancy (hours) for the total residential building stock . . . . .	198
Figure 189. ComStock compared to CBECS annual national electricity consumption . . . . .	199
Figure 190. ComStock compared to CBECS annual national electricity consumption by census division . . . .	199
Figure 191. ComStock compared to CBECS annual national electricity consumption by building type . . . . .	200
Figure 192. ComStock compared to CBECS annual national natural gas consumption . . . . .	200
Figure 193. ComStock compared to CBECS annual natural gas consumption by census division . . . . .	201
Figure 194. ComStock compared to CBECS annual natural gas consumption by building type . . . . .	201
Figure 195. ComStock compared to EIA-861M annual electricity consumption . . . . .	202
Figure 196. ComStock compared to EIA-861M annual electricity consumption by census division . . . . .	202
Figure 197. ComStock compared to EIA-861M monthly electricity consumption by division for East North Central . . . . .	203
Figure 198. ComStock compared to EIA-861M monthly electricity consumption by division for East South Central . . . . .	203
Figure 199. ComStock compared to EIA-861M monthly electricity consumption by division for Middle At- lantic . . . . .	203
Figure 200. ComStock compared to EIA-861M monthly electricity consumption by division for Mountain . .	204
Figure 201. ComStock compared to EIA-861M monthly electricity consumption by division for New England	204
Figure 202. ComStock compared to EIA-861M monthly electricity consumption by division for Pacific . . . .	204
Figure 203. ComStock compared to EIA-861M monthly electricity consumption by division for South Atlantic	205
Figure 204. ComStock compared to EIA-861M monthly electricity consumption by division for West North Central . . . . .	205
Figure 205. ComStock compared to EIA-861M monthly electricity consumption by division for West South Central . . . . .	205
Figure 206. ComStock compared to EIA-176 natural gas consumption . . . . .	206
Figure 207. ComStock compared to EIA-176 natural gas consumption by division . . . . .	206
Figure 208. ComStock compared to EIA-176 monthly natural gas consumption by division for East North Central . . . . .	206
Figure 209. ComStock compared to EIA-176 monthly natural gas consumption by division for East South Central . . . . .	207
Figure 210. ComStock compared to EIA-176 monthly natural gas consumption by division for Middle Atlantic	207
Figure 211. ComStock compared to EIA-176 monthly natural gas consumption by division for Mountain . . .	207
Figure 212. ComStock compared to EIA-176 monthly natural gas consumption by division for New England .	208
Figure 213. ComStock compared to EIA-176 monthly natural gas consumption by division for Pacific . . . .	208
Figure 214. ComStock compared to EIA-176 monthly natural gas consumption by division for South Atlantic	208
Figure 215. ComStock compared to EIA-176 monthly natural gas consumption by division for West North Central . . . . .	209
Figure 216. ComStock compared to EIA-176 monthly natural gas consumption by division for West South Central . . . . .	209

Figure 217. Comparison of AMI EUI distributions using different filtering methods . . . . .	211
Figure 218. Comparison of EUI distributions between regional AMI data and the corresponding census division from CBECS 2012 . . . . .	212
Figure 219. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for small offices . . . . .	214
Figure 220. Small office weekday seasonal average annual normalized day type comparison by end use . . . .	215
Figure 221. Small office weekday seasonal average day type comparison by end use . . . . .	216
Figure 222. Small office weekend seasonal average annual normalized day type comparison by end use . . . .	217
Figure 223. Small office weekend seasonal average day type comparison by end use . . . . .	218
Figure 224. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for medium offices . . . . .	220
Figure 225. Medium office weekday seasonal average annual normalized day type comparison by end use . .	221
Figure 226. Medium office weekday seasonal average day type comparison by end use . . . . .	222
Figure 227. Medium office weekend seasonal average annual normalized day type comparison by end use . .	223
Figure 228. Medium office weekend seasonal average day type comparison by end use . . . . .	224
Figure 229. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for large offices . . . . .	226
Figure 230. Large office weekday seasonal average annual normalized day type comparison by end use . . . .	227
Figure 231. Large office weekday seasonal average day type comparison by end use . . . . .	228
Figure 232. Large office weekend seasonal average annual normalized day type comparison by end use . . . .	229
Figure 233. Large office weekend seasonal average day type comparison by end use . . . . .	230
Figure 234. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for retail buildings . . . . .	232
Figure 235. Retail weekday seasonal average annual normalized day type comparison by end use . . . . .	233
Figure 236. Retail weekday seasonal average day type comparison by end use . . . . .	234
Figure 237. Retail weekend seasonal average annual normalized day type comparison by end use . . . . .	235
Figure 238. Retail weekend seasonal average day type comparison by end use . . . . .	236
Figure 239. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for strip malls . . . . .	238
Figure 240. Strip mall weekday seasonal average annual normalized day type comparison by end use . . . .	239
Figure 241. Strip mall weekday seasonal average day type comparison by end use . . . . .	240
Figure 242. Strip mall weekend seasonal average annual normalized day type comparison by end use . . . .	241
Figure 243. Strip mall weekend seasonal average day type comparison by end use . . . . .	242
Figure 244. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for warehouses . . . . .	244
Figure 245. Warehouse weekday seasonal average annual normalized day type comparison by end use . . . .	245
Figure 246. Warehouse weekday seasonal average day type comparison by end use . . . . .	246
Figure 247. Warehouse weekend seasonal average annual normalized day type comparison by end use . . . .	247
Figure 248. Warehouse weekend seasonal average day type comparison by end use . . . . .	248
Figure 249. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for primary schools . . . . .	250

Figure 250. Primary school weekday seasonal average annual normalized day type comparison by end use . . .	251
Figure 251. Primary school weekday seasonal average day type comparison by end use . . . . .	252
Figure 252. Primary school weekend seasonal average annual normalized day type comparison by end use . . .	253
Figure 253. Primary school weekend seasonal average day type comparison by end use . . . . .	254
Figure 254. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for sec- ondary schools . . . . .	256
Figure 255. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for full- service restaurants . . . . .	258
Figure 256. Full-service restaurant weekday seasonal average annual normalized day type comparison by end use . . . . .	259
Figure 257. Full-service restaurant weekday seasonal average day type comparison by end use . . . . .	260
Figure 258. Full-service restaurant weekend seasonal average annual normalized day type comparison by end use . . . . .	261
Figure 259. Full-service restaurant weekend seasonal average day type comparison by end use . . . . .	262
Figure 260. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for quick- service restaurants . . . . .	264
Figure 261. Quick-service restaurant weekday seasonal average annual normalized day type comparison by end use . . . . .	265
Figure 262. Quick-service restaurant weekday seasonal average day type comparison by end use . . . . .	266
Figure 263. Quick-service restaurant weekend seasonal average annual normalized day type comparison by end use . . . . .	267
Figure 264. Quick-service restaurant weekend seasonal average day type comparison by end use . . . . .	268
Figure 265. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for small hotels . . . . .	270
Figure 266. Small hotel weekday seasonal average annual normalized day type comparison by end use . . . .	271
Figure 267. Small hotel weekday seasonal average day type comparison by end use . . . . .	272
Figure 268. Small hotel weekend seasonal average annual normalized day type comparison by end use . . . .	273
Figure 269. Small hotel weekend seasonal average day type comparison by end use . . . . .	274
Figure 270. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for large hotels . . . . .	276
Figure 271. Large hotel weekday seasonal average annual normalized day type comparison by end use . . . .	277
Figure 272. Large hotel weekday seasonal average day type comparison by end use . . . . .	278
Figure 273. Large hotel weekend seasonal average annual normalized day type comparison by end use . . . .	279
Figure 274. Large hotel weekend seasonal average day type comparison by end use . . . . .	280
Figure 275. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for hospitals	282
Figure 276. Hospital weekday seasonal average annual normalized day type comparison by end use . . . . .	283
Figure 277. Hospital weekday seasonal average day type comparison by end use . . . . .	284
Figure 278. Hospital weekend seasonal average annual normalized day type comparison by end use . . . . .	285
Figure 279. Hospital weekend seasonal average day type comparison by end use . . . . .	286
Figure 280. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for outpa- tient buildings . . . . .	288



Figure 281. Outpatient weekday seasonal average annual normalized day type comparison by end use . . . . .	289
Figure 282. Outpatient weekday seasonal average day type comparison by end use . . . . .	290
Figure 283. Outpatient weekend seasonal average annual normalized day type comparison by end use . . . . .	291
Figure 284. Outpatient weekend seasonal average day type comparison by end use . . . . .	292
Figure 285. Normalized shape QOI comparison for small office. . . . .	294
Figure 286. Normalized shape QOI error for small office. . . . .	294
Figure 287. Magnitude QOI comparison for small office. . . . .	295
Figure 288. Magnitude QOI error for small office. . . . .	295
Figure 289. Timing QOI comparison for small office. . . . .	296
Figure 290. Timing QOI error (hours) for small office. . . . .	296
Figure 291. Normalized shape QOI comparison for medium office. . . . .	297
Figure 292. Normalized shape QOI error for medium office. . . . .	297
Figure 293. Magnitude QOI comparison for medium office. . . . .	298
Figure 294. Magnitude QOI error for medium office. . . . .	298
Figure 295. Timing QOI comparison for medium office. . . . .	299
Figure 296. Timing QOI error (hours) for medium office. . . . .	299
Figure 297. Normalized shape QOI comparison for large office. . . . .	300
Figure 298. Normalized shape QOI error for large office. . . . .	300
Figure 299. Magnitude QOI comparison for large office. . . . .	301
Figure 300. Magnitude QOI error for large office. . . . .	301
Figure 301. Timing QOI comparison for large office. . . . .	302
Figure 302. Timing QOI error (hours) for large office. . . . .	302
Figure 303. Normalized shape QOI comparison for retail. . . . .	303
Figure 304. Normalized shape QOI error for retail. . . . .	303
Figure 305. Magnitude QOI comparison for retail. . . . .	304
Figure 306. Magnitude QOI error for retail. . . . .	304
Figure 307. Timing QOI comparison for retail. . . . .	305
Figure 308. Timing QOI error (hours) for retail. . . . .	305
Figure 309. Normalized shape QOI comparison for strip mall. . . . .	306
Figure 310. Normalized shape QOI error for strip mall. . . . .	306
Figure 311. Magnitude QOI comparison for strip mall. . . . .	307
Figure 312. Magnitude QOI error for strip mall. . . . .	307
Figure 313. Timing QOI comparison for strip mall. . . . .	308
Figure 314. Timing QOI error (hours) for strip mall. . . . .	308
Figure 315. Normalized shape QOI comparison for warehouse. . . . .	309
Figure 316. Normalized shape QOI error for warehouse. . . . .	309
Figure 317. Magnitude QOI comparison for warehouse. . . . .	310
Figure 318. Magnitude QOI error for warehouse. . . . .	310



Figure 319. Timing QOI comparison for warehouse. . . . .	311
Figure 320. Timing QOI error (hours) for warehouse. . . . .	311
Figure 321. Normalized shape QOI comparison for primary school. . . . .	312
Figure 322. Normalized shape QOI error for primary school. . . . .	312
Figure 323. Magnitude QOI comparison for primary school. . . . .	313
Figure 324. Magnitude QOI error for primary school. . . . .	313
Figure 325. Timing QOI comparison for primary school. . . . .	314
Figure 326. Timing QOI error (hours) for primary school. . . . .	314
Figure 327. Normalized shape QOI comparison for full service restaurant. . . . .	316
Figure 328. Normalized shape QOI error for full service restaurant. . . . .	316
Figure 329. Magnitude QOI comparison for full service restaurant. . . . .	317
Figure 330. Magnitude QOI error for full service restaurant. . . . .	317
Figure 331. Timing QOI comparison for full service restaurant. . . . .	318
Figure 332. Timing QOI error (hours) for full service restaurant. . . . .	318
Figure 333. Normalized shape QOI comparison for quick service restaurant. . . . .	319
Figure 334. Normalized shape QOI error for quick service restaurant. . . . .	319
Figure 335. Magnitude QOI comparison for quick service restaurant. . . . .	320
Figure 336. Magnitude QOI error for quick service restaurant. . . . .	320
Figure 337. Timing QOI comparison for full service restaurant. . . . .	321
Figure 338. Timing QOI error (hours) for quick service restaurant. . . . .	321
Figure 339. Normalized shape QOI comparison for small hotel. . . . .	322
Figure 340. Normalized shape QOI error for small hotel. . . . .	322
Figure 341. Magnitude QOI comparison for small hotel. . . . .	323
Figure 342. Magnitude QOI error for small hotel. . . . .	323
Figure 343. Timing QOI comparison for small hotel. . . . .	324
Figure 344. Timing QOI error (hours) for small hotel. . . . .	324
Figure 345. Normalized shape QOI comparison for large hotel. . . . .	325
Figure 346. Normalized shape QOI error for large hotel. . . . .	325
Figure 347. Magnitude QOI comparison for large hotel. . . . .	326
Figure 348. Magnitude QOI error for large hotel. . . . .	326
Figure 349. Timing QOI comparison for large hotel. . . . .	327
Figure 350. Timing QOI error (hours) for large hotel. . . . .	327
Figure 351. Normalized shape QOI comparison for hospital. . . . .	328
Figure 352. Normalized shape QOI error for hospital. . . . .	328
Figure 353. Magnitude QOI comparison for hospital. . . . .	329
Figure 354. Magnitude QOI error for hospital. . . . .	329
Figure 355. Timing QOI comparison for hospital. . . . .	330
Figure 356. Timing QOI error (hours) for hospital. . . . .	330

Figure 357. Normalized shape QOI comparison for outpatient. . . . .	331
Figure 358. Normalized shape QOI error for outpatient. . . . .	331
Figure 359. Magnitude QOI comparison for outpatient. . . . .	332
Figure 360. Magnitude QOI error for outpatient. . . . .	332
Figure 361. Timing QOI comparison for outpatient. . . . .	333
Figure 362. Timing QOI error (hours) for outpatient. . . . .	333
Figure 363. Three different ways for accessing the dataset . . . . .	334
Figure 364. An example screenshot from the comstock.nrel.gov data viewer . . . . .	336
Figure 365. Before and after calibration: ResStock comparisons to annual EIA sales data . . . . .	344
Figure 366. Before and after calibration: ResStock comparisons to LRD average daily profiles . . . . .	345
Figure 367. Before and after calibration: ComStock comparisons to Fort Collins AMI average daily profiles . . . . .	345
Figure 368. Visual convergence of ERCOT loads by the number of models used to construct the timeseries . . . . .	346
Figure 369. Sample convergence plot for the magnitude building level QOIs . . . . .	347
Figure 370. Comparison of ResStock modeled load to ComEd residential AMI data for an example week in each season . . . . .	359
Figure 371. Comparison of ResStock modeled load to City of Fort Collins residential AMI data for an exam- ple week in each season . . . . .	360
Figure 372. Comparison of ResStock modeled load to Seattle City Light residential AMI data for an example week in each season . . . . .	361
Figure 373. Comparison of ResStock modeled load to EPB Chattanooga residential AMI data for an example week in each season . . . . .	362
Figure 374. Comparison of ResStock modeled load to City of Tallahassee residential AMI data for an exam- ple week in each season . . . . .	363
Figure 375. Comparison of ResStock modeled load to Horry Electric Cooperative residential AMI data for an example week in each season . . . . .	364
Figure 376. Comparison of ResStock modeled load to VEIC residential AMI data for an example week in each season . . . . .	365
Figure 377. Comparison of ResStock modeled load to Cherryland Electric Co-op residential AMI data for an example week in each season . . . . .	366
Figure 378. Comparison of ResStock modeled load to ComEd residential AMI data for daily sums of elec- tricity per unit . . . . .	367
Figure 379. Comparison of ResStock modeled load to City of Fort Collins residential AMI data for daily sums of electricity per unit . . . . .	367
Figure 380. Comparison of ResStock modeled load to Seattle City Light residential AMI data for daily sums of electricity per unit . . . . .	367
Figure 381. Comparison of ResStock modeled load to EPB Chattanooga residential AMI data for daily sums of electricity per unit . . . . .	368
Figure 382. Comparison of ResStock modeled load to City of Tallahassee residential AMI data for daily sums of electricity per unit . . . . .	368
Figure 383. Comparison of ResStock modeled load to Horry Electric Cooperative residential AMI data for daily sums of electricity per unit . . . . .	368
Figure 384. Comparison of ResStock modeled load to VEIC residential AMI data for daily sums of electric- ity per unit . . . . .	369

Figure 385. Comparison of ResStock modeled load to Cherryland Electric Co-op residential AMI data for daily sums of electricity per unit . . . . .	369
Figure 386. Submetered EULPs by month—RBSAM . . . . .	371
Figure 387. Submetered EULPs by month—HEMS . . . . .	372
Figure 388. Submetered EULPs by month—Mass. RES 1 . . . . .	373
Figure 389. Submetered EULPs by month—Pecan Street . . . . .	374
Figure 390. Submetered EULPs by month—FSEC . . . . .	375
Figure 391. Impact of new residential schedule generator: one home, typical week . . . . .	377
Figure 392. Impact of new residential schedule generator: one home, average week . . . . .	377
Figure 393. Impact of new residential schedule generator: 1,000 homes, typical week, hourly resolution . . . .	378
Figure 394. Impact of new residential schedule generator: 1,000 homes, typical week, 10-minute resolution . .	378
Figure 395. Impact of new residential schedule generator on DHW schedules: one home, typical week . . . .	379
Figure 396. Impact of new residential schedule generator on DHW schedules: one home, average week . . . .	379
Figure 397. Impact of new residential schedule generator on DHW schedules: 1,000 homes, typical week . . .	380
Figure 398. Impact of new residential schedule generator on DHW schedules: 1,000 homes, average week . .	380
Figure 399. ResStock AL monthly electric sales before and after correction compared to 2018 AL sales reported in EIA Form 861M (left). ResStock AL monthly natural gas energy compared to 2018 AL natural gas energy reported in EIA Form 176 (right). . . . .	381
Figure 400. ResStock AR monthly electric sales before and after correction compared to 2018 AR sales reported in EIA Form 861M (left). ResStock AR monthly natural gas energy compared to 2018 AR natural gas energy reported in EIA Form 176 (right). . . . .	382
Figure 401. ResStock AZ monthly electric sales before and after correction compared to 2018 AZ sales reported in EIA Form 861M (left). ResStock AZ monthly natural gas energy compared to 2018 AZ natural gas energy reported in EIA Form 176 (right). . . . .	382
Figure 402. ResStock CA monthly electric sales before and after correction compared to 2018 CA sales reported in EIA Form 861M (left). ResStock CA monthly natural gas energy compared to 2018 CA natural gas energy reported in EIA Form 176 (right). . . . .	383
Figure 403. ResStock CO monthly electric sales before and after correction compared to 2018 CO sales reported in EIA Form 861M (left). ResStock CO monthly natural gas energy compared to 2018 CO natural gas energy reported in EIA Form 176 (right). . . . .	383
Figure 404. ResStock CT monthly electric sales before and after correction compared to 2018 CT sales reported in EIA Form 861M (left). ResStock CT monthly natural gas energy compared to 2018 CT natural gas energy reported in EIA Form 176 (right). . . . .	384
Figure 405. ResStock DC monthly electric sales before and after correction compared to 2018 DC sales reported in EIA Form 861M (left). ResStock DC monthly natural gas energy compared to 2018 DC natural gas energy reported in EIA Form 176 (right). . . . .	384
Figure 406. ResStock DE monthly electric sales before and after correction compared to 2018 DE sales reported in EIA Form 861M (left). ResStock DE monthly natural gas energy compared to 2018 DE natural gas energy reported in EIA Form 176 (right). . . . .	385
Figure 407. ResStock FL monthly electric sales before and after correction compared to 2018 FL sales reported in EIA Form 861M (left). ResStock FL monthly natural gas energy compared to 2018 FL natural gas energy reported in EIA Form 176 (right). . . . .	385

Figure 408. ResStock GA monthly electric sales before and after correction compared to 2018 GA sales reported in EIA Form 861M (left). ResStock GA monthly natural gas energy compared to 2018 GA natural gas energy reported in EIA Form 176 (right). . . . .	386
Figure 409. ResStock IA monthly electric sales before and after correction compared to 2018 IA sales reported in EIA Form 861M (left). ResStock IA monthly natural gas energy compared to 2018 IA natural gas energy reported in EIA Form 176 (right). . . . .	386
Figure 410. ResStock ID monthly electric sales before and after correction compared to 2018 ID sales reported in EIA Form 861M (left). ResStock ID monthly natural gas energy compared to 2018 ID natural gas energy reported in EIA Form 176 (right). . . . .	387
Figure 411. ResStock IL monthly electric sales before and after correction compared to 2018 IL sales reported in EIA Form 861M (left). ResStock IL monthly natural gas energy compared to 2018 IL natural gas energy reported in EIA Form 176 (right). . . . .	387
Figure 412. ResStock IN monthly electric sales before and after correction compared to 2018 IN sales reported in EIA Form 861M (left). ResStock IN monthly natural gas energy compared to 2018 IN natural gas energy reported in EIA Form 176 (right). . . . .	388
Figure 413. ResStock KS monthly electric sales before and after correction compared to 2018 KS sales reported in EIA Form 861M (left). ResStock KS monthly natural gas energy compared to 2018 KS natural gas energy reported in EIA Form 176 (right). . . . .	388
Figure 414. ResStock KY monthly electric sales before and after correction compared to 2018 KY sales reported in EIA Form 861M (left). ResStock KY monthly natural gas energy compared to 2018 KY natural gas energy reported in EIA Form 176 (right). . . . .	389
Figure 415. ResStock LA monthly electric sales before and after correction compared to 2018 LA sales reported in EIA Form 861M (left). ResStock LA monthly natural gas energy compared to 2018 LA natural gas energy reported in EIA Form 176 (right). . . . .	389
Figure 416. ResStock MA monthly electric sales before and after correction compared to 2018 MA sales reported in EIA Form 861M (left). ResStock MA monthly natural gas energy compared to 2018 MA natural gas energy reported in EIA Form 176 (right). . . . .	390
Figure 417. ResStock MD monthly electric sales before and after correction compared to 2018 MD sales reported in EIA Form 861M (left). ResStock MD monthly natural gas energy compared to 2018 MD natural gas energy reported in EIA Form 176 (right). . . . .	390
Figure 418. ResStock ME monthly electric sales before and after correction compared to 2018 ME sales reported in EIA Form 861M (left). ResStock ME monthly natural gas energy compared to 2018 ME natural gas energy reported in EIA Form 176 (right). . . . .	391
Figure 419. ResStock MI monthly electric sales before and after correction compared to 2018 MI sales reported in EIA Form 861M (left). ResStock MI monthly natural gas energy compared to 2018 MI natural gas energy reported in EIA Form 176 (right). . . . .	391
Figure 420. ResStock MN monthly electric sales before and after correction compared to 2018 MN sales reported in EIA Form 861M (left). ResStock MN monthly natural gas energy compared to 2018 MN natural gas energy reported in EIA Form 176 (right). . . . .	392
Figure 421. ResStock MO monthly electric sales before and after correction compared to 2018 MO sales reported in EIA Form 861M (left). ResStock MO monthly natural gas energy compared to 2018 MO natural gas energy reported in EIA Form 176 (right). . . . .	392
Figure 422. ResStock MS monthly electric sales before and after correction compared to 2018 MS sales reported in EIA Form 861M (left). ResStock MS monthly natural gas energy compared to 2018 MS natural gas energy reported in EIA Form 176 (right). . . . .	393
Figure 423. ResStock MT monthly electric sales before and after correction compared to 2018 MT sales reported in EIA Form 861M (left). ResStock MT monthly natural gas energy compared to 2018 MT natural gas energy reported in EIA Form 176 (right). . . . .	393

Figure 424. ResStock NC monthly electric sales before and after correction compared to 2018 NC sales reported in EIA Form 861M (left). ResStock NC monthly natural gas energy compared to 2018 NC natural gas energy reported in EIA Form 176 (right). . . . .	394
Figure 425. ResStock ND monthly electric sales before and after correction compared to 2018 ND sales reported in EIA Form 861M (left). ResStock ND monthly natural gas energy compared to 2018 ND natural gas energy reported in EIA Form 176 (right). . . . .	394
Figure 426. ResStock NE monthly electric sales before and after correction compared to 2018 NE sales reported in EIA Form 861M (left). ResStock NE monthly natural gas energy compared to 2018 NE natural gas energy reported in EIA Form 176 (right). . . . .	395
Figure 427. ResStock NH monthly electric sales before and after correction compared to 2018 NH sales reported in EIA Form 861M (left). ResStock NH monthly natural gas energy compared to 2018 NH natural gas energy reported in EIA Form 176 (right). . . . .	395
Figure 428. ResStock NJ monthly electric sales before and after correction compared to 2018 NJ sales reported in EIA Form 861M (left). ResStock NJ monthly natural gas energy compared to 2018 NJ natural gas energy reported in EIA Form 176 (right). . . . .	396
Figure 429. ResStock NM monthly electric sales before and after correction compared to 2018 NM sales reported in EIA Form 861M (left). ResStock NM monthly natural gas energy compared to 2018 NM natural gas energy reported in EIA Form 176 (right). . . . .	396
Figure 430. ResStock NV monthly electric sales before and after correction compared to 2018 NV sales reported in EIA Form 861M (left). ResStock NV monthly natural gas energy compared to 2018 NV natural gas energy reported in EIA Form 176 (right). . . . .	397
Figure 431. ResStock NY monthly electric sales before and after correction compared to 2018 NY sales reported in EIA Form 861M (left). ResStock NY monthly natural gas energy compared to 2018 NY natural gas energy reported in EIA Form 176 (right). . . . .	397
Figure 432. ResStock OH monthly electric sales before and after correction compared to 2018 OH sales reported in EIA Form 861M (left). ResStock OH monthly natural gas energy compared to 2018 OH natural gas energy reported in EIA Form 176 (right). . . . .	398
Figure 433. ResStock OK monthly electric sales before and after correction compared to 2018 OK sales reported in EIA Form 861M (left). ResStock OK monthly natural gas energy compared to 2018 OK natural gas energy reported in EIA Form 176 (right). . . . .	398
Figure 434. ResStock OR monthly electric sales before and after correction compared to 2018 OR sales reported in EIA Form 861M (left). ResStock OR monthly natural gas energy compared to 2018 OR natural gas energy reported in EIA Form 176 (right). . . . .	399
Figure 435. ResStock PA monthly electric sales before and after correction compared to 2018 PA sales reported in EIA Form 861M (left). ResStock PA monthly natural gas energy compared to 2018 PA natural gas energy reported in EIA Form 176 (right). . . . .	399
Figure 436. ResStock RI monthly electric sales before and after correction compared to 2018 RI sales reported in EIA Form 861M (left). ResStock RI monthly natural gas energy compared to 2018 RI natural gas energy reported in EIA Form 176 (right). . . . .	400
Figure 437. ResStock SC monthly electric sales before and after correction compared to 2018 SC sales reported in EIA Form 861M (left). ResStock SC monthly natural gas energy compared to 2018 SC natural gas energy reported in EIA Form 176 (right). . . . .	400
Figure 438. ResStock SD monthly electric sales before and after correction compared to 2018 SD sales reported in EIA Form 861M (left). ResStock SD monthly natural gas energy compared to 2018 SD natural gas energy reported in EIA Form 176 (right). . . . .	401
Figure 439. ResStock TN monthly electric sales before and after correction compared to 2018 TN sales reported in EIA Form 861M (left). ResStock TN monthly natural gas energy compared to 2018 TN natural gas energy reported in EIA Form 176 (right). . . . .	401

Figure 440. ResStock TX monthly electric sales before and after correction compared to 2018 TX sales reported in EIA Form 861M (left). ResStock TX monthly natural gas energy compared to 2018 TX natural gas energy reported in EIA Form 176 (right). . . . .	402
Figure 441. ResStock UT monthly electric sales before and after correction compared to 2018 UT sales reported in EIA Form 861M (left). ResStock UT monthly natural gas energy compared to 2018 UT natural gas energy reported in EIA Form 176 (right). . . . .	402
Figure 442. ResStock VA monthly electric sales before and after correction compared to 2018 VA sales reported in EIA Form 861M (left). ResStock VA monthly natural gas energy compared to 2018 VA natural gas energy reported in EIA Form 176 (right). . . . .	403
Figure 443. ResStock VT monthly electric sales before and after correction compared to 2018 VT sales reported in EIA Form 861M (left). ResStock VT monthly natural gas energy compared to 2018 VT natural gas energy reported in EIA Form 176 (right). . . . .	403
Figure 444. ResStock WA monthly electric sales before and after correction compared to 2018 WA sales reported in EIA Form 861M (left). ResStock WA monthly natural gas energy compared to 2018 WA natural gas energy reported in EIA Form 176 (right). . . . .	404
Figure 445. ResStock WI monthly electric sales before and after correction compared to 2018 WI sales reported in EIA Form 861M (left). ResStock WI monthly natural gas energy compared to 2018 WI natural gas energy reported in EIA Form 176 (right). . . . .	404
Figure 446. ResStock WV monthly electric sales before and after correction compared to 2018 WV sales reported in EIA Form 861M (left). ResStock WV monthly natural gas energy compared to 2018 WV natural gas energy reported in EIA Form 176 (right). . . . .	405
Figure 447. ResStock WY monthly electric sales before and after correction compared to 2018 WY sales reported in EIA Form 861M (left). ResStock WY monthly natural gas energy compared to 2018 WY natural gas energy reported in EIA Form 176 (right). . . . .	405



## List of Tables

Table 1.	Quantities of Interest . . . . .	6
Table 2.	Major Data Sources Used in ResStock . . . . .	7
Table 3.	Major Data Sources Used in ComStock . . . . .	12
Table 4.	Annual EUI Threshold by Building Type . . . . .	20
Table 5.	Procured Commercial End Use Datasets . . . . .	22
Table 6.	Commercial Data for Weather-Driven End Uses . . . . .	22
Table 7.	Commercial Data for Schedule-Driven End Uses . . . . .	23
Table 8.	Calibration Data Types and Typical Resolution and Characteristics . . . . .	24
Table 9.	Residential Calibration Datasets . . . . .	25
Table 10.	Commercial Calibration Datasets . . . . .	33
Table 11.	Key Parameters to Characterize the Building Load Shape From the Time Domain . . . . .	50
Table 12.	Parameter Selection for Sensitivity Analyses . . . . .	61
Table 13.	Updated ResStock Floor Areas . . . . .	83
Table 14.	Room Air-Conditioner Efficiency as a Function of Age . . . . .	96
Table 15.	List of Heating and Cooling Setpoint Offset Schedules Applied in ResStock . . . . .	102
Table 16.	Building Counts With Thermostat Data by Building Type . . . . .	135
Table 17.	Exterior Lighting Power Density for Commercial Buildings . . . . .	141
Table 18.	Nighttime HVAC Analysis BAS Counts . . . . .	143
Table 19.	Nighttime HVAC Modes from BAS Data . . . . .	143
Table 20.	Commercial Lighting Power Density Comparison by Building Type . . . . .	148
Table 21.	Effective Useful Life of Major Commercial Building Systems . . . . .	151
Table 22.	Commercial Equipment Lifetime Weibull Distribution Parameters . . . . .	154
Table 23.	Residential AMI Data Season Definitions . . . . .	175
Table 24.	Description of Each ResStock Model Run . . . . .	188
Table 25.	Uncertainty of ComStock Energy QOIs by Census Division . . . . .	338
Table 26.	Uncertainty of ComStock Energy QOIs by Building Type . . . . .	338
Table 27.	Uncertainty of ComStock Timing QOIs by Census Division . . . . .	338
Table 28.	Uncertainty of ComStock Timing QOIs by Building Type . . . . .	339
Table 29.	Uncertainty of ResStock Energy QOIs by Census Division . . . . .	340
Table 30.	Uncertainty of ResStock Energy QOIs by Building Type . . . . .	340
Table 31.	Uncertainty of ResStock Timing QOIs by Census Division . . . . .	340
Table 32.	Uncertainty of ResStock Timing QOIs by Building Type . . . . .	341



# 1 Introduction

## 1.1 Project Overview

### 1.1.1 Motivation and Background

End-use load profiles (EULPs) describe *how* and *when* energy is used. EULPs are critically important to utilities, public utility commissions, state energy offices, and other stakeholders. Applications focus on understanding how efficiency, demand response, and other distributed energy resources are valued and used in research and development prioritization, utility resource and distribution system planning, and state and local energy planning and regulations. Consequently, high-quality EULPs are also critical for widespread adoption of grid-interactive efficient buildings (GEBs).<sup>3</sup> For example, EULPs can be used to accurately forecast energy savings in buildings or to identify energy activities that can be shifted to different times of the day. For a full discussion of the market needs and use cases addressed by this project, see Mims Frick et al. (2019).

Publicly available EULPs have had limited application because of age and incomplete geographic representation (Frick, Eckman, and Goldman 2017; Frick 2019). To help fill this gap, the U.S. Department of Energy (DOE) funded a three-year project—*End-Use Load Profiles for the U.S. Building Stock*—that culminated in the release of a publicly available dataset of simulated EULPs for the stock of residential and commercial buildings in the contiguous United States.<sup>4</sup> The project team included researchers from the National Renewable Energy Laboratory (NREL), Lawrence Berkeley National Laboratory (LBNL), and Argonne National Laboratory (ANL). The Electric Power Research Institute assisted the project team with utility data outreach. Another project partner, Northeast Energy Efficiency Partnerships, received funding from the New York State Energy Research and Development Authority and the Massachusetts Clean Energy Center to engage with stakeholders in the Northeast, develop a data inventory and needs assessment for the Northeast, and assist with data gathering and outreach within the Northeast (Titus and McChalicher 2021). The project was guided by an extensive technical advisory group (TAG, members listed in Appendix A).

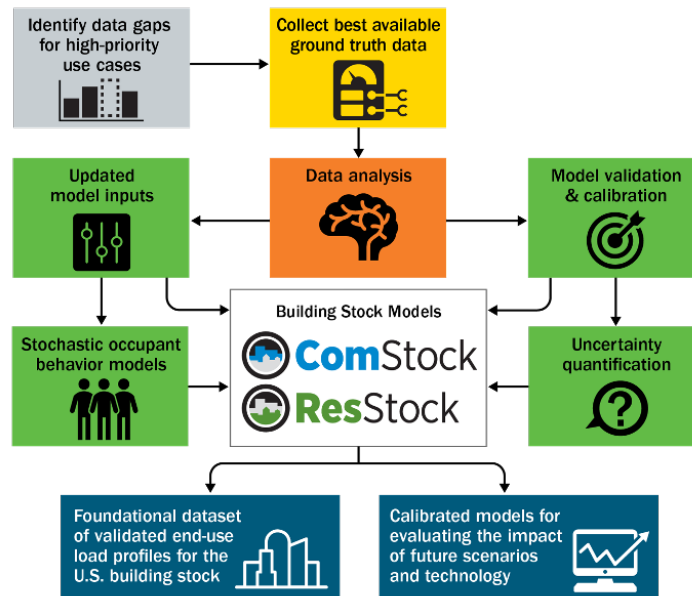
### 1.1.2 Project Approach and Components

The primary focus of the *End-Use Load Profiles for the U.S. Building Stock* project was calibrating and validating the EULP outputs from models of the U.S. residential and commercial building stocks—ResStock<sup>TM</sup> and ComStock<sup>TM</sup>. A variety of empirical ground truth datasets, including utility meter data from more than 2.3 million customers, various end-use submetering datasets, and other public and private datasets related to energy use in buildings, were used in the calibration and validation effort. The major components of the project are summarized in Figure 1.

One important question to address is: “why didn’t this project use direct submetering of a statistically representative sample of residential and commercial buildings to develop EULPs?” The first reason is that the costs were estimated to be an order of magnitude higher than the already significant available budget. These cost estimates were an extrapolation of the budget and scope of the competitively bid Northwest End Use Load Research Project recently commissioned by the Northwest Energy Efficiency Alliance (NEEA 2021b, NEEA 2021a). The second reason is that by using the selected approach, in addition to generating EULPs that represent the existing building stock, we are generating calibrated models of the building stock. These models are valuable because they can later be used to perform what-if analyses to estimate the impacts of various potential changes to the building stock to inform public- and private-sector decision-making.

<sup>3</sup>See <https://www.energy.gov/eere/buildings/grid-interactive-efficient-buildings> for more information on GEBs.

<sup>4</sup>As of October 28, 2021, the dataset is available at <https://www.nrel.gov/buildings/end-use-load-profiles.html>



**Figure 1. Major components of the *End-Use Load Profiles for the U.S. Building Stock* project**

### *Market Needs, Use Cases, and Data Gaps*

The first year of the project was focused on identifying and prioritizing use cases for EULPs, defining EULP data requirements for those use cases, defining the data needed for model calibration/validation, and identifying major data gaps. We assembled a TAG to assist with these tasks. Throughout the project, the TAG met quarterly to provide input on these activities (in year one), and to provide feedback on the calibration/validation process and the dataset publication plan (in years two and three). The TAG was composed of 92 individuals representing 61 organizations, including stakeholders from electric utilities, independent system operators (ISOs) and regional transmission organizations (RTOs), public utility commissions, state and local governments, consulting firms, software companies, academic institutions, nongovernmental organizations representing utilities and regional efficiency groups, and DOE. Geographically, our TAG members are located in 20 states that cover seven of the eight North American Electric Reliability Corporation regions, with many members conducting work regionally or nationally. A full list of TAG members can be found in Appendix A. The results of the project's first year are published in *End-Use Load Profiles for the U.S. Building Stock: Market Needs, Use Cases, and Data Gaps* (Mims Frick et al. 2019).

### *Acquisition of Data for Calibration/Validation*

The next major component of the project—spanning all three years—was acquiring empirical data for calibration and validation of the ResStock and ComStock models. We obtained access to a range of measured data, including utility meter data from more than 2.3 million customers, utility meter metadata, various end-use submetering datasets, and other public and private datasets related to energy use in buildings. The types of data collected, collection process, and data processing methodology are discussed in Section 2.3 of this report.

### *New Residential Stochastic Occupant Behavior Model*

The goal of this project was to produce EULPs at both the *aggregate* and *individual* building scales. Aggregate profiles represent the total profile for an end use in one or more customer segments in a utility territory or other region. Individual profiles represent real building or housing unit patterns, complete with the normal spikes and variability present in individual buildings and housing units. This is particularly important at the housing unit level; in large commercial and multifamily buildings, loads driven by stochastic occupant behavior are smoothed out to some degree, because of the larger number of occupants and the lesser degree of control that occupants have over end-use loads. To improve the realism of individual housing unit load profiles, we developed a new stochastic occupant behavior simulator and integrated it into ResStock, as discussed in Section 3.2.1. On the commercial side, we explored ways to improve the representation of occupant behavior in commercial buildings, but ultimately did not

have sufficient data to implement any improved methods. We did derive commercial building operation variability from advanced metering infrastructure (AMI) data and integrate this variability into ComStock (Section 3.3.17).

#### *Model Calibration, Validation, and Uncertainty Quantification*

The remainder of the project, spanning years two and three, was focused on ResStock and ComStock calibration, validation, and uncertainty quantification. These topics are the main subject of this report. An overview of the ResStock and ComStock approach to building stock modeling is presented in Section 2.2. Sections 2.3 and 2.5 present the methodology used for calibration, validation, and uncertainty quantification. Section 3 presents the changes made to ResStock and ComStock model inputs over the course of the calibration process. Section 4 presents the final results of model validation, consistent with the publicly released dataset. Section 5 presents a brief discussion and concluding remarks.

#### *Project Outcomes*

The culmination of this three-year project was the release of a publicly available dataset of simulated EULPs for the stock of residential and commercial buildings in the contiguous United States. In Section 4.3 we describe the published dataset and the three different ways to access it. Beyond the dataset of EULPs, we also published the set of calibrated building energy models in the OpenStudio® model (.osm) format, to enable our industry and researcher audience to model the impact of future scenarios and technology adoption.

In addition to this document, the project website provides instructions on how to access and download the various dataset formats, along with a frequently asked questions (FAQ) section. An accompanying report, *End-Use Load Profiles for the U.S. Building Stock: Applications and Opportunities*, provides discussion of applications and opportunities for using the dataset, with examples relevant for utilities, policymakers, technology developers, and researchers (Frick and Pigman 2022).

Ultimately, the outcomes of this project will help a broad audience—electric utilities, ISO/RTOs, public utility commissions, state and local governments, consulting firms, software companies, academic institutions, nongovernmental organizations, and DOE—make critical decisions about prioritizing research and development, utility resource and distribution system planning, and state and local energy planning and regulation. Additionally, the project outcomes can be a foundation for future work to develop *end-use savings shapes* that describe the difference in energy consumption between a baseline building and a building with an energy efficiency, electrification, or demand flexibility measure applied.

## 2 Methodology

### 2.1 Overall Approach to Calibration, Validation, and Uncertainty Quantification

#### 2.1.1 Definitions and Framework

Our approach to calibration, validation, and uncertainty quantification for this project is based on the framework presented in *Assessing the Reliability of Complex Models: Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification* by the National Research Council of the National Academies (National Research Council 2012). The authors of the document examined practices for verification, validation, and uncertainty quantification (VVUQ) of large-scale computational simulations across several research communities and identified common concepts, terms, approaches, tools, and best practices of VVUQ. We adopt the terminology of this document, as directly taken from Section 1.2 of the National Research Council (2012):

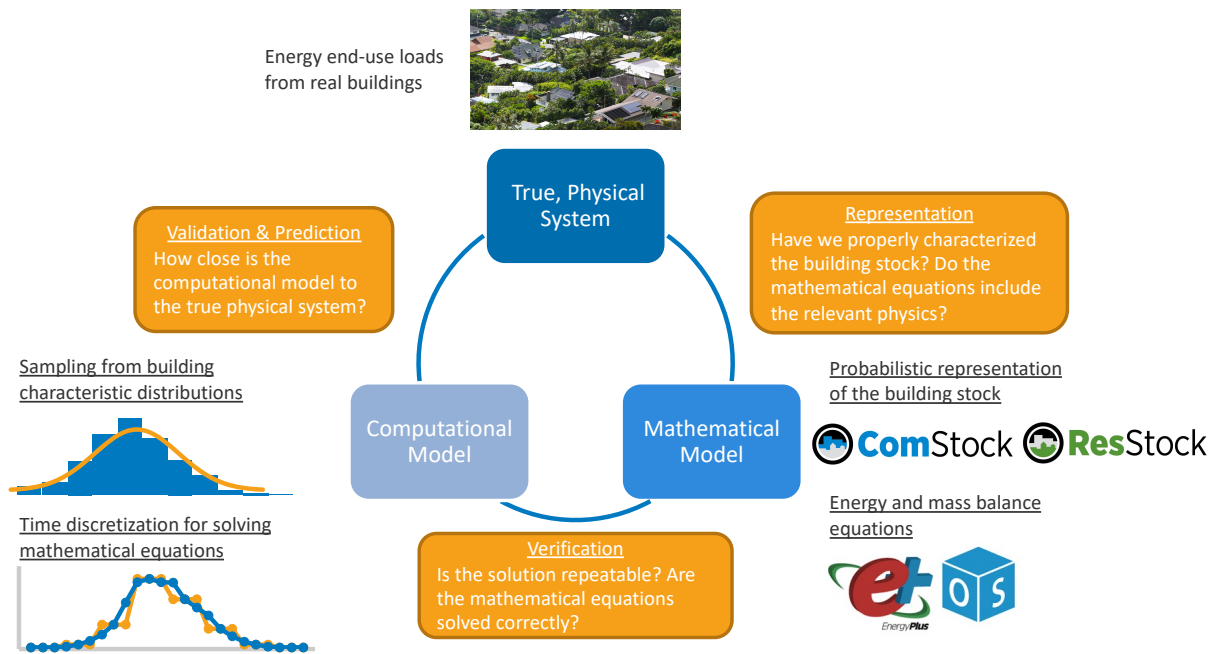
- *Verification*. The process of determining how accurately a computer program (“code”) correctly solves the equations of the mathematical model. This includes code verification (determining whether the code correctly implements the intended algorithms) and solution verification (determining the accuracy with which the algorithms solve the mathematical model’s equations for specified quantities of interest (QOI)).
- *Validation*. The process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model (from *Guide for the Verification and Validation of Computational Fluid Dynamics Simulations* 1998).
- *Uncertainty quantification (UQ)*. The process of quantifying uncertainties associated with model calculations of true, physical QOI, with the goals of accounting for all sources of uncertainty and quantifying the contributions of specific sources to the overall uncertainty.

We add to this our own definitions for model calibration and output correction:

- *Calibration*. The process of using empirical data to inform changes to model input parameters or changes to model structure.
- *Output correction*. The process of using a mathematical *correction model* to tune or correct the outputs of ResStock based on empirical data, with the goal of reducing any remaining discrepancies between model outputs and available empirical data after the calibration process is complete. For this project, output correction only applies to ResStock, because empirical data necessary for ComStock were not available. See Section 3.2.10 for a full description of the output correction model.

Figure 2 shows how these concepts relate to the true, physical system (end-use loads from real buildings), the mathematical model (represented by ComStock, ResStock, EnergyPlus, and OpenStudio in our case), and the computational model (sampling schemes and the numerical methods in EnergyPlus for solving building energy and mass balance questions). Because ComStock and ResStock are not mathematical models, but rather probabilistic representations of the building stock, they serve the role of providing input distributions to the mathematical model (EnergyPlus and OpenStudio) in our framework.

Extensive work has been completed on verification of accuracy of the EnergyPlus simulation engine (Judkoff and Neymark 2013, Judkoff et al. 2017). Thus, this project does not include verification. We do comment on one aspect of verification: how the number of ComStock or ResStock samples used to calculate model outputs affects the uncertainty of those outputs (e.g., estimates of peak demand in low population counties will have greater uncertainty because they are assigned fewer samples in the national-scale ComStock and ResStock simulations; see *Uncertainty Due to Model Sample Size*).



**Figure 2. Verification, validation, and prediction as they relate to the true, physical system, the mathematical model (ResStock, ComStock, EnergyPlus, and OpenStudio in our case), and the computational model (sampling schemes and numerical method to solve building energy and mass balance questions in our case). (Adapted from National Research Council 2012)**

### 2.1.2 Quantities of Interest

The calibration and validation process involves comparing stock model outputs to empirical data. For timeseries data, industry standards include the normalized mean bias error (NMBE) and coefficient of variation of root mean squared error metrics (CV(RMSE)) presented in ASHRAE Guideline 14 (ASHRAE 2014). NMBE and CV(RMSE) can be useful metrics, but they weight every time period (month or hour of the year) equally. For the use cases of EULPs identified with the TAG, certain hours (e.g., hours of peak demand) are much more important than other hours. Thus, for this project we developed a specific set of 12 metrics—quantities of interest (QOI)—designed to reflect the most important values for the primary use cases of EULPs. These QOIs are described in Table 1. By developing these QOIs before calibration started and tracking them throughout the calibration process, we were able to evaluate our progress and prioritize model changes based on metrics that have the most significance to the users of the final dataset. These QOIs are the primary way that we judge the model outputs, and are used throughout this document in presenting both the calibration/validation results and uncertainty quantification results.

#### Definition of Seasons in QOIs

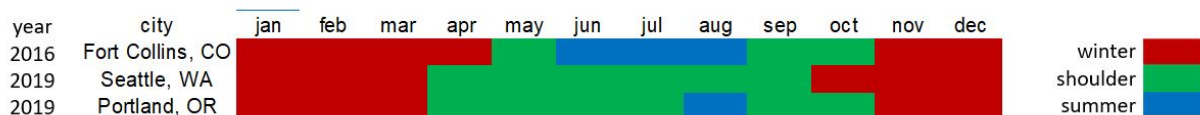
We use the terms “summer,” “shoulder,” and “winter” in these QOI descriptions to make interpretation intuitive. Because the QOIs are calculated for regions with different climates, we created a definition that could be applied to any climate. The choice of temperature cutoffs was based on research into the points at which, as a group, buildings transition from heating to no conditioning to cooling.

- For each calendar month, calculate average daily temperature
- Season = winter if average daily temperature  $< 55^{\circ}\text{F}$
- Season = shoulder if  $55^{\circ}\text{F} \leq \text{average daily temperature} \leq 70^{\circ}\text{F}$
- Season = summer if average daily temperature  $> 70^{\circ}\text{F}$

**Table 1. Quantities of Interest (QOI):** This set of 12 QOI was designed to reflect the most important model outputs for the primary use cases of EULPs

QOI	Example Units
Annual energy use	kWh
Daily minimum base load, average of all days in season, summer	kW (avg. over 1 hour)
Daily minimum base load, average of all days in season, shoulder	kW (avg. over 1 hour)
Daily minimum base load, average of all days in season, winter	kW (avg. over 1 hour)
Daily peak load, average of all days in season, summer	kW (avg. over 1 hour)
Daily peak load, average of all days in season, winter	kW (avg. over 1 hour)
Daily peak load, average of top 10 days of season, summer	kW (avg. over 1 hour)
Daily peak load, average of top 10 days of season, winter	kW (avg. over 1 hour)
Daily peak timing, average of all days in season, summer	Hour/minute of day
Daily peak timing, average of all days in season, winter	Hour/minute of day
Daily peak timing, average of top 10 days of season, summer	Hour/minute of day
Daily peak timing, average of top 10 days of season, winter	Hour/minute of day

As shown in Figure 3, the season assigned to each calendar month can vary depending on the location and the weather during a particular year. Also, as exemplified by Seattle in 2019, a region does not necessarily have all three seasons (Seattle has no summer in 2019).



**Figure 3. Season assignments for three example regions**

### 2.1.3 Organization of the Methodology Section

The remainder of the Methodology section is organized as follows. Section 2.2 describes the methodology of ResStock and ComStock, including data sources, sampling techniques, physics-simulation engines, and model outputs. Section 2.3 describes the methodology used for the calibration and validation procedure, including what empirical data were used and how they were obtained, our multidimensional region-by-region calibration strategy, and an end-use transferability study. Section 2.4 describes how the input weather data were developed and Section 2.5 describes our methodology for uncertainty quantification via propagation of input uncertainty ranges through surrogate models.

Documentation of the calibration process itself—the improvements made to the models—is presented in Section 3, Model Calibration. The results of validation are presented in Section 4, Results. The results of uncertainty quantification are presented in Section 4.4.

## 2.2 Methodology: ResStock and ComStock

ResStock and ComStock are physics-based simulation models developed to represent the energy use and energy saving potential of residential and commercial building stocks with high granularity at national, regional, and local scales. ResStock and ComStock are models that have been developed and maintained by NREL since 2014 and 2016, respectively.

Compared to other building stock models, ResStock and ComStock typically use a large number of representative models—10,000s or 100,000s, depending on the application—to represent the building stock with high fidelity. Unlike some urban building energy modeling approaches, ResStock and ComStock do not attempt to generate a physics model for every building, but rather use a relatively large number of statistically sampled models to represent the building stock with a realistic diversity of building characteristics. The major data sources for ResStock and ComStock are listed in Table 2 and Table 3, respectively.



**Table 2. Major Data Sources Used in ResStock (as of the publication of this report and EULP dataset v1.0)**

<b>ResStock Parameter</b>	<b>Data Source</b>
American Housing Survey (AHS) Region	Spatial definitions are from U.S. Census 2010.; Core Based Statistical Area (CBSA) data based on the February 2013 CBSA delineation file.
ASHRAE IECC Climate Zone 2004	Unit counts are from the American Community Survey 5-yr 2016.; Spatial definitions are from U.S. Census 2010.; Climate zone data are from ASHRAE 169 2006, IECC 2004, and M.C. Baechler 2015.
Bathroom Spot Vent Hour	Same as occupancy schedule from Wilson et al. “Building America House Simulation Protocols” 2014.
Bedrooms	2017 AHS microdata.; Building type categorization based on U.S. EIA 2009 Residential Energy Consumption Survey (RECS).
Building America Climate Zone	Unit counts are from the American Community Survey 5-yr 2016.; Spatial definitions are from U.S. Census 2010.; Climate zone data are from ASHRAE 169 2006, IECC 2004, and M.C. Baechler 2015.; ISO and RTO regions are from EIA Form 861.; CBSA data based on the Feb 2013 CBSA delineation file.
CEC Climate Zone	Census tract definitions are from U.S. Census 2010.; Zip code definitions are from the end of Q2 2020.; The climate zone to zip codes in California is from the California Energy Commission website.
Ceiling Fan	Wilson et al. “Building America House Simulation Protocols” 2014, national average used as saturation.
Census Division	Spatial definitions are from U.S. Census 2010.
Census Division RECS	U.S. EIA 2015 RECS codebook.
Census Region	Spatial definitions are from U.S. Census 2010.
Clothes Dryer	2017 AHS microdata; CBSA data based on the Feb 2013 CBSA delineation file.
Clothes Washer	U.S. EIA 2009 RECS microdata.
Clothes Washer Presence	2017 AHS microdata; CBSA data based on the Feb 2013 CBSA delineation file.
Cooking Range	(Fuel type) RECS 2009; (Usage) engineering judgment.
Cooling Setpoint	U.S. EIA 2009 RECS microdata.
Cooling Setpoint Has Offset	U.S. EIA 2009 RECS microdata.
Cooling Setpoint Offset Magnitude	U.S. EIA 2009 RECS microdata.
Cooling Setpoint Offset Period	U.S. EIA 2009 RECS microdata.
Corridor	Engineering judgment.
County	Unit counts are from the American Community Survey 5-yr 2016.; Spatial definitions are from U.S. Census 2010.
Dehumidifier	Not applicable (dehumidifiers are not explicitly modeled separate from plug loads)
Dishwasher	U.S. EIA 2009 RECS microdata.
Door Area	Engineering judgment.
Doors	Engineering judgment.
Ducts	(duct insulation as a function of location) IECC 2009; (leakage distribution) Lucas and Cole, ‘Impacts of the 2009 IECC for Residential Buildings at State Level’, 2009; (duct location) Wilson et al., ‘Building America House Simulation Protocols’, 2014
Electric Vehicle	Not applicable (electric vehicle charging is not currently modeled separate from plug loads)
Geometry Attic Type	U.S. EIA 2009 RECS microdata.
Geometry Building Horizontal Location multifamily (MF)	Calculated directly from other distributions.
Geometry Building Horizontal Location single-family attached (SFA)	Calculated directly from other distributions.



**Table 2: Major Data Sources Used in ResStock (Continued)**

<b>ResStock Parameter</b>	<b>Data Source</b>
Geometry Building Level MF	Calculated directly from other distributions.
Geometry Building Number Units MF	U.S. EIA 2009 RECS microdata.
Geometry Building Number Units SFA	U.S. EIA 2009 RECS microdata.
Geometry Building Type ACS	5-yr 2017 Public Use Microdata Samples (PUMS). IPUMS USA, University of Minnesota, <a href="http://www.ipums.org">www.ipums.org</a> .
Geometry Building Type Height	Calculated directly from other distributions.
Geometry Building Type RECS	5-yr 2017 PUMS. IPUMS USA, University of Minnesota, <a href="http://www.ipums.org">www.ipums.org</a> .
Geometry Floor Area	2017 AHS microdata.
Geometry Floor Area Bin	The sample counts and sample weights are constructed using U.S. EIA 2009 RECS microdata.; Geometry Floor Area bins are from the UNITSIZE field of the 2017 AHS.
Geometry Foundation Type	U.S. EIA 2009 RECS microdata.
Geometry Garage	U.S. EIA 2009 RECS microdata.
Geometry Stories	U.S. EIA 2009 RECS microdata.
Geometry Stories Low Rise	Calculated directly from other distributions.
Geometry Wall Exterior Finish	Homeland Infrastructure Foundation-Level Data (HIFLD) Parcel data.
Geometry Wall Type	HIFLD Parcel data.
Geometry Wall Type And Exterior Finish	HIFLD Parcel data.
HVAC Cooling Efficiency	The sample counts and sample weights are constructed using U.S. EIA 2009 RECS microdata.; Efficiency data based on CAC-ASHP-shipments-table.tsv, room_AC_efficiency_vs_age.tsv and expanded_HESC_HVAC_efficiencies.tsv combined with age of equipment data from RECS.
HVAC Cooling Type	The sample counts and sample weights are constructed using U.S. EIA 2009 RECS microdata.
HVAC Has Ducts	The sample counts and sample weights are constructed using U.S. EIA 2009 RECS microdata.
HVAC Has Shared System	The sample counts and sample weights are constructed using U.S. EIA 2009 RECS microdata.
HVAC Heating Efficiency	The sample counts and sample weights are constructed using U.S. EIA 2009 RECS microdata.; Shipment data based on CAC-ASHP-shipments-table.tsv and furnace-shipments-table.tsv.; Efficiency data based on expanded_HESC_HVAC_efficiencies.tsv combined with age of equipment data from RECS.
HVAC Heating Type	U.S. EIA 2009 RECS microdata.
HVAC Heating Type and Fuel	Calculated directly from other distributions.
HVAC Shared Efficiencies	The sample counts and sample weights are constructed using U.S. EIA 2009 RECS microdata.
Has solar photovoltaics (PV)	ACS population and RiDER data on PV installation that combines LBNL's 2020 Tracking the Sun and Wood Mackenzie's 2020 Q4 PV report.
Heating Fuel	5-yr 2017 PUMS. IPUMS USA, University of Minnesota, <a href="http://www.ipums.org">www.ipums.org</a> .
Heating Setpoint	U.S. EIA 2009 RECS microdata.
Heating Setpoint Has Offset	U.S. EIA 2009 RECS microdata.

**Table 2: Major Data Sources Used in ResStock (Continued)**

<b>ResStock Parameter</b>	<b>Data Source</b>
Heating Setpoint Offset Magnitude	U.S. EIA 2009 RECS microdata.
Heating Setpoint Offset Period	U.S. EIA 2009 RECS microdata.
Holiday Lighting	Not applicable (holiday lighting is not currently modeled separate from other exterior lighting)
Hot Water Distribution	Engineering judgment.
Hot Water Fixtures	Engineering judgment.
ISO RTO Region	Spatial definitions are from U.S. Census 2010.; ISO and RTO regions are from EIA Form 861.
Infiltration	Distributions are based on the cumulative distribution functions from the Residential Diagnostics Database (ResDB), <a href="http://resdb.lbl.gov/">http://resdb.lbl.gov/</a> .
Insulation Ceiling	NEEA Residential Building Stock Assessment, 2012; Nettleton, G.; Edwards, J. (2012). Data Collection-Data Characterization Summary, NorthernSTAR Building America Partnership, Building Technologies Program. Washington, D.C.; Derived from Home Innovation Research Labs 1982-2007 Data
Insulation Floor	Derived from Home Innovation Research Labs 1982-2007 Data; (pre-1980) Engineering judgment
Insulation Foundation Wall	Derived from Home Innovation Research Labs 1982-2007 Data; (pre-1980) Engineering judgment
Insulation Roof	Derived from Home Innovation Research Labs 1982-2007 Data; NEEA Residential Building Stock Assessment, 2012
Insulation Slab	Derived from Home Innovation Research Labs 1982-2007 Data; (pre-1980) Engineering judgment
Insulation Wall	Derived from Home Innovation Research Labs 1982-2007 Data; Ritschard et al. Single-Family Heating and Cooling Requirements; Nettleton, G., Edwards, J. (2012). Data Collection-Data Characterization Summary, NorthernSTAR Building America Partnership, Building Technologies Program. Washington, D.C.
Interior Shading	ANSI/RESNET/ICC 301 Standard.
Lighting	U.S. EIA 2015 RECS microdata.; 2019 Energy Savings Forecast of Solid-State Lighting in General Illumination Applications. <a href="https://www.energy.gov/sites/prod/files/2019/12/f69/2019_ssl-energy-savings-forecast.pdf">https://www.energy.gov/sites/prod/files/2019/12/f69/2019_ssl-energy-savings-forecast.pdf</a> .
Lighting Interior Use	Not applicable; this parameter for adding diversity to lighting usage patterns is not currently used.
Lighting Other Use	Not applicable; this parameter for adding diversity to lighting usage patterns is not currently used.
Location Region	Spatial definitions are from U.S. Census 2010.
Mechanical Ventilation	Engineering judgment.
Misc Extra Refrigerator	U.S. EIA 2009 RECS microdata.; Age of refrigerator converted to efficiency levels using ENERGY STAR® shipment-weighted efficiencies by year data from Home Energy Score.
Misc Freezer	U.S. EIA 2009 RECS microdata.
Misc Gas Fireplace	Wilson et al. "Building America House Simulation Protocols" 2014, national average fraction used for saturation.
Misc Gas Grill	Wilson et al. "Building America House Simulation Protocols" 2014, national average fraction used for saturation.
Misc Gas Lighting	Wilson et al. "Building America House Simulation Protocols" 2014, national average fraction used for saturation.
Misc Hot Tub Spa	U.S. EIA 2009 RECS microdata.
Misc Pool	U.S. EIA 2009 RECS microdata.

**Table 2: Major Data Sources Used in ResStock (Continued)**

<b>ResStock Parameter</b>	<b>Data Source</b>
Misc Pool Heater	U.S. EIA 2009 RECS microdata.
Misc Pool Pump	Wilson et al. “Building America House Simulation Protocols” 2014, national average fraction used for saturation.
Misc Well Pump	Wilson et al. “Building America House Simulation Protocols” 2014, national average fraction used for saturation.
Natural Ventilation	Wilson et al. “Building America House Simulation Protocols” 2014.
Neighbors	OpenStreetMap data queried by Radiant Labs for Multi-Family and Single-Family Attached.; Engineering judgment for others.
Occupants	5-yr 2017 PUMS. IPUMS USA, University of Minnesota, <a href="http://www.ipums.org">www.ipums.org</a> .
Orientation	OpenStreetMap data queried by Radiant Labs.
Overhangs	Not applicable; all homes are assumed to not have window overhangs other than eaves.
PUMA	Unit counts are from the American Community Survey 5-yr 2016.; Spatial definitions are from U.S. Census 2010.
PV Orientation	LBNL’s 2020 Tracking the Sun.
PV System Size	LBNL’s 2020 Tracking the Sun.
Plug Load Diversity	Engineering judgment, calibration.
Plug Loads	U.S. EIA 2015 RECS microdata.
ReEDS Balancing Area	Regional Energy Deployment System (ReEDS) Model Documentation (Ho et al. 2021)
Radiant Barrier	Not applicable; all homes are assumed to not have attic radiant barriers installed.
Range Spot Vent Hour	Derived from national average cooking range schedule in Wilson et al. “Building America House Simulation Protocols” 2014.
Refrigerator	U.S. EIA 2009 Residential Energy Consumption Survey (RECS) microdata.; Age of refrigerator converted to efficiency levels using ENERGY STAR shipment-weighted efficiencies by year.
Roof Material	U.S. EIA 2009 Residential Energy Consumption Survey (RECS) microdata.
Solar Hot Water	Not applicable; all homes are assumed to not have solar water heating
State	Spatial definitions are from U.S. Census 2010.
Usage Level	Engineering judgment, calibration of variation.
Vacancy Status	5-yr 2016 PUMS. IPUMS USA, University of Minnesota, <a href="http://www.ipums.org">www.ipums.org</a> .
Vintage	5-yr 2017 PUMS. IPUMS USA, University of Minnesota, <a href="http://www.ipums.org">www.ipums.org</a> .
Vintage ACS	5-yr 2017 PUMS. IPUMS USA, University of Minnesota, <a href="http://www.ipums.org">www.ipums.org</a> .
Water Heater Efficiency	U.S. EIA 2009 RECS microdata.; (Heat pump water heaters) 2016-17 RBSA II for CR06 (WA and OR) and Butzbaugh et al. 2017. US HPWH Market Transformation. for other regions.
Water Heater Fuel	U.S. EIA 2009 RECS microdata.
Water Heater In Unit	U.S. EIA 2009 RECS microdata.
Window Areas	2016-17 Residential Building Stock Assessment (RBSA) II microdata.
Windows	U.S. EIA 2015 RECS microdata.; Source of storm windows: D&R International, Ltd. ‘Residential Windows and Window Coverings: A Detailed View of the Installed Base and User Behavior’ 2013. <a href="https://www.energy.gov/sites/prod/files/2013/11/f5/residential_windows_coverings.pdf">https://www.energy.gov/sites/prod/files/2013/11/f5/residential_windows_coverings.pdf</a> . Source of low-e distribution is based on engineering judgement, informed by high-level sales trends observed in Ducker Worldwide studies of the U.S. Market for Windows, Doors and Skylights.

The ResStock probability distributions for each of these parameters are located at [https://github.com/NREL/resstock/tree/eulp\\_final/project\\_national/housing\\_characteristics](https://github.com/NREL/resstock/tree/eulp_final/project_national/housing_characteristics) (EULP dataset version) and <https://github.com/NREL/>

resstock/tree/develop/project\_national/housing\_characteristics (latest developmental version). See the comments included at the bottom of each file for additional assumptions about how the probability distributions were derived.

For both ResStock and ComStock, the general methodology is as follows:

1. **Stock characterization.** Conditional probability distributions for building stock characteristics are queried from data sources (e.g., distribution of “year structure built” as a function of location and “building type”). Parameters common across data sources, such as geographic location, building type, and vintage, are used to combine and map between the disparate data sources. Geographic resolution for queried distributions varies in scale—for example, from counties (around 3,000) to climate zones (16)—so various geospatial data sources are used to map between geographic resolutions. The conditional probability distributions take the form of a hierarchical tree of dependencies.
2. **Sampling.** The parameter space defined by the conditional probability distributions is sampled. ResStock currently uses deterministic quota sampling, with random combination of non-correlated parameters. ComStock currently uses quasi-random sampling using Sobol’ sequences. The number of samples is an input that can be customized. For U.S. national-scale runs of ResStock, we typically use 550,000 samples to represent 133,172,057 dwelling units (approximately 1:242). For U.S. national scale runs of ComStock, we currently use 350,000 samples to represent 1.8 million commercial buildings (approximately 1:5). The appropriate ratio of samples to buildings or dwelling units was initially determined through convergence testing for national-scale applications (Wilson et al. 2017); however, the appropriate ratio for different applications and scales is the subject of ongoing research (see Section 5.1.3 for more on sample size convergence testing).
3. **Physics simulation.** The samples are used to construct physics simulation models using a simulation engine of choice. NREL typically uses the EnergyPlus® simulation engine for this purpose (Crawley et al. 2001). Model construction and articulation is facilitated by the OpenStudio® software development kit (Roth, Goldwasser, and Parker 2016) and associated commercial and residential modeling workflows (e.g., OpenStudio-Standards: NREL et al. 2021 and OpenStudio-HPXML: NREL 2021b).
4. **Model outputs and post-processing.** Model outputs include both annual and hourly or subhourly timeseries energy use outputs for each sample for major and minor end uses (electricity and on-site natural gas, propane, and fuel oil use). Outputs for each sample also include HVAC system capacities and hours the heating and cooling setpoints were not met. Optional outputs also include timeseries indoor zone temperatures (e.g., for analyzing thermal comfort and resilience).

**Table 3. Major Data Sources Used in ComStock**

<b>Data Sources</b>	<b>ComStock Inputs</b>
2015 U.S. Lighting Market Characterization	Interior/exterior lighting power density.
ASHRAE 62.1	Ventilation rates.
ASHRAE 90.1	Nominal efficiency levels for lighting, HVAC and envelope properties.
ASHRAE Service Life and Maintenance Cost Database	Distribution shapes for equipment lifespans.
Building Codes Assistance Project	Timeline of energy code adoption by state.
DOE Prototype Buildings	Lighting power density, efficiency, occupancy.
CoStar Real Estate Data (2017)	Fine geospatial resolution for buildings that are part of an active real estate market: Building type, vintage, floor area, number of stories, location
Database for Energy Efficiency Resources (DEER)	Building system lifespans. For Buildings in California, schedules and power densities for interior lighting and interior equipment, building program, HVAC system controls and efficiencies, ventilation rates, infiltration rate.
Commercial Building Stock Assessment 4 (2019) Final Report	Lighting power density, lifespan of windows.
DOE Commercial Prototype Buildings	Loads, efficiency, occupancy, space type ratio and zone definition.
EIA Commercial Building 2012 Energy Consumption Survey (CBECS) 2012	Coarse geospatial resolution: Building type, vintage, floor area, number of stories, energy sources, conditioned floor space, HVAC and water heating equipment, hours of operation, aspect ratio, window-to-wall ratio, thermostat setback saturation, data center saturation, heating (space and water) saturation, other equipment saturation.
Guidehouse Commercial Fenestration Market Study	Window characteristics and window-to-wall ratio (WWR).
Homeland Security Infrastructure Plan (HSIP) Gold 2012 Database	Fine geospatial resolution for buildings that are not part of an active real estate market (hospital, schools): Building type, number of beds (hospitals), enrollment (schools), location
Proprietary End-Use Submeter Data	Schedules and power densities for interior lighting and interior equipment, HVAC operation patterns, thermostat setpoints.
Strip Mall Restaurant Surveying	Percentage of restaurants in strip malls.
U.S. Census Bureau Public Use Microdata Sample (PUMS), 5-yr 2017.	Heating fuel distributions are inferred from distributions of heating fuel for residential dwelling units.

## 2.3 Methodology: Calibration and Validation

As defined in Section 2.1.1, *calibration* is the process of using empirical data to inform model changes, and *validation* is the process of evaluating how accurately a model represents the real world. Thus, calibration and validation are closely related and are used together to iteratively improve a model's accuracy or reduce its uncertainty. This section provides background on large-scale building stock model calibration and validation and introduces the various approaches used for calibration of *individual* building energy models. We then describe several approaches to EULP modeling calibration that exist in industry and the literature. We then introduce our project's calibration philosophy. Next, we describe how we obtained, processed, and analyzed the empirical data we used for validation and calibration, before describing each of the empirical data sources used for residential and commercial calibration. Finally, we describe how calibration was divided into regional stages and the multifaceted approach we used to evaluate validation progress. This section also includes an evaluation of the region-to-region transferability of residential end uses, followed by the introduction of metrics used to evaluate the variability of commercial building operation schedules. We also present a residential end-use transferability study completed to evaluate how non-weather driven end-use behavior varies from region to region.

### 2.3.1 Background

The use of large numbers of bottom-up physics-simulation models to represent the buildings in a region or city (sometimes called urban building energy modeling, or UBE) is a relatively new field (Reinhart and Davila 2016). The application of physics-based building stock models to generate hourly EULPs has been studied even less. Davila (2017) reviewed UBE approaches in the literature and found that the lack of UBE validation and calibration studies was a major gap. Taylor et al. (2019) similarly found that there has been little to no work studying calibration of building stock models, particularly at multiple scales. Restrictions on access to measured building energy use data and unknown thermal properties of buildings were identified as two major barriers to UBE calibration (Reinhart and Davila 2016). Our team's work on ResStock is an early example of building stock model calibration, although prior to this project, it was limited to calibration of *annual* end-use energy (Wilson et al. 2016) and one application at a city/utility scale with hourly resolution (Cochran et al. 2021).

Members of our project team and our project sponsor, DOE's Building Technologies Office, represented the United States by participating in the International Energy Agency (IEA) Energy in Buildings and Communities (EBC) *Annex 70 – Building Energy Epidemiology: Analysis of real building energy use at scale*. As part of this work, we convened with an international community of building stock modelers and shared knowledge on the state of the art of building stock model calibration, validation, and uncertainty quantification. This work is discussed in Langevin et al. (2020) and Fennell et al. (2021).

### 2.3.2 Approaches to Building Energy Model Calibration

In this section, we discuss various approaches to building energy modeling calibration. We divide calibration approaches into two categories: *output* calibration and *input* calibration. Models may use just one approach or a combination of both approaches.

#### *Input Calibration*

*Input* calibration—where uncertain inputs to a building simulation program are tuned to improve the match between model output and measured data—is the type of calibration most typically used in the building energy modeling community. Reddy (2006) describes it as “an art form that inevitably relies on user knowledge, past experience, statistical expertise, engineering judgment, and an abundance of trial and error.” Because it is labor intensive when performed manually, this type of calibration is typically only used for individual building energy models (not stock models or UBE), and is typically only used for informing energy efficiency retrofits of large existing buildings.

ASHRAE Guideline 14-2014 provides criteria for the goodness of fit for model calibration (ASHRAE 2014). For example, the whole-building calibrated simulation performance path requires that the building energy model have an NMBE of 5% and a CV(RMSE) of 15% for calibration to monthly data, and an NMBE of 10% and a CV(RMSE) of 30% for calibration to hourly data. However, because input calibration is almost always an underdetermined problem—there are multiple calibration parameters available that affect modeled end uses in similar ways (e.g., thermostat setpoint and air leakage rate)—tuning model inputs without a basis in observed data is prone to result in getting the right answer for the wrong reason (sometimes called “input error” or “input-side error”; Garrett and New 2016). This is problematic for physics-based models that are intended for making predictions about the impact of physical changes to building characteristics, especially when the studied changes are the same as the calibration



parameters (e.g., thermostat setpoint or air leakage rate from the example above). Judkoff et al. (2016) described this issue and presented a method to test model calibration techniques.

### *Automated Input Calibration*

Methods for automating model input calibration using optimization techniques have emerged in the past decade. One example is Autotune, which achieved calibration results similar to an expert modeler in a fraction of the time (New et al. 2012, Chaudhary et al. 2016). Robertson et al. (2013) adapted the Building Energy Simulation Test for Existing Homes (BESTEST-Ex) framework developed by Judkoff et al. (2010) and applied it to automated model calibration techniques. They concluded that the automated calibration optimization problem is “still significantly underdetermined,” even when calibrating to hourly data, and is thus still likely to get the right answer for the wrong reason.

### *Output Calibration*

In contrast to input calibration, *output* calibration involves modification of simulation outputs to achieve a better match with empirical validation data. The modification could be as simple as applying scaling factors (colloquially known as “fudge factors” or “correction coefficients”) to true up model outputs, or could use more sophisticated statistical or machine learning models (Valovcin et al. 2014). The modifications could be at annual, monthly, or hourly resolution, depending on the validation data available. Output modifications could be for whole-building electricity or gas usage, or for specific end uses, again depending on the validation data available.

Output calibration is appealing in that it is agnostic to the cause of error and therefore avoids the issue of getting the right answer for the wrong reason. It can also be applied to what-if scenario predictions, which would assume that the cause and magnitude of error is the same for any scenario predictions. Similar to pure statistical models, it is prudent to cross-validate the output calibration model using a training dataset to derive the scaling factors or correction model and a test dataset to evaluate the model’s performance and avoid overfitting.

Though we mostly rely on input calibration for this EULP project, we use output calibration in a targeted manner only for the residential sector EULP dataset (empirical data necessary for the commercial sector were not available). Throughout the rest of this report, we refer to this step as “output correction” to distinguish it from the input calibration procedure. See Section 3.2.10 for a full description of the ResStock output correction model.

## **2.3.3 Load Profile Calibration Examples**

In this section, we highlight the few examples of EULP model calibration that have been used by researchers and the utility industry. Representatives from the teams that worked on the 2006 California Commercial End-Use Survey (CEUS) and 2019 California Investor-Owned Utility Electricity Load Shapes study served on our TAG and provided valuable feedback throughout the project.

### *Input Calibration Examples*

#### **2006 California Commercial End-Use Survey**

The 2006 CEUS was a study conducted by Itron, Inc. and funded by the California Energy Commission to create EULPs for the commercial building stock in California (Itron 2006). The study involved the creation of detailed energy models for 485 commercial buildings; individual, mostly manual, input calibration of those models to utility data; and then weighting of those building models to represent the entire stock. Most buildings had only monthly utility data available, but 17% had hourly AMI data and 17% also had short-term metering of some end uses.

#### **2019 California Investor-Owned Utility Electricity Load Shapes**

In 2018–2020, the California Energy Commission contracted with ADM Associates to conduct an update to the load shapes used by California Energy Commission’s Hourly Electric Load Model, which is used for forecasting the load of California’s investor-owned utilities (Baroiant et al. 2019). The approach used for creating commercial building load shapes was similar to the approach we used but on a different scale. First, internal load shapes and magnitudes were created by modifying data from the Commercial End-Use Survey (Itron 2006) (input calibration). Next, these internal loads were substituted into DOE’s Commercial Prototype Building models. These models were run, and regressions for the weather-dependent loads were created from the output data. A regression model of the residual was created for each building type to account for the difference between average AMI load shapes and the modeled load shapes (output calibration).



The approach used for residential load shapes did not use building energy modeling. For non-weather-dependent loads, a variety of sources of end-use load shapes were used to build up a composite total load profile. For weather-dependent loads, a change-point type model was first used to break the average AMI load profiles into weather-dependent and weather-independent portions. Next a regression model of the weather-dependent load was created as a function of temperature and weekday/weekend. Finally, they used output calibration in the form of a correction model to fill in final missing loads and true-up error relative to the AMI data.

#### *Output Calibration Examples*

##### **Pacific Northwest National Laboratory's BEND Model**

One example of output calibration that is highly relevant to our work is the scaling factor-based calibration approach used for the Pacific Northwest National Laboratory's building energy demand model (BEND model). BEND is an EnergyPlus-based physics-simulation building stock model used to generate aggregate hourly load profiles (Taylor et al. 2019). Using 2007 U.S. Energy Information Administration (EIA) data, the authors calculated 288 scaling factors for each combination of 12 months and 24 hours, for the residential and commercial sectors in each electricity balancing authority. The scaling factors were calculated using two data sources:

1. Monthly total electricity sales by utility company and sector (residential, commercial, industrial, and other) from the EIA's Form EIA-861M 2007 detailed data (EIA 2007).
2. Hourly total electricity by balancing authority from Federal Energy Regulatory Commission (FERC) Form 714 data (FERC 2007).

The authors note that the two datasets are not entirely consistent (i.e., if you aggregate the hourly data to monthly and the monthly data across sectors). When applied to the raw output from the BEND EnergyPlus models, the output calibration results in commercial sector scaling factors generally in the 0–25 range for most balancing authorities (one balancing authority has scaling factors that approach 100 (i.e., 9,900% error) in July and August) and residential sector scaling factors generally in the 0.1–3 range. The scaling factor approach performs relatively well when trained with 2007 data and evaluated on 2010 data, with seasonal peak demand estimates by balancing authority ranging from a 17% underestimate to a 37% overestimate. The authors acknowledge that the scaling factor approach is limited in that it does not account for how year-to-year variations in weather affect peak loads, with error in peak load being highly correlated with year-to-year temperature differences (Taylor et al. 2019). Nevertheless, the simplicity of the approach is appealing, and the study provides several insights that we apply in our own development of a residential output correction model.

##### **EnergyPATHWAYS**

Evolved Energy Research's application of the EnergyPATHWAYS model for NREL's Electrification Futures Study is another example of using correction factors to reconcile the difference between a bottom-up modeled electricity load shape and top-down historical data, again using FERC Form 714 data (Mai et al. 2018). The correction factors used to reconcile these differences were not published, but the authors note that the need for reconciliation is due to a lack of adequate temporal and spatial granularity in the constituent end-use profiles, which were a mix of measured and simulated profiles.

##### **ResStock and ComStock for LA100**

One example of output calibration that used a more elaborate approach was NREL's application of ResStock and ComStock for the Los Angeles 100% Renewable Energy Study (LA100). In this case, after a substantial effort on manual input calibration at the stock scale, a variety of optimization and machine learning techniques were developed and applied to true up predicted EULPs against measured sector total hourly load research data obtained from the Los Angeles Department of Water and Power. The correction models were trained with 2016 simulation output and 2016 load research data and found to perform well when applied to 2017 simulations and evaluated against 2017 load research data (Hale et al. 2021). The resulting 8,760 (hourly) correction timeseries was added as an artificial "calibration" end use to the stack of other modeled end uses. The correction models were primarily used to fill in missing nighttime loads from unknown end uses in ComStock outputs and to correct time-shifted ResStock cooling energy outputs.

Our team found that the complexity of the correction models made them challenging to work with. The development of these hourly machine-learning-based correction models provided valuable lessons learned for our team, but for

the national-scale project that is the subject of this report, they are of limited use because they require hourly load research data, which we do not have nationwide. However, utility industry users of our EULPs may be interested in applying this type of output calibration to our EULP dataset using hourly utility load data to which they do have access.

### 2.3.4 Our Calibration Philosophy

The calibration philosophy developed by our team to guide this project is rooted in a variety of sources. NREL's historical contributions to the field of building energy simulation engine validation and testing date back to the early 1980s (Judkoff, Wortman, and Burch 1982) and were further developed over subsequent years with BESTEST (Judkoff and Neymark 1995), ASHRAE Standard 140 (Judkoff and Neymark 2006), and BESTEST-Ex (Judkoff et al. 2010). More recently, NREL pioneered work on evaluating the accuracy of building energy simulations at scale, using empirical field data from large numbers of buildings (Polly, Kruis, and Roberts 2011) and statistical analysis of the biases and variance in building energy simulation outputs (Valovcin et al. 2014). This history provides a foundation for our project's philosophy on calibration and model validation.

We complement this historical domain expertise with the cross-domain state-of-the-art framing of verification, validation, and uncertainty quantification (VVUQ) for complex computational modeling (National Research Council 2012), as described in Section 2.1.1. Additionally, our philosophy is informed by Bayesian methods. Bayesian calibration has mainly been applied to individual building model calibration (Muehleisen and Bergerson 2016). Although we do not use Bayesian inference for calibration, our team includes such expertise and it informs our approach to calibration and uncertainty quantification.

Finally, our calibration philosophy is grounded in the immediate needs of the 60 real-world use cases for EULPs defined by our TAG in 2019 (Mims Frick et al. 2019). The wide variety of use cases reinforced the fact that we were not just producing a one-time load shape dataset, but calibrating very broad-purpose models that would be asked in the future to evaluate a wide range of what-if scenarios about changes to the building stock involving energy efficiency, electrification, demand flexibility, and shifting weather patterns.

Our TAG includes individual subject matter experts and organizations involved in many of the major end-use metering studies (Larson et al. 2014, Navigant 2018, NEEA 2021b, NEEA 2021a), end-use surveys (KEMA-XENERGY 2004, Itron 2006, KEMA 2010, DNV GL 2021, EIA 2021f), and load profile modeling studies (Itron 2006, Baroiant et al. 2019). These individuals and organizations contributed their expertise throughout the three years of the project.

Together, these three sources formed the environment in which our calibration philosophy was developed. The philosophy can be summarized by these principles:

1. **Prioritize model input calibration;** output calibration is a last resort.
2. **Manually update inputs based on data,** as opposed to automatically tuning uncertain inputs to improve fit.
3. **Propagate input uncertainties** to quantify output uncertainty ranges.
4. **Calibrate/validate in multiple dimensions at once** to avoid overfitting.
5. **Report out validation results and uncertainty** so dataset and model users can decide if model accuracy and precision are sufficient for their use case.

The objective behind these principles is that we are striving to get the “why” right so we can develop accurate EULPs and ask questions about changes to the building stock (e.g., savings load shapes).

### 2.3.5 Obtaining and Processing Empirical Data

Identifying data available for use was a key effort in the project's first year, as documented in Mims Frick et al. (2019). We knew about or quickly found many types of useful data that were readily available, such as EIA and census data, and several of the residential end-use datasets. Working with our TAG and their networks, we were able to find and gain access to additional data sources for moderate levels of effort.

Two types of data ultimately warranted more extensive search efforts: utility AMI data and commercial end-use data. Utility AMI interval data were a known data need from the start of the project. Through our survey of available data sources in year one of the project, we identified commercial end-use data as the primary data gap that would need to be filled from to-be-determined sources (Mims Frick et al. 2019). The effort to gather these two types of

data is described in the following sections. Then we describe the work done to process the data to make it usable for calibration and validation.

### ***AMI Data Gathering Process***

Our data gathering process is documented more thoroughly in a paper published in the proceedings of the 2020 ACEEE Summer Study on Energy Efficiency in Buildings (Present et al. 2020). This effort included considerable outreach work, including outreach through our TAG, DOE, and through industry organizations such as National Rural Electric Cooperative Association (NRECA) and Bonneville Power Administration (BPA). Project partner Electric Power Research Institute (EPRI) also assisted the project team with utility data outreach. 2016 AMI data from the ComEd dataset was purchased through ComEd's publicly available Anonymous Data Service product (ComEd 2018). We ultimately gained access to eight large datasets of residential AMI data, including more than 2.3 million meters of timeseries electricity data, and ten large datasets of commercial AMI data, including around 60,000 buildings (served by more than 200,000 meters) of timeseries electricity data. These datasets varied considerably in size, data quality, and other factors, making some more useful to the project than others.

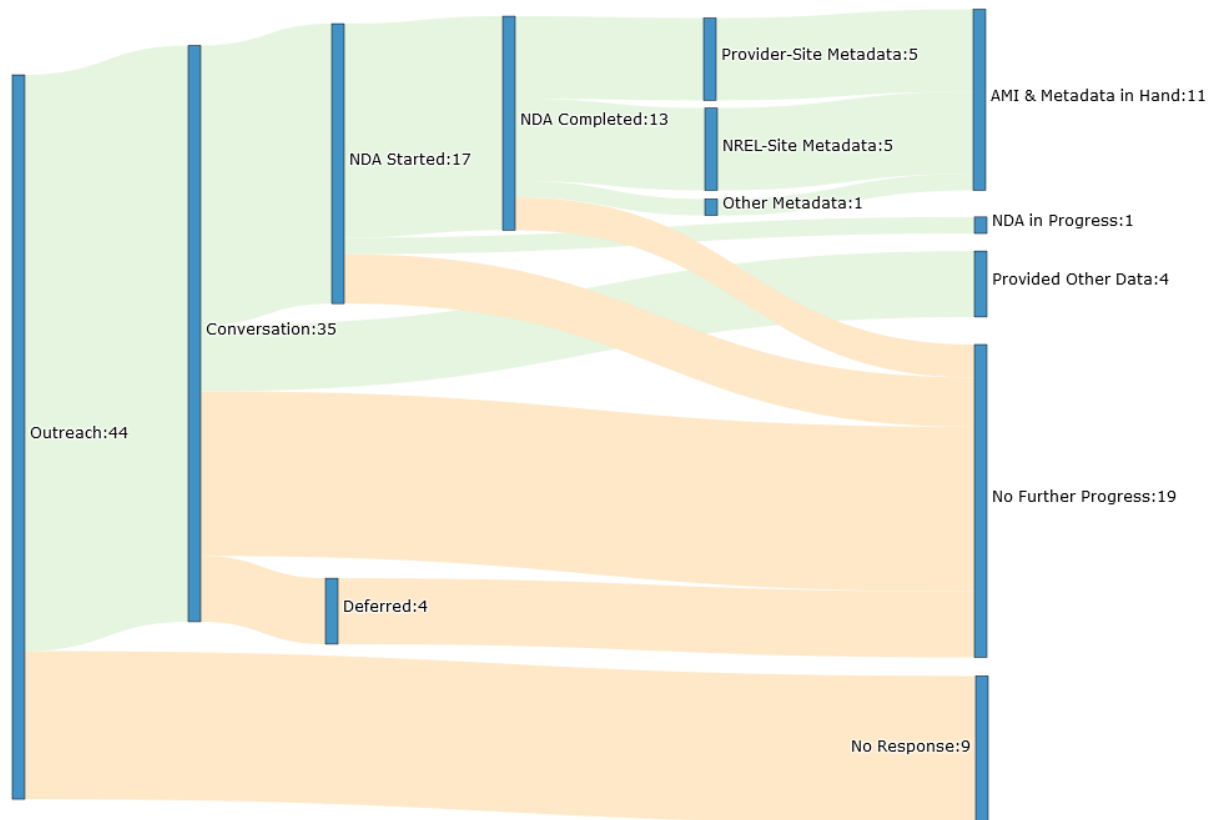
### **AMI Metadata**

Utility meter data are made more useful if they can be associated with information about the building or housing unit the meter is attached to, such as location, building type or use, floor area, etc. We refer to this information as *metadata*. For residential datasets, we worked with data at the dwelling unit level using whatever metadata the utility had available. We were most interested in having information on building type (e.g., single-family vs. multifamily) and location (ZIP or county, so that we could use the correct weather data). Most utilities were able to provide one or both of these, or in some cases the utility territory was small enough to be covered by just one weather station. We also requested any information available on heating fuel and vintage, which were much less commonly available.

For commercial datasets, the project necessitated a higher standard for metadata. This was in large part because the data came in at the meter level, whereas our project was working at the building level, which could include any number of meters. Most utilities had meters grouped by account, not by building, if they were grouped at all. An account could include meters at many buildings, and a building could include meters at many accounts. Additionally, because of the very large range of square footage in commercial buildings, for commercial calibration and validation we worked primarily with energy use intensity (EUI), that is, energy use per square foot of floor area, rather than building total consumption. And, there is a much larger variety of use types for commercial building than residential. These factors together meant that to use a commercial dataset we needed to be able to group meters into buildings, and to have reliable information on, at least, building type and square footage for those buildings.

### **AMI Data Outreach**

Figure 4 illustrates how the AMI data and metadata outreach conversations progressed. The time it took us to work through the specific data availability, data processing abilities, and privacy requirements for each utility greatly extended the time frame of getting some of the datasets. All told, for the ten commercial building AMI datasets that we were able to procure for this project, the mean time from initial outreach to the utility to having a set of data and appropriate metadata in hand was 405 days. The minimum was 62, the median was 391, and the maximum was 648. Many of the datasets required revision after the initial transfer for a variety of reasons, and that time is not counted in the numbers above. The level of usability we ended up with varied considerably. Some of these ten datasets were ultimately not of sufficient size or quality to be used for calibration; others, in particular those that had been used for previous analysis by other parties, were much more robust.



**Figure 4. Sankey diagram of commercial AMI data outreach**

### *Data Processing*

For each AMI dataset, we cleaned, filtered, and aggregated the data to a level that is suitable for comparison with the simulated results from the building stock models. The data processing procedure includes error detection, outlier removal, conversion to local standard time, data statistics extraction, and aggregation to hourly frequency and other desired data dimensions.

### **Residential AMI Data Processing**

For the residential calibration, for each dataset we calculated the timeseries of average hourly electricity consumption per housing unit and standard error at each timestamp. These timeseries may be broken out by building types and primary heating fuel types depending on the availability of these data dimensions. Because the AMI data are measured, they are subject to errors related to both AMI hardware and software, including measurement inaccuracy, and meter and data transmission outages. Additionally, AMI meters may be deployed incrementally and changed out periodically, resulting in a changing number of data points over time. For these reasons, the number of valid AMI readings at a given timestamp vary, and a standard error is calculated for each timestamp. Standard error is used to understand how the number of samples affects uncertainty of the mean hourly electricity consumption per housing unit.

### **Residential AMI Data Filtering**

To screen out errors and outliers on the ground truth data to the extent possible, we implemented data quality checks and data filtering algorithms. To start, we removed suspected measurement errors, including negative and zero readings. When available, we used the typical current rating of residential AMI meters to establish upper bounds for the data, and we subsequently removed the data outside of those bounds. For example, most AMI meters are rated for 200A residential service, which gives an upper bound of 48 kW at a 240V circuit. We experimented with filtering

using the common  $1.5 \times$  interquartile range (IQR) outlier removal method, but found the 48-kW bound based on the metering infrastructure to be more robust.

### **Commercial AMI Data Processing**

For the commercial data we computed the timeseries of average EUI per square foot of building floor area, broken out by building use type. Compared to the residential sector, commercial buildings have a greater variation in the makeup of building characteristics, occupant behavior, and the mode of building operation, all of which affect energy use profiles. Comparing the EUI by building type was our attempt to normalize the comparisons for two of the key sources of difference: building type and building area. However, to achieve this, the commercial AMI data processing included additional steps to fulfill the need for more robust metadata than most utilities had available.

The team licensed a database of commercial real estate data from the CoStar Group (referred to as CoStar data henceforth) and joined it to meter data from utilities by developing an address matching script. This address matching script by default assigned all meters with the same service address to the building at that address, along with the properties of that building from the CoStar data (building type, square footage, vintage, and number of stories). This approach left us with better metadata than the utilities were generally able to provide, but was not without its limitations. For example, the CoStar data do not include government-built, owned, and operated buildings as well as certain other omissions. Additionally, utility addresses and CoStar addresses were not always similar enough for our script to handle, even when using approximate string matching techniques.

The team made adjustments to this approach depending on the unique aggregation and anonymization rules of each utility. For several utilities, we added an additional step of assigning each building to a metadata “bin,” which included a small range of building sizes and vintages within one building type. We developed an auto-binning script to do this task subject to a minimum number of buildings per bin set by the utility. We were then sent individual building timeseries energy consumption along with which bin the building was in, but did not have the building’s unique metadata. Other utilities went a step further and provided only the number of buildings, total energy consumption at each timestamp, and total square footage in each bin. These all worked for our purposes if the starting dataset was very clean, but made filtering questionable data much more challenging. In some cases we had to discard entire bins of data, or spend extensive time working with the utility to ensure the best data quality possible.

### **Commercial AMI Data Filtering**

To filter the commercial AMI data, we first generated a summary statistic for each building or building bin in the dataset. We then removed buildings or bins based on the following criteria: (1) the annual EUI must be greater than the threshold value for their building type (see Table 4), (2) the data must have at least 2,000 non-zero timestamps, (3) the building or bin must be at least 500 ft<sup>2</sup> in floor area per building, and (4) the annual EUI must be within the 3x median EUI bounds for their building type, which are constructed based on the buildings meeting the first three criteria. These data filtering criteria were developed based on expert judgment, including input from the TAG, and a balance between removing extremely problematic data points and retaining a large enough sample size for the ground truth data to remain useful. They were validated using the Xcel Energy dataset. Because addresses were available for this dataset, specific outliers could be spot-checked. A forthcoming publication will detail this methodology (DeWitt et al. 2022). The annual building EUI used in the filtering process is adjusted by assuming an average value for all timestamps that have an invalid reading. This is calculated by taking the hourly average EUI (based on non-zero readings) to the number of hours in the data year.

For each of the figures in Section 4 where AMI is compared to ComStock, two dashed lines are shown to represent the 80% confidence interval of the AMI mean at each timestep.

### **Impact of the COVID-19 Pandemic on Commercial AMI Data**

Early in the project, the plan was to send a team member to each of the data-providing utilities to work directly with utility staff to perform the address matching between electric meter service addresses and the CoStar data. This plan was designed to enable our team to identify and address any unique challenges with each utility’s data system and to spot check results to ensure that the matching was as complete as possible within the utility’s anonymization and aggregation requirements. Unfortunately, just as this process was getting underway, the COVID-19 pandemic began. Utility staff were under tremendous pressure just to keep operations running smoothly, let alone work with researchers, and NREL travel was not allowed.

**Table 4. Minimum Annual EUI Threshold by Building Type, Set to Match the 5th Percentile of 2012 CBECS Electricity Consumption**

<b>Building Type</b>	<b>Minimum Annual EUI Threshold (kWh/ft<sup>2</sup>)</b>
Warehouse	1.0
Retail	1.6
Strip mall	1.6
Other	2.0
School (primary, secondary)	2.4
Office (small, medium, large)	2.5
Outpatient	2.7
Hotel (small, large)	2.8
Restaurant (full service, quick service)	6.5
Hospital	11.4

The preferred approach to working within these constraints was for the utilities to send only the addresses to NREL so that NREL could perform the address matching to assign metadata to group meters by building and assign metadata to those buildings. In many cases, NREL was required to delete addresses before receiving the AMI data to ensure anonymization, retaining only anonymized IDs matched with building type, approximate size, and approximate vintage. In these cases, the team did not have the ability to troubleshoot and verify questionable EUIs. In cases where NREL was able to have both the addresses and energy consumption for a length of time, greater error checking was possible.

Unfortunately, multiple utilities were unable to send even the service addresses alone due to privacy concerns or other policies. The second best approach then to working within these constraints was for NREL to send the address matching code to the utilities, and then to remotely walk utility staff through the process of running the code. Although the team worked through several unique challenges, there were limitations to the amount and type of verification that could be done due to utility staff availability or cybersecurity constraints.

The overall impact of COVID-19 on commercial AMI was significant. It caused delays in receiving data, which necessitated pauses in the calibration work. Mistakes or miscommunication that likely would not have happened while working in-person caused further delays because identifying, diagnosing, and fixing the problems all took time. Sometimes, issues were not identified until after the data had been used to inform input calibration, which then had to be corrected.

The biggest impact of all was that meter-to-building mappings could not be evaluated as thoroughly as planned, leading to uncertainty in the quality of the AMI data. If address matching for a building was extremely bad (missing 9 of 10 meters, for example), the outlier detection process would likely identify and remove that building. However, if address matching for a building was slightly bad (missing 2 of 10 meters, for example), the EUI for that building would appear slightly lower than reality, but not low enough to be caught by outlier detection techniques. This uncertainty necessitated comparison of the AMI data against other established data sources, and ultimately required a re-examination of how the AMI would be compared to the ComStock results.

#### *Commercial End-Use Data Gathering Process*

The commercial calibration effort required substantial end-use timeseries data to inform behavior-driven inputs to the models such as schedules, setpoints, and occupancy, which posed its own set of challenges. The limited availability of commercial end-use timeseries data was identified as an area of focus early in the project which sparked an extensive outreach effort by our team to identify and obtain data of this type. With a three-year project and the calibration effort spanning years two and three, there was not sufficient time to embark on new end-use metering studies, which are multiyear efforts on their own. Instead, we identified a range of unconventional sources of commercial end-use data that already existed, most of which was not collected for the purpose of studying end-use load shapes (e.g., building automation system data). Before embarking on this effort, we determined minimum sample sizes that would be required for commercial end uses and building types to achieve a desired level of precision: 20% precision with 80% confidence for heating and cooling end uses, and 30% precision with 80% confidence for other end uses. These minimum sample sizes were vetted by a group of end-use metering subject matter experts from our TAG and can be found in Table 5 for weather-driven end uses and Table 6 for schedule-driven end uses.



### Commercial End-Use Data Outreach

With input from our TAG, NEEP, the Massachusetts Energy Efficiency Advisory Council, our national laboratory colleagues, and outreach at conferences such as DOE's Better Buildings Summit and the International Energy Program Evaluation Conference, we contacted 63 varied organizations on the topic of commercial end-use timeseries data, of which we received responses from 50. Many organizations did not have the type of data we were looking for. Many organizations did have such data, but similar to the commercial AMI data, there were often challenges related to obtaining adequate metadata. Figure 5 illustrates how the commercial end-use data outreach conversations progressed. Data quality was often an issue, along with the non-trivial time requirements for companies to gather, organize, and transfer the data in a useful format. Ultimately, we received data from 10 of these organizations, primarily BAS and submetered end use data, often at a cost to compensate organizations for their time gathering, organizing, and transferring the data (Present et al. 2020). We achieved the previously determined minimum sample sizes in almost all of the categories, and greatly surpassed the minimums in many categories, as shown in Table 5 for weather-driven end uses and Table 6 for schedule-driven end uses. There were still issues that needed to be considered in these datasets, such as bias and data gaps, but overall the purchased datasets were critical for the commercial calibration effort. We summarize the procured datasets in Table 5. Many of the datasets were provided on the condition of being anonymous, so we do not list the sources by name.

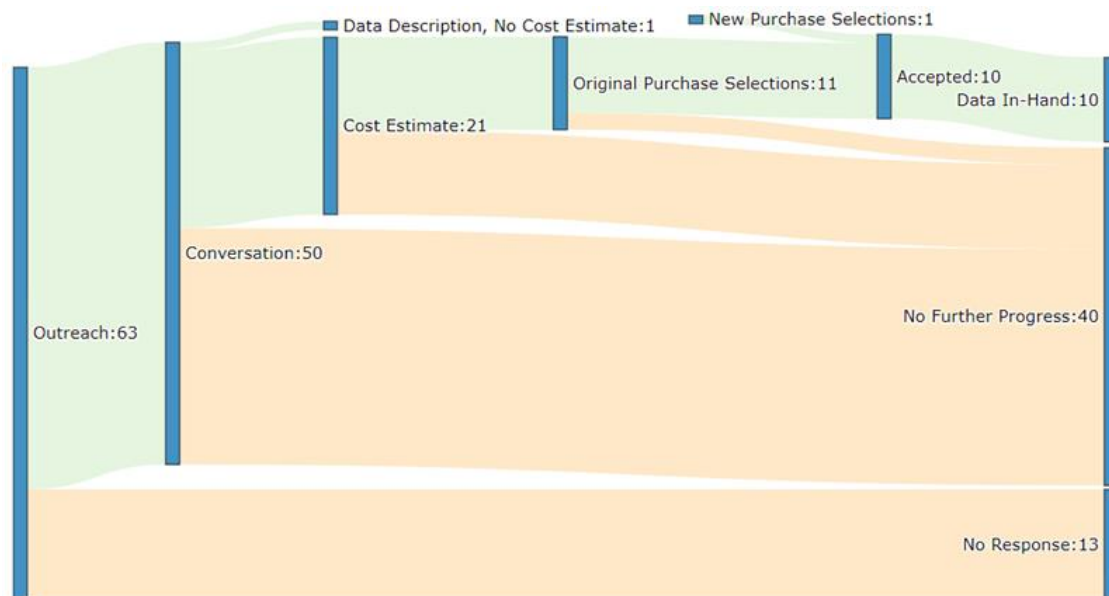


Figure 5. Sankey diagram of commercial end-use data outreach



**Table 5. Procured Commercial End Use Datasets**

<b>Company</b>	<b>Summary</b>
Anonymous 1	Submetered timeseries energy data by end use and BAS data for thousands of commercial buildings including restaurants, retail, offices, hotels, and schools.
Anonymous 2	Submetered timeseries BAS data for thousands of commercial buildings including restaurants, retail, offices, hotels, healthcare, warehouses and schools.
Anonymous 3	BAS data for multifamily building common areas.
Anonymous 4	Submetered timeseries energy data by end use for several buildings including warehouses, offices, a grocery store, a hospital, a hotel, restaurants, and retail.
Anonymous 5	Submetered BAS timeseries data and energy data for several grocery stores and hospitals.
Anonymous 6	Submetered BAS timeseries data for several multifamily buildings.
Anonymous 7	Submetered timeseries energy data and BAS data for several schools and some offices.
Anonymous 8	Submetered timeseries energy data by end use and BAS data for an office building.
Anonymous 9	Plug load survey and timeseries energy data for several schools and some administrative buildings.
Anonymous 10	Submetered plug load energy use data for a hospital.

**Table 6. Commercial End-Use Data Outreach Results for Weather-Driven End Uses**

<b>End Use</b>	<b>Proposed Minimum Sample Size<sup>1</sup></b>	<b>Pursued Sample Size<sup>2</sup></b>	<b>Procured Sample Size<sup>3</sup></b>
Heating	48	6,218	4,119
Cooling	48	6,598	3,957
Fans	21	2,497	645
Pumps	21	500	95
Heat Rejection	21	21	51
Humidification	21	27	53
Heat Recovery	21	22	49
Refrigeration	21	1,076	1,080
Exterior Lighting	21	846	627

<sup>1</sup> Proposed minimum sample size targets for successful project completion, presented at a subject matter expert webinar in August 2019.

<sup>2</sup> Counts include in-hand data, data we expected to be able to get for free from other research institutions, and data to which we intended to purchase access—the potential purchase sample sizes were based on vendor rough estimates obtained during market outreach.

<sup>3</sup> Procured sample size includes data in hand and data that was contracted for procurement.

**Table 7. Commercial End-Use Data Outreach Results for Schedule-Driven End Uses**

	Proposed Minimum Sample Size <sup>1</sup>				Pursued Sample Size <sup>2</sup>				Procured Sample Size <sup>3</sup>			
	Interior Light-ing	Plug and Pro-cess	Service Water Heat-ing	Cook-ing	Interior Light-ing	Plug and Pro-cess	Service Water Heat-ing	Cook-ing	Interior Light-ing	Plug and Pro-cess	Service Water Heat-ing	Cook-ing
<b>Hospital Outpatient</b>	21	21	0	n/a	103	2 <sup>4</sup>	0	0	76	4 <sup>4</sup>	0	0
<b>Primary School</b>	21	21	0	n/a	281	285	0	2	178	332	6	9
<b>Secondary School</b>												
<b>Full-Service Restaurant</b>	21	21	0	n/a	760	196	316	2618	159	56	58	687
<b>Fast-Service Restaurant</b>												
<b>Retail Strip Mall</b>	21	21	0	n/a	1046	214	106	0	908	30	99	18
<b>Small Hotel</b>	21	21	0	n/a	53	5 <sup>4</sup>	0	1	42	2 <sup>4</sup>	1	0
<b>Large Hotel</b>												
<b>Warehouse</b>	21	21	0	n/a	270	25	0	0	132	36	15	1
<b>Sm. Office</b>	21	21	0	n/a	337	270	1	0	133	91	10	8
<b>Med. Office</b>												
<b>Large Office</b>												

<sup>1</sup> Proposed minimum sample size targets for successful project completion, presented at a subject matter expert webinar in August 2019. No goals were set for the cooking end use or the multifamily building type.

<sup>2</sup> Counts include in-hand data, data we expected to be able to get for free from other research institutions, and data to which we intended to purchase access—the potential purchase sample sizes were based on vendor rough estimates obtained during market outreach.

<sup>3</sup> Procured sample size includes data in hand and data that was contracted for procurement.

<sup>4</sup> Procured data sample size is less than the proposed minimum sample size.

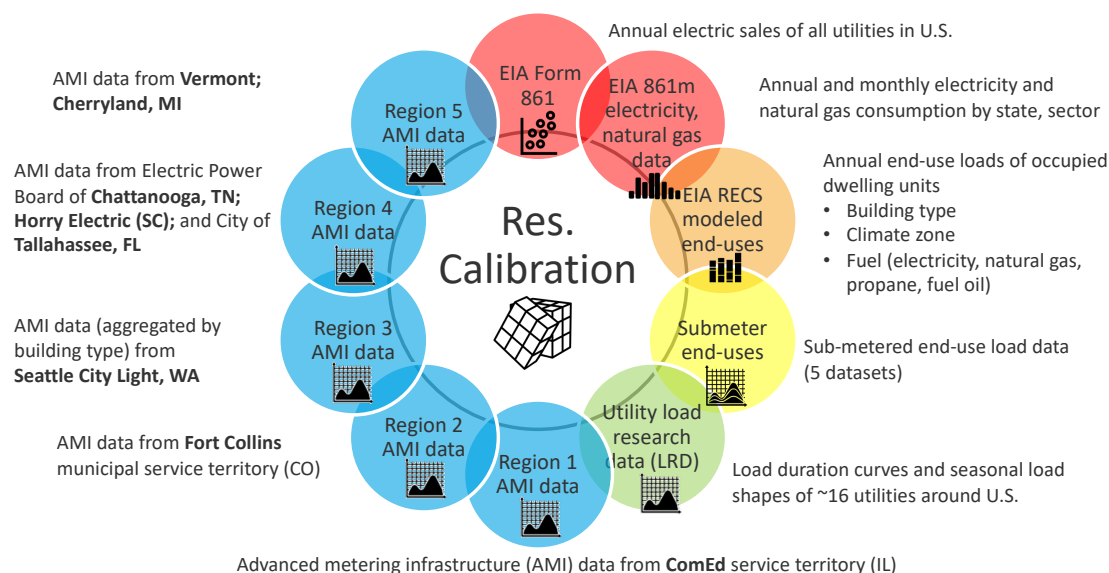
### 2.3.6 Empirical Data Used for Residential Calibration and Validation

As with any validation effort, the empirical data used to evaluate model outputs is the most important component. Here we describe the various datasets we used for calibration and validation for the residential building stock model. Section 2.3.5 describes the data collection and processing effort.

Table 8 introduces the major *types* of data we used for calibration and the temporal, geographic, and end-use resolution of each, as well as whether they can provide insights on specific building types (i.e., metadata) and on the diversity of load. Figure 6 summarizes the 10 main data sources used for residential calibration. Table 9 lists the full set of datasets used for residential calibration and validation and the relevant characteristics of each.

**Table 8. Calibration data types and typical resolution and characteristics of each (residential and commercial). Each type of calibration data covers different dimensions; no one data source provides information at all dimensions and scales.**

Type of Data	Temporal Resolution	Geographic Resolution	Sectoral/End-Use Resolution	Meta-data	Load Diversity
EIA 861 electricity sales	Annual	All utilities	Whole sector	No	No
EIA 861M electricity sales	Monthly	All states	Whole sector	No	No
Load research	Hourly	Several regions	Whole sector	No	No
AMI interval	Hourly	Several regions	Whole building/meter	Yes	Yes
End-use metering	Hourly	Several regions	End uses/circuits	Yes	Yes



**Figure 6. Residential calibration used 10 main dimensions or data sources; like a 10-sided Rubik's cube, changes made to improve the fit in one dimension needed to be evaluated across the other nine dimensions**

**Table 9. Description of Residential Calibration Datasets**

Type	Name	Location	Description	Metadata	Main Comparison Type
Survey	Form EIA-861	Nationally available by utility	Annual electric sales of all utilities in the United States	N/A	Scatterplot
Survey	Form EIA-861M	Nationally available by state	Monthly electricity sales consumption by state and sector	N/A	Line plots for each state
Survey	Form EIA-176	Nationally available by state	Monthly natural gas consumption by state and sector	N/A	Line plots for each state
Survey	EIA RECS 2009, 2015	National coverage	Annual electricity and fuel use for surveyed homes; modeled end-use breakouts	Surveyed housing characteristics	End use scatterplots
End use	NEEA 2011 RBSA Metering	WA, OR, ID, MT	Circuit-level submetering study	Surveyed housing characteristics	End-use transferability, Region 3 end-use stacked area plots
End use	NEEA 2019 HEMS Metering	WA, OR, ID, MT	Circuit-level submetering study	Surveyed housing characteristics	End-use transferability, Region 3 end-use stacked area plots
End use	Mass RES 1 Load Shape Study	Massachusetts	Aggregate load shapes by end use only	N/A	End-use transferability
End use	Pecan St. Dataport	Austin, TX (most samples)	Circuit-level submetering study	Some housing characteristics	End-use transferability
End use	FSEC PDR Metering	Florida	Circuit-level submetering study	Some housing characteristics	End-use transferability
LRD	Load research data (LRD)	16 utilities (see map)	Publicly available load research data from various utilities	Customer class (varies)	Day plots by season
AMI	ComEd	Northern Illinois including Chicago	Residential AMI interval data for most of ComEd territory; supplemented with ComEd LRD (Region 1)	Customer class	Day plots by season; QOIs
AMI	Fort Collins Utilities	Fort Collins, CO	Residential AMI interval data for Fort Collins Utilities (Region 2)	Single-family vs. multifamily	Day plots by season; QOIs
AMI	Seattle City Light	Seattle, WA	Aggregate residential AMI data by building type (Region 3)	Single-family vs. multifamily	Day plots by season; QOIs
AMI	EPB	Chattanooga, TN	Residential AMI interval data (Region 4)	Single-family vs. multifamily	Day plots by season; QOIs
AMI	City of Tallahassee Utilities	Tallahassee, FL	Residential AMI interval data (Region 4)	Single-family vs. multifamily	Day plots by season; QOIs
AMI	Horry Electric Cooperative	Horry County, SC, including Myrtle Beach	Residential AMI interval data (Region 4)	Single-family vs. multifamily	Day plots by season; QOIs
AMI	VEIC/Green Mountain Power	Vermont	Residential AMI interval data (Region 5)	Single-family vs. multifamily	Day plots by season; QOIs
AMI	Cherryland Electric Cooperative	Michigan	Residential AMI interval data (Region 5)	Single-family vs. multifamily	Day plots by season; QOIs

### *Residential EIA Survey Data*

#### **Form EIA-861—Annual Electricity Sales**

The EIA collects mandatory reporting data from all electric utilities (approximately 3,300) on an annual basis, for its Annual Electric Power Industry Report (EIA 2021b). About two-thirds of utilities are required to report annual electricity sales and number of customers by sector (residential, commercial, and industrial), via Form EIA-861. About one-third of utilities complete the short Form EIA-861S, which does not require breaking out sales by sector. EIA estimates the customer sector breakdown for these utilities. The Form EIA-861 data provide the best available data on annual electricity use by utility and sector. We also used Form EIA-861 data to map the buildings we simulate to utility companies, based on the service territory county lists collected via EIA-861 (accounting for counties served by multiple utilities). Because ResStock represents the entire residential sector, we were able to directly compare ResStock annual electricity use estimates against EIA-861 data by utility, in the form of scatterplots. Section 4, Results, includes these scatterplot figures from before and after our input calibration effort.

#### **Form EIA-861M—Monthly Electricity Sales**

Form EIA-861M is similar to Form EIA-861 in that it collects data on electricity sales by sector, but the data are collected on a monthly basis instead of annual, but only for a “statistically chosen sample of electric utilities”. EIA imputes monthly sales, revenue, and customer counts for the utilities that are not surveyed, which are added as a state-level “adjustment”, to develop aggregated state-level estimates of monthly electricity sales, revenue, and customer counts by sector (EIA 2021c). We compare these empirical data on monthly electricity sales data by state to outputs from ResStock (see Section 4, Results, and Appendix F). We plot the adjustment in a different color to illustrate the fact that there is some uncertainty in these sales data.

#### **Form EIA-176—Monthly Natural Gas Sales**

Form EIA-176 (EIA 2021e) collects data on natural gas sales by sector. Similar to the EIA-861M comparisons for electricity, we compare the empirical residential sector annual and monthly natural gas sales data by state to outputs from ResStock (see Appendix F).

#### **EIA Residential Energy Consumption Survey**

RECS is a survey that EIA administers to a nationally representative sample of housing units. RECS collects energy-related information on the U.S. residential building stock, including housing unit characteristics, household demographics, and usage patterns, as well as metered energy usage for the surveyed housing units obtained directly from utility companies (EIA 2021f). RECS includes modeled disaggregation of the energy usage into constituent end uses.

RECS microdata have been used extensively for deriving ResStock model inputs including probability distributions of HVAC system types, setpoints and setback behavior, water heater characteristics, presence of appliances, multi-family building characteristics, and plug loads. 2015 RECS surveyed 5,600 households, which limits the geographic resolution with which it can be used, so we have primarily used 2009 RECS (12,083 households) for deriving these probability distributions for ResStock, although 2015 RECS is used for some fields (see Table 2). 2020 RECS data (target of 18,000 households) are expected to become available in 2022.

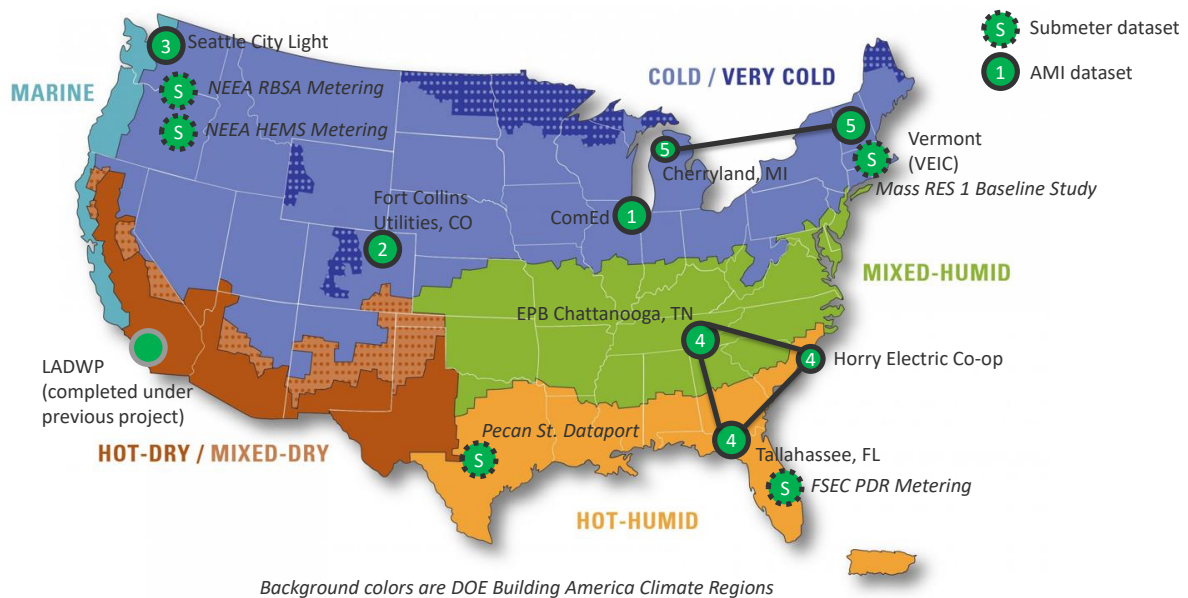
We previously used annual energy consumption data from 2009 RECS to calibrate and validate an earlier version of ResStock for single-family detached homes at an annual level (Wilson et al. 2017). For EULP validation, we primarily use the modeled end-use breakouts from both 2009 and 2015 RECS as comparison points for annual end use outputs from ResStock, typically aggregated by building type and/or region.

#### **Residential Utility AMI Datasets**

Figure 7 shows a map of the utility AMI datasets used for residential calibration. Each AMI dataset is briefly described below.

#### **ComEd**

2016 AMI data from the ComEd dataset was purchased through ComEd’s Anonymous Data Service product (ComEd 2018) for residential Region 1 calibration. ComEd delivers electricity to about 3.6 million residential customers in Northern Illinois, including the city of Chicago. In 2016 and 2017, AMI meters were not fully rolled out across ComEd territory, so we used the AMI data in conjunction with 2012, 2016, and 2018 load research data



**Figure 7.** This map shows the locations covered by the utility AMI and submetering datasets used for the residential calibration and validation effort. The numbered markers show how the eight AMI datasets were grouped into five regional phases. The LADWP marker indicates that calibration was previously completed for the Los Angeles Department of Water and Power. The “S” markers show the locations of the four primary end-use submetering datasets that were used for residential calibration and validation.

(LRD) available from ComEd. Both the AMI and LRD datasets tag each meter with the customer class, which for residential customers tells us whether the customer is in a single-family or multifamily building, as well as whether the customer uses electricity as the primary heating fuel. The ComEd AMI data also includes the ZIP code and ZIP+4 code for each customer.<sup>5</sup> The ComEd AMI data were unique in that it was the only AMI dataset that is publicly available for purchase (ComEd 2018). One significant limitation of the ComEd AMI data is that the anonymous meter IDs are reset for every month of data, making it impossible to understand how individual customers use electricity from month-to-month (Elevate 2017).

Based on data in ResStock from PUMS (Ruggles et al. 2021), about 66% of ComEd residential customers are in the single-family non-electric heating class, and 17% are in the multifamily non-electric heating class. The remaining 17% use electric heat and are about equally split between single-family and multifamily customers.

### Fort Collins Utilities

2018 data from the Fort Collins dataset, which covers the city of Fort Collins, CO, was used for residential Region 2 calibration. The municipal utility serves about 65,000 residential customers. We used building type metadata from the Larimer County tax assessor database to disaggregate customers into single-family and multifamily cohorts.

Based on data in ResStock from PUMS (Ruggles et al. 2021), about 66% of residential customers are in single-family detached homes, with the remainder split across single-family attached (i.e., townhomes and duplexes; 8%), multifamily 2–4 units (7%), and multifamily 5+ units (20%). About 70% of residential customers use natural gas for space heating, with 24% using electricity and the remainder using propane or other sources.

### Seattle City Light

2019 AMI data from the Seattle City Light dataset was used for residential Region 3 calibration. The municipal utility serves about 411,000 residential customers in Seattle, WA, and parts of adjacent suburbs. NREL used

<sup>5</sup>Subject to the 15/15 anonymization protocol: “Each dataset must contain at least 15 customers, and not one of the 15 can have 15% or more of the usage within that dataset.”



anonymized aggregate load profiles by residential building type (single-family, multifamily, etc.) that Seattle City Light facilitated based on an 8% sample of residential customers.

Compared to previous regional AMI datasets, Seattle had a higher percentage of multifamily customers (48%) and a higher percentage of customers using electricity for space heating (55%), based on data queried for ResStock from PUMS (Ruggles et al. 2021).

### **EPB**

2019 AMI data from the EPB dataset was used for residential Region 4 calibration. The municipal utility, formerly known as the Electric Power Board of Chattanooga, serves about 165,000 residential customers in Tennessee and Georgia.

Based on data in ResStock from PUMS (Ruggles et al. 2021), about 71% of residential customers are in single-family detached homes, with the remainder split across single-family attached (i.e., townhomes and duplexes; 2%), multifamily 2–4 units (7%), multifamily 5+ units (10%), and mobile homes (10%). About 70% of residential customers use electricity for space heating, with 22% using natural gas, and the remainder using propane (6%) or other sources.

### **Tallahassee**

2019 AMI data from the Tallahassee dataset was used for residential Region 4 calibration. The municipal utility serves around 102,000 residential customers in Tallahassee, FL.

Based on data in ResStock from PUMS (Ruggles et al. 2021), about 53% of residential customers are in single-family detached homes, with the remainder split across single-family attached (8%), multifamily 2–4 units (10%), multifamily 5+ units (23%), and mobile homes (7%). About 89% of residential customers use electricity for space heating, with 8% using natural gas, and the remainder using propane (3%) or other sources.

### **Horry Electric Cooperative**

2018 AMI data from the Horry Electric Cooperative dataset was used for residential Region 4 calibration. The cooperative serves around 69,000 residential members covering most of Horry County, SC, including the city of Myrtle Beach. A large fraction of customers are located near the coast.

Based on data in ResStock from PUMS (Ruggles et al. 2021), about 57% of residential members are in single-family detached homes, with the remainder split across single-family attached (5%), multifamily 2–4 units (5%), multifamily 5+ units (18%), and mobile homes (15%). Compared to other dataset locations, Horry County has a much higher fraction of housing units that are unoccupied vacation or vacant homes (66% of units in multifamily 5+ unit buildings). About 95% of residential customers use electricity for space heating, with 3% using natural gas, and the remainder using propane (2%) or other sources.

### **Vermont Energy Investment Corporation**

2018 AMI data from the Vermont Energy Investment Corporation (VEIC) dataset was used for residential Region 5 calibration. VEIC operates Efficiency Vermont, the statewide energy efficiency program for Vermont. The dataset includes AMI data from Green Mountain Power, which serves about 222,000 residential customers across most of the state of Vermont.

Based on data in ResStock from PUMS (Ruggles et al. 2021), about 65% of residential customers are in single-family detached homes, with the remainder split across single-family attached (3%), multifamily 2–4 units (14%), multifamily 5+ units (10%), and mobile homes (8%). Only 6% of residential customers use electricity for primary space heating, with 43% using fuel oil, 16% using natural gas, 16% using propane, and 18% using wood or other sources. Roughly 20% of housing units are vacant, vacation, or seasonally occupied. The VEIC data are overall much more rural than previous dataset locations.



### **Cherryland Electric Cooperative**

2019 AMI data from the Cherryland dataset was used for residential Region 5 calibration. Cherryland Electric Cooperative serves around 33,000 residential customers across six counties in the northwest region of Michigan's lower peninsula.

Based on data in ResStock from PUMS (Ruggles et al. 2021), about 78% of residential customers are in single-family detached homes, with the remainder split across single-family attached (2%), multifamily 2–4 units (4%), multifamily 5+ units (7%), and mobile homes (8%). About 12% of residential customers use electricity for primary space heating, with 56% using natural gas, 20% using propane, and the remainder using wood or other sources. Roughly 30% of housing units are vacant, vacation, or seasonally occupied. Like VEIC, the Cherryland data are overall much more rural than previous dataset locations.

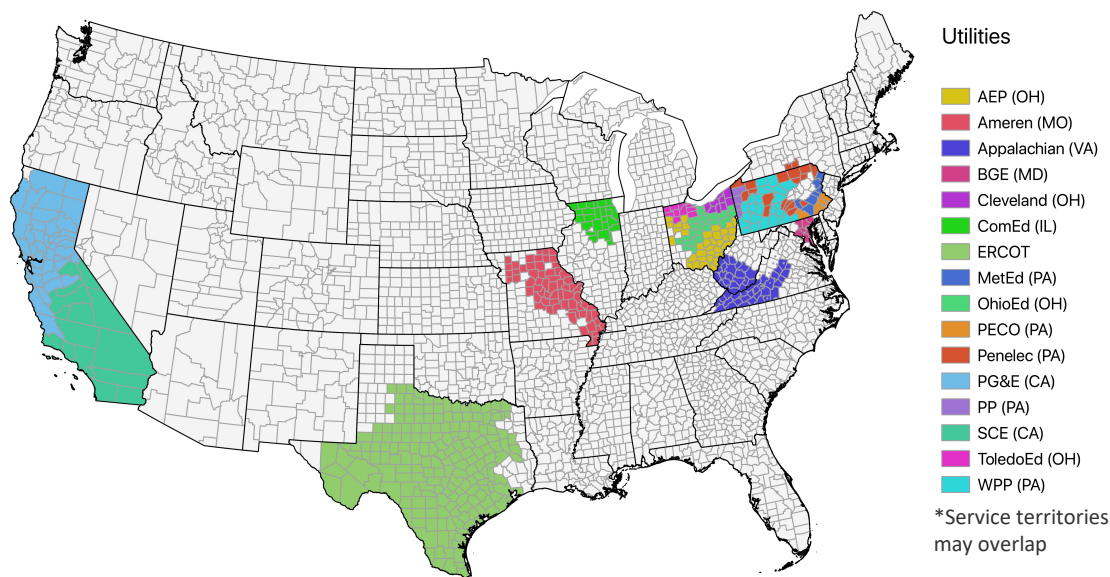
### **Los Angeles Department of Water and Power**

Under a previous project, we had access to AMI data for a subset of LADWP residential customers, load research data representing all residential customers (2016 and 2017), and monthly billing data from most customers (Hale et al. 2021). LADWP serves around 1.3 million residential customers primarily within the city of Los Angeles, CA. The dataset was joined with Los Angeles County Assessor data to obtain and building type metadata. This was used for a calibration effort specific to Los Angeles that was incorporated with the national ResStock model as part of this project.

#### *Utility Load Research Data*

Utility LRD is similar to AMI interval data in that it is hourly load data, but instead of being hourly load data for individual customers or meters, it is an aggregate load shape for an entire customer class (e.g., residential). LRD is typically developed by aggregating interval meters from a representative sample of customers, although in some cases it may be modeled data (e.g., from regression-based load forecasting models). LRD may be called different names by different entities, such as customer class load profiles, dynamic load profiles, or “Backcasted (Actual) Load Profiles—Historical” (ERCOT 2021). Utilities or other entities may make the LRD publicly available on their website, particularly in states with deregulated electricity where retail energy suppliers operate. In other cases, LRD can be obtained by requesting it from the utility. Because LRD is only a single aggregate profile for each customer class, there are not typically privacy concerns that prevent them from being shared. Some LRD is a customer average profile (e.g., with units of “kW per customer”) while other LRD is a customer aggregate profile (e.g., with units of “MW”), so we needed to convert some of the datasets. There was some uncertainty in numbers of customers served by Penn Power and Western Penn Power (subsidiaries of FirstEnergy), which we believe caused a greater discrepancy in LRD comparisons for those two utilities.

Early in the project we had access to a collection of LRD that NREL had obtained in past years, and 2012 was the year for which we had the most LRD. Later in the project, we updated our LRD comparisons with a new collection of LRD for 2018, which was the primary year we used for AMI comparisons. This resulted in a slightly smaller set of utilities, but allowed for comparison of a more recent year. Figure 8 shows a map of the 2018 utility LRD used for residential validation.



**Figure 8.** This map shows the publicly available utility load research datasets used for the residential calibration and validation effort. The utility territories are mapped by counties served based on EIA Form 861 data (EIA 2018c).

#### *Residential End Use Datasets*

The map in Figure 7 shows the locations of the five residential submetering datasets used for residential calibration. Each dataset is briefly described below.

#### **NEEA 2011 Residential Building Stock Assessment Metering (“RBSAM”)**

NEEA’s 2011 RBSA Metering study (Larson et al. 2014) was part of a larger 2011 Residential Building Stock Assessment study of 1,400 homes in the Pacific Northwest (NEEA 2021c). End-use data were collected for 101 homes from April 2012 to July 2014. The sample was intended to be representative of the housing stock across the region (Washington, Oregon, Idaho, and western Montana). Lighting and miscellaneous plug loads were not disaggregated because they are often found on the same circuit and would require many plug-level meters to fully separate. However, lighting usage was measured in a subset of the homes. Building characteristics and energy audit data are available for all of the 101 homes because they were part of the larger RBSA study. The RBSAM dataset is publicly available at NEEA (2021).

#### **NEEA 2019 HEMS Metering (“HEMS”)**

NEEA’s Home Energy Monitoring Study (HEMS) is an ongoing end-use load research study that is similar to the 2011 RBSA Metering study, but is expected to eventually have a much larger sample size of 400 homes (NEEA 2021b). At the time of this report, one full year of data (2019) was available for a sample of 60 homes. Like RBSAM, the sample was intended to be representative of the housing stock across the region (Washington, Oregon, Idaho, and western Montana), although there was intentional oversampling of homes heated with electric heat pumps. Lighting and miscellaneous plug loads were not to be disaggregated. Building characteristics and energy audit data are available for all of the homes. The HEMS dataset is publicly available at NEEA 2021b.

#### **Massachusetts RES 1 Load Shape Study (“Mass RES 1”)**

The RES 1 Baseline Load Shape Study was prepared by Navigant for the Electric and Gas Program Administrators of Massachusetts (Navigant 2018). On-site metering of 25 end uses was conducted at 356 sites across Massachusetts, between May 2017 and April 2018. Like the NEEA studies, the sampling approach was intended to be representative. One of our TAG members was able to provide the project with quantitative aggregate end-use load shapes. Of the end-use datasets we had access to, this was the only one where we did not have data from individual homes, only

aggregates. This does not make it any less useful for studying transferability, but does mean that we could not use it to understand the diversity of behavior across homes. The Mass RES 1 dataset is not publicly available, but load shapes can be found in the report (Navigant 2018).

#### **Pecan Street Dataport (“Pecan Street”)**

The Pecan Street Dataport dataset that we used included end-use submetered data from 998 homes collected between 2011 and 2014. Most of the homes were located in Austin, TX, although some were in other locations. Unlike the NEEA and Massachusetts datasets, Pecan Street is not intended to have samples that are representative of a certain location; many of the samples are newer homes concentrated in the Mueller neighborhood in Austin. Access to Pecan Street Dataport is available for purchase (with free and discounted licenses available for university researchers) (Pecan Street Inc. 2021).

#### **FSEC PDR Metering (“FSEC”)**

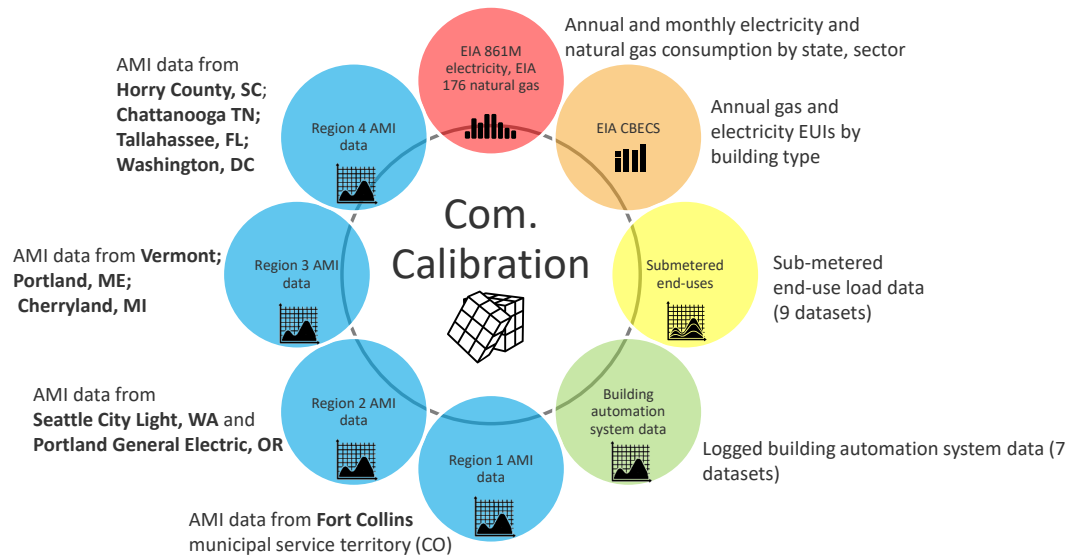
These data were collected by the Florida Solar Energy Center (FSEC) from 2012 to 2016 in partnership with Florida Power and Light, an investor-owned utility in Florida, with funding from the Building America program through DOE’s Building Technologies Office (Parker et al. 2016; Fenaughty, Parker, and Martin 2017). The data were collected in 56 homes that were recruited for a phased deep retrofit program. The sample was not intended to be representative, and may be biased towards households interested in participating in a deep energy retrofit program. We only use the pre-retrofit data. The data are publicly available in a dashboard (Building America Partnership for Improved Residential Construction 2021).

#### **ecobee Donate Your Data (“ecobee”)**

ecobee is a manufacturer of smart thermostats. NREL is a research partner with ecobee’s Donate Your Data program, where thermostat owners can opt to share their anonymized 5-minute resolution thermostat data with researchers. Although ecobee data provide a variety of useful data on occupants’ setpoint preferences, setback behavior, and home/away status, the data cannot be directly used to derive EULPs. This is because the thermostats record run time of equipment but do not measure power draw. Even if one had access to the rated power draw (or capacity and efficiency) of the equipment, the actual power draw of single-speed air conditioners varies with outdoor conditions. However, although run time data cannot be used to derive the magnitude of peak demand or shape of the cooling end use, it can provide insights into the timing of peak demand (Meier et al. 2019). Use of ecobee for EULP is also limited because it is not a representative sample of cooling demand; because it is a smart thermostat with motion sensing, setpoint control patterns likely differ from the population at large. Meier et al. (2019) provide an excellent summary of what smart thermostat data can and cannot be used for. For this project, we primarily used the ecobee data as a comparison point with the self-reported thermostat setpoint data in RECS, and thus do not include it in our list of end-use datasets.

### 2.3.7 Empirical Data Used for Commercial Calibration and Validation

Here we describe the various datasets we used for calibration and validation for the commercial building stock model. Section 2.3.5 describes the data collection and processing effort. Figure 9 summarizes the eight dimensions used for commercial calibration. Table 10 lists the full set of datasets used for commercial calibration and validation and the relevant characteristics of each.



**Figure 9. Commercial calibration used eight main dimensions or data sources; like an eight-sided Rubik's cube, changes made to improve the fit in one dimension needed to be evaluated across the other seven dimensions**

**Table 10. Description of Commercial Calibration Datasets**

Type	Name	Location	Description	Metadata	Main comparison type
Bench- marking	Building Performance Database (BPD)	Several states	Annual Electric and Gas Consumption	Area and building type	EUIs
Bench- marking	DC Bench- marking	District of Columbia	Annual Electric and Gas Consumption	Area and building type	EUIs
Bench- marking	Seattle Benchmarking	Seattle, WA	Annual Electric and Gas Consumption	Area and building type	EUIs
Survey	CBSA	Pacific North-west	Annual Electric and Gas Consumption	Area and building type	EUIs
Survey	EIA CBECS 2012	National	Commercial Survey and Annual Energy Consumption	Surveyed characteristics	EUIs
AMI	Fort Collins Utilities	Fort Collins, CO	Commercial AMI interval data (Region 1)	Area and building type	Seasonal day plots, load profile characteristics, EUIs
AMI	Seattle City Light	Seattle, WA	Commercial AMI interval data aggregated by building type (Region 2a)	Area and building type	Seasonal day plots, load profile characteristics, EUIs
AMI	Portland General Electric	Portland, OR	Commercial AMI interval data (Region 2b)	Area and building type	Seasonal day plots, load profile characteristics, EUIs
AMI	Efficiency Maine	Portland, ME	Commercial AMI interval data (Region 3a)	Area and building type	Seasonal day plots, load profile characteristics, EUIs
AMI	VEIC	Vermont	Commercial AMI interval data (Region 3b)	Area and building type	Seasonal day plots, load profile characteristics, EUIs
AMI	Cherryland	Wisconsin	Commercial AMI interval data (Region 3c)	Area and building type	Seasonal day plots, load profile characteristics, EUIs
AMI	PEPCO	District of Columbia	Commercial AMI interval data aggregated by building type (Region 4a)	Area and building type	Seasonal day plots
AMI	EPB	Chattanooga, TN	Commercial AMI interval data (Region 4b)	Area and building type	Seasonal day plots, load profile characteristics, EUIs
AMI	City of Tallahassee	Tallahassee, FL	Commercial AMI interval data (Region 4c)	Area and building type	Seasonal day plots, load profile characteristics
AMI	Horry Electric Cooperative	Horry County, SC	Commercial AMI interval data (Region 4d)	Area and building type	Seasonal day plots, load profile characteristics
Utility	Xcel Energy	8 states	Commercial monthly billing data	Area and building type	EUIs, outlier detection
End use	Multiple anonymous	50 states	Commercial end-use interval data	area and building type	Thermostats, interior lighting, interior equipment
BAS	Multiple anonymous	Several states	Commercial end-use interval and BAS data	Area and building type	Thermostats, HVAC operation

## *Commercial EIA Data*

### **CBECS**

The 2012 CBECS is a national sample survey that collects energy-related information on the U.S. commercial building stock, including building characteristics, equipment saturation, and metered energy usage (EIA 2012). CBECS was used extensively for deriving ComStock model inputs, including representative distributions of HVAC system types, general equipment, data center prevalence in offices, heating equipment (space and water), and unoccupied thermostat setback prevalence. Additionally, the energy consumption data were used for annual commercial calibration, as well as for checking the reasonableness of the AMI datasets through annual EUI comparisons. CBECS was uniquely appropriate for these use cases as it is intended to provide a representative distribution of buildings, whereas other datasets are not necessarily meant to be representative. One important note is that the CBECS data are from 2012, while the national annual ComStock run being compared to CBECS represents the building stock in 2018. For comparison purposes, the building area of ComStock is scaled to match CBECS by building type, so we expect that differences caused by growth of the building stock between 2012 and 2018 are minimal. However, some building-stock studies, such as the CBSA in the Pacific Northwest, indicate that there have been significant changes to the stock in that time frame, particularly the high adoption of LED lighting. These changes are not reflected in the 2012 CBECS data, but are included in the 2018 ComStock models. Thus some difference between the data and model are expected.

### **Form EIA-861M—Monthly Electricity Sales**

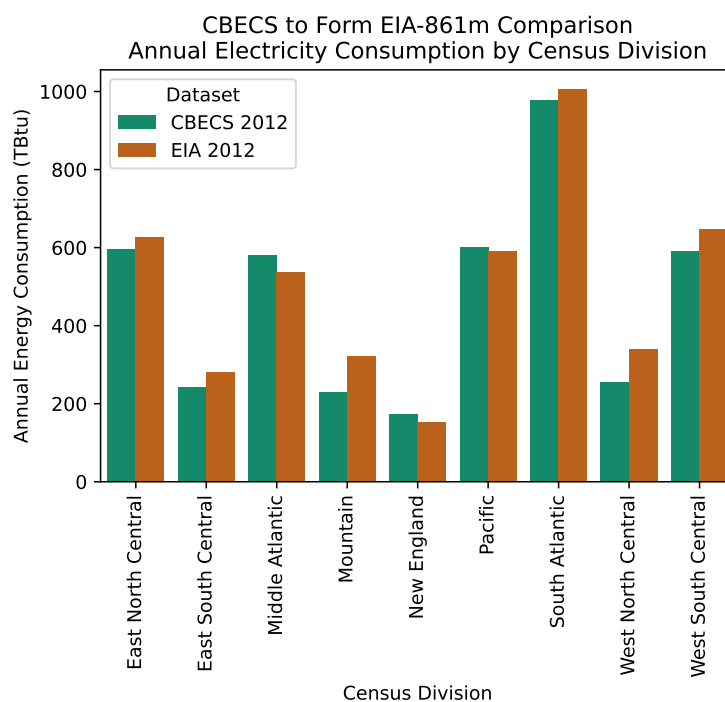
As described in Section 2.3.6, Form EIA-861M collects monthly data on electricity sales by sector. We compare these empirical commercial sector monthly electricity sales data by census division to outputs from ComStock (see Section 4.2). ComStock only represents about 70% of the commercial sector floor area (focused on the most common building types), so in order to compare to sector totals from Form EIA-861M, we need to include some representation of the buildings that ComStock does not represent. In the comparison graphics, this missing piece, labeled “ComStock Gap Model” is created by taking the annual electricity consumption for these buildings from 2012 CBECS and distributing it across months in the same proportion as ComStock outputs.

One important point to note with the Form EIA-861 data is that utilities are responsible for assigning their metered data into sectors (residential, commercial, industrial, etc.). A comparison of CBECS data and Form EIA-861 at the census division scale (Figure 10) shows some discrepancy between the two datasets, but the difference is small.

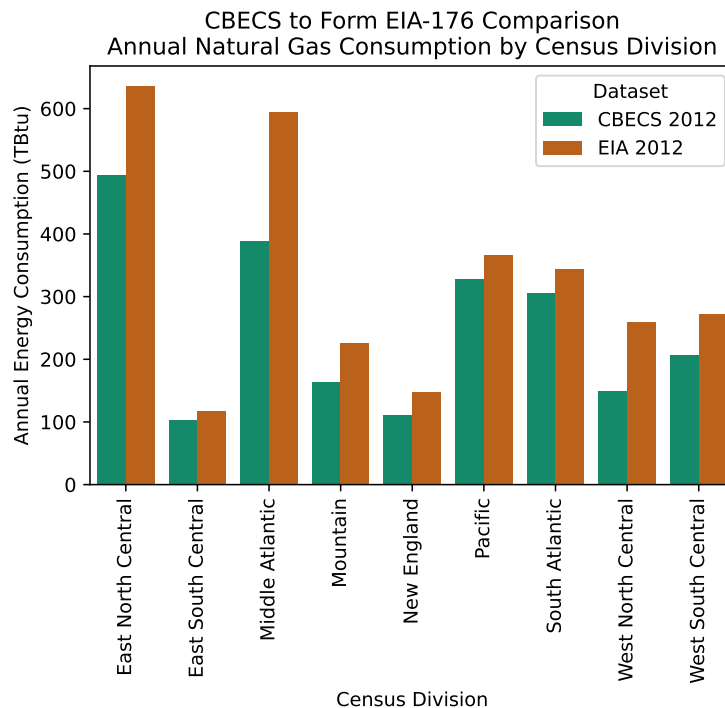
### **Form EIA-176—Monthly Natural Gas Sales**

Form EIA-176 (EIA 2021e) collects data on natural gas sales by sector. We compare these empirical commercial sector monthly natural gas sales data by census division to outputs from ComStock (see Section 4.2). ComStock only represents about 70% of the commercial sector floor area (focused on the most common building types), so in order to compare to sector totals from Form EIA-176, we need to include some representation of the buildings that ComStock does not represent. In the comparison graphics, this missing piece, labeled “ComStock Gap Model” is created by taking the annual natural gas consumption for these buildings from 2012 CBECS and distributing it across months in the same proportion as ComStock outputs.

One important point to note with the Form EIA-176 natural gas data is that utilities are responsible for assigning their data into sectors (residential, commercial, industrial, etc.). A comparison of CBECS data and Form EIA-176 at the census division scale (11) shows a large discrepancy between the two datasets in some cases. A discussion with EIA CBECS staff indicates that although EIA provides guidance on sector classification, the discrepancy could be due to a difference in what CBECS and the utilities submitting Form EIA-176 consider commercial buildings.



**Figure 10. Comparison of 2012 CBECS and 2012 Form EIA-861M commercial sector electricity consumption by census division**

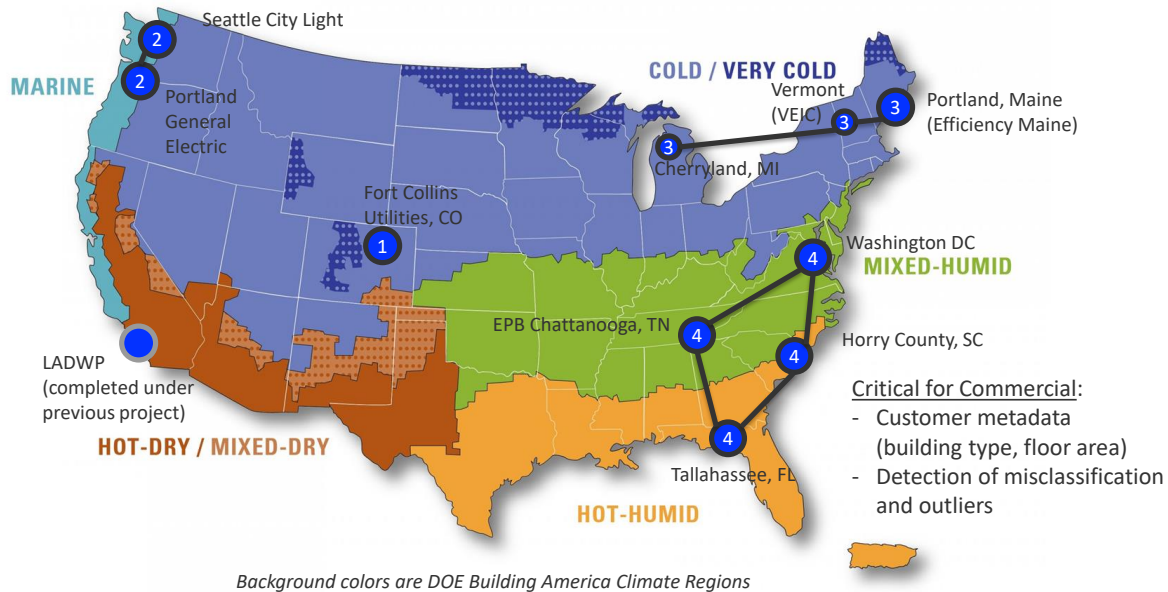


**Figure 11. Comparison of 2012 CBECS and 2012 Form EIA-176 commercial sector natural gas consumption by census division**



### Commercial Utility AMI Datasets

Figure 12 shows a map of the utility AMI datasets used for commercial calibration. Each AMI dataset and source of AMI metadata is briefly described below.



**Figure 12.** This map shows the utility AMI datasets used for the commercial calibration and validation effort. The numbered markers show how the ten AMI datasets were grouped into four regional phases. The LADWP marker indicates that calibration was previously completed for the Los Angeles Department of Water and Power.

#### Fort Collins Utilities

The Fort Collins AMI dataset, which covers the city of Fort Collins, CO, was used for Commercial Region 1 calibration. The dataset includes area and building type classification, as well as building address. Overall, there were 1,722 unique building samples for use with commercial calibration. Because we had addresses for each meter and the full set of metered data for the utility, we were able to ascertain that the meter-to-building mapping was done thoroughly and could assess outliers manually. We found this dataset to be the most reliable of the commercial AMI datasets because it had reasonable energy behavior that aligned with data from other sources. Because of the known high-quality meter-to-building matching and good agreement with other empirical data sources, this dataset was most trusted AMI for commercial calibration.

#### Seattle City Light

The Seattle City Light AMI dataset, which covers the city of Seattle, WA, was used for Commercial Region 2a calibration. We developed code that Seattle City Light staff used to join their meter data with the CoStar real estate database to derive area and building type classifications. They then shared the anonymized aggregate AMI load profiles by building type with NREL and provided technical support to NREL in the use of their data. Overall, there were 8,592 unique building samples for use with commercial calibration, covering all ComStock building types except for secondary schools.

#### Portland General Electric

The PGE AMI dataset, which covers the city of Portland, OR, was used for Commercial Region 2b calibration. The dataset was joined with the CoStar real estate database to derive area and building type classification. Overall, there were 41,423 unique building samples for use with commercial calibration, covering all ComStock building types except for secondary schools.

### **Efficiency Maine**

The Efficiency Maine dataset, which covers a subset of the city of Portland, ME, was used for Commercial Region 3a calibration. The dataset was joined with the CoStar real estate database to derive area and building type classification. Overall, there were 215 unique building samples for use with commercial calibration.

### **Vermont Energy Investment Corporation**

The VEIC dataset, which covers a subset of the state of Vermont, was used for Commercial Region 3b calibration. VEIC operates Efficiency Vermont, the statewide energy efficiency program for Vermont. The dataset was joined with the CoStar real estate database to derive area and building type classification. Overall, there were 1,824 unique building samples for use with commercial calibration.

### **Cherryland Electric Cooperative**

The Cherryland dataset, which covers a subset of the state of Michigan, was used for Commercial Region 3c calibration. The dataset was joined with the CoStar real estate database to derive area and building type classification. Overall, there were 239 unique building samples for use with commercial calibration.

### **PEPCO**

The PEPCO dataset, which covers a subset of the District of Columbia, was used for Commercial Region 4a calibration. The dataset was joined with the CoStar real estate database to derive area and building type classification. The utility provided data binned to meet the applicable privacy requirements. Overall, there were 3,898 unique building samples provided in 518 aggregated 'bins' for use with commercial calibration.

### **EPB**

The EPB dataset, which covers a subset of Chattanooga, TN, was used for Commercial Region 4b calibration. The dataset was joined with the CoStar real estate database to derive area and building type classification. Overall, there were 5,130 unique building samples for use with commercial calibration.

### **Tallahassee**

The Tallahassee dataset, which covers a subset of the city of Tallahassee, FL, was used for Commercial Region 4c calibration. The dataset was joined with the CoStar real estate database to derive area and building type classification. Overall, there were 4,573 unique building samples for use with commercial calibration.

### **Horry Electric Cooperative**

The Horry Electric Cooperative dataset, which covers a subset of Horry County, SC, was used for Commercial Region 4d calibration. The dataset was joined with the CoStar real estate database to derive area and building type classification. Overall, there were 419 unique building samples for use with commercial calibration.

### **Los Angeles Department of Water and Power**

During the previously completed LA100 project, we had access to commercial AMI data from a subset of LADWP customers and monthly billing data from most customers (Hale et al. 2021). The dataset was joined with the local tax assessor's database to derive area and building type classification. Many of the input calibration changes made during the LA100 project were incorporated into ResStock and ComStock before the EULP project started. So although this region was not one of the regions for which QOIs were calculated for the EULP project, some calibration of ResStock and ComStock have been performed for this region.

### *Other Commercial Datasets*

#### **Xcel Energy**

The Xcel Energy monthly billing dataset, which covers most Xcel Energy customers across eight states, was used for EUI comparisons and testing of different outlier detection methods for misclassified and unrealistic data. The dataset was joined with the CoStar real estate database to derive area and building type classification. Overall, there were 59,896 unique building samples for use with commercial calibration.

## **BPD**

The BPD dataset, which comprises several benchmarking datasets across the United States, was used for EUI comparisons against ComStock models as well as checks to the AMI datasets for reasonableness. The dataset includes annual gas and electricity consumption for 287,856 buildings, with building type and area, from multiple climate regions across the United States, spanning from 2006 to 2018 (*Building Performance Database (BPD)*). Because of the nature of the BPD dataset, we do not know how representative the buildings are of the population.

## **Other Benchmarking Data**

Several other benchmarking datasets were used for EUI comparisons against ComStock models as well as checks to the AMI datasets for reasonableness. The D.C. benchmarking dataset covers Washington, D.C., and includes annual electricity and gas consumption for 4,466 meters with building type and area. The Seattle benchmarking dataset covers Seattle, WA, and includes annual electricity and gas consumption for 3,582 unique building IDs with building type and area. Because benchmarking datasets only apply to buildings over a certain size, we know that these datasets are not representative of the overall population of buildings in the areas covered.

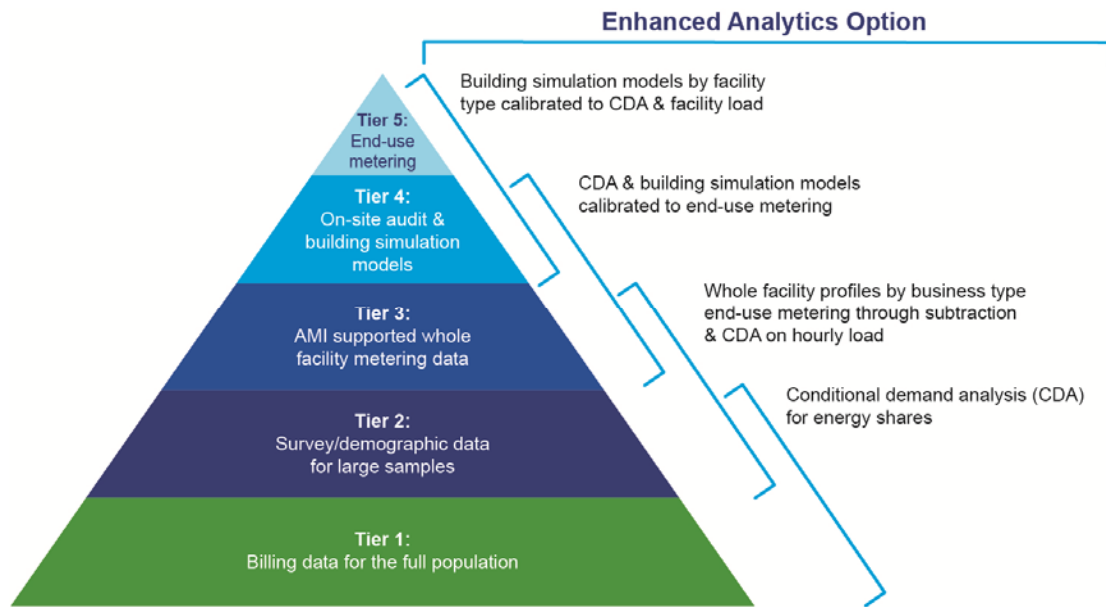
## **CBSA**

NEEA's CBSA dataset contains 932 unique sites throughout the Pacific Northwest region and includes annual electricity and natural gas consumption with building type and area. This dataset was intentionally designed to be statistically representative of the population of buildings in the area covered.

### **2.3.8 Calibration and Validation in Multiple Dimensions**

As presented in the preceding section, the empirical data sources we used cover a variety of scales and geographies. No one data source provides a full picture against which model outputs could be compared. For example, end-use metering data are an ideal source of empirical data on hourly or subhourly energy use for specific end uses, but are typically only available for small samples—tens or hundreds—of buildings in specific locations, and may exhibit bias in terms of not representing the population of buildings at large in a region and outside the region. Furthermore, tens or hundreds of samples is not large enough to fully diversify loads, especially on peak days, so these small sample datasets provide less confidence for calibrating the timing and magnitude of peaks. In contrast, interval meter data for a large sample of building meters provides empirical data on peak timing and magnitude with high confidence, but does not provide empirical end-use data. Both of these empirical data types are only available for a relatively small number of locations across the United States, but they can be supplemented with electricity sales data collected by EIA, which is only monthly (some utilities; all states) or annual (all utilities).

Several of our TAG members from DNV GL and the EPRI published a white paper in 2014 that reviewed methods for developing end-use data and proposed a tiered strategy for developing end-use load shape data (Puckett et al. 2014). The strategy, illustrated in Figure 13, uses building simulation models that are calibrated against end-use metering data from a sample of buildings and billing or interval meter data for the full population of buildings. This proposed strategy was a natural fit for the challenge of calibrating national building stock models at the multiple dimensions and scales of data we had available.



**Figure 13. The tiered strategy for developing end-use load shape data proposed by DNV GL, EPRI, and ORNL informed our project’s multidimensional approach to calibration (Image source: Puckett et al. 2014)**

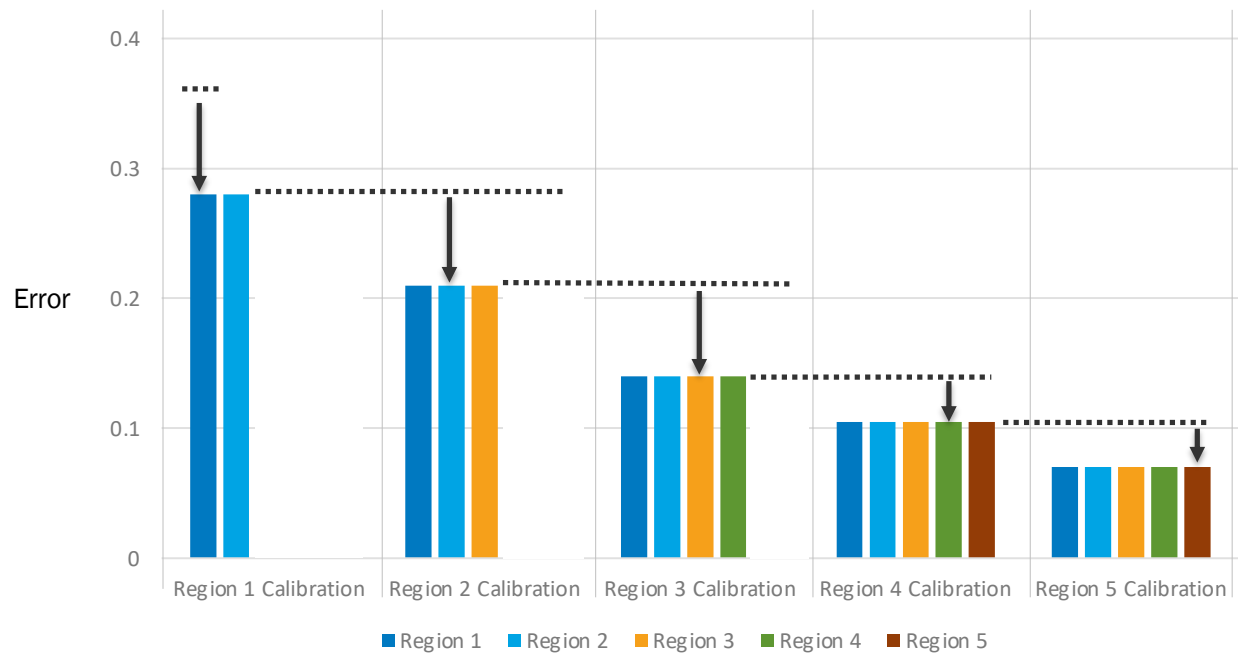
### *Region-by-Region Calibration Strategy*

In implementing the multidimensional calibration approach, we chose to divide the calibration effort into phases, each focused on a particular region of the contiguous United States for which we obtained AMI interval data. This is primarily for practical reasons; it took months of conversations to develop nondisclosure agreements with utilities sharing AMI interval data for the project, and additional time to transfer, process, and analyze each utility dataset (see Section 2.3.5). Thus, the utility AMI data became available slowly over the course of the three years of the project. We ended up with ten commercial sector AMI datasets, grouped into four regions with similar climates, and eight residential sector AMI datasets, grouped into five regions with similar climates. See Figures 7 and 12 for maps of the AMI dataset regions. Over the course of two years, we spent three to four months on each commercial or residential regional calibration effort. Each regional phase gave an opportunity to investigate model output discrepancies of particular importance to that region (e.g., electric heating in the Southeast). After each regional phase, we summarized the changes made, impact on QOIs, and overall progress, and reported to DOE and our TAG.

In all cases, the changes made to model inputs were applied nationally, though regional impacts would often differ because of differences in climate or technology saturation. In other words, we specifically did *not* seek to create five distinct regional versions of ComStock and ResStock, but rather a single version of ComStock and ResStock for the United States. Thus, when we completed each regional phase, we were not finished with that region; indeed, we often had more work to improve the model fit in that region. We continued to monitor progress on regional QOIs as we progressed through the four or five regions. The idealized form of this region-by-region approach is shown in Figure 14.

### **The Rubik’s Cube Analogy**

As presented in our calibration philosophy, our focus on getting the “why” right led us to prioritize input calibration over output calibration and manual changes to inputs based on data over automated input calibration. If we were using automated optimization algorithms for calibration (or were only using statistical output calibration) then we could use a cost function that minimized model errors across all QOIs, dimensions, and scales at once. However, because we divided the calibration effort into regional phases and because we were manually updating model inputs with changes applied nationally, it was necessary to regularly evaluate how the changes we were making to improve model fit (QOIs) in one region were affecting QOIs in other regions and against calibration datasets with other scales or dimensions. We use the analogy of a Rubik’s Cube—any change you make to match up colors on one side of the



**Figure 14. This idealized graph illustrates how we focused on reducing modeling errors one region at a time, but because most model changes affected all regions, we continuously evaluated the impact of each change on previous regions and nationally available metrics. The calibration efforts for earlier regions often create better starting points for later regions. Improvements made in the later regional phases also tend to improve results for the earlier regions.**

cube affects all other sides of the cube. The changes to the cube must be strategic and keep in mind how they affect all sides of the cube.

### 2.3.9 End-Use Transferability Study

Because submetered hourly end-use datasets are only available in a few locations, we wanted to study how *transferable* these data are from one location to another. When considering transferability, it is useful to differentiate between weather-dependent loads (heating and cooling) and non-weather dependent loads (everything else). Previous work by KEMA for the Northwest Power and Conservation Council and NEEP developed transferability ratings for different residential and non-residential end uses. Non-weather-dependent end uses like lighting and appliances are typically given “high” transferability ratings, and HVAC-related end uses are given “low” ratings, with some end uses like water heating and pool pumps getting a “medium” rating (KEMA 2009).

Although weather-dependent loads are not directly transferable between regions, this project heavily relies on the fact that we use physics-based heat transfer simulations along with robust weather data and data on building shell, equipment, and thermostat setpoint characteristics to predict heating and cooling energy usage across many climates. In this section, we focus on non-weather dependent end uses and evaluating how they differ across datasets from different locations.

#### *Commercial End-Use Transferability*

We did not have enough commercial end-use data by building type and location to study transferability of commercial end uses. However, we were able to use AMI data to compare overall hours of operation between regions (Section 3.3.17). There were not major differences in distributions of hours of operation across the AMI datasets.

#### *Residential End-Use Transferability*

The residential end-use transferability study compared end-use data from five datasets to determine the degree to which end-use usage patterns vary across datasets. The five datasets are described in Section 2.3.6.<sup>6</sup> If the usage pattern was generally the same across all five datasets, then we concluded that usage patterns do not vary strongly with location, and we applied usage schedules as model inputs that are the same nationwide. If the usage pattern is different across the five datasets, then this prompted us to look into possible explanations, which could range from data quality or analysis issues to dataset sampling bias to a true difference in usage pattern in households in different regions.

### Comparison Process

The residential end-use comparisons involved substantial effort to standardize end-use categories across datasets, so that the comparisons could be as precise as possible. Sample sizes for each dataset vary based on how many homes have each end use; i.e., not all homes have an electric cooking range. Some end uses, like pool and spa heaters, were only measured in a subset of the datasets. The results of the comparisons are shown in Figures 15–24. The shaded areas indicate the 80% confidence interval of the mean ( $\text{mean} \pm 1.28 \times \text{standard error}$ ), so larger shaded areas indicate either fewer samples in that dataset or larger variation (across homes or from day to day).

In addition to the five residential end-use datasets, we include ResStock output (as they were at the start of the project) and time-use diary data from the American Time Use Survey (“ATUS”) in the comparisons. Note that ATUS is not measured energy use, but rather survey respondents’ self-reported record of their activities over the course of 24 hours. The “doing laundry” activity may include loading and unloading machines and folding clothes, the timing of which does not directly correspond to when the clothes washer uses energy or hot water. The primary reason that ATUS is included in comparisons here is that we later used ATUS as the primary data source for the residential stochastic occupancy simulator that was developed under this project to generate schedules for occupant driven loads (see Section 3.2.1 and Chen et al. 2022 for details).

### Comparison Results

In most cases, the majority of the shaded areas shown in Figures 15–24 are overlapping, suggesting that we can say with high confidence that the normalized load shape mean values are close to each other. There are several exceptions to the alignment between datasets. One exception is that the Pecan Street plug load shape has a morning peak, whereas the other datasets peak in the evening (Figure 19). We suspect Pecan Street has some degree of sampling bias because many of the homes are in the same neighborhood and may include large plug loads such as Level 1

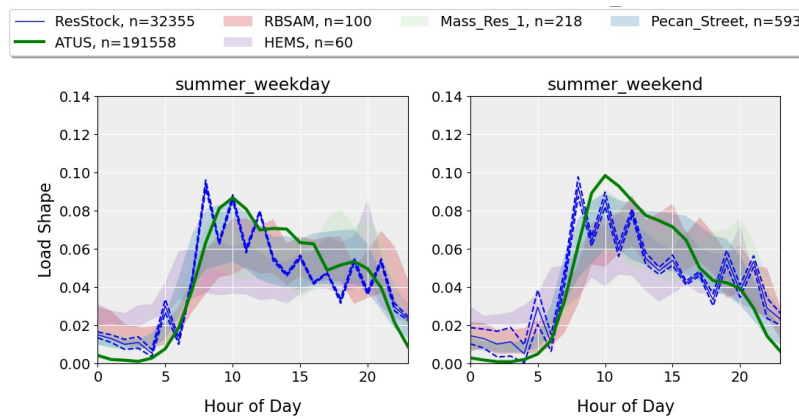
<sup>6</sup>We are grateful to our data sharing partner Southern Company, who provided access to 1-min resolution data from their Smart Neighborhood project in Birmingham, AL. Unfortunately, we did not have time to process and analyze the high-resolution data in time for the transferability study, but we are working with the data for another project and hope to include it in future comparisons.



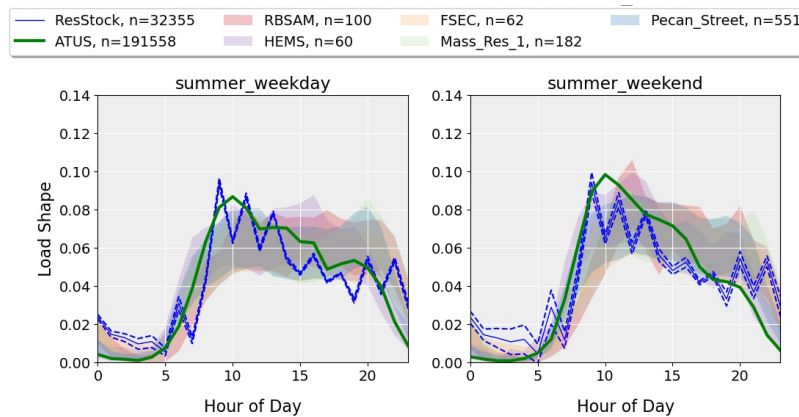
electric vehicle chargers not on a separate circuit. Section 3.2.1 discusses a change made to the plug load schedule used in ResStock. The plug load schedule is also now modified by the new stochastic schedule generator that was developed in Section 3.2.1.

Another misalignment is in the lighting comparison (Figure 22), where Mass RES 1 shows higher midday lighting use, and RBSAM shows a higher evening peak. Differences in lighting are to be expected because different locations have different local sunrise and sunset times based on their latitude, longitude, and time zone. The lighting algorithm in ResStock already accounts for sunrise and sunset times that vary with location, but we did not investigate further to see how much of the misalignment the algorithm explains. The lighting schedule is also now modified by the new stochastic schedule generator that was developed in Section 3.2.1, *Residential Stochastic Occupant-Driven Loads*.

Finally, we see major differences in pool and spa pump load shapes; however, sample sizes for these shapes are much lower because fewer homes have pools (Figure 23). One hypothesis is that pools are common in some datasets (FSEC, Mass, Pecan Street) but spas/hot tubs are more common in the RBSAM and HEMS datasets from the Northwest, which leads to a difference in shapes. Ultimately, we did not make any model changes based on this comparison. Similarly, pool and spa heater shapes were only available from two datasets (Figure 24), and we did not make model changes based on the comparison.

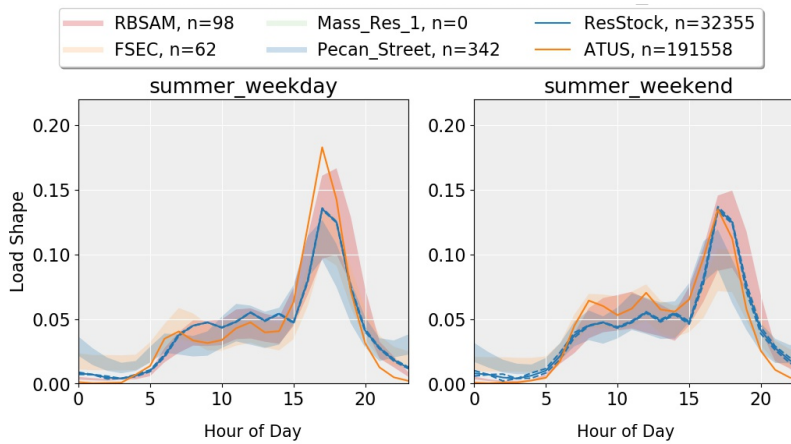


**Figure 15.** Clothes washer normalized load shapes are compared across datasets, to ResStock output (before calibration) and to frequency of ATUS-reported activities labeled “laundry” The shaded area indicates the 80% confidence interval of the mean ( $\text{mean} \pm 1.28 \times \text{standard error}$ ).

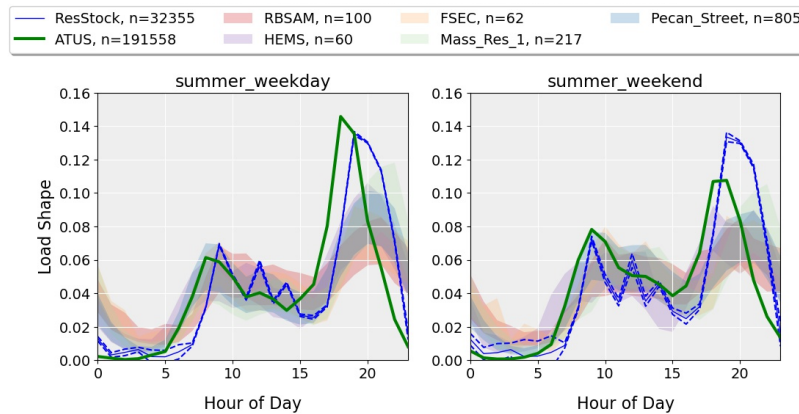


**Figure 16.** Clothes dryer normalized load shapes are compared across datasets, to ResStock output (before calibration), and to the frequency of ATUS-reported activities labeled “laundry” The shaded area indicates the 80% confidence interval of the mean ( $\text{mean} \pm 1.28 \times \text{standard error}$ ).

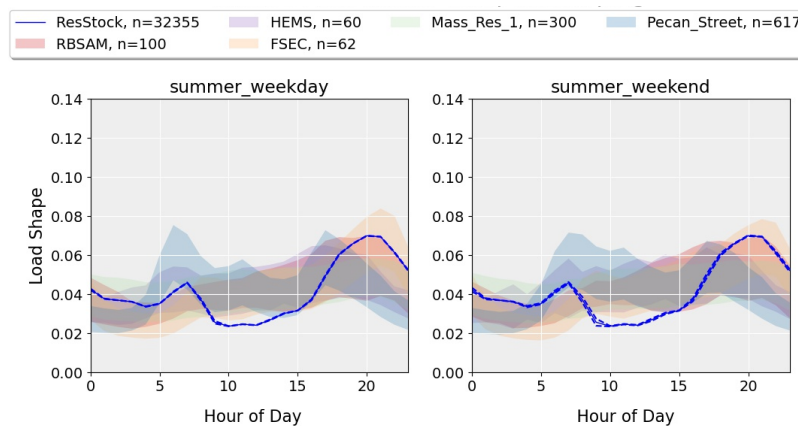




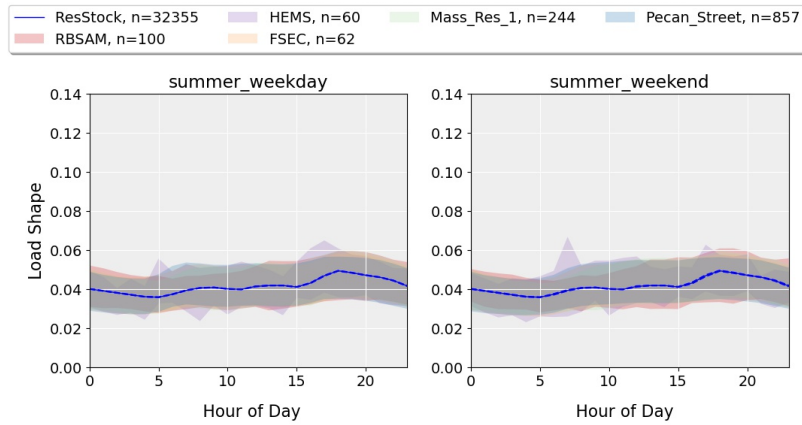
**Figure 17.** Cooking range normalized load shapes are compared across datasets, to ResStock output (before calibration), and to the frequency of ATUS-reported activities labeled “Food and drink preparation” The shaded area indicates the 80% confidence interval of the mean ( $\text{mean} \pm 1.28 \times \text{standard error}$ ).



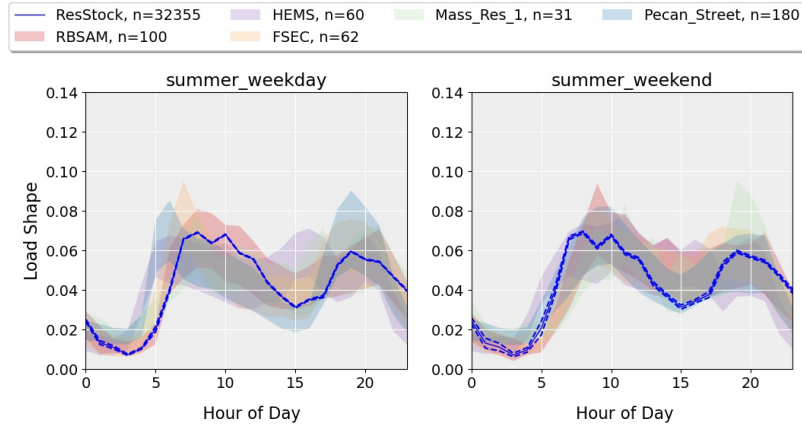
**Figure 18.** Dishwasher normalized load shapes are compared across datasets, to ResStock output (before calibration), and to the frequency of ATUS-reported activities labeled “Kitchen and food clean-up” The shaded area indicates the 80% confidence interval of the mean ( $\text{mean} \pm 1.28 \times \text{standard error}$ ).



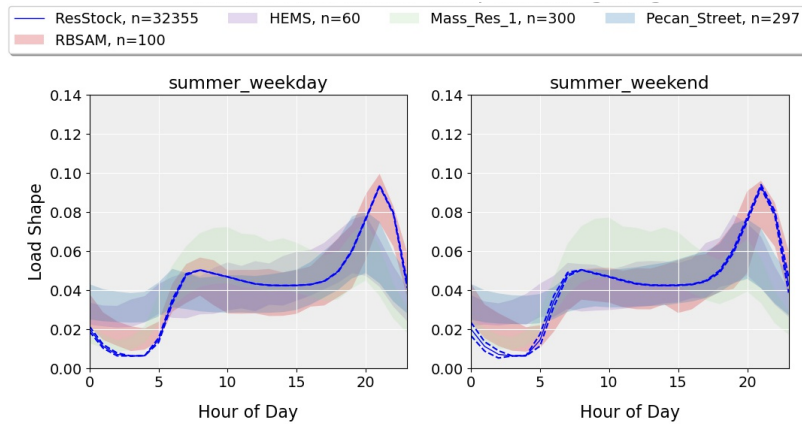
**Figure 19.** Plug load normalized load shapes are compared across datasets and to ResStock output (before calibration) The shaded area indicates the 80% confidence interval of the mean ( $\text{mean} \pm 1.28 \times \text{standard error}$ ).



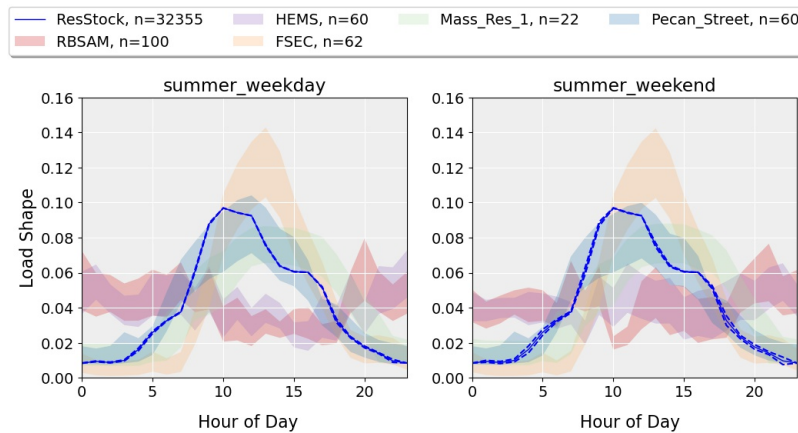
**Figure 20. Refrigerator normalized load shapes are compared across datasets and to ResStock output (before calibration) The shaded area indicates the 80% confidence interval of the mean ( $\text{mean} \pm 1.28 \times \text{standard error}$ ).**



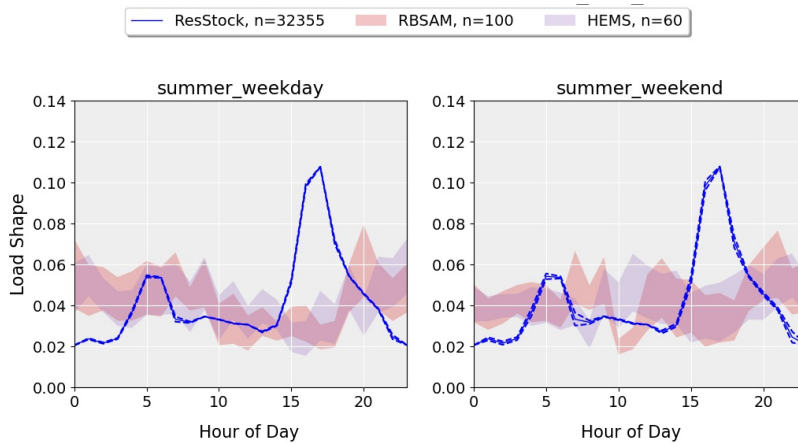
**Figure 21. Water heating normalized load shapes are compared across datasets and to ResStock output (before calibration) The shaded area indicates the 80% confidence interval of the mean ( $\text{mean} \pm 1.28 \times \text{standard error}$ ).**



**Figure 22. Lighting normalized load shapes are compared across datasets and to ResStock output (before calibration) The shaded area indicates the 80% confidence interval of the mean ( $\text{mean} \pm 1.28 \times \text{standard error}$ ).**



**Figure 23. Pool and spa pump normalized load shapes are compared across datasets and to ResStock output for ComEd territory (before calibration) The shaded area indicates the 80% confidence interval of the mean ( $\text{mean} \pm 1.28 \times \text{standard error}$ ).**



**Figure 24. Pool and spa heater normalized load shapes are compared across datasets and to ResStock output (before calibration) The shaded area indicates the 80% confidence interval of the mean ( $\text{mean} \pm 1.28 \times \text{standard error}$ ).**

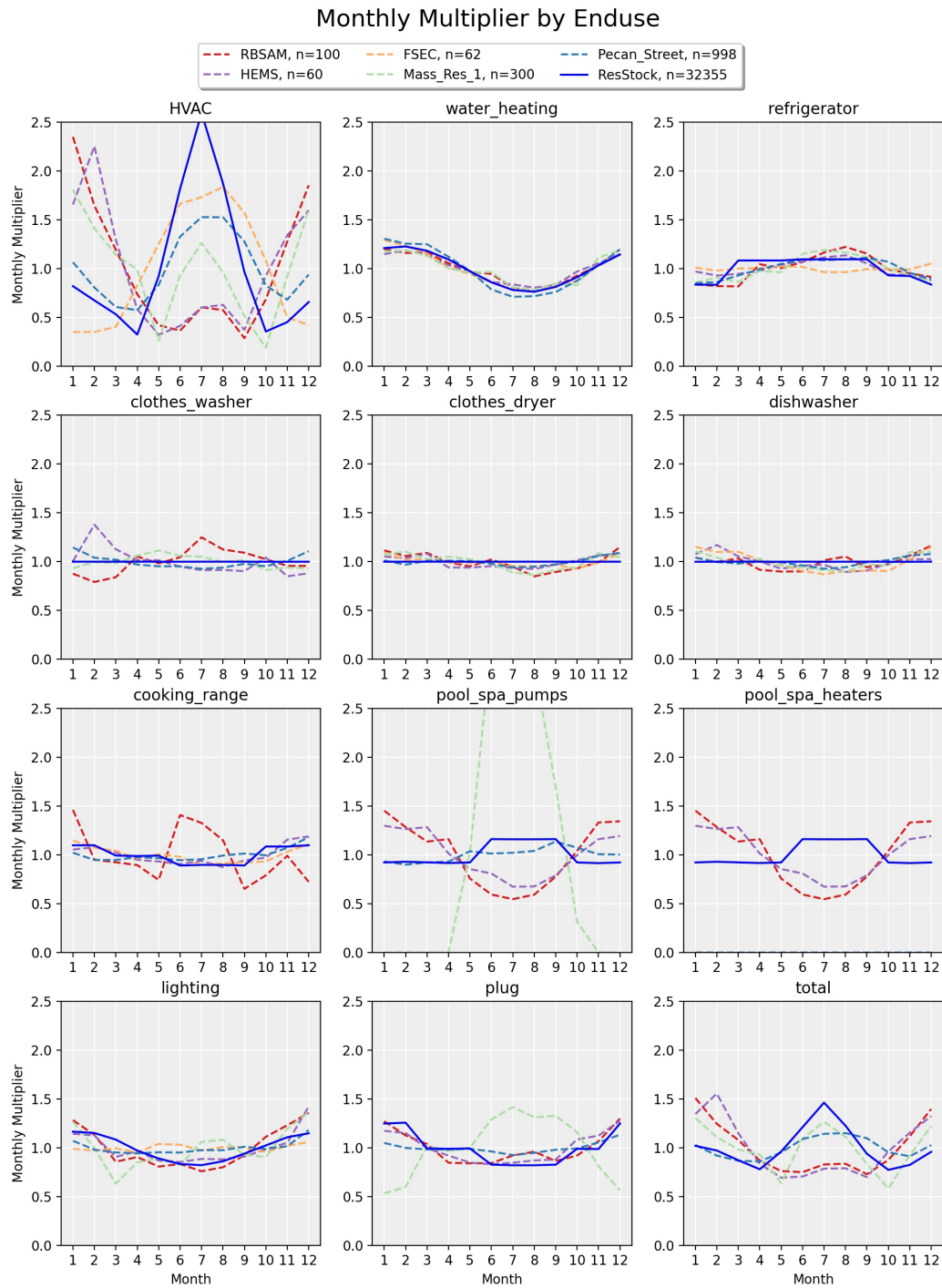
From the comparisons shown in Figures 15–24, we conclude that, with some exceptions, the *shapes* of non-weather-dependent end uses can be considered transferable between regions, although the magnitudes of these end uses will vary depending on factors such as the saturation of end-use equipment (e.g., electric clothes dryers) in a given location.

### Monthly Variation in End-Use Loads

Another outcome of the residential end use transferability study was a comparison of how normalized end-use loads vary from month to month, which are typically used as model inputs called “monthly multipliers.” Figure 25 shows the results of this comparison, along with the monthly multipliers from ResStock before any calibration work. Based on these comparisons we averaged the monthly multipliers across datasets for clothes washer, clothes dryer, cooking range, and dishwasher usage schedules, as shown in Figure 26. These multipliers were implemented in the new stochastic schedule generator (see Section 3.2.1) by slightly lengthening or shortening event durations to achieve correct monthly usage.

**Additional End-Use Dataset Analysis**

Appendix D includes supplemental figures from our analysis of the five residential end-use datasets, showing average 24-hour profiles for each month for each end use category measured in the five datasets.



**Figure 25. Normalized monthly electricity use was compared for the five datasets across all residential end-use categories**

# End-Use Load Profiles for the U.S. Building Stock: Methodology and Results of Model Calibration, Validation, and Uncertainty Quantification



**Figure 26. Monthly appliance end-use multipliers for ResStock were based on averages across six end-use datasets. Clothes dryer, dishwasher, and range usage generally is lower in summer and higher in winter. Clothes washer usage shows less of a clear trend. The monthly FSEC multipliers for cooking range were not included in the average because it was a clear outlier and the FSEC PDR dataset was the least statistically representative sample of homes.**



### 2.3.10 Methods to Understand Variability in Commercial Building Operation

Whereas the new residential stochastic occupant behavior model had access to detailed time-use data on occupant activities in homes, there was not a similarly detailed data source for commercial building occupant behavior, so we had to take a different approach. In this section, we briefly introduce a two-path approach (time-domain and frequency-domain, as shown in Figure 27) developed through this project to characterize building load profile variability. More details of the methodology and experiment on a real building dataset can be found in Li et al. 2021a. For this aspect of the project, we focused on commercial building loads, because variability in residential loads was addressed in Section 3.2.1. The time-domain analysis presented here was directly used to develop the hours of operation schedule distributions presented in Section 3.3.17. The frequency-domain method presented below was not applied in this project due to time constraints, but we plan to apply it for variability comparisons in the future.

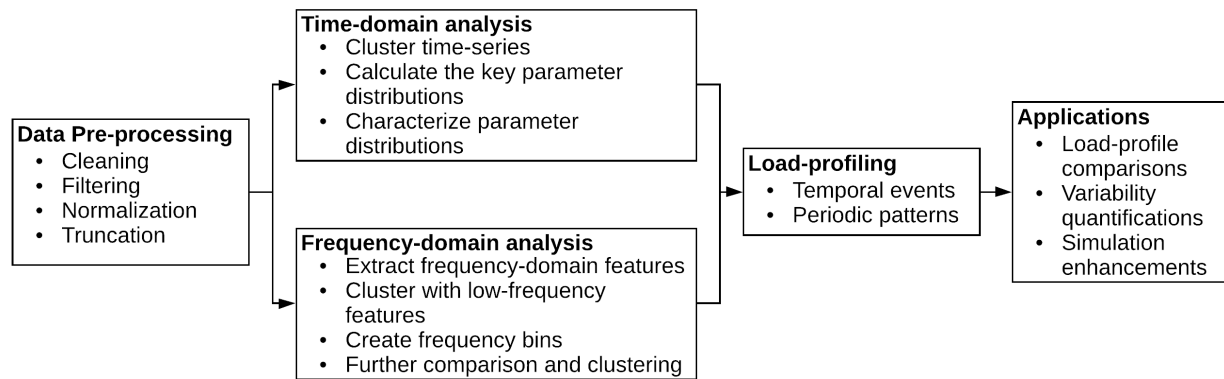


Figure 27. Two-path approach to characterize load profile variability

#### Time Domain

Building electric loads often show a clear periodic behavior. Time-domain characterization refers to the analysis of building load's periodic patterns with respect to time. For example, the load of an office building rises in the morning, peaks around noon, starts to decrease in the afternoon, and returns to the base load at night. To characterize the building load shape, we used the method described in Price 2010, which includes nine key parameters: two loads (base load, peak load); four times (morning rise, high-load start, high-load finish, afternoon fall finish); and three time intervals (rise time, high-load duration, fall time), as shown in Figure 28 and defined in Table 11.

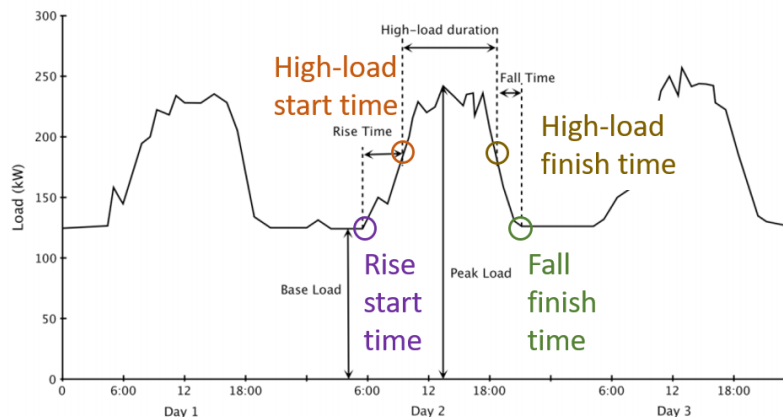


Figure 28. Key parameters to characterize building load shape from the time domain

As shown in Table 11, we used the 97.5th and 2.5th percentile of daily load to define the peak and base load, because it is possible that in some buildings, the very highest 15-minute data point is substantially higher than any other data point, which could be an outlier or due to some extreme events. Excluding the most extreme values results in both more stable and more relevant values, rather than the absolute maximum and minimum values (Price 2010).

**Table 11. Key Parameters to Characterize the Building Load Shape From the Time Domain**

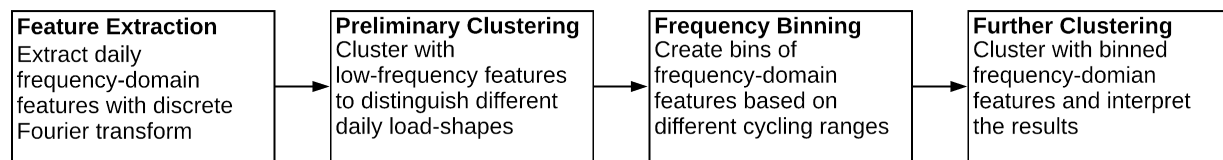
Parameters	Definition
Peak load	97.5 percentile of daily load measurements with a given sampling rate, such as 1-hour, 15-minute etc. of a specific day
Base load	2.5 percentile of daily load measurements
Rise start time	The latest time in the morning when the load is less than: $\text{base load} + 0.05 \times (\text{peak load} - \text{base load})$
High load start time	The earliest time during each day when the load is more than halfway to the 97.5 percentile load for the day
High load finish time	The latest time during each day when the load is more than halfway to the 97.5 percentile load for the day
Fall finish time	The earliest time in the afternoon when the load is less than: $\text{base load} + 0.05 \times (\text{peak load} - \text{base load})$
Rise time	The duration it takes for the load to increase halfway to the peak load; the time interval between rise start time and high load start time
High load duration	The duration the load stays above the halfway mark; the time interval between the high load start time and high load finish time
Fall time	The duration it takes the load to fall from the halfway point back to the base load; the time interval between the high load finish time and fall finish time

Based on the base load and peak load, we defined the four time points that are critical to describe the building load curve. Rise start time characterizes when the building load starts to increase in the morning, which might be due to the operator turning on building services such as air conditioning. High load start time describes when the building enters full-load operation. High load finish time and fall finish time define when the building load starts to decrease and return to the base load, respectively. Built upon the four time points, we defined three time intervals: rise time, high load duration, and fall time, as illustrated in Figure 28 and defined in Table 11.

The last step is to determine the distribution of each key parameter, the shape of which represents the variability by day or building.

#### *Frequency Domain*

Frequency-domain characterization refers to the analysis of a signal's periodic patterns (i.e., cycling and amplitude) with respect to frequency. Theoretically, any timeseries signal can be converted to a frequency spectrum with a mathematical transform operation. In this project, we selected the discrete Fourier transform (DFT), which is a commonly used method for discrete timeseries such as electric load profiles. An overall workflow of frequency-domain analysis is shown in Figure 29.



**Figure 29. Workflow of the frequency-domain load profile characterization**

Because we are interested in understanding the variability introduced by occupant-driven loads on a daily basis, we truncated load profiles into individual days before applying the DFT. To mitigate the spectral leakage issue due to the truncation, we applied Hann window functions to the daily load profiles before extracting the frequency spectrum with DFT. The frequency spectrum extracted from daily load profiles contains plenty of unique frequency components, especially for load profiles with granular temporal resolution (e.g., 1-minute to 15-minute intervals). For example, a daily load profile with 10-minute sampling interval has 144 timestamps, which yields 72 unique frequency components. Many of those frequency components are very close to each other and do not have distinct physical meaning. Therefore, binning neighboring frequency components not only groups features with similar

physical meanings together, but also reduces the dimension of the data, which benefits the subsequent clustering analysis. For clustering, we used the  $k$ -means clustering algorithm on the binned frequency features. Please refer to Li et al. 2021a for details of the binning and clustering analysis configurations.

### Example Application

Here, we demonstrate the frequency-domain load profile variability analysis with an example dataset, which consists of 188 small- to medium-size office buildings in California. Figure 30 shows two daily load profiles and their frequency spectrum. The Example 1 load profile has a smooth daily curve; its consumption started to increase at about 03:00 and reduced to the base load at about 21:00. This daily pattern is reflected by the frequency spectrum, where the low frequency (2.3e-5 Hz, which corresponds to a 12-hour cycle) has a large amplitude. In contrast, the Example 2 load profile has many short-term spikes during the 10:00 to 18:00 period in addition to the daily cycle. Those variations were captured in the frequency spectrum graph as well. The frequency-domain features are powerful indicators of the frequencies of short-term cycles and their corresponding magnitudes, which can help us characterize and distinguish load profiles.

We can cluster the load profile with the frequency components to understand what the typical load profiles are. Figure 31 shows two load profiles in a week and their daily binned frequency feature distribution boxplots. The timeseries plot shows that Example A has high short-term variability, which is captured by Bin 1 (0.5—0.75 hours) in the boxplot. In comparison, both the mean value and the interquartile range of Bin 1 in Example B are lower than those of Example A, which is because the load profile in Example B has lower working hour demands and lower short-term fluctuations. The examples prove that the binned frequency features are good metrics for quantifying load-profile variability, as they not only indicate the magnitude of the variations, but also pinpoint the frequency components.

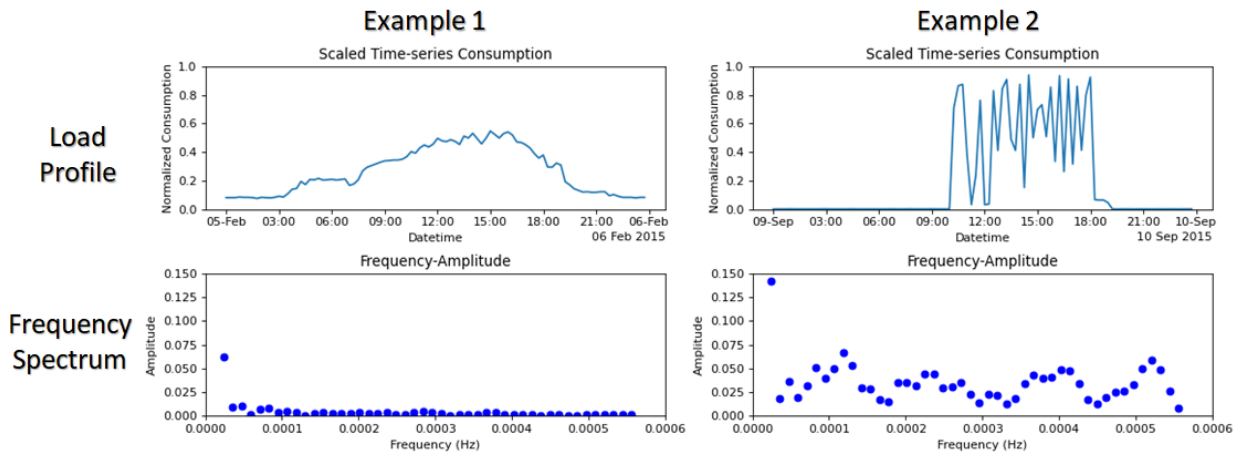
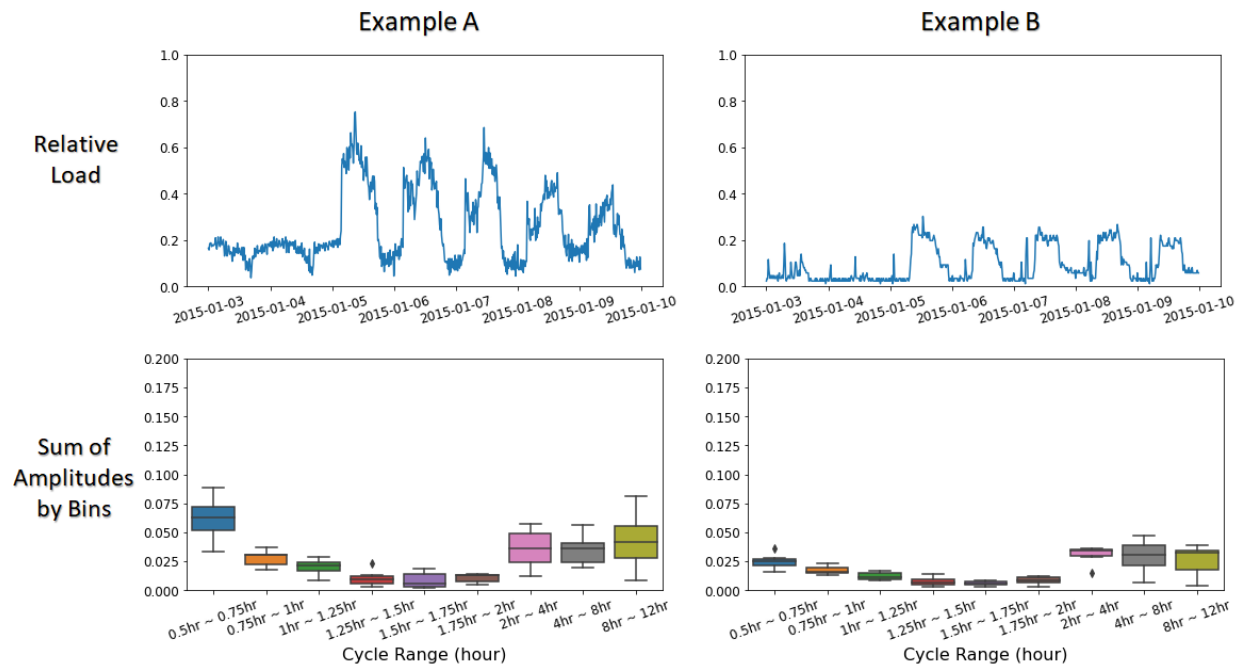


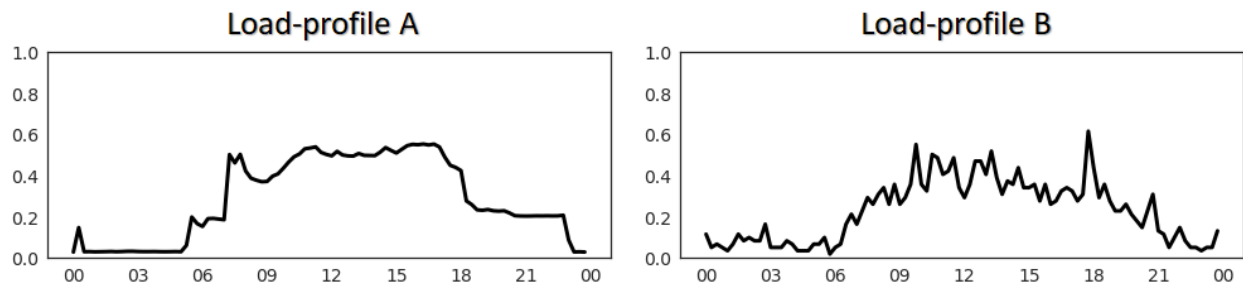
Figure 30. Example frequency spectrums extracted from daily load profiles



**Figure 31. Example load profiles and the binned daily frequency features**

### Discussion

The two-path variability characterization method consists of time-domain and frequency-domain analysis, where both can provide unique insights. Time-domain analysis is good at identifying key temporal events, such as the high-load start and end timestamp, and the high-load duration. Frequency-domain analysis is good at identifying the amplitudes and duration periodic patterns. They are complementary to each other. For example, consider the two real daily load profiles shown in Figure 32. Time-domain analysis might consider them as the same type in terms of the load shape, whereas frequency-domain could distinguish between them by the durations and amplitudes of the periodic patterns. On the other hand, time-domain analysis could easily identify the start and end time of the high load, whereas a frequency-domain analysis could not.



**Figure 32. Two example daily load profiles**

The two-path variability characterization method was demonstrated with a dataset of real buildings in California. Some of the time-domain analysis findings were adopted to enhance the hours of operation schedule distributions presented in Section 3.3.17. The frequency-domain method was not applied in this project due to time constraints, but we plan to apply it for variability comparisons in the future.

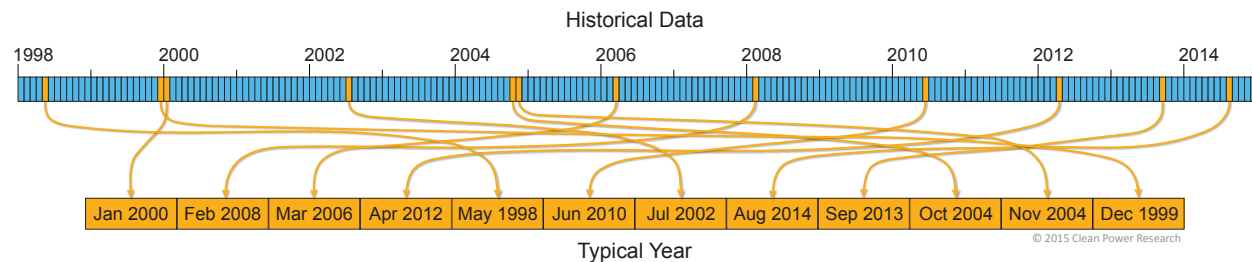
## 2.4 Methodology: Weather Data

Because weather drives much of the variation in electricity demand, with periods of peak demand typically occurring on the hottest or coldest days in a location, it was essential for this project to have accurate weather data to use as simulation inputs. This section describes how we developed the sets of historical weather data used for calibration, validation, and production of the final published dataset.

### 2.4.1 Background on Historical Weather Data

#### Importance of Historical Weather Data

Building energy simulations have traditionally used typical meteorological year (TMY) weather data, based on 30-year climate normals, for building design and retrofit analysis (Wilcox and Marion 2008). Figure 33 illustrates how TMY weather data files are constructed: each month is selected from a historical year for which that month most closely matches that long-term average weather of that month (Kankiewicz and Novotny 2015). However, when comparing simulation outputs to hourly empirical data, it is important to simulate the building models using historical weather data that aligns with the year of the empirical data. These historical weather data files are often called actual meteorological year (AMY) weather data. AMY weather is also important because load profiles generated with AMY weather can be aggregated across weather station regions, whereas load profiles generated with TMY weather should not be aggregated across weather station regions because of the use of non-coincident weather results in artificially muted peaks.



**Figure 33. Illustration of how typical meteorological year (TMY) weather data files are constructed: each month is selected from a historical year for which that month most closely matches that long-term average weather of that month. Source: Kankiewicz and Novotny (2015).**

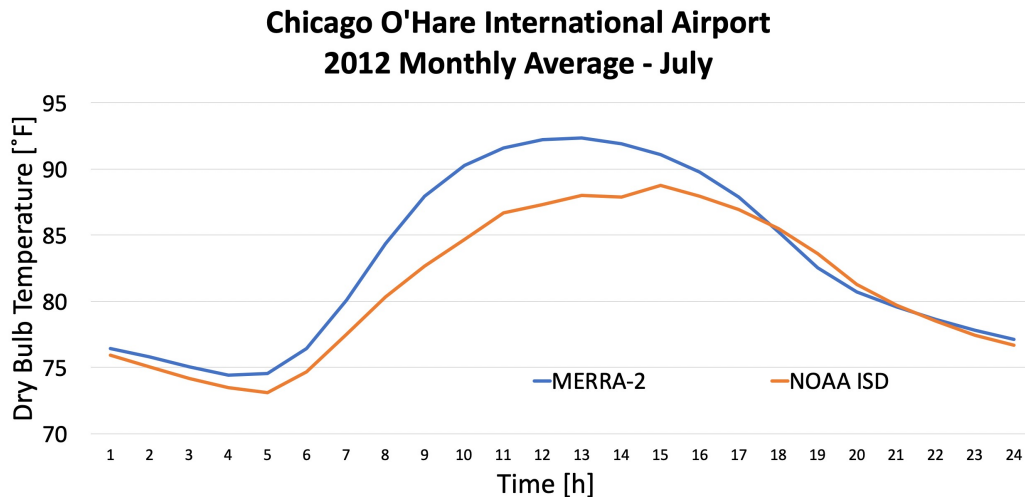
For this project, we constructed our own historical AMY weather data files using ground-based measurements from the Integrated Surface Database (ISD) project of the National Oceanic and Atmospheric Administration (NOAA) National Climatic Data Center (Smith, Lott, and Vose 2011) for most variables (temperature, humidity, wind speed/direction, and pressure) and satellite-derived solar radiation data from NREL's National Solar Radiation Database (NSRDB) (NREL 2021a). Ground-based solar radiation data are not widely collected, so using satellite-derived solar radiation data is standard practice for both the solar industry and building energy modelers using historical weather data. Kankiewicz and Novotny (2015) provide some history of the use of satellite-derived solar radiation data.

#### Limitations of Satellite-Based Weather Data

NREL's NSRDB provides historical weather data—fully derived from satellite data—for more than 1 million grid points across the United States with 4-km by 4-km spatial resolution (NREL 2021a). However, we found that the non-solar radiation variables in the NSRDB, including temperature, humidity, pressure, wind speed/direction, and precipitation, were not accurate enough to use for our building energy simulations. These non-solar radiation variables in the NSRDB are sourced from satellite-based NASA Modern-Era Retrospective analysis for Research and Applications (MERRA) datasets. The MERRA variables are at 40 km<sup>2</sup> spatial resolution, and are interpolated to 4 km<sup>2</sup> for the NSRDB (NREL 2021a).

Early on in this project we explored using the MERRA-based temperature data from the NSRDB as well as MERRA-2 temperature data, which has some minor improvements over MERRA. We realized that the satellite-derived temperature data exhibited some major discrepancies with ground-based measurements, including major differences in the magnitude and timing of temperature peaks. Figure 34 shows an example of this temperature discrepancy by comparing the July 2012 average 24-hour profile for dry-bulb temperature measured at the weather station at

Chicago's O'Hare International Airport with the dry-bulb temperature for that latitude and longitude from MERRA-2. The MERRA-2 temperature data are significantly hotter (up to five degrees) in the morning and early afternoon, and peak about three hours sooner than the ground-measured temperatures. Temperature is not as important for solar modeling (the primary application of the NSRDB), but is very important for building energy modeling. Our findings were consistent with others in the building energy modeling community who have compared MERRA and MERRA-2 satellite-reanalysis data to ground measurements (Huang 2018).



**Figure 34. Comparison of MERRA-2 and NOAA ISD weather: July 2012 average 24-hour profile for dry-bulb temperature measured at the Chicago O'Hare International Airport weather station is compared with the dry-bulb temperature for that latitude and longitude from MERRA-2 (satellite reanalysis). The MERRA-2 temperature data are about five degrees hotter and peak about three hours sooner than the ground-measured temperatures.**

Although the MERRA-derived temperature data have finer spatial resolution than ground-measured data, the errors that would be introduced by the discrepancies described above would outweigh any benefit of using finer-resolution MERRA temperature data. Additionally, we noticed some temperature anomalies along the coast of Los Angeles, which we hypothesized were due to how the 40 km<sup>2</sup> lined up with the coast line (i.e., urban areas on the coast could be assigned temperatures derived from a satellite image spanning from the coast to 40 km out in the ocean).

#### 2.4.2 Process to Generate Weather Data Files

To produce the weather data necessary for this project, we developed a process to combine ground-based measurements from NOAA ISD (Smith, Lott, and Vose 2011) for most variables (temperature, humidity, wind speed and direction, and pressure) and satellite-derived solar radiation data from NSRDB (NREL 2021a). This was done for the more than 1,200 weather station locations that are used for ComStock and ResStock. The majority of the stations we used are consistent with the set of stations used for the TMY3 dataset, but we sometimes needed to use nearby stations because of missing data. Figure 35 shows a map of the station locations.

#### Treatment of Missing Data

Depending on the historical year, some of the 1,200 stations might not provide data at an acceptable level of quality. For each year, and for each station, we started by looking for weather data in the NOAA ISD. Typically, out of 8,760 hourly timesteps for 5 weather variables (dry-bulb temperature, relative humidity, wind direction, wind speed, atmospheric pressure), some timesteps would be missing. Each variable is treated separately. In cases where the data gap is less than three consecutive hours, a linear interpolation is applied to fill the gap. In cases where the gap is longer than three hours, we apply a multistep gap-filling technique.



### **Multistep Gap-Filling Technique**

The first step of the gap-filling technique was to search for data from the same location in the MesoWest database (Horel et al. 2002). MesoWest provides data from the national weather station network similar to NOAA ISD, as well as from additional networks. If gap-filling data are not located in MesoWest for the same location, we search for data from nearby stations (within five miles) in the MesoWest database. If hourly gaps are still present after searching for nearby stations in the MesoWest database, then we check if the gap includes either more than 100 consecutive hours or more than 300 hours total in the whole year, for any variable. If that is the case, then the selected station is discarded. If not, then we apply a naive-forecast method, in which each hour in the gap is filled up with the corresponding hour from a representative 24-hour period created by averaging 48 hours of data before and after the gap.

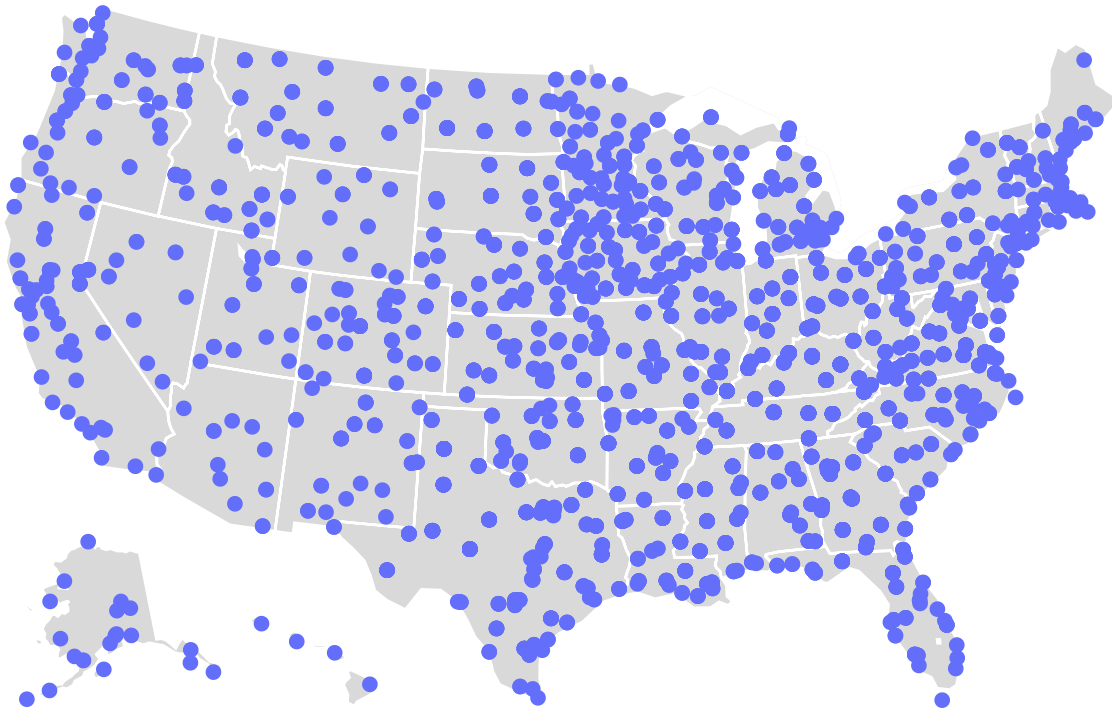
After all gaps are filled, anomalously high or low values are detected and removed using predefined thresholds. For example, if dry-bulb temperature values below  $-50^{\circ}\text{C}$  or above  $60^{\circ}\text{C}$  are detected, the values are removed and a linear interpolation is applied. Finally, a psychrometric check is applied to make sure that all the variables together respect psychrometric relationships.

### **Adding Solar Radiation Data**

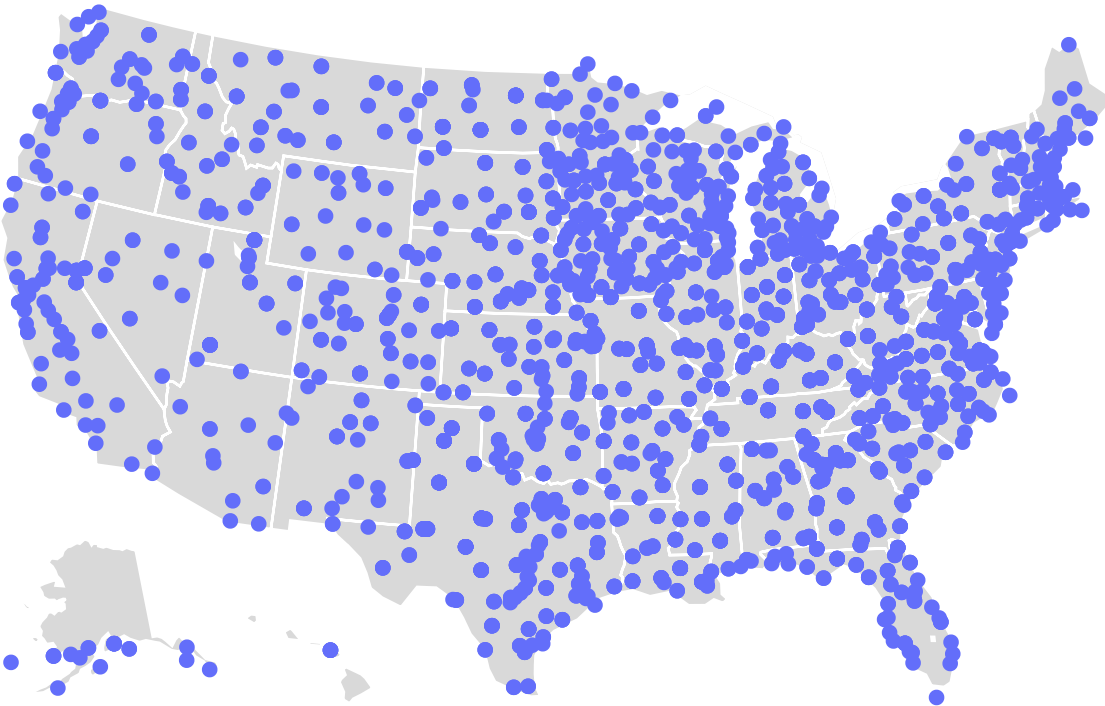
At this point, the following variables are retrieved from the NSRDB for the station location and year: Global Horizontal Irradiation, Diffuse Horizontal Irradiation, Direct Normal Irradiation, Sky Cover, and Surface Albedo (NREL 2021a). These data are modeled data derived from satellite imagery (not measured data, as is the case for the NOAA ISD and MesoWest data) and quality control steps are already applied, so there is no need to further inspect the downloaded data. All of the variables are then merged into a single EnergyPlus weather (EPW) file, which is the weather file format used for EnergyPlus simulations. It is important to note that EPW files require 29 variables, as mentioned in the EnergyPlus documentation (Big Ladder Software 2021). For the 19 minor variables not retrieved from NOAA ISD, MesoWest, or NSRDB, default values are used, as indicated by the EnergyPlus documentation. We do not expect these 19 variables to significantly affect building energy consumption, as reported in the EnergyPlus documentation (Big Ladder Software 2021). This process is repeated for each year and for each of the 1,200 considered locations. For each year, data for 1,000 stations are retrieved, while about 200 locations are fully discarded for lack of sufficient data quality or availability.

### **Assigning Buildings to Weather Stations**

For ResStock and ComStock simulations, we assign each county to a weather station with the following methodology. For each U.S. county, we check if the county contains a station for which we obtained data. If not, we assign the county to the station (among those we retrieved data for in the needed year) that is closest to the county's population centroid. Stations in the same ASHRAE/IECC climate zone are prioritized. If the selected station is in a different time zone than the county it represents, then the weather timestamps are shifted into local standard time for the county. This method led to an average distance between population centroid and the assigned weather station of around 23 miles, with a maximum around 120 miles, for the contiguous United States. All building and housing units in a given county will use that county's associated weather data for simulations.



**Figure 35. Map of the weather stations used for ComStock and ResStock (TMY3). Stations in Alaska and Hawaii are currently only used for ComStock. This map shows locations of the 889 stations for the TMY3 version of the dataset.**



**Figure 36. Map of the weather stations used for ComStock and ResStock (2018). Stations in Alaska and Hawaii are currently only used for ComStock. This map shows the locations of the 1,035 stations for the 2018 version of the dataset; AMY weather sometimes uses nearby stations because of missing data at the 1,200 stations.**

## 2.5 Methodology: Uncertainty Quantification

As defined in Section 2.1.1, uncertainty quantification (UQ) is defined as “the process of quantifying uncertainties associated with model calculations of true, physical QOI, with the goals of accounting for all sources of uncertainty and quantifying the contributions of specific sources to the overall uncertainty.” A recent review on the treatment of uncertainty in building stock energy models conducted by Fennell et al. (2021) concluded that neither uncertainty analysis nor sensitivity analysis are common practice. This review was a product of IEA EBC Annex 70<sup>7</sup>, which is conducting parallel work to address the treatment of UQ in building stock energy modeling; NREL researchers have been active participants in that effort. Despite its importance, UQ of building stock energy models is an emerging field, and there has never been a UQ study applied to a building stock energy model as large or complex as ResStock or ComStock. For the purposes of this project, we seek to develop an approach that captures a majority of the sources of uncertainty while being computationally tractable with our available resources. Our motivation is multifold: (1) to increase stakeholder confidence in the published load profiles; (2) to communicate uncertainty around each of the QOIs; and (3) to advance the practice of uncertainty quantification for building stock energy modeling as well as other complex, high-dimensional physics-based models. Note that the methodology presented in this section is solely about quantifying the uncertainty in model inputs and outputs. Quantification of the uncertainty and potential bias in the empirical load data used for validation is discussed in Sections 2.3.5 and 4.2.2, and resulted in the confidence intervals included in the validation graphics in Section 4.

### 2.5.1 Overview of Uncertainty Propagation Approach

For the EULP project, we developed an approach to propagate uncertainty in our input variables through the models to estimate the resulting output uncertainties. The goal of our UQ effort is to be able to translate uncertainty in our inputs used for model initialization into uncertainty ranges around each of our QOIs in the output. Although there are other potential sources of error, as discussed in Section 2.1.1, we limit our analysis to input uncertainty propagation for a few reasons. Based on the judgment of the project team, we think that the uncertainty in the stock-level input parameters is likely many orders of magnitude higher than other sources of uncertainty, such as the specification of the heat transfer equations in EnergyPlus, for the stock-level QOIs that are the main output of this project. Secondly, as discussed above, UQ of building stock energy models is a nascent field, and there are not yet standard procedures for capturing all sources of uncertainty comprehensively in an assessment (Fennell et al. 2021). Finally, ResStock and ComStock are complex physics-based models that provide rich output detail, but which require significant computational resources to run. More traditional uncertainty quantification techniques require tens or hundreds of thousands of model runs to understand the uncertainty space; in the case of ResStock and ComStock this would translate to hundreds of millions of EnergyPlus runs, which is computationally intractable even with access to high-performance computing resources.

To address this final point, we develop a surrogate (or “emulator”) model for both ResStock and ComStock. Functionally, this is a machine learning model that trains on the inputs and outputs for specific ResStock and ComStock runs to create a quick-running version of the models. Surrogate models are useful in a range of contexts where computational efficiency is prioritized, but they cannot fully replace physics-based models, particularly when answering questions that explore a new parameter space (e.g., building technology upgrade analysis). However, for the purposes of UQ for EULP, we think that a surrogate model is appropriate for QOI uncertainty range estimation, and that the uncertainty introduced by the surrogate is less significant than the potential uncertainty in the input parameters that we seek to quantify.

Our approach is summarized in the following steps:

1. **Parameter Selection:** select the input parameters for propagation based on their significance to the output as well as their levels of uncertainty
2. **Input Uncertainty Characterization:** for each selected parameter, determine the uncertainty distribution
3. **Surrogate Development:** develop emulator or “surrogate” models for ResStock and ComStock
4. **Uncertainty Propagation:** sample from the input uncertainty ranges and run surrogates with the sampled inputs to propagate input uncertainty to output QOIs

<sup>7</sup><https://energyepidemiology.org/>

### 2.5.2 Selection of Input Parameters and Development of Input Uncertainty Ranges

The inputs of ResStock and ComStock are structured differently, with ResStock having more parameters defined through probability distributions and ComStock having more parameters selected via OpenStudio-Standards. Therefore, the selection of input parameters also differs. Based on the developed surrogate models for ComStock and ResStock, important inputs (features) of the surrogate model are identified and input uncertainty range is decided by the statistical distribution of the important features. After deciding those features, the estimators of  $\mu$  and  $\sigma$  are calculated for each selected feature to decide the range of perturbation range for each input. To be more specific, the range of inputs uncertainty multiplier is  $[(1 - c\sigma/\mu), (1 + c\sigma/\mu)]$ , where  $c$  is set to 0.4 for the results presented in this report. We also ran  $c$  values of 0.1 and 0.2. Based on the uncertainty range, a Monte Carlo simulation is conducted to repeatedly generate random inputs based on the calculated range.

### 2.5.3 Development of Surrogate Models

Surrogate model development for building energy modeling is an active area of research, as detailed in the literature review of Zhang et al. 2021. Despite this active interest from the research community, there's no standard for algorithm selection, especially for building stock energy modeling applications as opposed to individual building energy modeling efforts. Because of this, the NREL and ANL buildings teams partnered with NREL's computational science center to leverage additional expertise in developing our approach.

The surrogate models serve as approximations for ResStock and ComStock in producing census division-level QOI results by building type and season. However, our approach actually produces a surrogate at the *individual building* level, and then aggregates the results. Because of the size and complexity of the surrogate model process, we develop a six module framework detailed in a separate publication (Zhang et al. 2021), but summarized here:

1. **ResStock & ComStock:** the ResStock and ComStock models themselves function as the first module, which feeds the surrogate development.
2. **Data Engineering:** ResStock and ComStock inputs and outputs serve as the training data for the surrogate. These datasets are large, so this module includes post-processing of results to prepare the data for the UQ tasks.
3. **Feature Engineering:** from all of the parameters in ResStock and ComStock, we selected and manipulated the most significant variables to the QOIs. These variables overlap with the inputs selected for uncertainty propagation, discussed above.
4. **Surrogate Training:** with the data properly prepared, we next trained the surrogate models. Separate models are trained by building type and location. During this process, we tried a variety of algorithms for the surrogate including random forest, Gaussian process, deep neural networks, decision tree, AutoML, and XGBoost.
5. **High-Performance Computing Initialization:** once we selected algorithms, we developed a module to run on high-performance computing (HPC) resources to handle all phases of the surrogate training, testing, and validation, including handling the substantial amount of data created. The surrogate models produce hourly electricity consumption load profiles for a full synthetic building stock of a region, tagged by building type and a range of characteristics; the results are similar in format to ResStock and ComStock. The HPC module can run on either NREL's Eagle supercomputer or Argonne's Bebop supercomputer.
6. **Quantity of Interest Post-Processing:** as a final module, we calculate QOI for our simulated building stock. For the project, we calculate QOIs by census division and building type, but they could also theoretically be calculated for other aggregations of buildings. In each Monte Carlo run or each set of randomly perturbed inputs, we calculate the QOIs including annual total energy, seasonal base load, seasonal peak load, seasonal peak hours, etc. With a number of runs of Monte Carlo simulation, same number of sets of QOIs are generated, and we further aggregate these QOIs into QOI uncertainty ranges by region and building type, which are the final outputs of the developed UQ methodology.

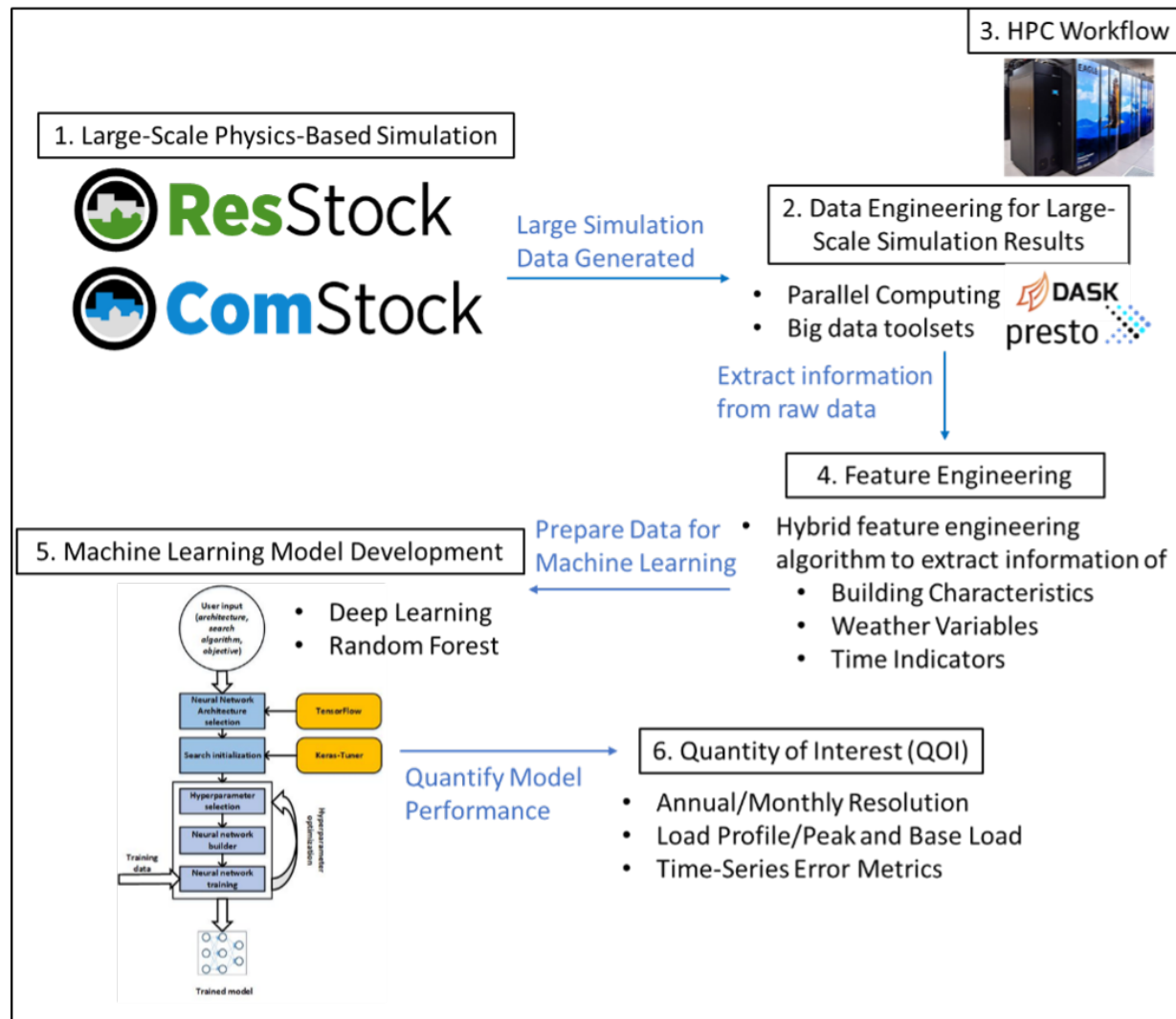


Figure 37. Graphical summary of the surrogate framework (From Zhang et al. 2021)

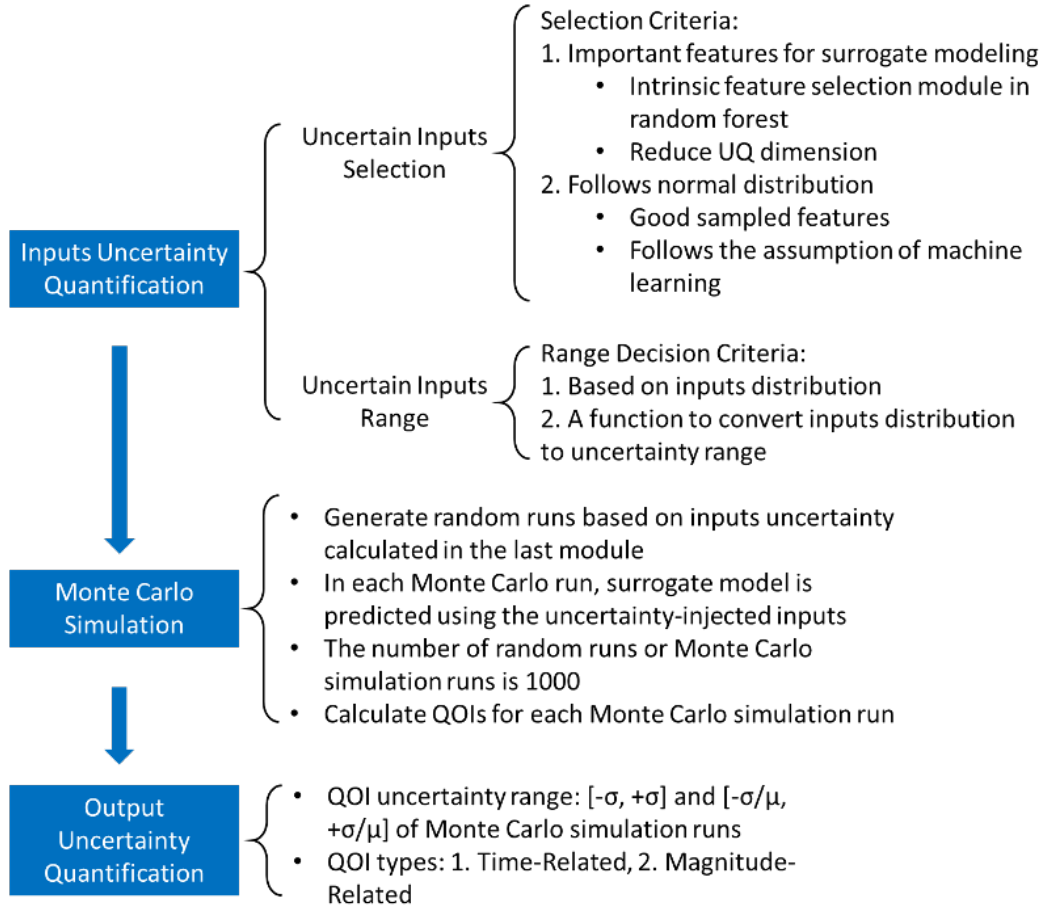
## 2.5.4 Application of Surrogate Models to Uncertainty Propagation

On top of the modules developed for the surrogate training, we layer on additional features to create a fully functional workflow for not only training/testing/validating the surrogate, but also propagating the input uncertainties. Functionally, this is taking the workflow listed in the above section and batching it tens of thousands of times for each of the sampled input uncertainty values. Each individual run of the workflow is called a “replicate” and the result is a single value for each of the QOIs (by building type, season, and census division). When run with different sampled input uncertainty values, each replicate run represents an alternate potential value of the output values. Looking across the range of each of these QOIs, we can understand the distribution of different alternate versions of that QOI, given the different combinations of the sampled input uncertainty. In summary, the full propagation process follows the following steps:

1. Process data from final national ResStock and ComStock runs
2. Train the surrogate models by census division (9) and building type (21) for a total of 189 trained surrogates
3. Batch run the surrogate to create 1,000 replicates (i.e., alternate national results) for each surrogate by sampling the input parameters and using the model to predict the electricity timeseries for each building within ResStock and ComStock with the alternate inputs

4. Aggregate each replicate timeseries result to the QOIs
5. Quantify QOI uncertainty ranges by building type and census division

A summary of the UQ process is shown in Figure 38.



**Figure 38. Monte Carlo simulation methodology**

The results of uncertainty propagation are presented in Section 4.4, Uncertainty Quantification Results.



## 3 Model Calibration

### 3.1 Sensitivity Analysis

#### 3.1.1 Methodology

##### Computational Expense of Traditional Sensitivity Analysis

Building energy models contain hundreds of variables with differing importance across building types that determine how and when a building uses energy. Many of the parameters are interdependent, such as the window properties and window-to-wall ratio. This project uses a parameter importance analysis to narrow down which influential parameters to adjust for calibration. This analysis is conceptually similar to a sensitivity analysis. Sensitivity analysis methods generate a sample space and evaluate a function against that sample space. In this case, the sample space is the set of inputs to the building energy models, and the function is the building stock energy model. As discussed in the previous section, ResStock and ComStock simulations are computationally expensive, and a full factorial analysis of even a small set of parameters with few options would be computationally prohibitive. Methods to sample the input parameter space greatly reduce the number of simulations required, but it would still be computationally intractable to run the full ComStock or ResStock model thousands of times. Given the computational expense and that ComStock and ResStock are already a statistical sample of buildings with representative distributions of input parameters, this project uses random forest models, trained with ComStock and ResStock results, to determine parameter importance. This is a novel approach for dealing with building stock energy modeling sensitivity, as most of the previous work uses smaller samples or stock models with more traditional, computationally intensive approaches (Fennell et al. 2021). We worked closely with the UQ subtask of IEA Annex 70, which comprises international experts in building stock energy modeling, to vet our approach. Significant parameters to adjust for calibration will be those whose features show up as highly significant in the random forest model.

##### Quantities of Interest (QOI)

Most building energy modeling sensitivity studies done to inform calibration use total annual energy use as the output variable (Coakley, Raftery, and Keane 2014, Reddy 2006b). As discussed in Section 2.1.2 on QOI, there are several characteristics of the output energy timeseries we are interested in to create accurate load profiles. This analysis uses electricity QOIs as the random forest model output variables, the same as the metrics used for calibration.

##### Parameter Selection

Only a subset of the hundreds of input parameters are particularly impactful, but this varies by building type, location, and other high-level stock characteristics. Several studies have put forward parameters to use for commercial building sensitivity analysis (Yang and Becerik-Gerber 2015, Xu 2012, Bertagnolio 2012). The parameters used for the ComStock analysis are drawn from these studies, as well as additional parameters to include a wider range of end uses and building types.

For ResStock, most of the key inputs to the model are articulated in the conditional probability distribution input network; furthermore, these are the same parameters we modify to calibrate ResStock. We therefore train the residential random forest models on the ResStock input distributions. To ensure that we did not overlook any key EnergyPlus inputs that are not articulated in the ResStock input network, our ANL partners performed an EnergyPlus-level sensitivity analysis in three sample locations of varying climate and stock characteristics using the Morris method, a more traditional sensitivity approach. This analysis found that all variables significant to the EnergyPlus models are articulated in the ResStock input distribution network.

**Table 12. Summary of how parameters were selected, and the parameter types, for the ComStock and ResStock sensitivity analyses**

	ComStock	ResStock
Parameter sources	Literature (Yang and Becerik-Gerber 2015, Xu 2012, Bertagnolio 2012), supplemented with additional analysis (Dahlhausen 2020)	Existing input probability distributions; verified using Morris method
Parameter types	Categorical and continuous; see Chapter 6 of Dahlhausen (2020) for details	Categorical; encoded on a case-by-case basis using engineering expertise

For ComStock, there are a mix of categorical and continuous parameters. Categorical parameters were encoded using a mix of one-hot encoding and mean encoding. For ResStock, the input distributions are categorical, and a mixture of quantitative and qualitative. Each individual input went through a variable encoding process using the engineering expertise of the NREL research team. For example, walls may be specified by detailed thermal properties of each component, or an overall thermal resistance, whereas equipment efficiency used a derivative of its performance rating. All values are normalized before model training. This analysis abstracts away many of the detailed model inputs to arrive at more general parameters for the analysis. A full list of building parameters used in the ComStock regression model is available in Dahlhausen (2020). This dissertation also discusses the qualities of building parameters for ComStock, including how to handle categorical vs. continuous parameters, implicit vs. explicit parameters, schedule parameters, and parameter interaction. For ResStock, the parameters used include all of the main inputs to ResStock as described in the input distribution network.<sup>8</sup> ResStock inputs that are attributional tagging only, and that do not directly describe physical characteristics of the housing stock (e.g., vintage, geography) are not included.

### **Selection of Random Forest Approach to Sensitivity**

For both ResStock and ComStock sensitivity, we train random forests using the respective input parameters of each model. Random forest models are often used for output prediction, and in this case, we train random forest models by calibration region to predict each of the QOIs. An output of the random forest approach is that we obtain a relative ranking of the significance of each of the inputs in predicting the output—each of the QOIs. We use the parameter importance metric to rank the significance of all the inputs for a given QOI. The relative weight of feature importance determines how strongly a given QOI is influenced by that feature. There are also many other models and methods that could have been used for this type of sensitivity work. Prior studies have typically used one-at-a-time or other simple linear regression methods (Reddy 2006b). Often, these are not able to properly capture parameter interaction, or can result in model overfitting if they use artificial features composed of multiple inputs to capture interaction. For these reasons, this analysis uses a random forest method (Breiman 2011), which is decent at capturing parameter interaction. We also avoid overfitting, because random forest is an ensemble method, where many different decision trees are fitted to the same data, and then the aggregated model is used. Implementation details are discussed in Dahlhausen (2020).

#### **3.1.2 Sensitivity Analysis Results**

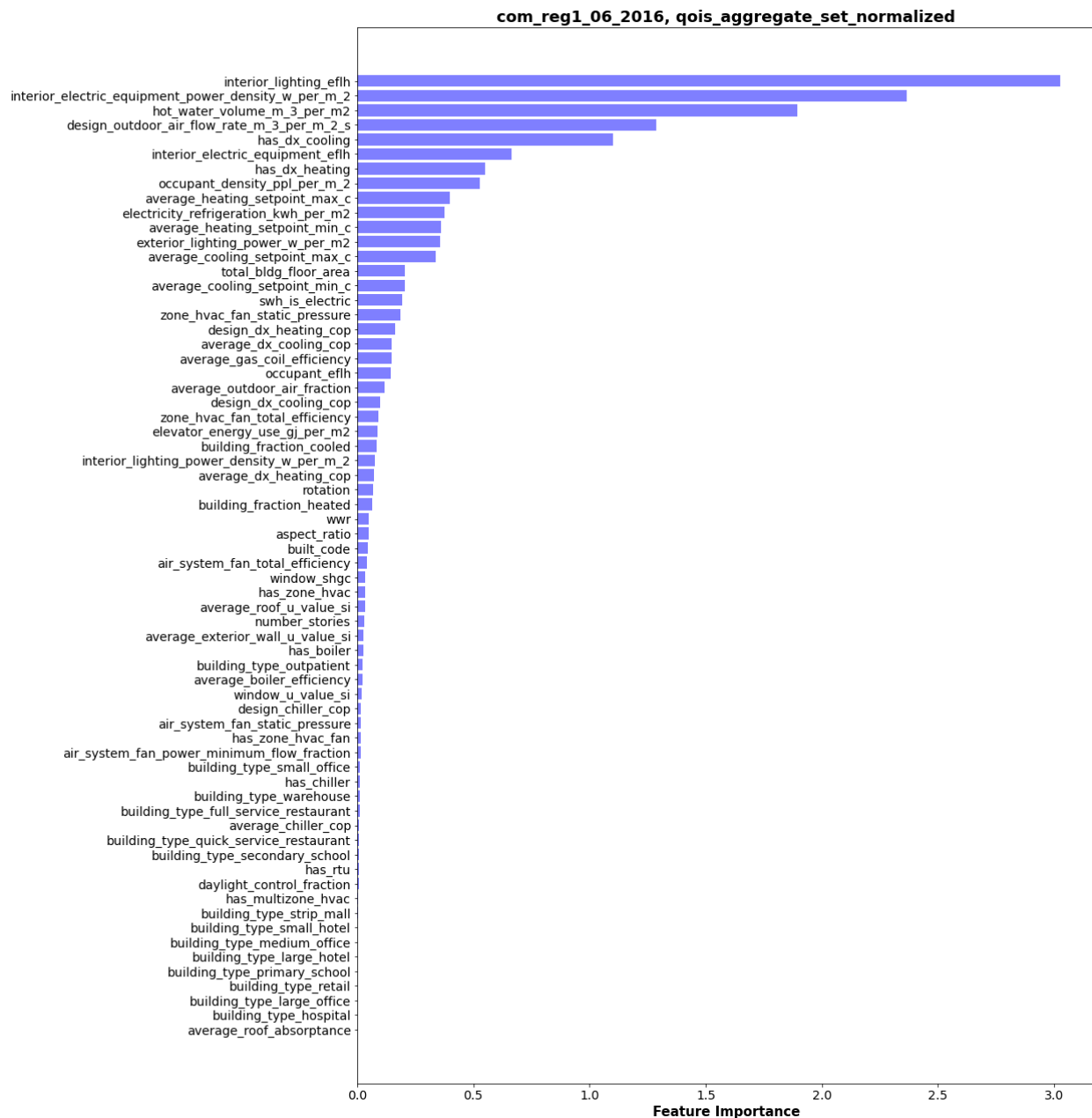
The results of the sensitivity analysis include rankings of feature importance (summing to 1) for each QOI, as well as the random forest model training and testing accuracy. QOIs can be calculated at the individual building level, such as influence of a given parameter to the summer peak demand magnitude for an individual given building, or in relation to the aggregate building stock, which is the parameter's importance to the overall peak of the electric grid. ComStock sensitivity results are normalized by floor area because building size is by far the most significant predictor of QOIs for commercial buildings. ResStock sensitivity results are presented on a per-housing-unit basis. To obtain an overall ranking, the feature importance is summed across all QOIs to get an overall average feature importance for each parameter. In the next two sections, we present samples of the sensitivity analysis for some regions to illustrate the types of insights that we gain from this analysis.

##### *ComStock Sensitivity Results*

Figure 39 shows the ranking of feature importance of ComStock model parameters for Fort Collins, CO. In this case, the individual building magnitude values are calculated in relation to the aggregate grid peak and normalized by floor area. Interior equipment and lighting densities and schedules, outdoor air flow rate, heating setpoints, occupant density, and the presence of rooftop units are the most significant parameters. This suggests special attention should be paid to internal load power densities and schedules, thermostats setpoints, and roof top units for calibrating the Fort Collins, CO, stock model.

A similar analysis for a national ComStock run is shown in Figure 40. In this case, buildings are grouped by climate zone to determine their peak QOI values in relation to a grid peak. There are a few notable differences at the national scale compared to results for a specific region. Because the nation contains a mix of heating-dominated and cooling-dominated climates, HVAC parameters show up as less significant because heating-related parameters are more significant in some regions and cooling-related parameters are more significant in other regions. Hot water use

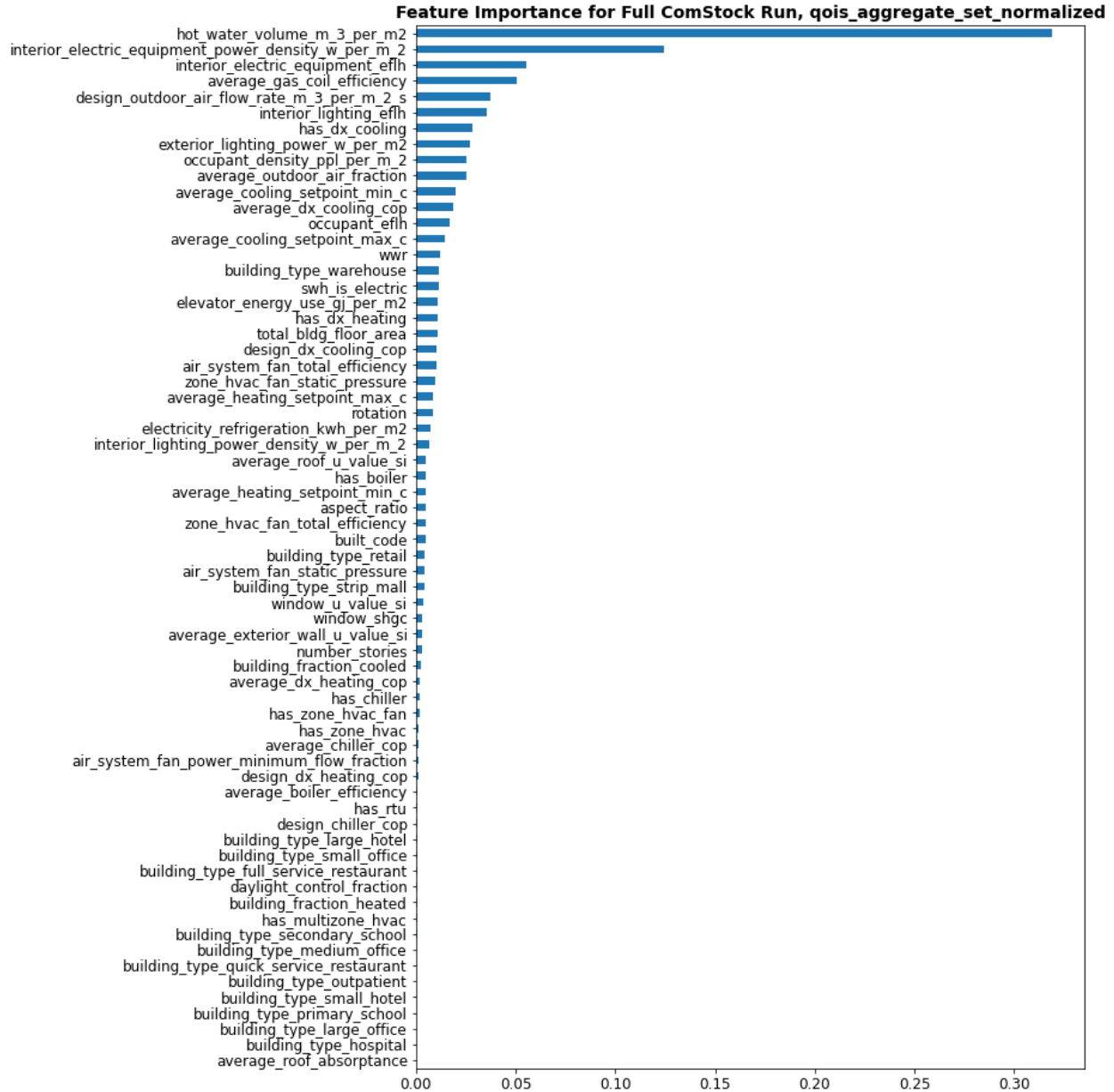
<sup>8</sup>[https://github.com/NREL/resstock/tree/develop/project\\_national/housing\\_characteristics](https://github.com/NREL/resstock/tree/develop/project_national/housing_characteristics)



**Figure 39. Weighted feature importance of input parameters, aggregated across QOIs, for Fort Collins, CO. Interior equipment and lighting densities and schedules, outdoor air flow rate, heating setpoints, occupant density, and the presence of rooftop units are the most significant parameters, which suggests that special attention should be paid to internal load power densities and schedules, thermostats setpoints, and roof top units for calibrating ComStock for Fort Collins, CO.**

density and gas coil efficiency show up as top parameters. In this case, these are not directly causal, but are highly correlated with other characteristics of a building. Hot water use is highly correlated with high equipment power density in restaurants, and the electric equipment parameter itself is not able to capture this because of the bimodal nature of the distribution, with restaurants having much higher equipment loads than other building types. Gas coil efficiency has two possible values, 78% or 80%, depending on the code year, making it the strongest correlate of building code year. Built code year itself is not as high, largely because retrofits of other equipment can occur in the model, so gas coil efficiency is a better marker of HVAC system age (in the model). The national sensitivity results

indicate that a better accounting of kitchen equipment—making it explicit in the model, and making sure we model code year and retrofit frequency accurately—would be important for calibrating the full stock model. As with the results for Fort Collins, CO, interior equipment and lighting densities/schedules, outdoor air flow rate, occupant density, and the presence of rooftop units were also highly significant parameters.

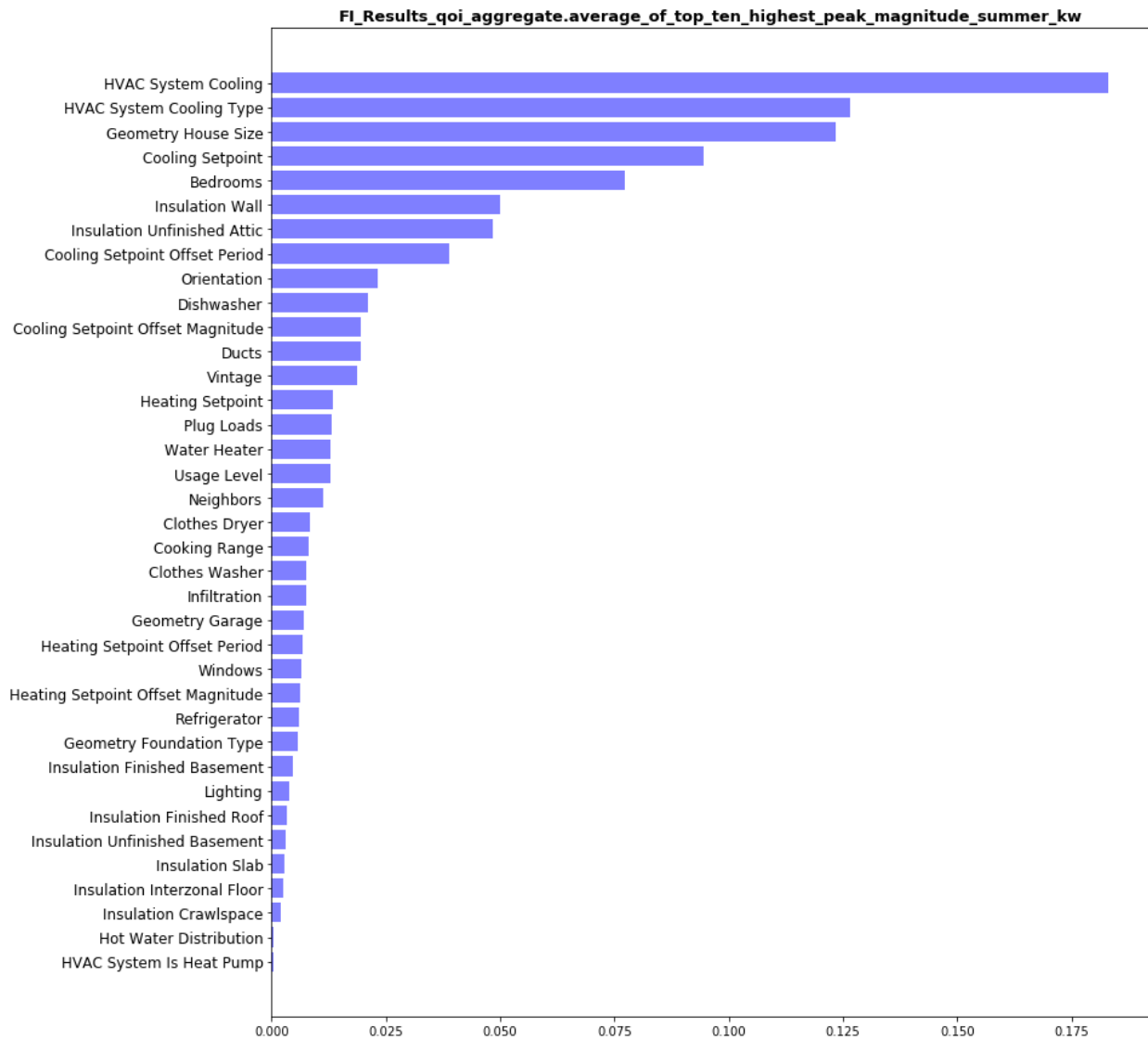


**Figure 40. Weighted feature importance of input parameters, aggregated across QOIs, for the full U.S. commercial building stock. Most significant are hot water use density, interior equipment and lighting densities/schedules, gas coil efficiency, outdoor air flow rate, occupant density, and the presence of rooftop units.**

### ResStock Sensitivity Results

In Figure 41 we show sample ResStock sensitivity results for the ComEd service territory, the first calibration region for ResStock. In this example, we highlight the relative feature importance of ResStock inputs that influence the top 10 peak summer days in aggregate. In this particular sample, we only show single-family detached homes that have non-electric heating systems (the most prevalent residential housing type in the ComEd service territory). In-line

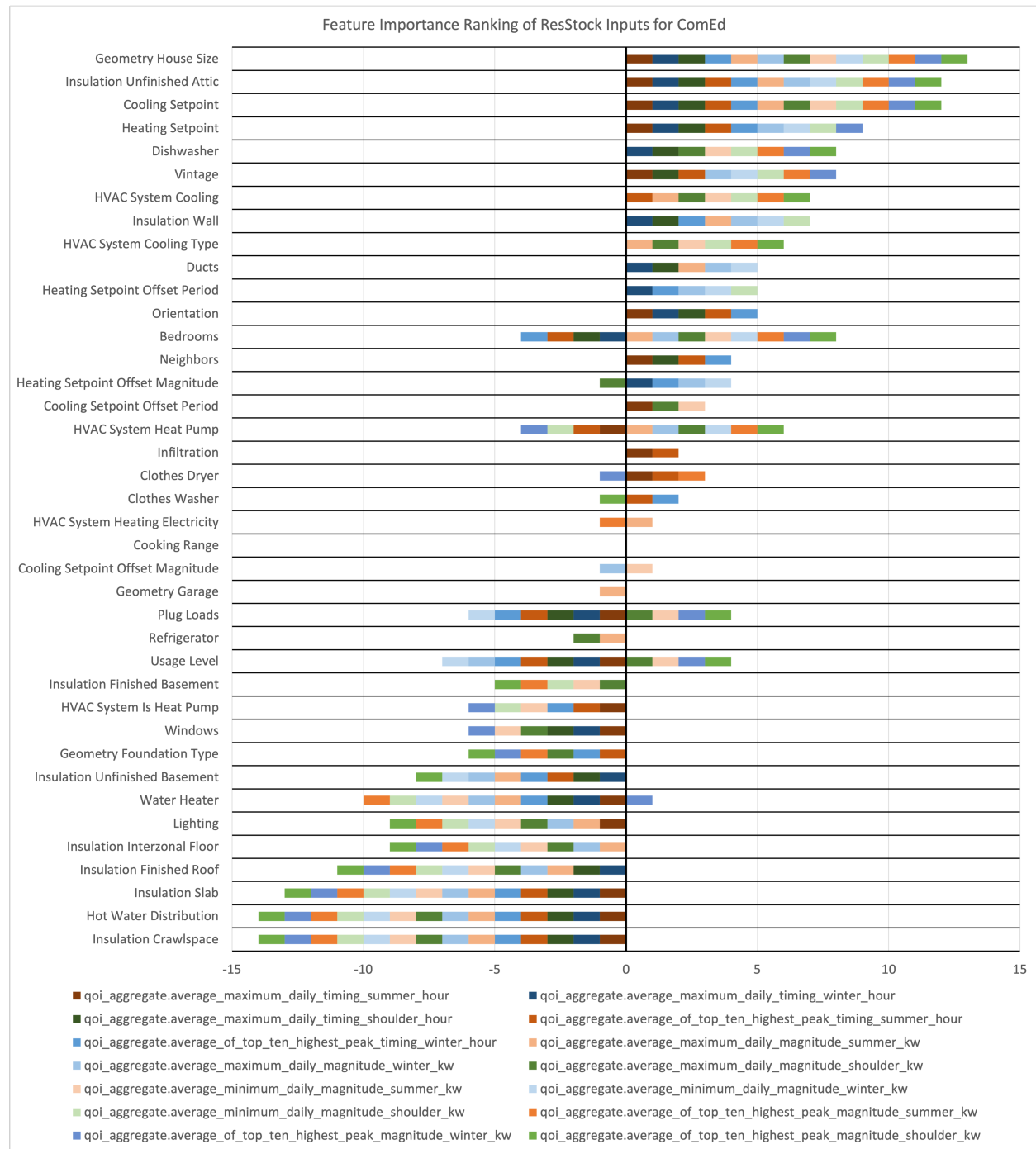
with the rest of calibration, the ComEd territory was disaggregated by customer class: Single-Family Non-Electric Heating; Single-Family Electric Heating; Multifamily Non-Electric Heating; and Multifamily Electric Heating.



**Figure 41. ResStock summer peak magnitude QOI sensitivity for single-family detached homes with non-electric space heating**

In this example, the top contributing ResStock inputs to summer peak magnitude make intuitive sense. The cooling system presence, type, and efficiency rank highest, followed by the floor area of the home and the cooling setpoint.

When we look at the average prevalence of the different parameters across all of the QOIs for ComEd (Figure 42), we can see the parameters of most significance for the residential sector. Floor area, attic insulation levels, and cooling setpoint are all important for more than 12 of the total 14 QOIs. In the Chicago area, space conditioning is a major end use due to both the cold winters as well as the hot summers, so it is unsurprising that the majority of the most significant QOIs are related to space conditioning. The feature importance ranking is weighted slightly more toward cooling than heating than would be expected given the climate, except that all the QOIs are *electricity* focused, and the vast majority of heating systems in the ComEd service territory use natural gas.



**Figure 42. Importance ranking of ResStock parameters across all 14 QOIs. For this summary plot, if a parameter is in either the top 10 or bottom 10 for a QOI in terms of feature importance, it receives a positive or negative count, respectively. This allows us to distinguish input parameters that are broadly important across a range of QOIs as well as the ones that have little to no significance.**

### 3.1.3 Application to Calibration and Validation

The sensitivity analysis served as one of many pieces of information during the calibration process. During calibration, if ResStock or ComStock had a discrepancy in estimating a certain QOI, the random forest sensitivity analysis could provide a ranked list of the input parameters that were most influential (i.e., where a change in the input would



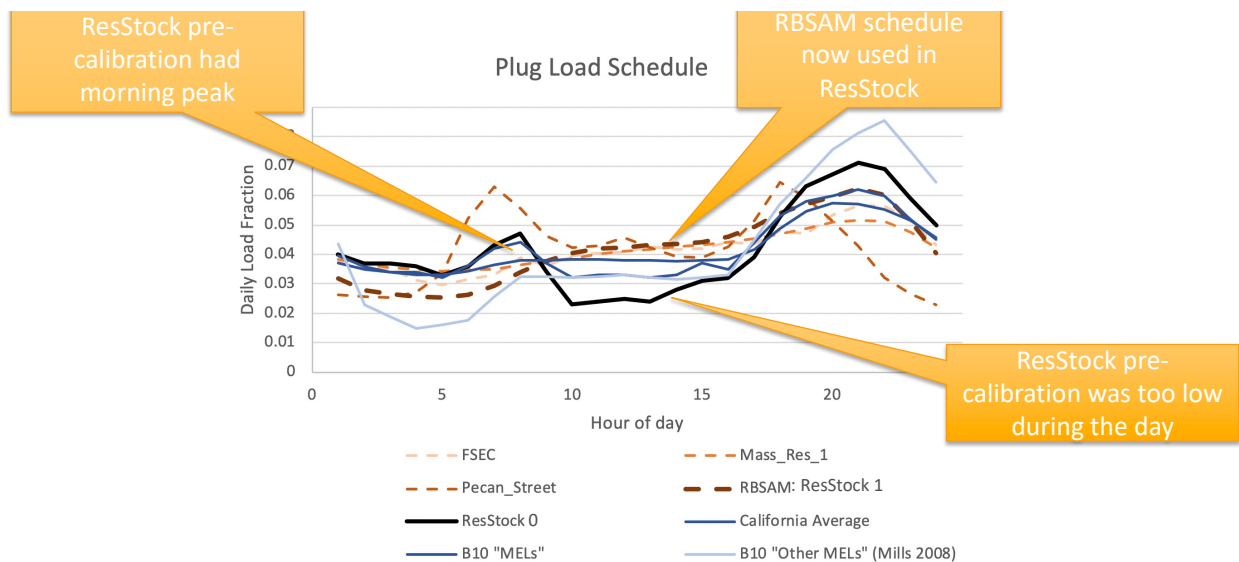
have the biggest impact on the QOI). However, having a parameter ranked as influential to a QOI didn't necessarily indicate either a problem or show *how* that parameter could be changed. It just indicted a list of possible inputs that could be changed to influence the output. For this reason, the sensitivity analysis was coupled with other techniques, such as comparison to AMI or submetered data, as well as the engineering experience of the project team. Additionally, as briefly discussed in the Methodology section, this sensitivity approach is limited in that it can only report sensitivity to parameters that we expose as inputs and use to train the random forest models. The Morris method analysis (discussed previously) gives us some confidence that we've exposed the relevant EnergyPlus inputs in our chosen parameters, but coupling this approach with domain knowledge and modeling experience is essential to get a full understanding of calibration parameter priorities.

## 3.2 Residential Model Updates

### 3.2.1 Occupant-Driven Schedule Updates

#### Plug Load Schedule

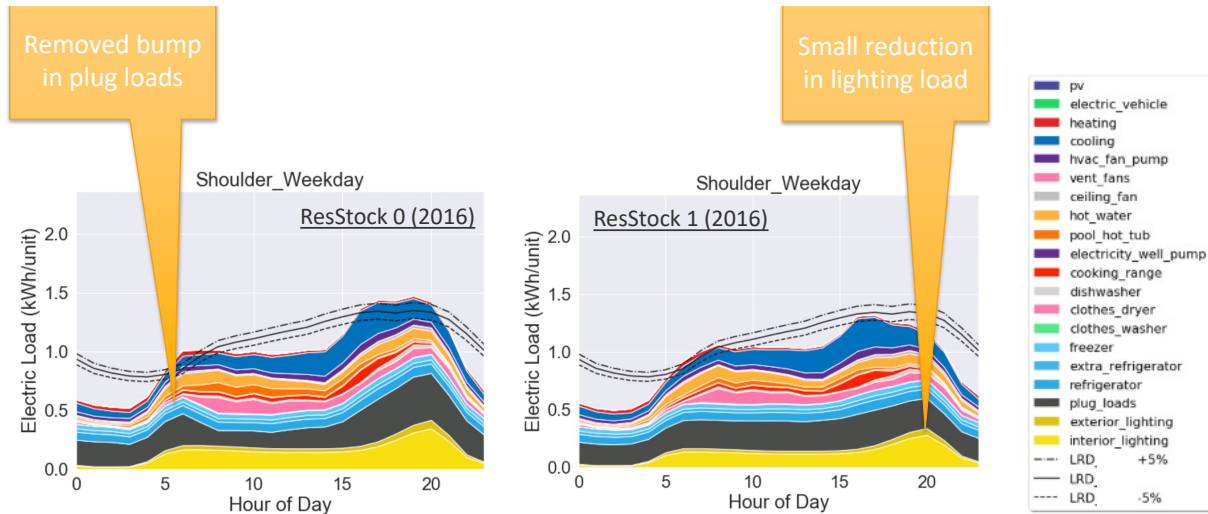
ResStock inherited a plug load schedule from the 2014 Building America House Simulation Protocols (Wilson et al. 2014), which includes the modified version of the plug load schedule from Mills (2008) that was also used in BEopt™ (Christensen et al. 2005). We noticed that ResStock outputs exhibited a faster morning ramp-up of load when compared to AMI data and end-use metered data. As part of our residential end-use transferability study, we compared plug load (also known as miscellaneous electric load, or MEL) schedules from a range of sources, as shown in Figure 43. FSEC, Pecan Street, Mass RES 1, RBSAM (Larson et al. 2014), and California Average (RASS) are all end-use submetering studies, and B10 “MELS” and “Other MELS” (Mills 2008) are for historical context.



**Figure 43.** We compared plug load (also known as miscellaneous electric load) schedules from a range of sources. FSEC, Pecan Street, Mass RES 1, RBSAM (Larson et al. 2014), and California Average (RASS) are all end-use submetering studies, and B10 “MELS” and “Other MELS” (Mills 2008) are for historical context. Of the end-use submetered sources, all except for Pecan Street are quite similar. We suspect Pecan Street has some degree of sampling bias because many of the homes are in the same neighborhood.

Of the end-use submetered sources shown in Figure 43, all except for Pecan Street are quite similar. We suspect Pecan Street has some degree of sampling bias because many of the homes are in the same neighborhood. As illustrated in Figure 19, there is substantial overlap between the 80% confidence interval ranges of the end-use plug load shapes. Remaining differences could be due to regional differences in plug load behavior or some other difference in the sets of sampled homes.

We selected the RBSAM (Larson et al. 2014) profile to use for all homes in ResStock, because we know that it is a representative sample of homes across four states. The impact of this change on the plug load profile on shoulder season weekdays in ComEd territory can be seen in Figure 44; the morning bump and daytime valley are smoothed out. We later used data from the American Time Use Survey to modify plug load schedules based on home vs. away occupancy patterns (see *New Residential Stochastic Occupant Behavior Model*) and also to shift occupancy patterns forward and backward in time based on regional patterns (e.g., time zones; see *Occupant Schedule Shifting Based on Time Use Data*).



**Figure 44.** This figure shows how the changes in lighting technology saturation and plug load schedule affect the EULP on shoulder season weekdays in ComEd territory. The morning bump and daytime valley in plug loads are smoothed out and there is a small reduction in lighting load. Note: final comparison is shown in Section 4.

#### *New Residential Stochastic Occupant Behavior Model*

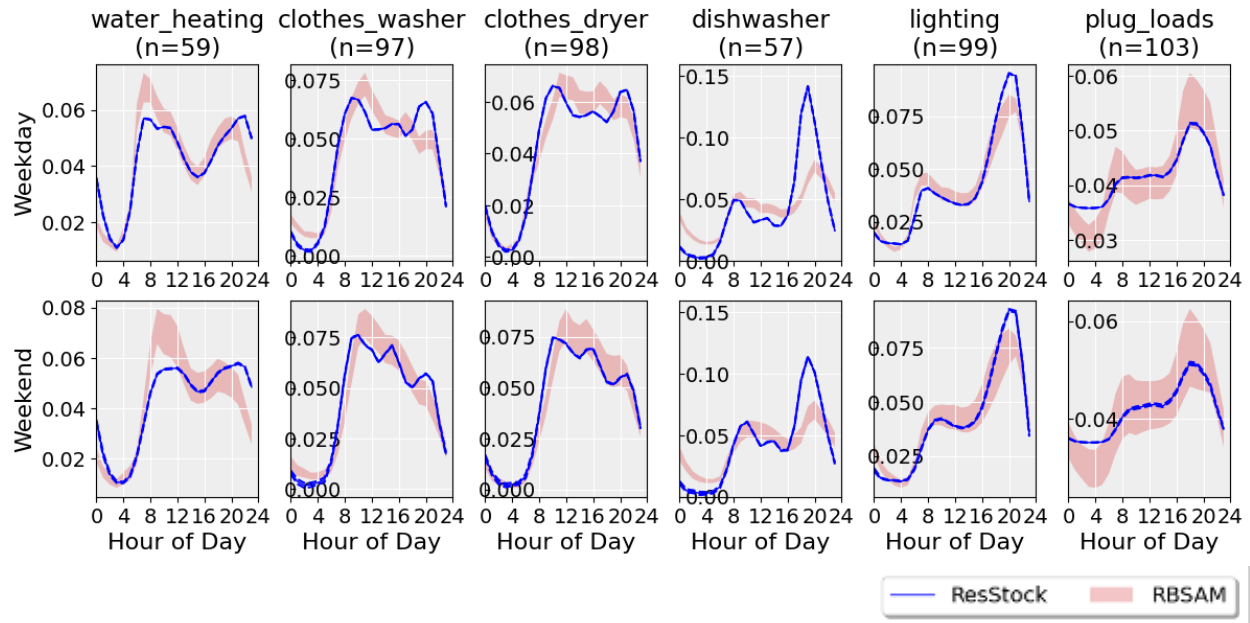
To improve the realism of individual housing unit load profiles, we developed a new residential stochastic occupant behavior model. The methodology for the residential stochastic occupant behavior model is documented in Chen et al. (2022). To develop this model, we first conducted an extensive literature review (Chen, Polly, and Wilson 2022). As with previous approaches in the literature, we rely on time-use survey data—in our case, the American Time Use Survey. We uniquely supplement the time use survey data with other data sources for the magnitude and duration of power and hot water draws, to model the effect of occupant behavior on electricity, gas, and hot water use consistently across a wide range of end uses: clothes washing, clothes drying, cooking, dishwashing, sinks, showers, baths, lighting, and plug loads. We reviewed the performance of three different behavior modeling approaches by comparing with ATUS source data. The best performing approach was a hybrid that uses Markov chains for activity selection, combined with probabilistic sampling of event durations. We developed an OpenStudio measure to integrate this stochastic behavior simulator into the ResStock workflow.

#### **Validation of Residential Stochastic Occupant Behavior Model**

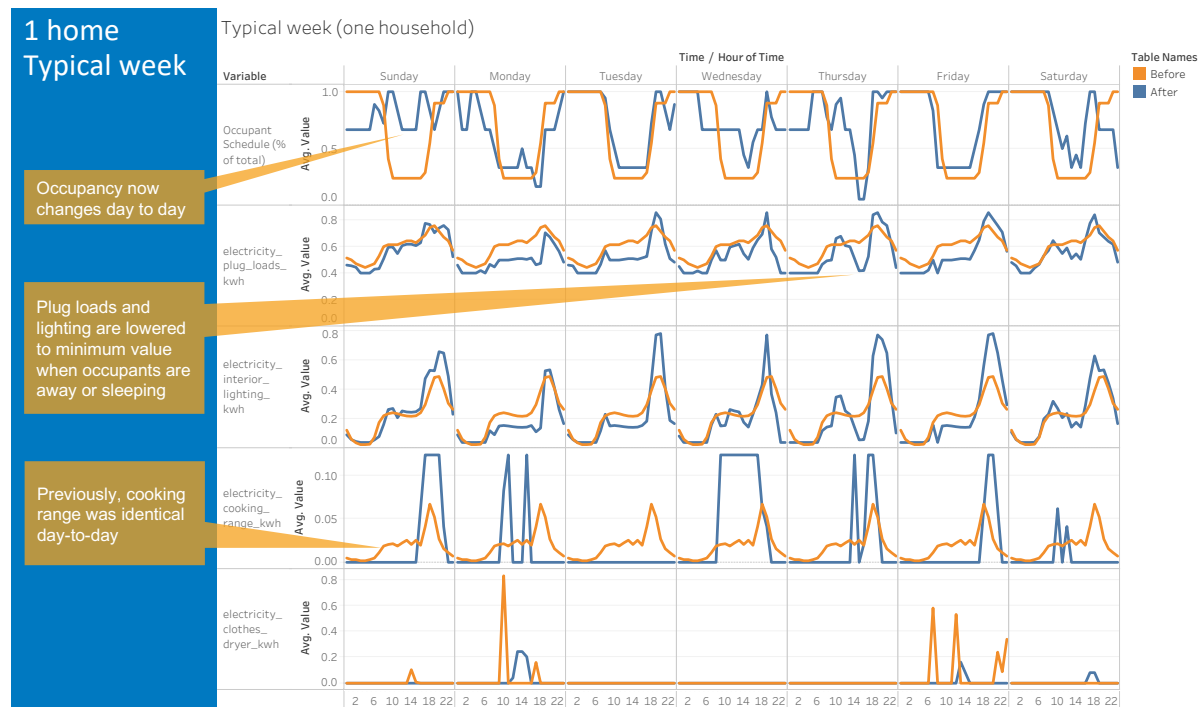
As shown in Figure 45, the simulated residential building loads were validated against circuit-level metering data from the Northwest Energy Efficiency Alliance’s 2011 Residential Building Stock Assessment (RBSA) Metering Study, which submetered end-use energy consumption of appliances and other circuits with 15-minute resolution in around 100 households (Larson et al. 2014). The largest discrepancy is that the ATUS-based model has a large spike in dishwasher usage in the evening, whereas RBSAM shows usage more spread out during the day and through the night. This could potentially be corrected by introducing a delay between ATUS-reported dishwashing activity and dishwasher energy and water use. See Chen et al. (2022) for more details on the residential stochastic occupant behavior model developed for this project.

#### **Impact of Residential Stochastic Occupant Behavior Model**

The impact of changing to the new stochastic schedule generator can best be seen by looking at EULPs before and after the change for a typical week in an example home (Figure 46). In the top row, we can see that the household’s home/away presence schedule now changes from day to day, which is also used to modify the previously static plug load (2nd row) and lighting (3rd row) schedules. The cooking range schedule was previously identical day-to-day, but now exhibits discrete cooking events (4th row).



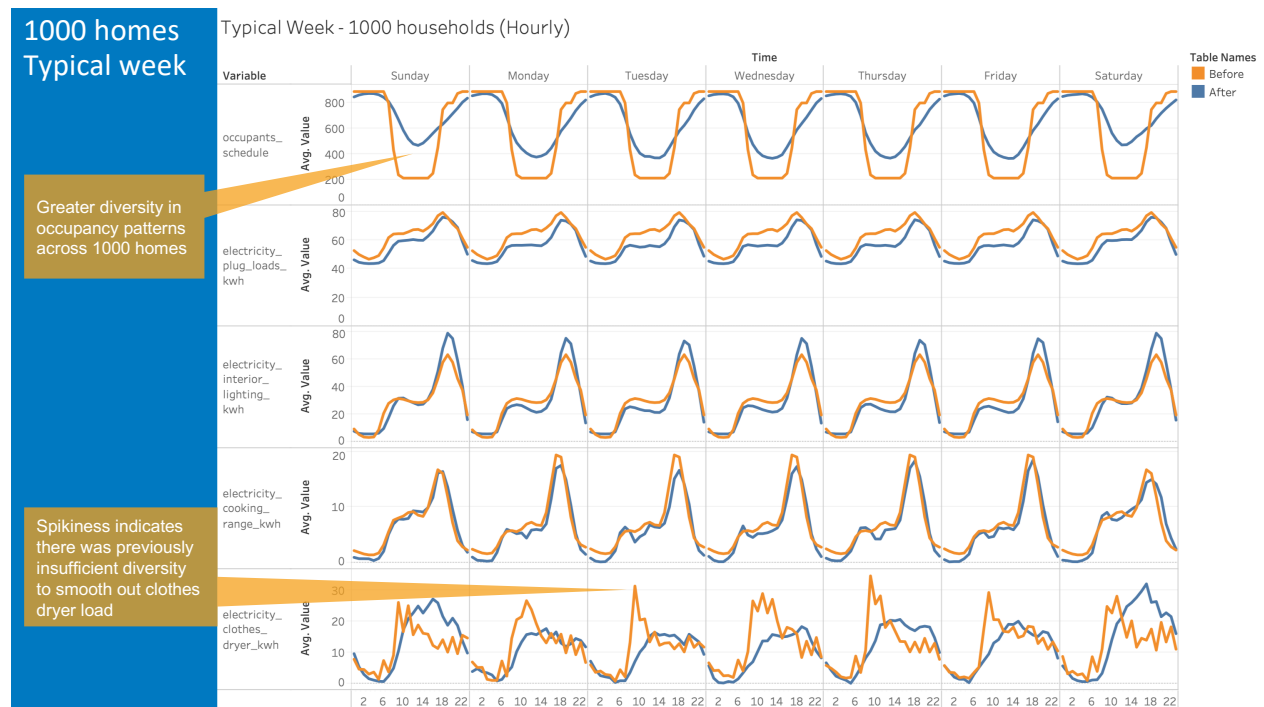
**Figure 45. Validation of residential stochastic occupant-driven loads: comparison of normalized average weekday and weekend profiles for ResStock simulation results and RBSAM measured data, which is shown as a 95% confidence interval of the mean ( $\text{mean} \pm 1.96 \times \text{standard error}$ ) for six end uses**



**Figure 46. Impact of new residential schedule generator on EULPs for a typical week in an example home**

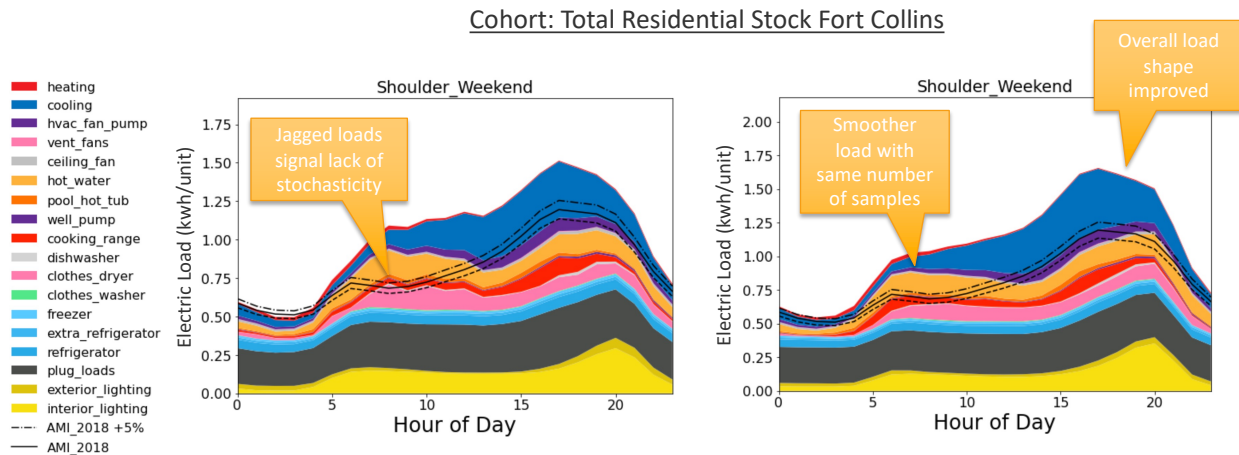
The impact at an aggregate scale of 1,000 example homes is shown in Figure 47. At this scale, the most important feature is that the new model avoids erroneous coincidence of clothes dryer loads, due to the previous clothes dryer schedules having insufficient diversity to smooth out the aggregated clothes dryer load shape (i.e., only 20 schedules

were used for 1,000 homes). Additional figures showing both typical and average weeks for one and 1,000 homes for these end uses and five domestic hot water schedules are included in Appendix E.



**Figure 47. Impact of new residential schedule generator on EULPs for a typical week in 1,000 example homes**

We also evaluated the impact of the new stochastic schedule generator on our comparisons against AMI data. Figure 48 shows before (left) and after (right) the change for an average shoulder season weekend 24-hour profile. We focused on a shoulder season comparison because occupant-driven loads are most clearly seen during shoulder seasons. There was not much change to the overall magnitude of each end use, which is expected; the main purpose of the new occupancy simulator was to improve the accuracy at an individual building or distribution transformer scale. We did see some of the jagged appliance end uses get smoothed out, which was desired and consistent with findings from the 1,000-home impact analysis in Figure 47. The schedule generator change accidentally introduced a bug that increased the magnitude of internal gains used for the design cooling load calculation used to size air-conditioner capacities. This did not significantly affect annual energy use, only peak demand (about 1% of hours). This was found and fixed during the Residential Region 3 phase of calibration (see *Air-Conditioner Sizing Fix*).



**Figure 48. Impact of new residential schedule generator on Fort Collins AMI comparison, before (left) and after (right) the change for an average shoulder season weekend 24-hour profile. Note: final comparison is shown in Section 4.**

#### *Occupant Schedule Shifting Based on Time Use Data*

One limitation of the residential stochastic occupant behavior model presented above is that it did not account for how occupant behavior patterns may vary in different parts of the country. Correlating occupant behavior with socio-economic/demographic household characteristics is an area of ongoing work for our team, but for this project, there were two specific geographic behavior differences we were aware of and sought to account for. One difference is behavior patterns of households in urban, suburban, or rural areas. After observing that residential load research data from Consolidated Edison, which serves New York City, tapered off later in the evening than other utilities, we suspected that this could be true of other urban areas as well. The other geographic behavior difference is the effect of location within a time zone, which has been shown to affect bedtimes (Giuntella and Mazzonna 2019).

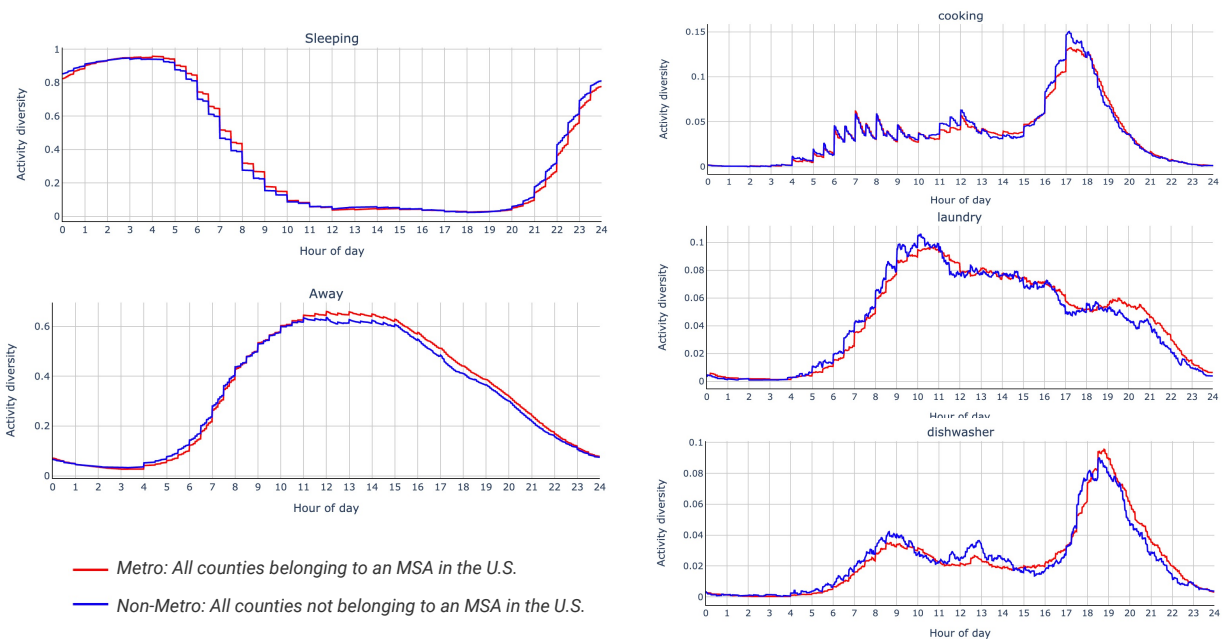
#### **Urban vs. Rural Behavior Patterns**

The residential end-use transferability analysis presented in Section 2.3.9 did not conclusively find differences between the five end-use datasets that could be attributed to regional differences in occupant behavior, so we found other data sources to address these geographic behavior differences. Using ATUS data, we investigated differences in behavior patterns between counties located in metropolitan statistical areas (MSA) and those elsewhere (Figure 49). MSA schedules appear to be shifted about 0–30 minutes later in local time, though we suspect that because MSAs and counties combine urban households with suburban and exurban households, they do not have adequate resolution to isolate the suspected behavior differences. In an attempt to isolate an extreme example, we compared sleep and away schedules for the borough of Manhattan to counties in New York State outside of the five boroughs of New York City, finding that Manhattan sleep schedules are about 30 minutes later and away schedules are about 60–90 minutes later than non-NYC New York (Figure 50). However, most counties do not isolate urban downtown areas like we could with the Manhattan example. Therefore, for the initial EULP dataset, we do not attempt to include the urban behavior effect, and we plan to investigate this further through work correlating behavior patterns to socio-demographic variables.

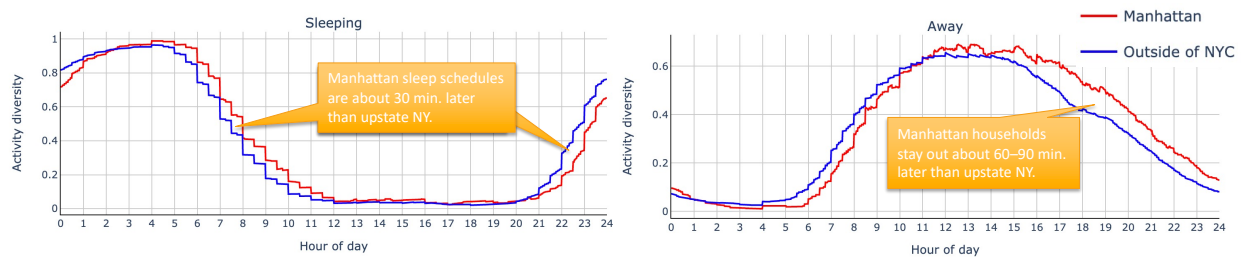
#### **Time Zone Effects**

We also investigated how ATUS data could be used to capture the effect that a household's location within a time zone has on the timing of occupant behavior. Giuntella and Mazzonna (2019) presented data from Jawbone sleep trackers and ATUS showing that location within a timezone affects average bedtime and sleep durations because of the timing of sunset. Figure 51 (left) shows that the average bedtime by county can differ by up to 19 minutes, based on data from Jawbone sleep trackers via The Washington Post (Ingraham 2019). This map also shows that some urban counties (such as the counties containing Los Angeles, Las Vegas, and Chicago) have later bedtimes than less urban counties within the same state. As discussed above, ATUS sample sizes limit our ability to capture county-





**Figure 49.** Using ATUS data, we compared occupant schedules between counties located in metropolitan statistical areas (MSA) and those elsewhere. MSA schedules are shifted about 0–30 min. later in local time.

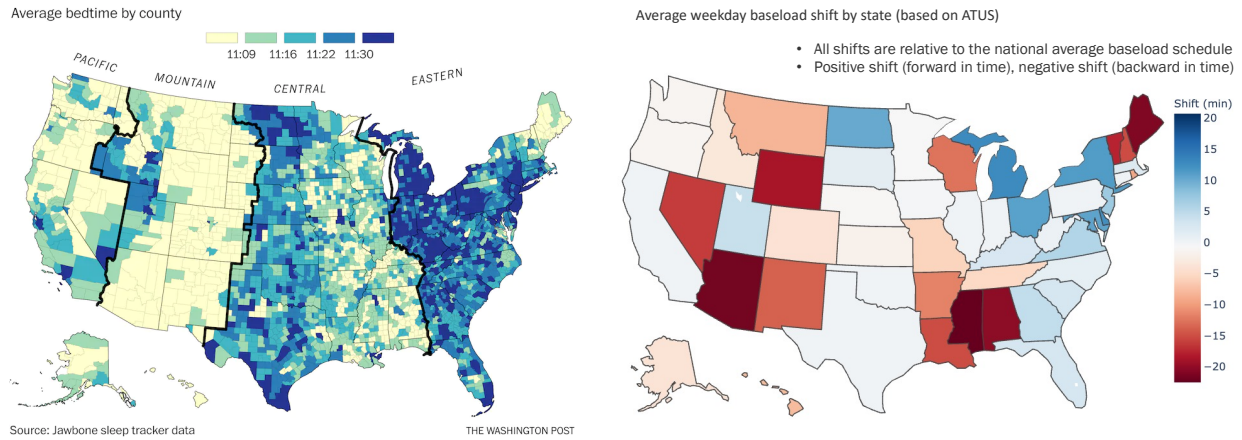


**Figure 50.** Using ATUS data, we compared occupant schedules between Manhattan and counties in New York State outside of the five boroughs of New York City. Manhattan schedules are shifted about 30 min. later for sleep and 60–90 min. later for returning home. Low sample sizes in ATUS make other activity comparisons difficult.

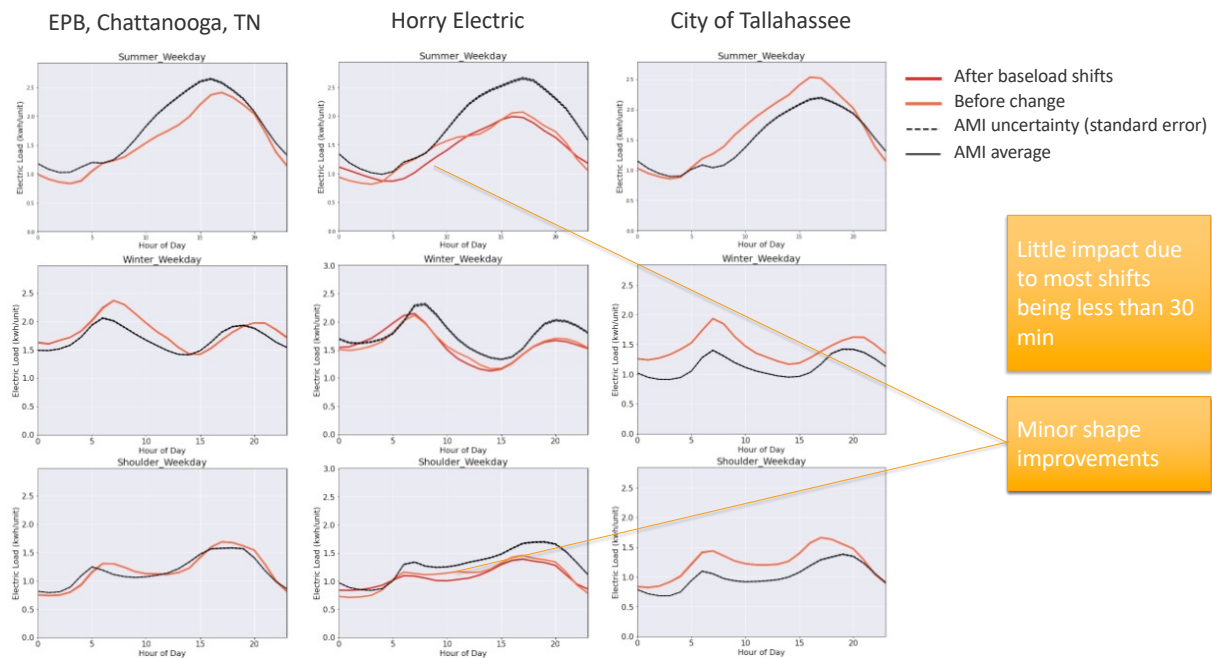
specific effects, so at this stage, we focused on deriving state-level average shifts in home/away/sleep schedule based on ATUS. ATUS sample sizes would not support using state-specific data to derive clusters for training the occupancy simulator Markov chains, so instead we derive simple state-average offsets that shift the start times for each cluster forward or backward in time relative to the national average home/away/sleep schedule (Figure 51, right). The impact of these schedule shifts was evaluated on the three AMI dataset comparisons for the Residential Region 4 focus area (Figure 52). Tennessee, South Carolina, and Florida all have relatively minor offsets ( $\pm 5$  minutes), so the effect is not very pronounced in the comparisons.



## End-Use Load Profiles for the U.S. Building Stock: Methodology and Results of Model Calibration, Validation, and Uncertainty Quantification



**Figure 51.** Left: Jawbone sleep tracker data show that the average bedtime by county can differ by 19 minutes, partially driven by how a county's location within a time zone affects the local sunset time. Source: Graphic from The Washington Post (Ingraham 2019). Right: We derived and applied average schedule shifting factors for each state from ATUS data.



**Figure 52.** The impact of ATUS-derived state average schedule shifts was evaluated on the three AMI dataset comparisons for the Residential Region 4 focus area. Tennessee, South Carolina, and Florida all have relatively minor offsets ( $\pm 5$  minutes), so the effect is not very pronounced in these comparisons. Note: final comparison is shown in Section 4.

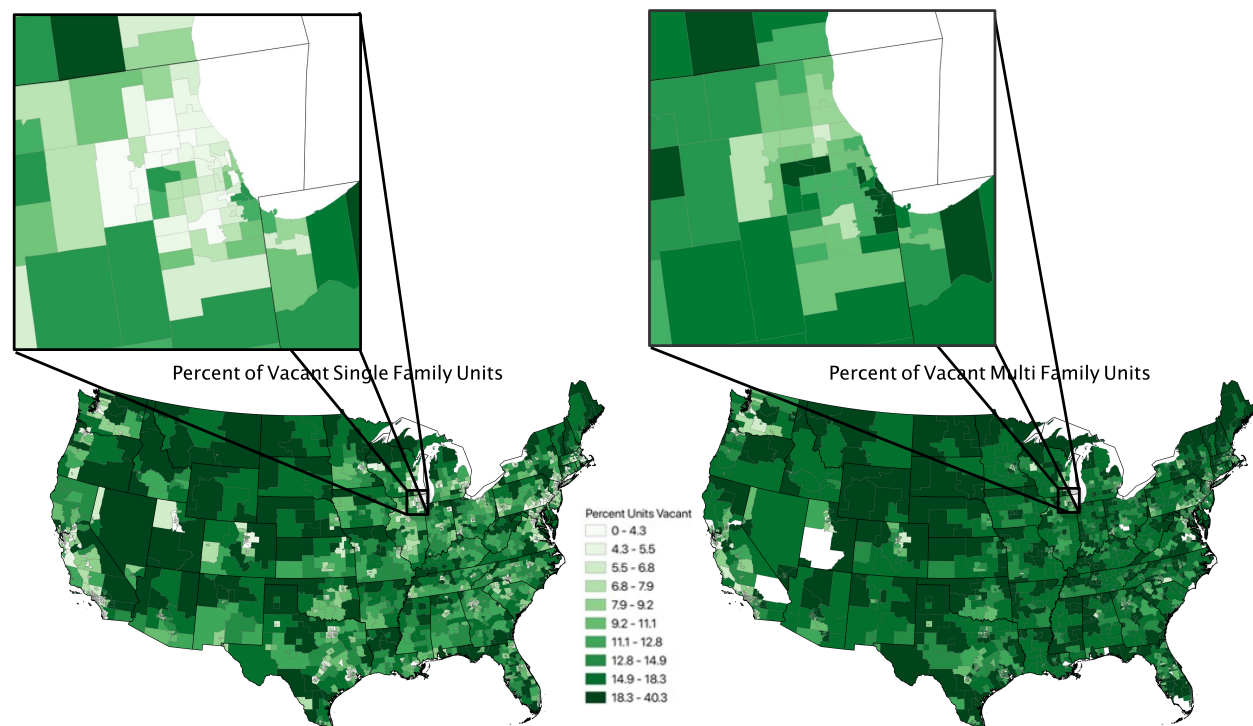
### Unoccupied Housing Units

Another occupancy-related update was the introduction of modeling some housing units as unoccupied. Prior to EULP calibration, ResStock modeled all residential dwelling units as occupied year-round. This was suitable for some analyses, and was consistent with the sampling frame of RECS; however, unoccupied housing units can be significant when modeling EULPs and comparing to utility- or state-level loads. PUMS 2016 shows approximately 14 million unoccupied housing units in the United States, or about 10% of all housing units. We therefore added

an occupancy status characteristic to ResStock. This characteristic has dependencies on housing type and location (PUMA) and has a distribution generated using data from ACS PUMS (Ruggles et al. 2021). Rates of unoccupied units vary considerably in different areas of the country, with higher rates in many rural and seasonally popular areas as well as in multifamily units in many urban areas. Figure 53 gives a sense of this variation. We did not attempt to explicitly model seasonal or occasional occupancy of vacation homes or short-term rentals because data to do so was not available.

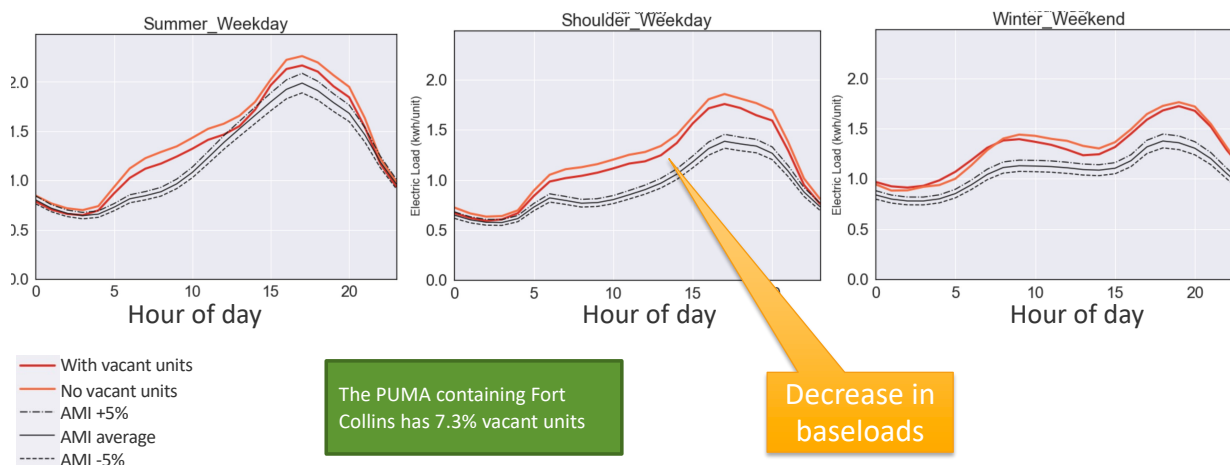
### Refinements to Unoccupied Housing Units

Modeling assumptions for unoccupied units were refined over the course of the calibration effort. During Region 2 of residential calibration, when unoccupied units were initially introduced, the only differences from occupied units were no occupant-driven appliance, lighting, and plug loads (see *New Residential Stochastic Occupant Behavior Model*), and no ceiling fan energy consumption. During Region 4 we added a static heating setpoint of 55°F (intended as a “don’t freeze the pipes” approach). Additionally, solar PV systems are not modeled for unoccupied units.

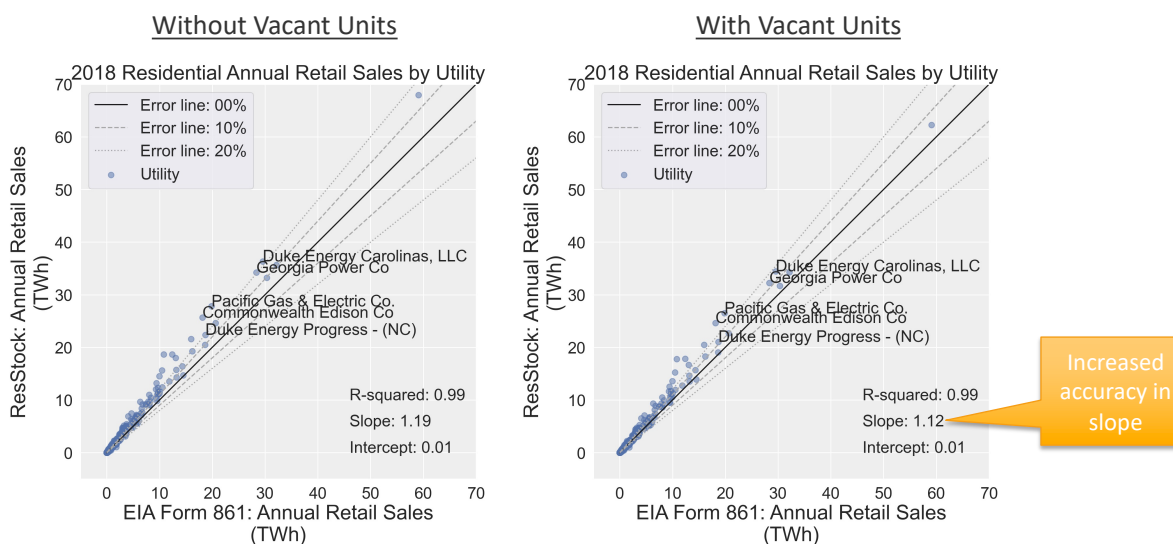


**Figure 53.** This figure shows how the rates of unoccupied/vacant units vary by PUMA for single family and multifamily units

Cohort: Single-Family Detached, Fort Collins



**Figure 54. Timeseries impact of introduction of unoccupied units in Region 2 residential calibration; this figure does not include the impact of later refinements to unoccupied units such as the change to heating setpoints**

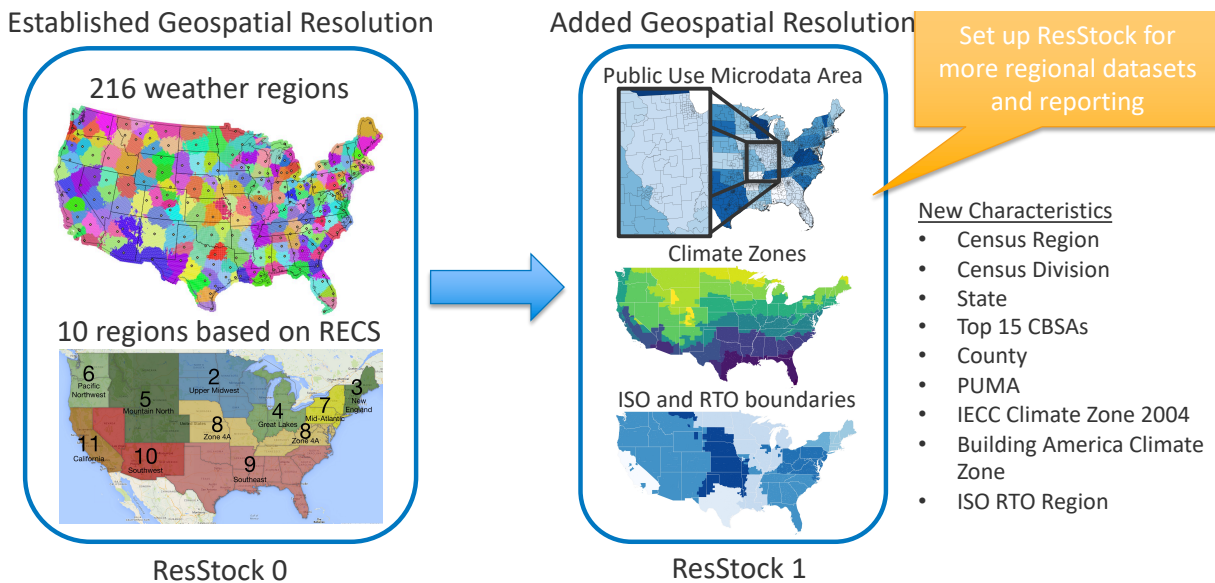


**Figure 55. Utility-scale annual impact of introduction of unoccupied units in Region 2 residential calibration; this figure does not include the impact of later refinements to unoccupied units such as the change to heating setpoints**

### 3.2.2 Geospatial and Weather Data Updates

#### Increased Geospatial Resolution

Prior to the EULP project, there were only two spatial descriptions in the ResStock housing characteristics: (1) 216 weather station regions, and (2) 10 “custom regions” based on 2009 RECS. One early change that we made was to create 10 new parameter distribution files that help spatially define the location of a sampled housing unit. Figure 56 shows the old and new spatial descriptions. Prior to this change, ResStock simulations were only natively described with the two spatial descriptions on the left; with some difficulty they could be aggregated to the state or utility level in post-processing.



**Figure 56. The geospatial resolution of ResStock was increased through the addition of 10 new spatial descriptions (listed on right)**

When ResStock was first developed circa 2015, we created the 10 custom regions by aggregating the 26 reportable domains (16 more-populous states and 10 groups of less-populous states) from 2009 RECS into 10 regions with similar climates. The aggregations were done in order to increase sample sizes when query 2009 RECS microdata with multiple dependencies (e.g., heating system type as a function of housing type, region, and vintage). In 2017, EIA added additional climate zone variables (e.g., 11 IECC climate zone groups) to the 2009 microdata, which can serve the same purpose as the custom regions (EIA 2013).

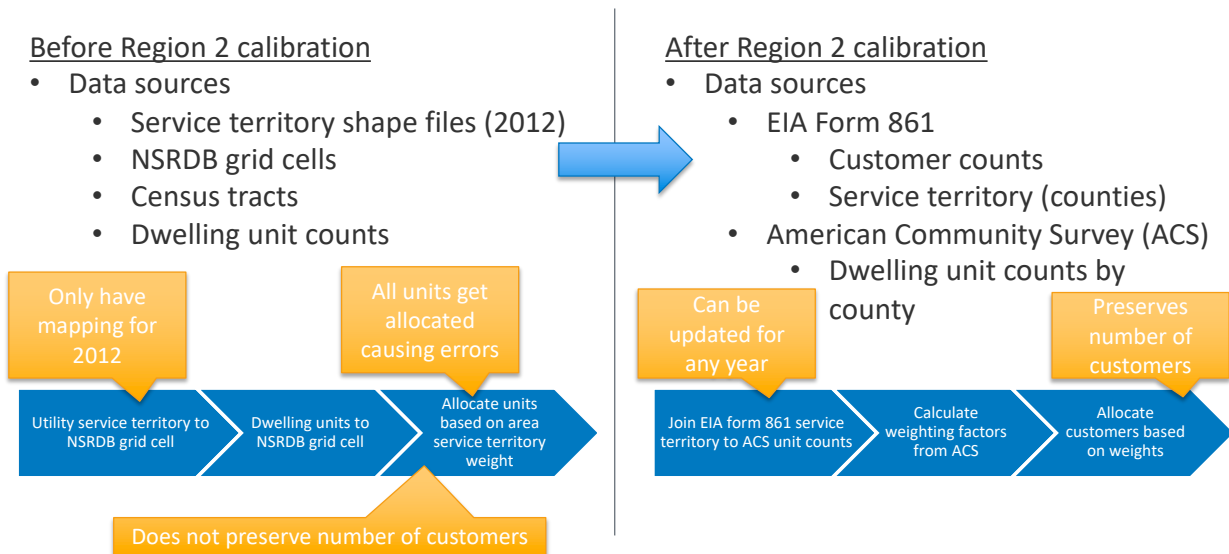
Data for the new spatial characteristics are based on the American Community Survey (ACS) 5-year 2016 tables B25001 and B25004. Unit counts and vacant unit counts are specified at the census tract level. The distributions are based on the occupied unit counts (i.e., table B25001 minus table B25004). Census tracts can be aggregated into counties or public use microdata areas (PUMAs), which are the two most granular new parameters. PUMAs are used for reporting of data from the U.S. Census Bureau’s ACS Public Use Microdata Sample (PUMS) (Ruggles et al. 2021). There are around 2,400 PUMAs, each of which contains at least 100,000 people. Unlike counties, PUMAs are designed to scale with population; they are smaller in urban areas and larger in rural areas, which makes them an ideal dependency parameter for querying other parameter distributions.

Previously, ACS was used to derive probability distributions for housing type and vintage for the 216 weather regions. The introduction of these new spatial parameters allowed us to use ACS PUMS (2017) to describe building type, vintage, and primary heating fuel with much finer resolution.<sup>9</sup> The new spatial characteristics also make it much easier to aggregate results up to the state, city, or utility level.

<sup>9</sup>The current parameter distributions for ResStock can be found on the ResStock github repository. Each tab-separated value (.tsv) file has comments at the bottom listing the source of the data and any assumptions made when deriving the distributions.

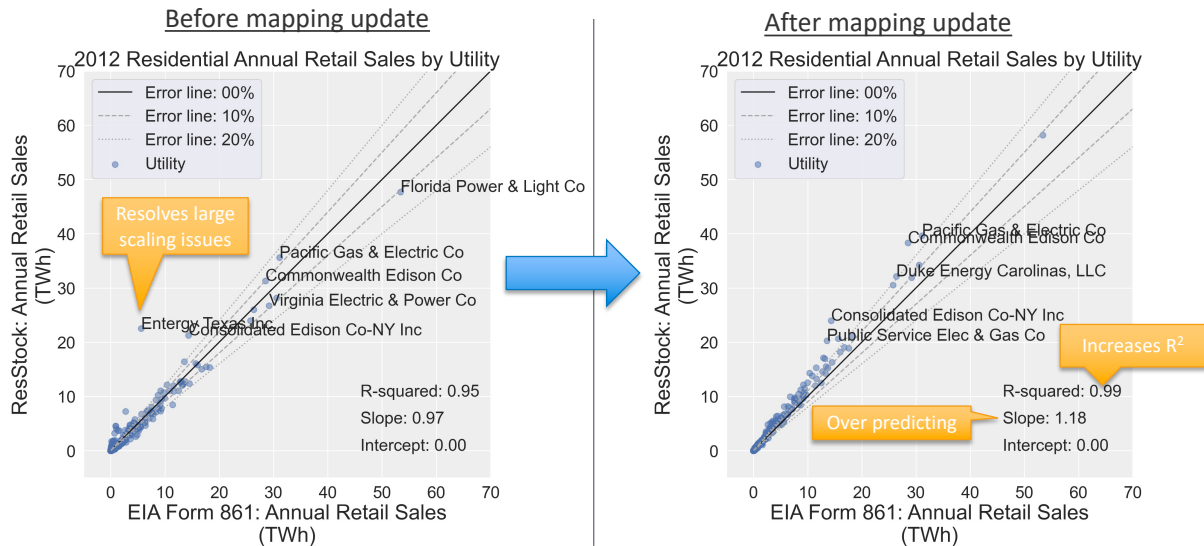
### Updated Mapping of Dwelling Units to Utilities

During Residential Region 2, we realized that inaccuracies in estimating the number of customers in each utility region was posing challenges for AMI, LRD, and EIA comparisons. To address this, we updated the method used to allocate dwelling units to utility territories. Previously we used a proprietary dataset with utility service territory shapefiles from 2012 in combination with dwelling unit count by census tract from ACS (Ruggles et al. 2021). The dwelling units were allocated to utility service territory using the 4-km by 4-km gridcells from NSRDB (NREL 2021a) as an intermediary. This was problematic because it was difficult to update for more recent years and did not preserve the number of customers in each service territory. The new method uses data on customer counts and service territories (counties served) from Form EIA-861 (EIA 2021b) in combination with dwelling unit counts by county from ACS (Ruggles et al. 2021). The new method can be updated for any year and, most importantly, preserves the number of customers in each utility service territory. The impact of this changes can be seen in Figure 58. Before the update, sales for several utilities, such as Entergy Texas, were severely overestimated. After the update, these anomalies are resolved and the  $R^2$  of a linear fit was improved from 0.95 to 0.99. This early change illuminated the fact that we were overestimating sales with a slope of 1.18, which was previously obscured (slope: 0.97) by the mapping inaccuracies.



**Figure 57. We updated the mapping of dwelling units to utilities so that it could be easily updated for any year and so that the number of customers for each utility is preserved**





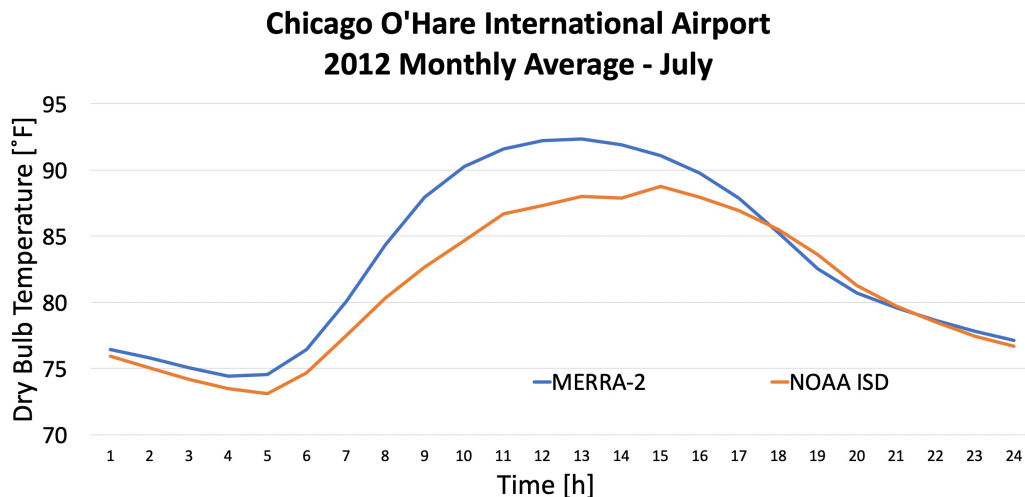
**Figure 58. Impact of updated mapping of dwelling units to utilities: Before the update, sales for several utilities were severely overestimated. After the update, these anomalies are resolved and the  $R^2$  of a linear fit is improved from 0.95 to 0.99. This early change illuminated the fact that we were overestimating sales with a slope of 1.18, which was previously obscured (slope: 0.97) by the mapping inaccuracies. Note: final comparison is shown in Section 4.**

#### Improved Historical Weather Data

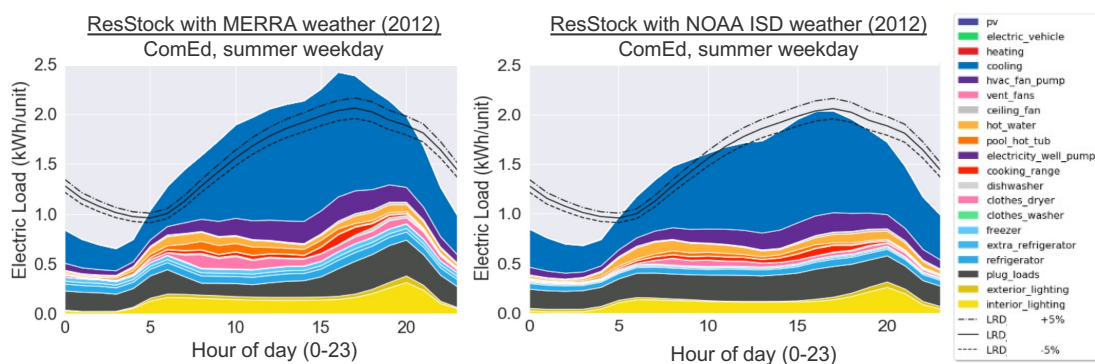
At the beginning of the EULP project, our standard practice was to run ResStock and ComStock with either TMY3 weather files or with historical weather data from the NSRDB, fully derived from satellite data. Early in the project, we realized that the satellite-derived temperature data exhibited large discrepancies with ground-measured temperature data for many locations. Figure 59 shows an example of this as seen in the average 24-hour temperature profile for O'Hare International Airport in July 2012.

As described in more detail in Section 2.4, we developed a process to combine ground station-measured temperature, humidity, wind speed/direction, and pressure data from the Integrated Surface Database (ISD) project of the National Oceanic and Atmospheric Administration (NOAA) National Climatic Data Center (Smith, Lott, and Vose 2011), with solar radiation data from the NSRDB. Figure 60 illustrates the impact that this change—from satellite (MERRA) to ground station (ISD) temperature data—had on a comparison between ResStock results and ComEd residential LRD for an average summer weekday in 2012. Before the change, the average summer weekday peak was overestimated by about 20%; after the change, it was within 2%.





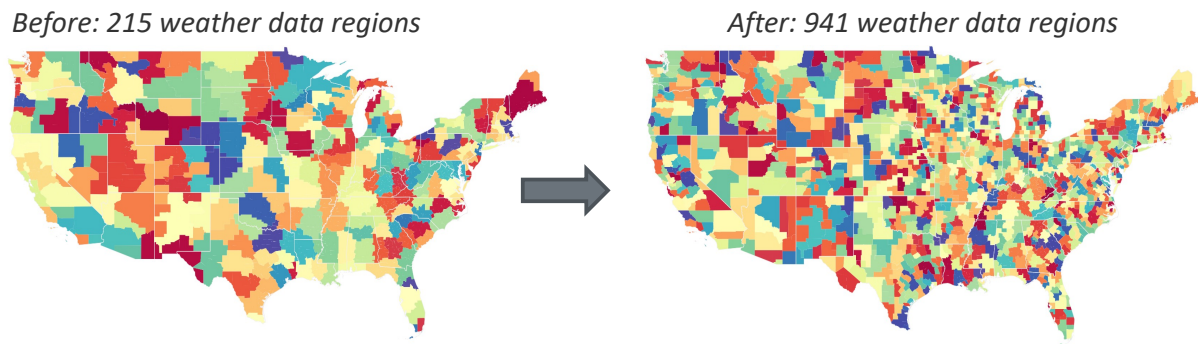
**Figure 59. Comparison of MERRA-2 and NOAA ISD weather:** July 2012 average 24-hour profile for dry-bulb temperature measured at Chicago's O'Hare International Airport weather station is compared with the dry-bulb temperature for that latitude and longitude from MERRA-2 (satellite reanalysis). The MERRA-2 temperature data are about five degrees hotter and peak about three hours sooner than the ground-measured temperatures.



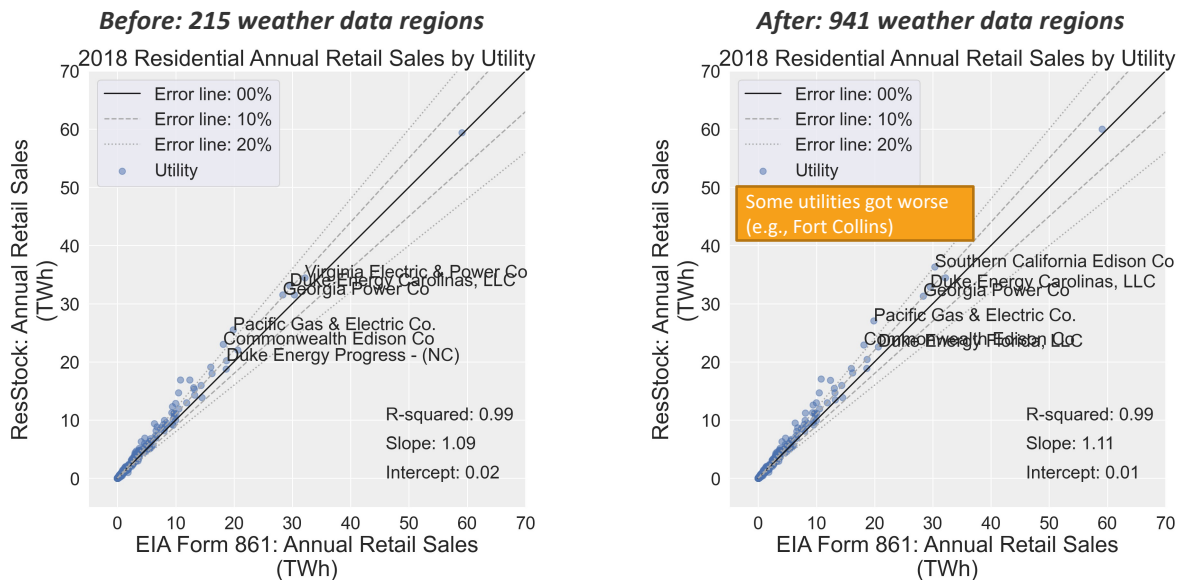
**Figure 60. Impact of weather data change:** Changing from satellite-based (MERRA) to ground-based (ISD) temperature data significantly improved the comparison between ResStock results and ComEd residential load data for an average summer 2012 weekday. (The “after” image includes several changes to the shape and magnitude of base load components, but the change to cooling load is still significant.) Note: final comparison is shown in Section 4.

### Increased Geographic Resolution of Weather Data

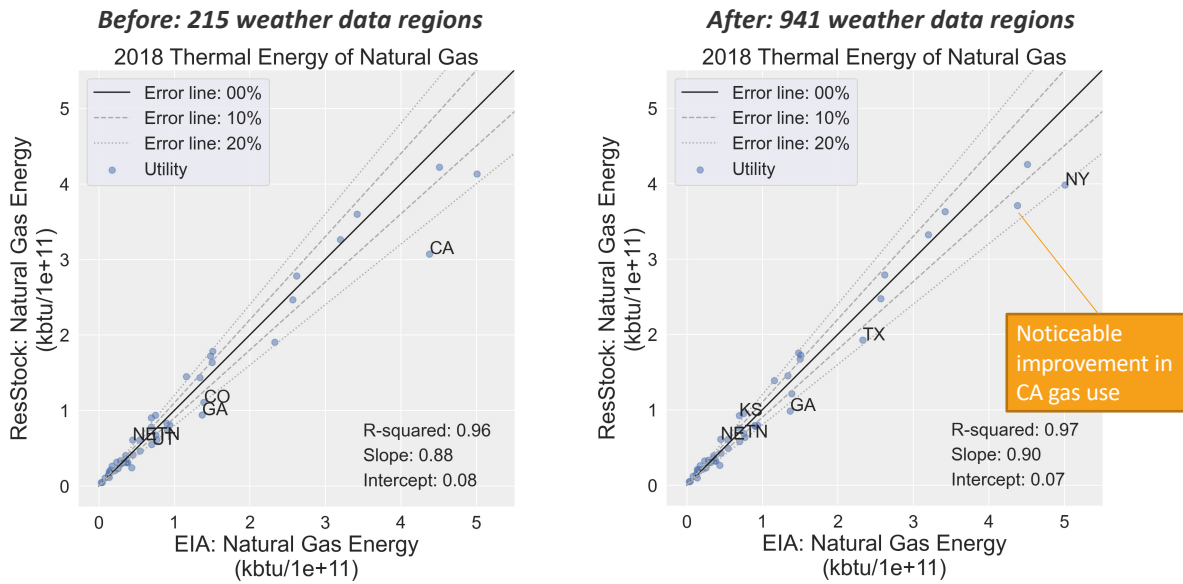
At the start of the EULP project, ResStock used 215 weather data regions. We were motivated to increase the number of weather stations to better capture climatic variation in areas like California, to increase the geographic resolution and therefore temporal fidelity of weather changes and sunrise/sunsets as they roll across parts of the electric grid, and also to align weather station locations between ResStock and ComStock. We adopted the weather data regions used for ComStock, which varies between 900 and 1,200 depending on the availability of data for a given year (AMY or TMY3). Figure 61 shows how the geographic resolution of weather regions changes. The assignment of housing units to weather stations is discussed in Section 2.4.2, *Assigning Buildings to Weather Stations*. We used EIA data comparisons to evaluate the impact of this change nationally, on an annual scale. There was not much change in electricity sales by utility (Figure 62), and agreement worsened for some utilities, which was also observed in the Fort Collins, CO, AMI comparison. There were improvements in the EIA natural gas comparison, particularly for California (Figure 63).



**Figure 61.** The number of weather stations used for ResStock was increased from 215 to 900–1,200 depending on the year



**Figure 62.** Impact of increased geographic resolution of weather data compared with EIA annual electricity sales by utility. Note: final comparison is shown in Section 4.

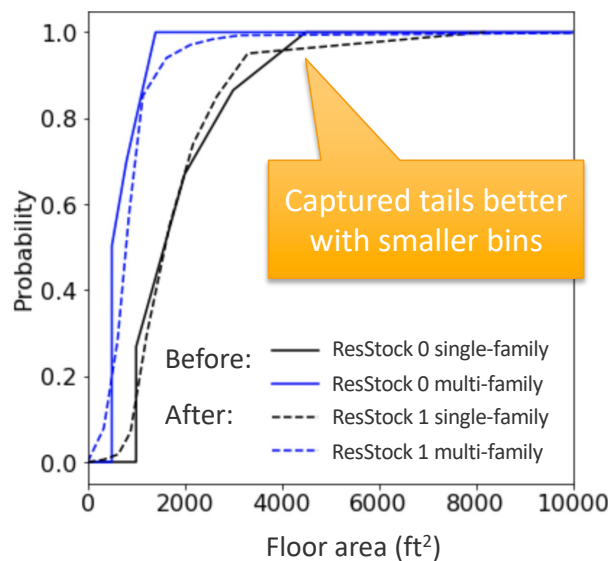


**Figure 63. Impact of increased geographic resolution of weather data compared with EIA annual residential natural gas sales by state. Note: final comparison is shown in Section 4.**

### 3.2.3 Geometry Updates

#### Increased Resolution of Floor Area Bins

Prior to EULP calibration, ResStock used four relatively coarse floor area bins as model inputs and as dependencies in other model inputs. During the EULP project, the team changed to using AHS 2017 data for floor area bins to capitalize on the greater sample size and geospatial resolution available in the AHS 2017 dataset (U.S. Census Bureau 2017). Other inputs that have floor area as a dependency were revised to use the AHS 2017 bins as inputs. RECS data have individual floor area values for individual samples, so it was possible to create dependencies in RECS-derived distributions using the AHS 2017 floor area data, which is available only in 9 discrete bins. These changes resulted in a better representation of the floor area in U.S. homes, as shown in Figure 64.



**Figure 64. Modeled floor area distributions by AHS 2017 floor area bin and dwelling unit type**

EnergyPlus, which ultimately runs the ResStock models as described in (section), does require individual floor area values rather than floor area bins as inputs. Accordingly, one floor area value was chosen for use in modeling dwelling units in each floor area bin, as shown in Table 13. Figure 68 shows the impact of the new floor area bins (combined with updated appliance saturation) for the 24-hour shoulder season comparison to ComEd multifamily customer load, with the most significant impact being a moderate increase in plug load and lighting magnitude.

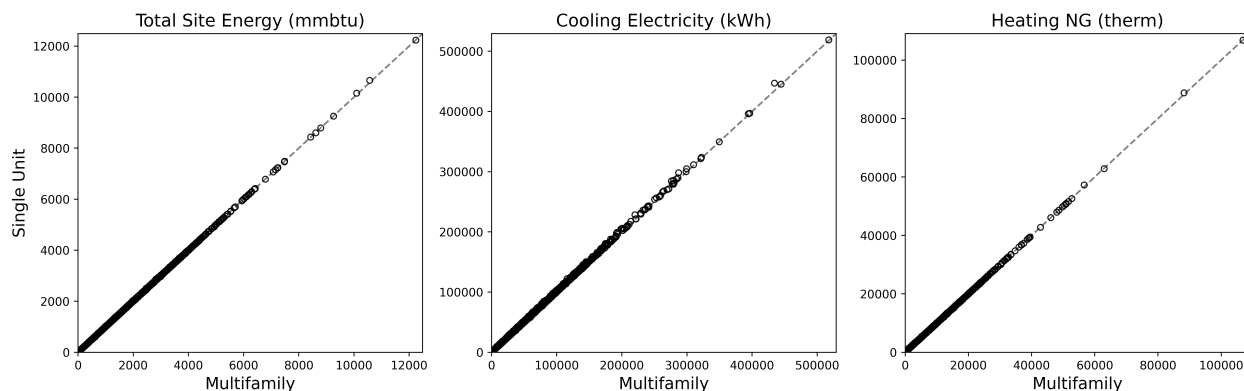
**Table 13. Updated ResStock floor area bins from AHS, and the average floor area values for each bin, which are used in the OpenStudio and EnergyPlus model articulation**

<b>AHS 2017 Floor Area Bin [ft<sup>2</sup>]</b>	<b>Modeled Floor Area: Single-Family De- tached [ft<sup>2</sup>]</b>	<b>Modeled Floor Area: Single-Family At- tached [ft<sup>2</sup>]</b>	<b>Modeled Floor Area: Multifamily [ft<sup>2</sup>]</b>
0–499	328	317	333
500–749	633	617	617
750–999	885	866	853
1,000–1,499	1,220	1,202	1,138
1,500–1,999	1,690	1,675	1,623
2,000–2,499	2,176	2,152	2,115
2,500–2,999	2,663	2,631	2,590
3,000–3,999	3,301	3,241	3,138
4,000+	8,194	13,414	12,291

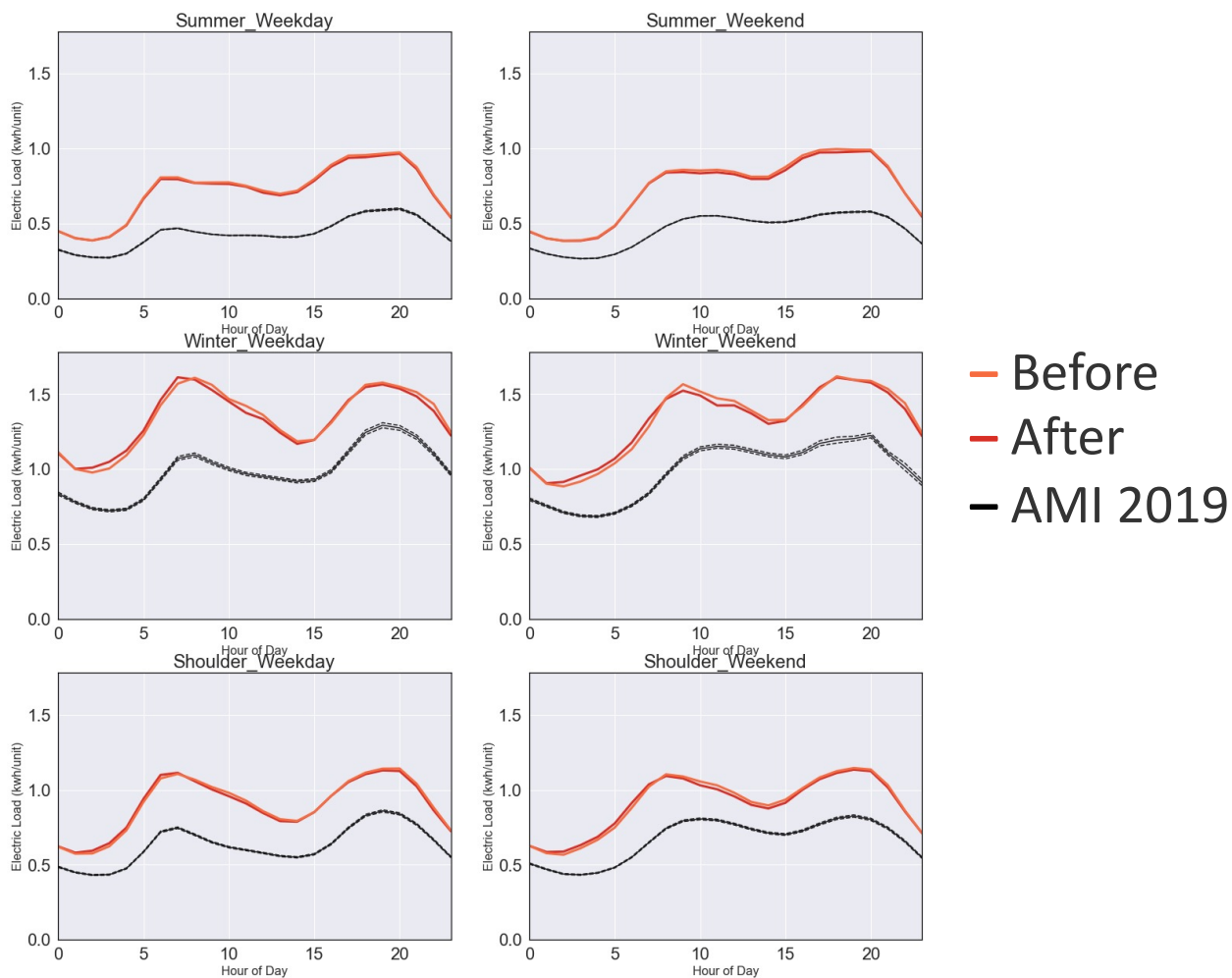
### *Faster Multifamily Unit Modeling*

When representation of multifamily and single-family attached buildings was added to ResStock in 2017, each sampled housing unit was modeled along with all the other units in the building, leading to a relatively small number of samples with very long simulation run times. Run times for multifamily and single-family attached building simulations could be orders of magnitude larger when compared to single-family detached homes, which introduces difficulties in estimating and scheduling jobs on high-performance computing resources. To address this, we transitioned to modeling individual units of multifamily and single-family attached buildings by representing shared surfaces as adiabatic. This approach improves computation times and provides more predictable and uniform resource requests for high-performance computing jobs, while also aligning with ResStock data sources, which are primarily defined in terms of dwelling units. Further, modeling individual dwelling units aligns with OpenStudio-HPXML and associated workflows such as Home Energy Score, Weatherization Assistance Program, and Energy Rating Index (Horowitz et al. 2019).

Because this improvement was driven by computation times and not by an observed error, it was desirable to have a negligible change in energy use, as shown in Figures 65 and 66. Heat transfer between shared walls and minor shading differences in buildings with eaves were the primary reasons for the observed changes in modeled energy use. Additionally, we can no longer explicitly model central HVAC systems serving multiple units, so we instead use the ANSI/RESNET/ICC 301-2019 (RESNET 2019) approach to calculate central HVAC energy and apportion it to individual units.



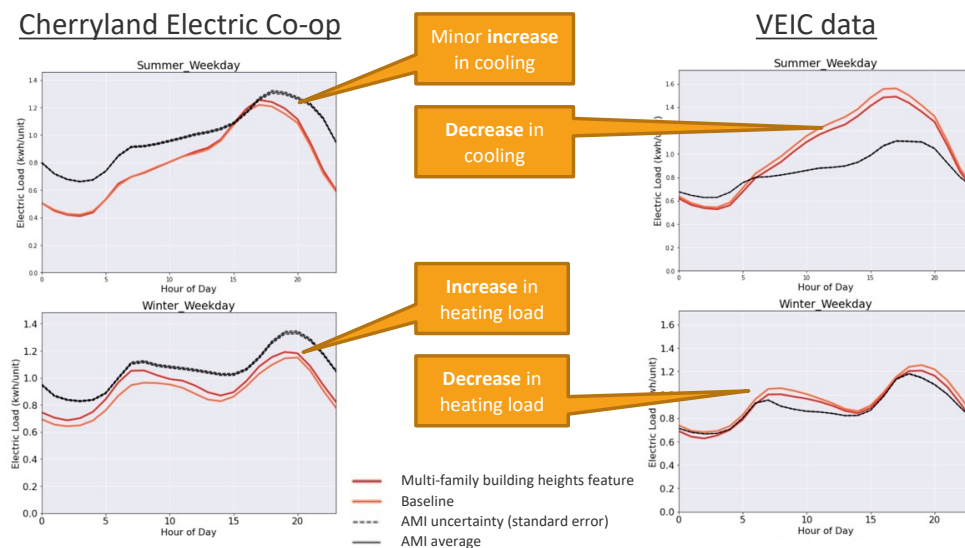
**Figure 65. Annual results of 10,000 multifamily buildings modeled using the whole building approach ('Multifamily') vs. the individual unit modeling approach ('Single Unit'). Changes in total site energy, cooling electricity, and heating electricity were minimal, with a median absolute site energy difference of 0.11%.**



**Figure 66. Effect of multifamily and single-family attached modeling speed improvements on simulated electric load profile for Seattle City Light. It was desirable to have a negligible change in energy use, as these changes were primarily introduced to improve speed and simplicity. Note: final comparison is shown in Section 4.**

### Adding Modeling of Dwelling Units in Mid-Rise and High-Rise Buildings

As discussed in the *Faster Multifamily Unit Modeling* section, we improved computation times for simulation of multifamily and single-family attached buildings by modeling individual dwelling units as opposed to entire buildings. This approach has also enabled more flexibility in describing multifamily and single-family attached inputs. One such change is allowing more realistic building heights for multifamily buildings, which was previously capped at three stories because we were only modeling low-rise residential buildings. Under the EULP project, we expanded ResStock to include modeling dwelling units in mid-rise and high-rise buildings; therefore, we expanded the number of story distributions using data from RECS 2009 (EIA 2013), so that units in multifamily buildings are modeled up to 21 stories. With this, we more accurately model unit locations within 3+ story multifamily buildings, because the distributions of unit level and horizontal positioning are influenced by the number of stories. This ultimately affects the breakdown between shared surfaces and exterior surfaces. Further, with higher unit heights, we more accurately capture the wind speed for wind-driven air infiltration for units in mid- and high-rise multifamily buildings. This change was made during Region 5 of residential calibration, so we evaluated the impact on the Cherryland, MI, and Vermont AMI comparisons (Figure 67). Unfortunately, both of these datasets contain fewer mid- and high-rise multifamily buildings relative to other regions, but nevertheless, we see an overall improvement for both summer and winter profiles.



**Figure 67. Impact of expanding multifamily building height options on AMI comparisons for the Cherryland, MI, and VEIC (Vermont) data areas. Note: final comparison is shown in Section 4.**

### 3.2.4 Appliance and Plug Load Updates

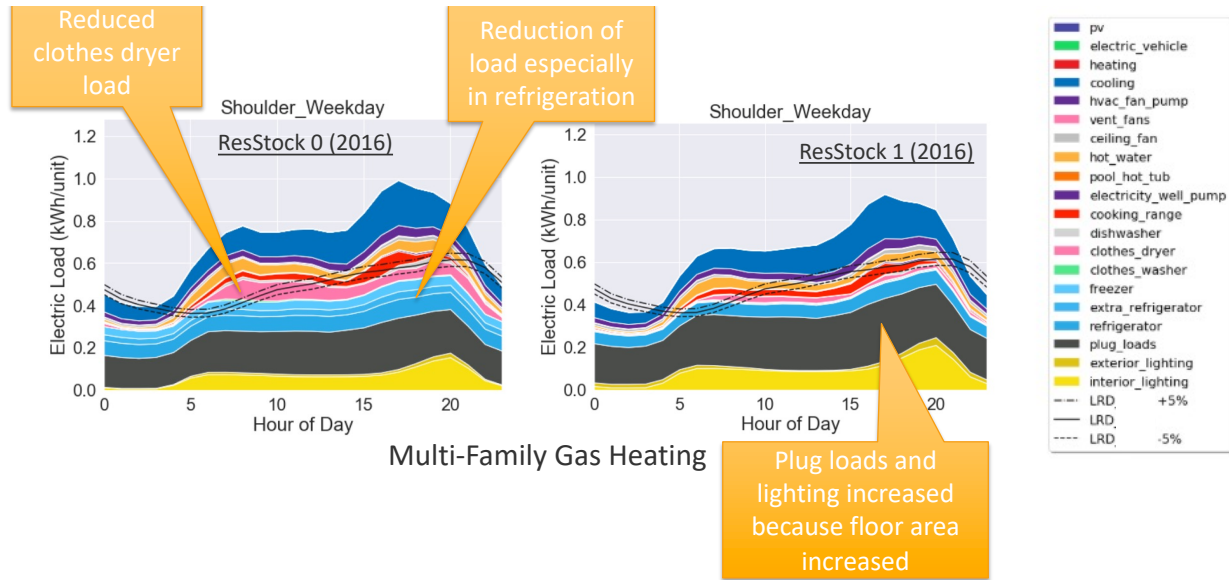
#### Major Appliance Saturation

Prior to EULP calibration, ResStock modeled dishwashers, clothes washers, and clothes dryers in all dwelling units. However, multifamily homes have a lower probability of having in-unit laundry facilities, dishwashers, extra refrigerators, and standalone freezers than single-family homes. This update added dependencies on housing type to the presence of major appliances—clothes washer, clothes dryer, dishwasher, refrigerator, extra refrigerator, and standalone freezer—in the ResStock models. Region was also added as a dependency to clothes washer, clothes dryer, dishwasher, and standalone freezer distributions.

We derived input distributions for refrigerator, extra refrigerator, and standalone freezer from RECS 2009, with custom region (Figure 56) as the spatial dependency for standalone freezers and dishwashers. Distributions for laundry appliances were based on AHS 2017 data (U.S. Census Bureau 2017). When using AHS data, each county in one of the top 15 U.S. core-based statistical areas (CBSAs) was provided with probabilities based on the data for that CBSA, and counties outside of CBSAs were provided with probabilities based on their census division. Figure 68 shows before (left) and after (right) the change, combined with an update to floor areas, for an average shoulder



season weekday 24-hour profile for multifamily gas-heated dwelling units in ComEd territory. The appliance saturation update significantly reduces the magnitude of clothes dryer, extra refrigerator, and standalone freezer load in multifamily homes.



**Figure 68. Impact of major appliance saturation and floor area updates on shoulder season profiles for multifamily gas-heated dwelling units in ComEd territory. The appliance saturation update significantly reduces the magnitude of clothes dryer, extra refrigerator, and standalone freezer load in multifamily homes, while the floor area update increases plug loads and lighting. Note: final comparison is shown in Section 4.**

#### Plug Load Equations by Housing Type

In comparisons between ResStock and RECS 2015 annual energy by end use, we observed that ResStock estimated higher plug load usage in multifamily and single-family attached homes compared to RECS 2015. End-use breakouts in RECS are modeled and therefore have higher uncertainty than measured end-use breakouts (EIA 2018a), but the discrepancy motivated us to introduce a dependence on housing type when calculating annual plug load electricity use for homes simulated with ResStock. We previously used a single equation (eq. 3.1), which was a regression of miscellaneous plug load electricity ( $E_{plug}$ ) from RECS 2015, which we define as televisions, microwaves, humidifiers, and other devices not elsewhere classified, on number of occupants ( $n_{occupants}$ ) and housing unit finished floor area ( $FFA$ ).  $n_{occupants}$  is sampled from probability distributions derived from PUMS as a function of PUMA, housing type, and number of bedrooms. Number of bedrooms is sampled from probability distributions derived from AHS as a function of housing type and floor area (U.S. Census Bureau 2017).

With this change, we moved to using a separate regression for each of the three housing types (eqs. 3.2–3.4). For the purpose of deriving and applying these regressions, SFD includes single-family detached and mobile homes, SFA includes single-family attached homes only, and MF includes units in multifamily 2–4 unit and 5+ unit buildings, using RECS definitions.

Before:

$$E_{plug} = 908.91 + 277.75 \cdot n_{occupants} + 0.39 \cdot FFA \quad (3.1)$$

After:

$$E_{plug,SFD} = 1146.95 + 296.94 \cdot n_{occupants} + 0.30 \cdot FFA \quad (3.2)$$

$$E_{plug,SFA} = 1395.84 + 136.53 \cdot n_{occupants} + 0.16 \cdot FFA \quad (3.3)$$

$$E_{plug,MF} = 875.22 + 184.11 \cdot n_{occupants} + 0.38 \cdot FFA \quad (3.4)$$

$$(3.5)$$

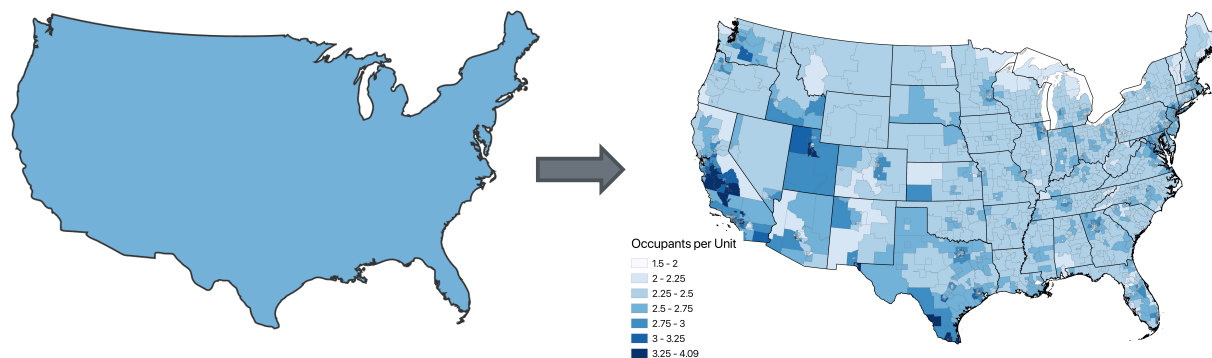
### Impact of Plug Load Equations by Housing Type

Using equation 3.1, the difference between ResStock and RECS 2015 total plug load electricity was -12%, 14%, and 21%, for SFD, SFA, and MF, respectively. Using the new equations 3.2–3.4, the errors were reduced to -9%, 2%, and 6%, respectively. We later refined plug load usage with a regional usage multiplier, which further reduced these errors (see *Regional Variation in Plug Load Usage*). Accuracy of plug load usage was also improved by adding geographic variation in the number of occupants (see *Household Size—Adding Geographic Dependency*) and increasing the resolution of floor area bins (see *Increased Resolution of Floor Area Bins*).

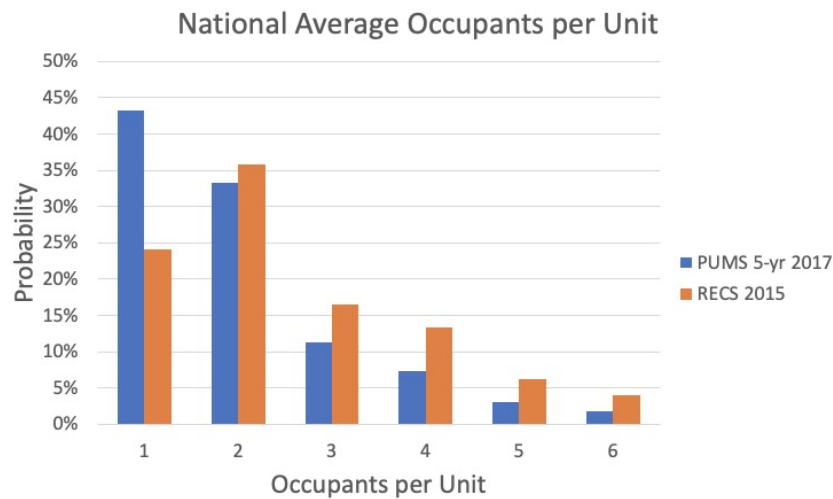
#### *Household Size—Adding Geographic Dependency*

Prior to EULP calibration, the ResStock distribution of the number of occupants per dwelling unit was taken from RECS 2015 and was dependent on housing type and number of bedrooms only; there was no dependency on location. However, PUMS data show considerable geographic variation in the number of occupants per dwelling unit across its 6 million samples. This variation across the 2,335 PUMAs is shown in Figure 69. Using the PUMS 2017 5-year data (Ruggles et al. 2021), we updated the ResStock household size distributions to have a dependency on PUMA in addition to the existing dependencies.

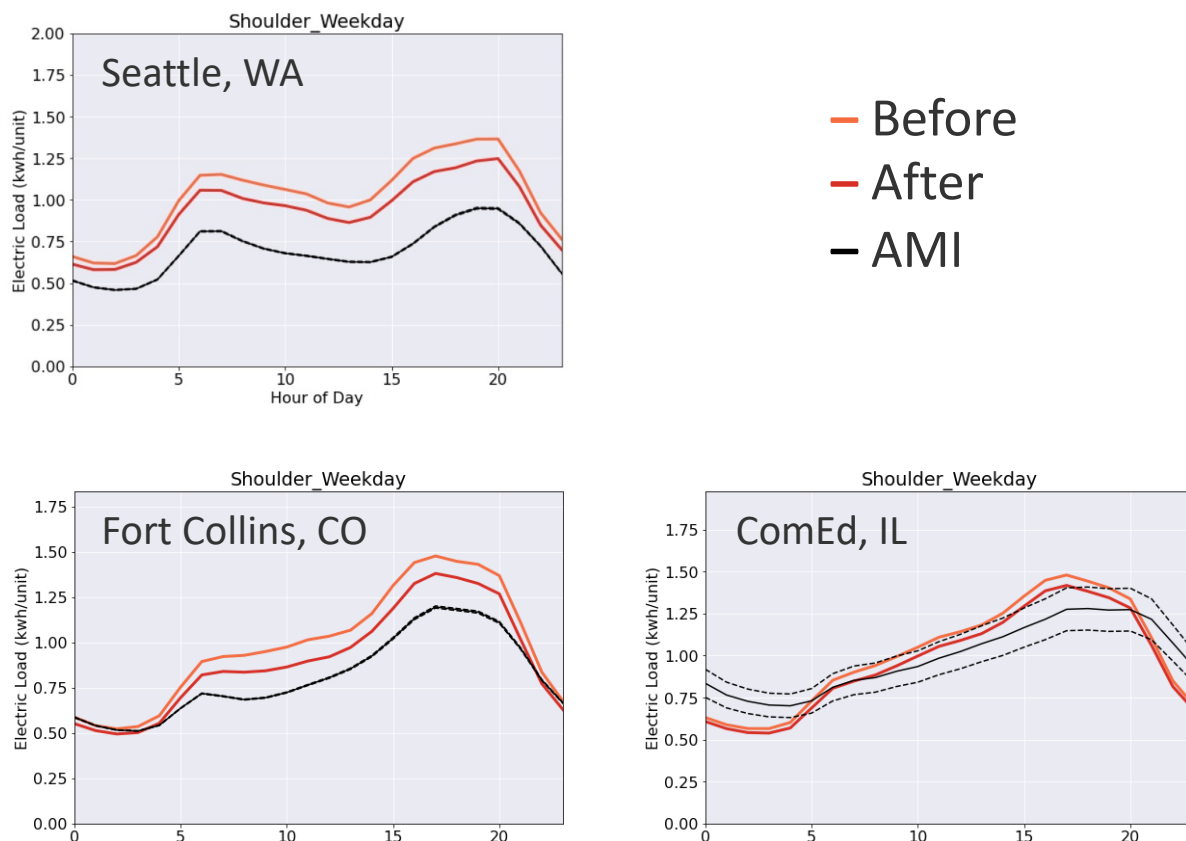
The household size impacts the usage of domestic hot water, appliances, and plug loads in our models. The addition of PUMA as a dependency therefore creates a greater geographic variation in these end uses. As shown in Figure 70, the PUMS 2017 data also show fewer occupants on average than the RECS 2015 data, causing a national reduction in baseload in addition to the increase in variation when the update to the distribution data source was made. Figure 71 shows the total residential electric load shoulder-season impact from the addition of the geographic dependency for household size, including the move from RECS 2015 to PUMS 2017.



**Figure 69. Geographic visualization of the average occupants per unit before and after adding geographic variation via the 2,335 PUMS regions**



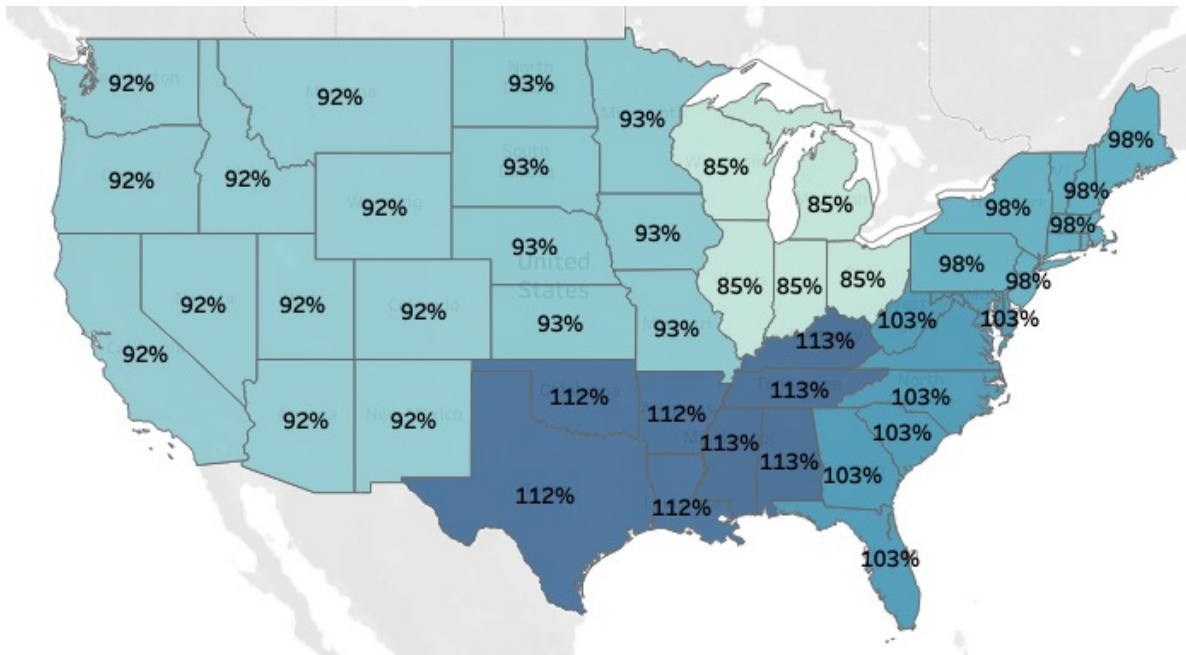
**Figure 70. National distributions of occupants per unit from RECS 2015 versus PUMS 2017. PUMS shows fewer occupants on average.**



**Figure 71. This figure shows the impact of using household sizes from PUMS 2017 with a geographic dependency rather than RECS 2015 and no geographic dependency. The impact on electrical load is a decrease in baseload due to the lower average household occupancy in the PUMS 2017 data. Note: final comparison is shown in Section 4.**

### *Regional Variation in Plug Load Usage*

After obtaining several AMI datasets from different regions, we observed variety in the magnitude of shoulder season baseload across the regions. After introducing plug load equations that account for housing type, we evaluated the resulting difference between the ResStock and RECS 2015 plug loads at a regional scale, finding that there was a strong regional dependence, with ResStock plug loads underestimated compared to RECS 2015 in the Southeast and overestimated in the Midwest and West (Figure 72). This regional dependence makes sense because we group together the RECS 2015 categories of televisions, microwaves, humidifiers, and other devices not elsewhere classified for modeling plug loads. We suspect that the regional variation is due to regional differences in usage of plug loads like dehumidifiers and fans. To address the regional variation, we introduced regional usage multipliers that simply multiply the result of equations 3.2–3.4 for each household by the regional factors shown in Figure 72. The impact of this change was evaluated in combination with the addition of regional variation in lighting efficiency and the addition of monthly appliance usage multipliers. The effect on the Seattle AMI comparison was minimal, with only a slight decrease in winter load (Figure 75).



**Figure 72. Ratio of RECS 2015 plug loads to aggregate ResStock plug loads calculated using equations 3.2–3.4, by census division; RECS 2015 plug loads are higher in the Southeast and lower in the Midwest and West**

### **3.2.5 Lighting Updates**

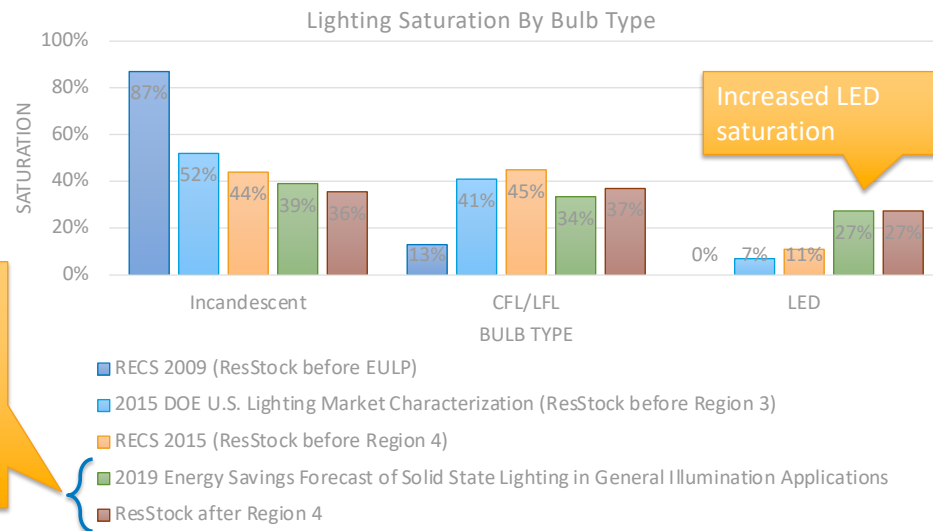
#### *Lighting Technology Saturation*

Prior to EULP calibration, ResStock was using data from 2009 RECS (EIA 2013) to estimate the saturation of lighting technologies in homes. Lighting is one of the fastest changing end uses due to the rapid adoption of LEDs, driven by technology advancement, regulations, and energy efficiency programs (Buccitelli et al. 2017), so we anticipated that this would be an important change for accurately modeling baseload.

As a simplification, ResStock models each home as having 100% of a given lighting type (incandescent, compact fluorescent [CFL]/linear fluorescent, or LED) even though real homes usually have a mix. During Region 1 of residential calibration, we increased the saturation of CFL and LED lamps from the 2009 RECS values to values from the 2015 U.S. Lighting Market Characterization Study (Buccitelli et al. 2017). Figure 73 compares saturation values between these two sources and shows that the RECS 2009 distributions used for ResStock before the beginning of EULP included too many incandescent bulbs and not enough CFLs and LEDs.

Comparison of national average lighting saturation to previous ResStock data sources →

ResStock is now a combination of RECS 2015 for regional diversity, but the LED saturations are adjusted to match forecasted 2019 levels



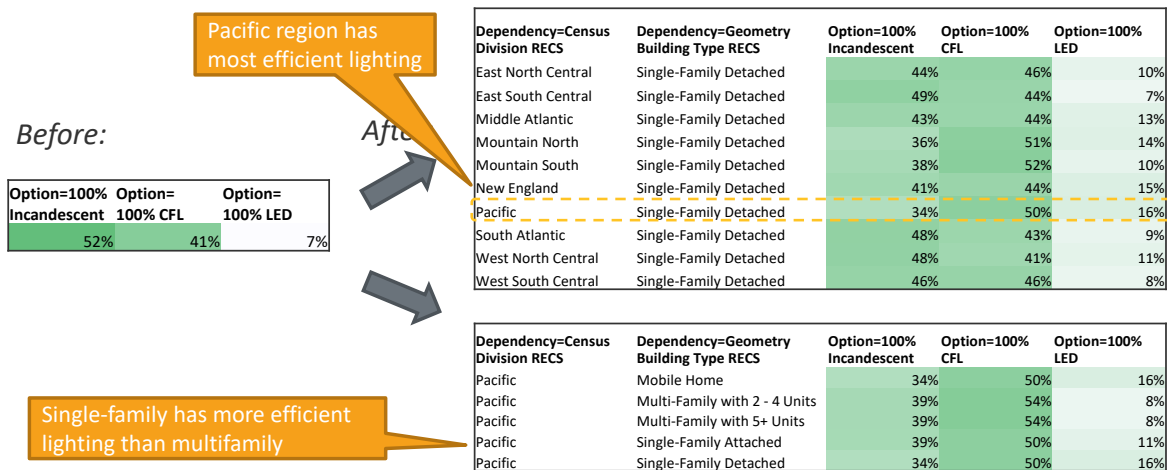
**Figure 73. Comparison of national average lighting efficiency distributions from four sources and ResStock**

The impact of the lighting technology saturation change on shoulder season weekdays in ComEd territory can be seen in Figure 44. The change is not very noticeable. We later refined this change, as described in the *Regional Variation in Lighting Efficiency* and *Further Update Lighting Technology Saturation* sections.

### Regional Variation in Lighting Efficiency

Early in the project, lighting efficiency distributions were updated based on the 2015 DOE U.S. Lighting Market Characterization (Buccitelli et al. 2017), the most recent robust data available.

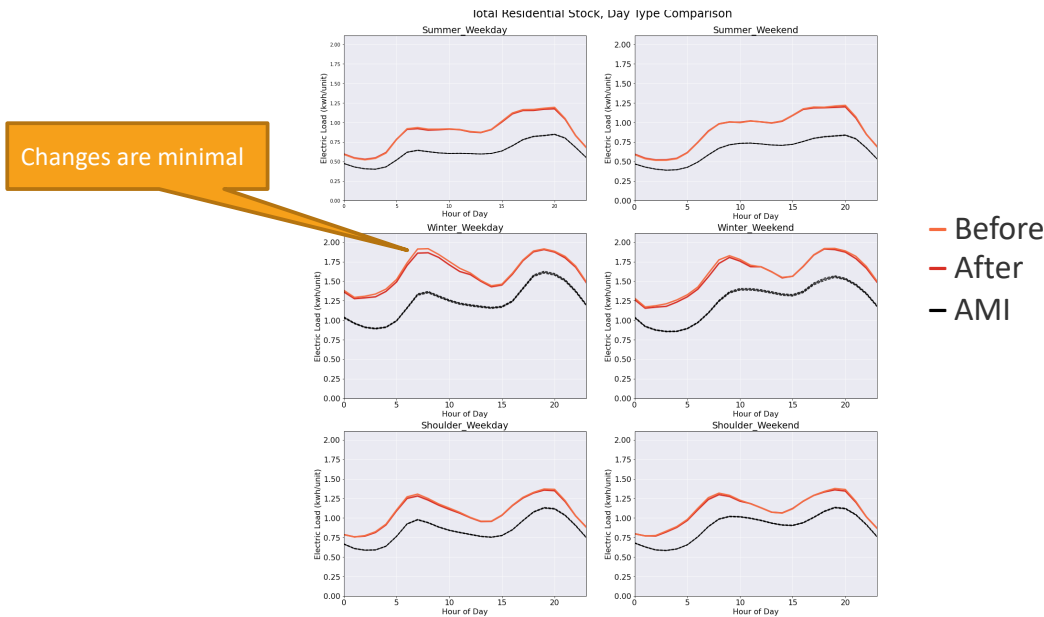
While some of the drivers of lighting efficiency distribution changes have been at the national level, many have been more limited in geographic scope, such as regulations at the state level and energy efficiency programs at the local and regional levels. This has led to considerable geographic variation in lighting efficiency distributions across the United States. Therefore, during Region 3 of residential calibration, a second update was made to use RECS 2015 data, which has distributions generally similar to the DOE report nationally as shown in Figure 73. Use of the RECS 2015 data allowed us to introduce lighting efficiency distribution dependencies on location (RECS census division) and housing type, whereas previous to this update the distributions were constant across all types of dwelling units and all locations across the United States (Figure 74).



**Figure 74. RECS 2015 was used to introduce lighting efficiency distributions that depend on census division and housing type. RECS 2015 data show that the Pacific census division has the lowest saturation of incandescent lighting, and saturation of LED lamps is higher in single-family homes than in multifamily homes.**

The electric load impact from these changes was minimal and not readily visible in calibration comparison plots at the utility level (Figure 75), although it is possible that changes were canceled out by the plug load and appliance changes also included in the comparison. We later refined this change, as described in the *Further Update Lighting Technology Saturation* section.





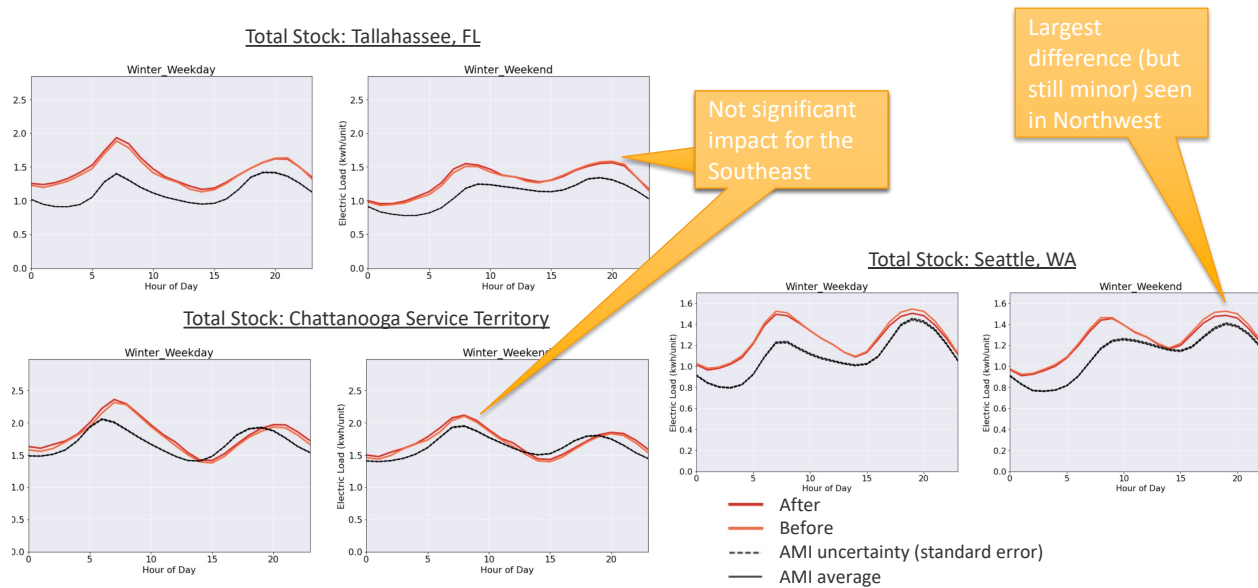
**Figure 75. Impact of adding regional variation in lighting technology saturation (p. 91), regional variation in plug loads (p. 89), and monthly appliance multipliers (p. 45) on Seattle AMI comparison during Region 3 of residential calibration. The impacts for Seattle are minimal, with a slight decrease most visible in winter. Note: final comparison is shown in Section 4.**

#### *Further Update Lighting Technology Saturation*

During this project, DOE published its *Energy Savings Forecast of Solid State Lighting in General Applications* (Yamada et al. 2019). Because of the quickly changing nature of lighting efficiency, we updated our distributions to reflect further changes since the 2015 report (Buccitelli et al. 2017) we had used for distributions starting in Region 1, and the RECS 2015 data we had started using in Region 3, to allow for our preferred dependencies.

Therefore, during Region 4 of residential calibration, we moved to a dual-source approach. We continued to use RECS 2015 data to capture regional diversity, but adjusted the LED saturation to match the forecasted 2019 levels from Yamada et al. (2019). Figure 73 shows a comparison between all five of these data sources and ResStock distributions at the national level, including a large increase from 7% LEDs in the 2015 report and 11% LEDs in the 2015 RECS data to 27% LEDs in the 2019 forecast and the distributions used in ResStock by the end of EULP.

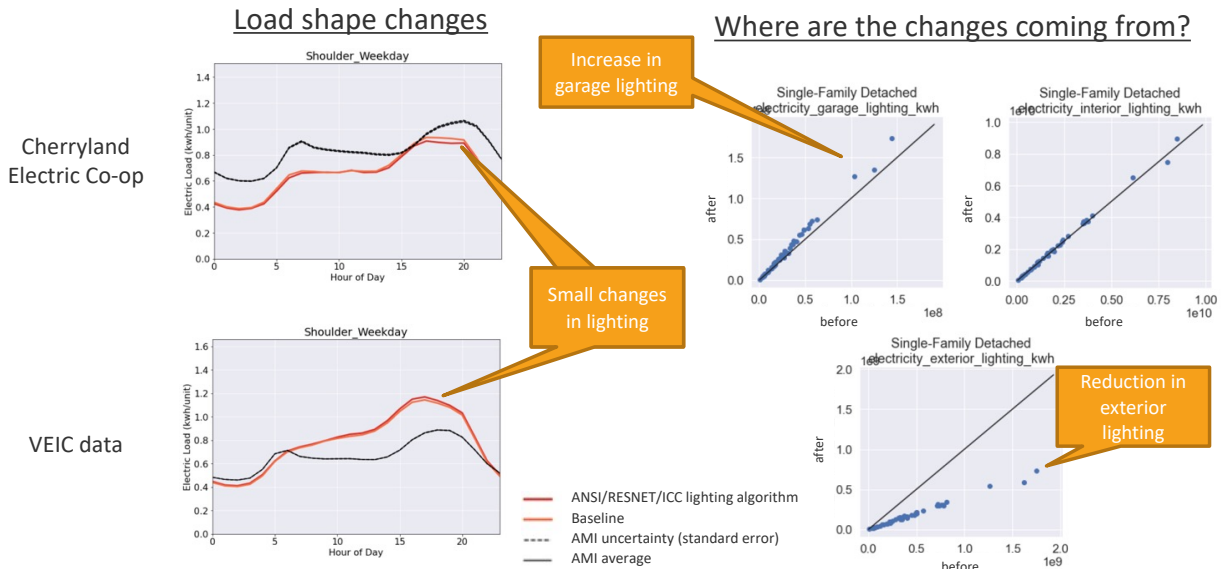
The electric load impact of this further update was relatively modest and most visible in the Northwest, as shown in Figure 76.



**Figure 76. Impact of lighting technology saturation change during Region 4 of calibration. The impacts are minor, with the greatest impacts occurring in the Northwest. Note: final comparison is shown in Section 4.**

### Updated Lighting Usage Equation

ResStock has historically modeled interior, exterior, and garage lighting loads using equations from the 2014 Building America House Simulation Protocols (Wilson et al. 2014). We updated the lighting equations to instead use the ANSI/RESNET/ICC 301-2019 approach (RESNET 2019). This update was not motivated by calibration, but rather to align lighting usage with OpenStudio-HPXML and associated workflows such as Home Energy Score, Weatherization Assistance Program, and Energy Rating Index (Horowitz et al. 2019), but nevertheless we wanted to evaluate any impact on AMI comparisons. As shown in Figure 77, the new approach for lighting has a minimal effect on the interior and the total overall lighting energy, while garage lighting energy increases and exterior lighting energy decreases.

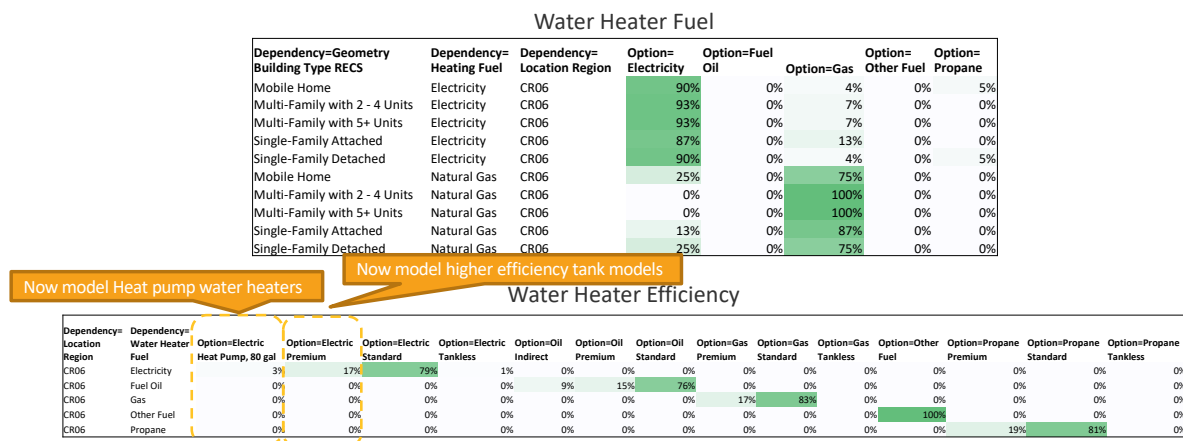


**Figure 77. Impact of updating lighting algorithm to the ANSI/RESNET/ICC 301-2019 approach—garage lighting increased slightly, interior lighting did not change, and exterior lighting decreased by about 50%. Note: final comparison is shown in Section 4.**

### 3.2.6 Water Heating Updates

#### Updated Water Heater Dependencies

Region 3 (Seattle) was the first regional phase where electric space and water heating was significant. To facilitate updates in this area, we separated the specification of water heater efficiency from the specification of water heater fuel type, while still preserving the relationship between these two characteristics. Before this change, the probability distribution of sampling water heater fuel type and efficiency options was dependent on the space heating fuel type and custom region (10 aggregations of RECS 2009 reportable domains; Figure 56) of a given home.



**Figure 78. Water heater distributions were split into separate distributions for fuel type (top) and efficiency (bottom), which facilitated the addition of higher efficiency fuel-fired and electric water heater options. Example distributions are shown here for Oregon and Washington (CR06).**

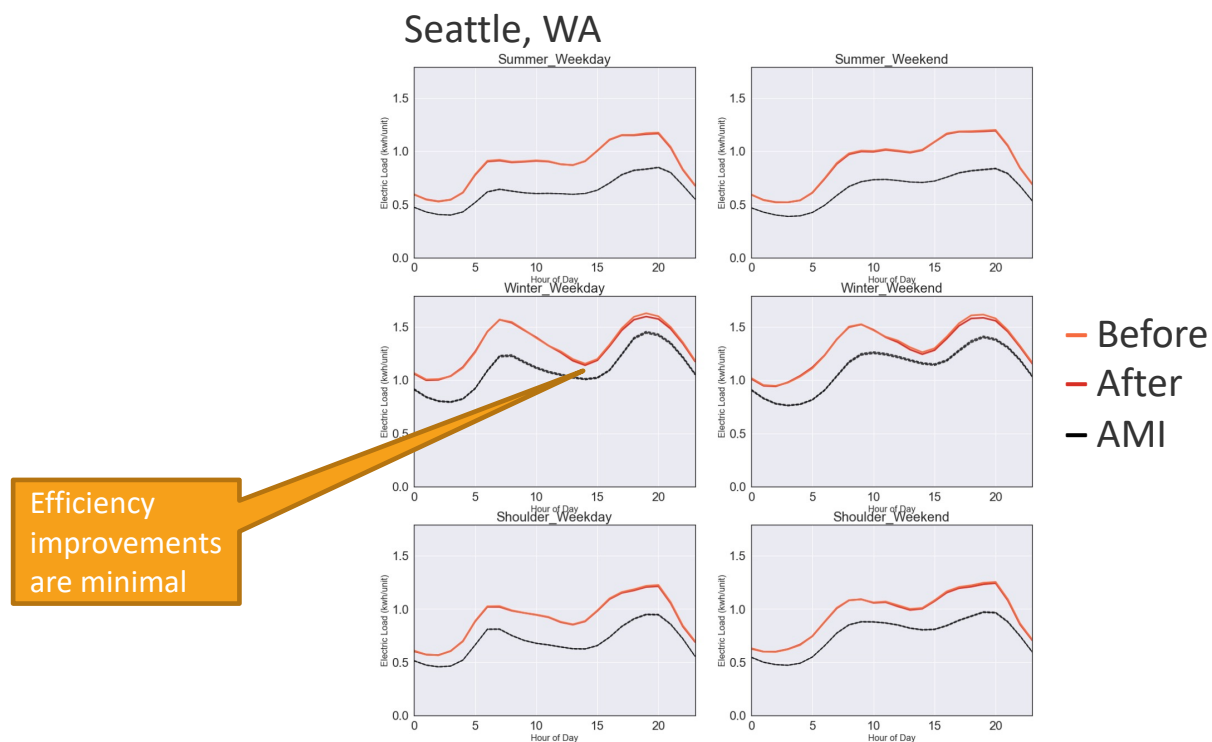
Figure 78 shows an example of the new probability distributions for Oregon and Washington (CR06). The housing type dependency is significant; in Oregon and Washington, single-family detached and mobile homes that use gas

for space heating have a 25% chance of using electricity for water heating, whereas that probability is zero for multifamily units in the region.

### *Adding Higher Efficiency Water Heaters*

The updated dependency structure for water heating characteristics facilitated the addition of higher efficiency fuel-fired and electric water heater options (Figure 78, bottom). Prior to this change, ResStock input distributions only included differentiation of water heater fuel type and whether the water heater was a tank or tankless, all queried from RECS 2009. All tank-type water heaters were modeled as “standard efficiency” (EF 0.92 electric, EF 0.59 gas, and EF 0.62 oil). After this change, we use the RECS 2009 field for whether a home’s water heater has an add-on insulation blanket to determine the saturation of “premium” water heaters (modeled as moderately more efficient; EF 0.95 electric, EF 0.67 gas, and EF 0.68 oil) in a given region and with a given fuel type.

RECS 2009 and 2015 do not have data on whether an electric water heater is a heat pump water heater (HPWH), so we add these in manually. There are no data on HPWH saturation available nationally. We know from 2016–2017 RBSA II study (NEEA 2021c) that HPWH saturation is greater in the Northwest than in other regions, so we use the RBSA II data to set the saturation of HPWHs at 3% of the installed electric water heater stock for Oregon and Washington and use a value of 0.5% for the rest of the United States, based on a value reported in Butzbaugh, Sandahl, and Baechler (2017). For the EULP dataset, HPWHs are modeled as 80-gallon units with performance approximately EF 2.40. Figure 79 shows that the overall impact of water heater efficiency updates on Seattle AMI comparisons was negligible; the increase in electric water heater efficiency may have been canceled out by an increase in the number of electric water heaters in single-family homes (which, on average, use more hot water than multifamily homes).



**Figure 79. The overall impact of water heater efficiency updates on Seattle AMI comparisons was negligible. Note: final comparison is shown in Section 4.**

### **3.2.7 HVAC Updates**

#### *Improving the Structure of HVAC Characteristics*

The probability distributions that ResStock samples to assign HVAC characteristics to housing units have historically been derived from RECS 2009 microdata (EIA 2013), which provides data on heating and cooling equipment fuels,

types, and ages (as a function of housing type, region, and vintage). We use efficiency vs. age distributions derived from AHRI data and other sources to convert equipment ages to efficiencies (LBNL 2020). Over years of developing ResStock, the conditional probability tables for HVAC had become quite complex and needed to be updated manually. For this project, we automated the table generation process and improved the structure of dependencies, simplifying the HVAC probability distributions from 37,071 rows across 17 tables to 2,416 rows in seven tables.

The benefits of this restructuring are that it makes future updates easier, makes simulation outputs more transparent and easier to understand, and simplifies the handling of heat pumps and HVAC systems serving multiple housing units. The change also separates the assignment of equipment type (furnace, boiler, heat pump, etc.) from efficiency level, which enables us to use data from the American Housing Survey (N=114,860) instead of RECS (N=12,083) for determining HVAC equipment types, eliminating issues with small sample sizes for some dependency combinations. The efficiency distributions are not completely independent; they still depend on equipment type. The impact of this change on empirical data validation (combined with updates to efficiency distributions) is shown in Figures 81 and 82.

### Updating HVAC Efficiency Distributions

**Table 14. Multiple Data Sources Used to Determine Room Air-Conditioner Efficiency as a Function of Age**

Shipment Year	Shipment-Weighted EER <sup>1</sup>	Federal Minimum EER <sup>2</sup>	ENERGY STAR Minimum EER <sup>3</sup>	Percent ENERGY STAR <sup>4</sup>
1996	9.08			
1997	9.09			
1998	9.08			
1999	9.07			
2000	9.3			
2000	9.3			
2001	9.63	9.79	10.88	
2002	9.75	9.79	10.88	
2003	9.75	9.79	10.88	
2004	9.71	9.79	10.88	
2005	9.95	9.79	10.88	
2006	10.02	9.79	10.88	
2007	9.81	9.79	10.88	
2008	9.93	9.79	10.88	
2009	9.93	9.79	10.88	
2010		9.79	10.88	33%
2011		9.79	10.88	62%
2012		9.79	10.88	58%
2013		9.79	10.88	72%
2014		11	11	50%
2015		11	11	54%
2016		11	12	38%
2017		11	12	34%
2018		11	12	42%

<sup>1</sup> Shipment-weighted values for Room AC comes from AHRI data from Home Energy Score documentation (1970–2008; 2009 is a copy of 2008 values; LBNL 2020)

<sup>2</sup> Federal minimum EER values from a descriptive paragraph at <https://appliance-standards.org/product/room-air-conditioners>.

<sup>3</sup> ENERGY STAR minimum EER values are simplifications based on ENERGY STAR 2020b (2014–2015) and ENERGY STAR 2020c (2016–2020). ENERGY STAR EER value simplification was based on listing at [data.energystar.gov](https://data.energystar.gov), where a majority of typical qualifying units were EER=12.0 (ENERGY STAR 2020d).

<sup>4</sup> Percent ENERGY STAR comes from data found on [energystar.gov](https://energystar.gov) (no data for 2009 and prior; 2019 and 2020 copied from 2018; ENERGY STAR 2020a)

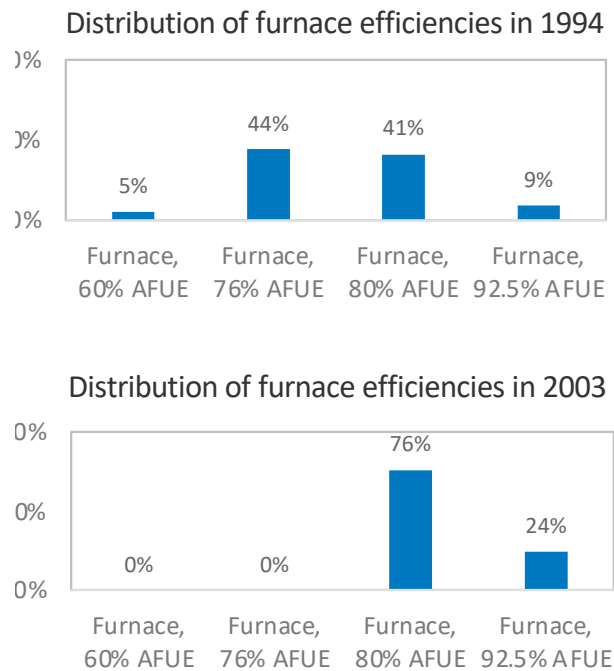
### Updating Room Air-Conditioner Efficiency Distributions

Prior to this change, ResStock used a simple assumption that one-third of all room air conditioners (i.e., window-mounted or through-the-wall packaged air conditioners) had an energy efficiency ratio (EER) of 8.5, and two-thirds

had an EER of 10.7 based on older ENERGY STAR data. We updated the distribution of room AC efficiencies using historical ENERGY STAR saturation data (ENERGY STAR 2020a), ENERGY STAR minimum efficiency values over time (ENERGY STAR 2020b, ENERGY STAR 2020c, ENERGY STAR 2020d), and federal minimum efficiency values over time. The data we used are shown in Table 14. The efficiency vs. age distributions are used in combination with equipment age distributions from RECS 2009 (using 2018 as the reference year for equipment age) to determine efficiency distributions for ResStock.

### Updating All HVAC and Refrigerator Efficiency Distributions

We also updated the reference year for equipment age from 2009 to 2018, which increased equipment efficiencies for all HVAC equipment and refrigerators. ResStock uses data from RECS 2009 on age distributions of equipment. The effect of this can be seen in the furnace efficiency distributions shown in Figure 80.



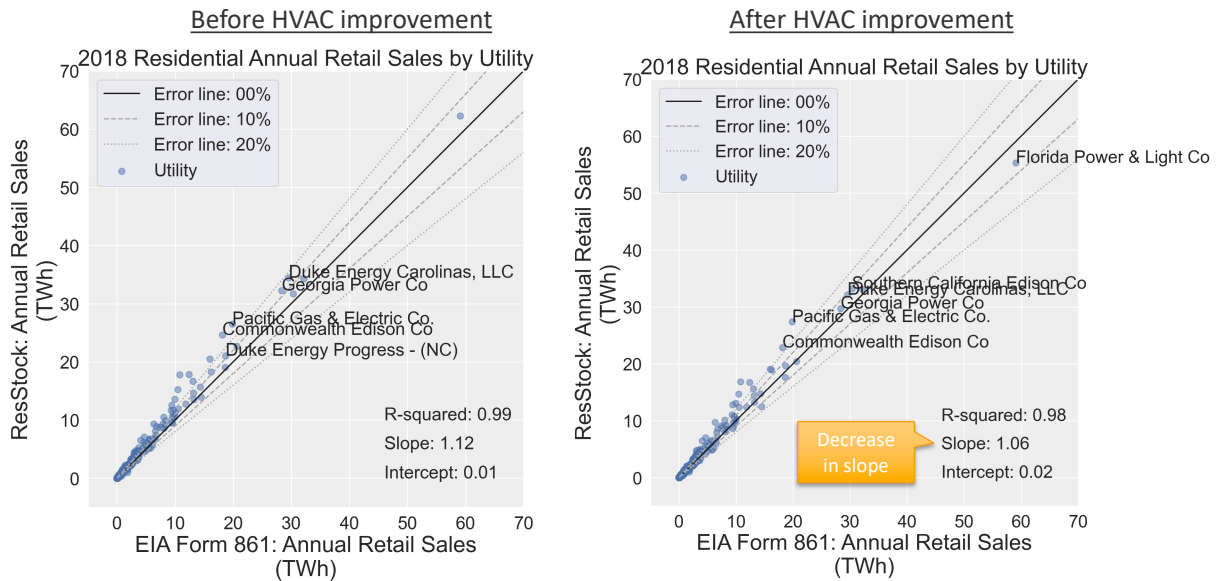
**Figure 80. Distribution of furnace efficiencies in 1994 and 2003. In 2009, a 15-year-old furnace was made in 1994, which used the top efficiency distribution. In 2018, a 15-year-old furnace was made in 2003, which used the bottom efficiency distribution. We updated the reference year for equipment age from 2009 to 2018, which increased equipment efficiencies for all HVAC equipment and refrigerators.**

### Impact of Updated HVAC Efficiency Distributions

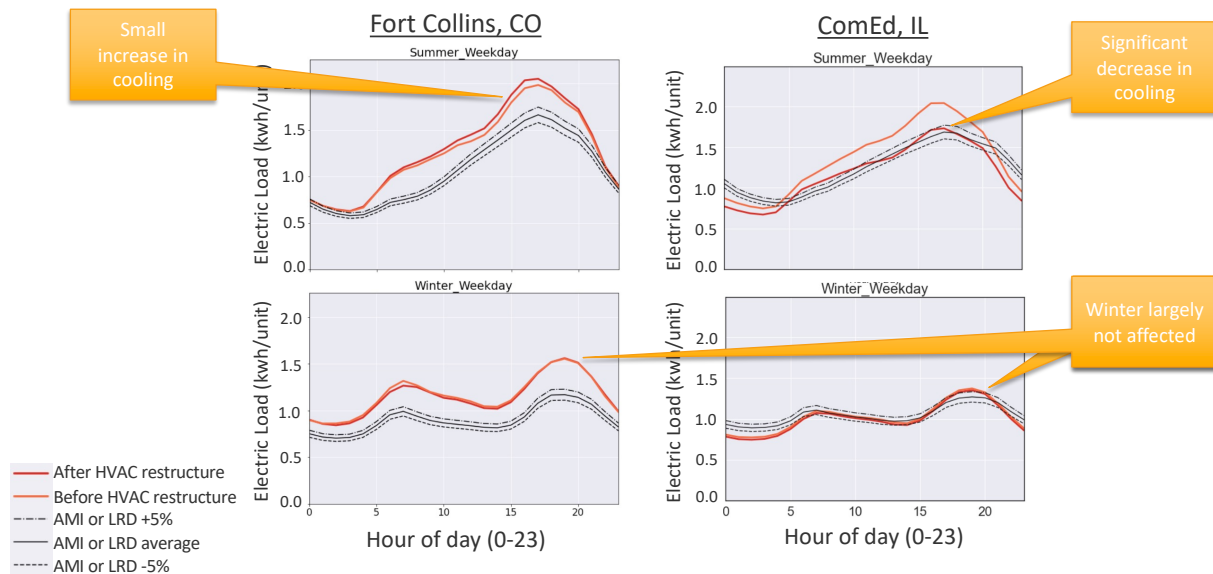
The impact of the improved HVAC efficiency distributions described above was evaluated along with the restructure of HVAC characteristics described in *Improving the Structure of HVAC Characteristics*. Figure 81 shows that the overestimation of annual electricity sales by utility for 2018 was reduced by half because of these changes. The slope of the modeled vs. measured line fit was reduced from 1.12 to 1.06.

Figure 82 shows the impact of these changes on the AMI comparison of average summer weekday 24-hour profiles for Fort Collins, CO, and ComEd (Illinois). We see a significant decrease in summer weekday demand in ComEd, an area with older housing stock and higher saturation of room air conditioners. We see a small increase in summer weekday demand for Fort Collins, an area with relatively newer housing stock. Electric heating was largely unaffected.





**Figure 81. Impact of updated HVAC efficiency distributions on annual electricity sales by utility—the update reduced the average overestimation by half, from a slope of 1.12 to 1.06. Note: final comparison is shown in Section 4.**



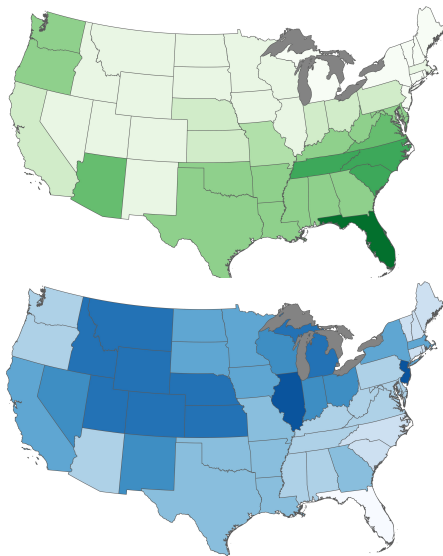
**Figure 82. Impact of updated HVAC efficiency distributions on AMI comparisons of average summer weekday 24-hour profiles for Fort Collins, CO, and ComEd (Illinois). We see a significant decrease in summer weekday demand in ComEd, an area with older housing stock and higher saturation of room ACs. We see a small increase in summer weekday demand for Fort Collins, an area with relatively newer housing stock. Electric heating was largely unaffected. Note: final comparison is shown in Section 4.**

#### Heating Fuel Type—Higher Geographic Resolution

During EULP residential calibration of Region 2, we revised the heating fuel geographic dependency from reportable domains (16 states and 10 groups of states) using RECS 2009 data to 2,400 PUMAs using PUMS 5-year 2016 data in order to increase the local accuracy of our results. The pre-EULP and post-Region 2 heating fuel distributions are shown in Figure 84. We used the state average values from PUMS for PUMAs with sample sizes below 10.

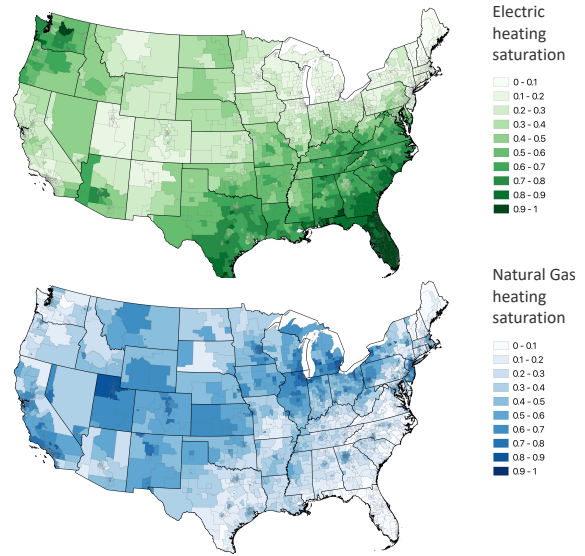
Before: the EULP project

Heating fuel described by groups of states

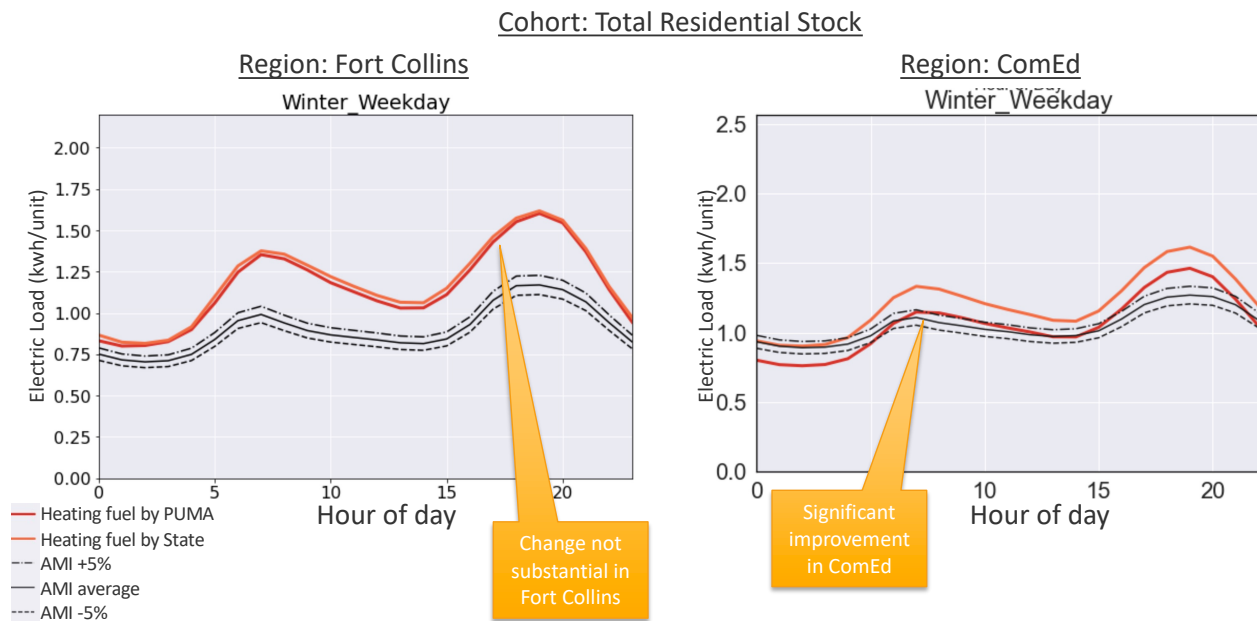


After: Calibration region 2

Heating fuel described by PUMA



**Figure 83. Saturations of heating fuels by state, as used by ResStock pre-EULP, and by PUMA, as used by ResStock starting in Region 2 of EULP residential calibration**

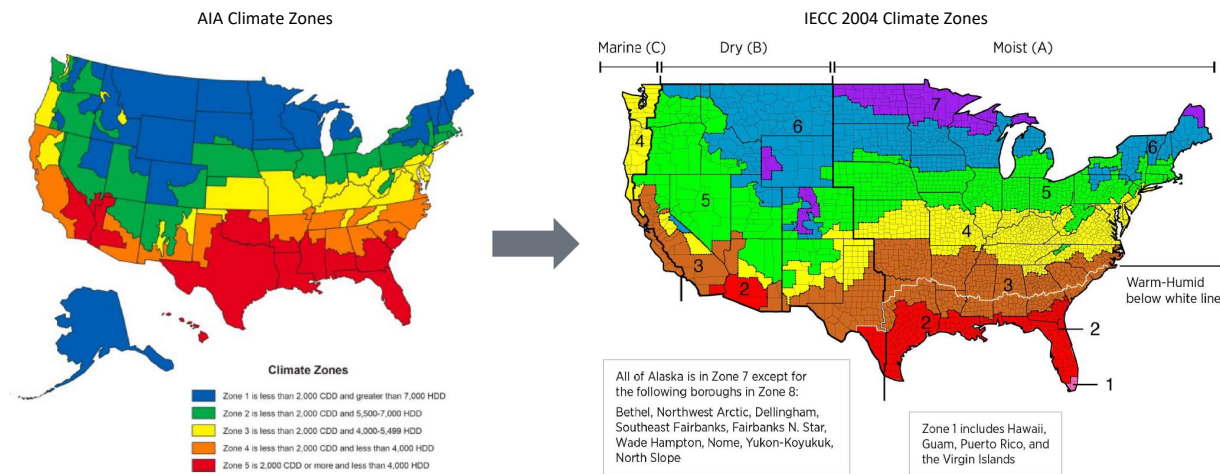


**Figure 84. Impact on electric load of moving from state to PUMA area for heating fuel geographic resolution, shown for the Fort Collins and ComEd utility service areas. Note: final comparison is shown in Section 4.**

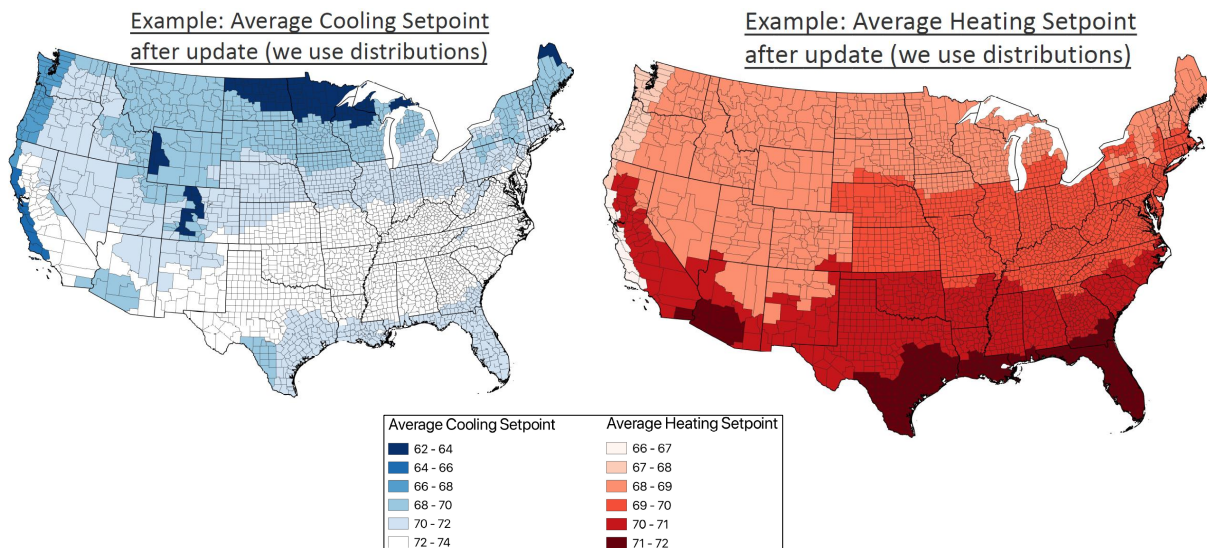
The impact of this change depended on how well an area's heating fuel distributions were represented by the larger state or regional average from RECS 2009 that was used previously. Two utility service area comparisons are shown in Figure 84—Fort Collins's service territory sees only minor change in overall electric load results, whereas the change in ComEd is more substantial.

### Accounting for Climate Zone Humidity in HVAC Setpoint Distributions

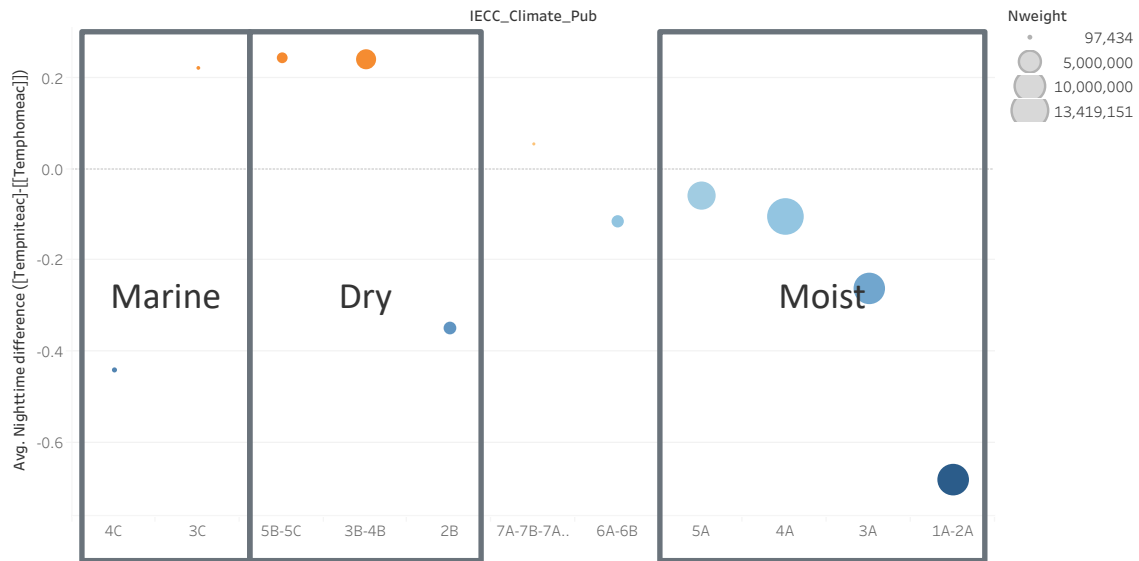
Prior to EULP calibration, heating and cooling setpoint-related distributions in ResStock were derived from RECS 2009 microdata (EIA 2013) and depended on the five American Institute of Architects (AIA) climate zones (Figure 85). The introduction of these distributions was discussed in Das, Wilson, and Williams (2021). During residential calibration of Region 2, we updated setpoints, setback magnitudes, and setback schedules for heating and cooling systems—still using RECS 2009 microdata, but depending instead on the IECC 2004 climate zones, which distinguish between moist, dry, and marine moisture regimes (Figure 85).



**Figure 85. AIA climate zones (left) and IECC 2004 climate zones (right). ResStock was updated to use 15 IECC 2004 climate and moisture zones to determine heating and cooling setpoint-related inputs.**



**Figure 86. Average heating and cooling setpoint values by IECC 2004 climate zone. Actual setpoint values are implemented as distributions within each climate zone that also depend on housing type and HVAC type. Although not many homes in Marine climates have air conditioning, it is notable that Marine climates (3C and 4C) have average cooling setpoints that are much lower than in Dry and Moist climates with similar degree day ranges (3A, 3B, 4A, and 4B). This pattern holds true for heating; climates 3C and 4C have lower average heating setpoints than 3A, 3B, 4A, and 4B. It is also notable that 5B (Dry) has lower average heating setpoint than the corresponding 5A (Moist).**



**Figure 87. Average difference between nighttime and home air-conditioning setpoints by IECC climate zone. These data are used to calculate the cooling setpoint offsets in ResStock.**

The primary advantage of these updates is to better capture the relationship of climate humidity with heating and cooling setpoint preferences. This relationship can be seen in Figure 86, where there are notable differences between average setpoints between Marine, Dry, and Moist zones within the same degree day range. We introduced the same IECC climate zone dependency for distributions of setback magnitude and timing behavior. Figure 87 demonstrates how nighttime cooling setpoint behavior is closely tied to both climate and humidity. On average, households in Dry climates shift their cooling setpoints higher at night, while households in Moist climates shift to a lower cooling setpoint at night.

Figures 89 and 90 provide examples of how loads changed in response to these updates, in combination with the increased setpoint schedule diversity, as discussed in the *HVAC Setpoint Schedule Diversity* section. In general, changes in overall magnitude can be attributed to the climate zone dependency for HVAC setpoints, while changes to the shape of the load profiles can be attributed to the schedule changes.

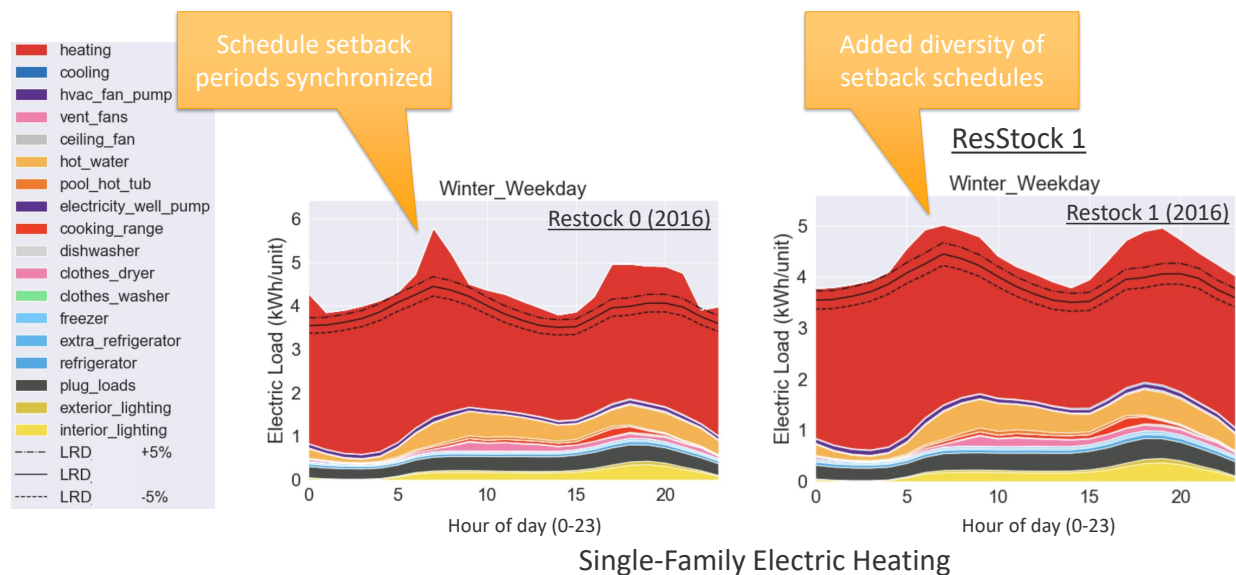
#### *HVAC Setpoint Schedule Diversity*

In 2019, ResStock switched from constant heating and cooling setpoint schedules to more advanced schedules that include setpoint offset behavior, as derived from RECS 2009 (EIA 2013). We introduced new options that include offset magnitude and offset schedules to describe the extent and time that homes apply thermostat offsets, respectively (Das, Wilson, and Williams 2021). These options were inferred from daytime occupied, daytime unoccupied, and nighttime temperatures reported in RECS 2009, and are dependent on location. Exact timing of the setpoint changes is not collected by RECS, so we assume the hours that correspond to nighttime and daytime unoccupied (i.e., working hours) periods. Heating setpoint offsets of 3°F, 6°F, and 12°F, and cooling setpoint offsets of 2°F, 5°F, and 9°F were used as the offset magnitude bins, and the offset schedule options are shown in Table 15. For homes that do not have a constant setpoint, these offset schedules and magnitudes are applied to the baseline distribution of constant setpoints.

Although this effort more accurately captured the hourly operation of HVAC systems, it introduced coincident peaks caused by the synchronized recovery periods at the end of each offset. To remedy this, we introduced more diversity in the offset schedules by staggering the timing up to  $\pm 2$  hours. This expanded the setpoint offset schedule options by evenly dividing each schedule probability (six cooling schedules and three heating schedules) to shift timing of the offset from -2 hours to +2 hours at increments of 1 hour. Increasing the setpoint schedule diversity helped flatten peaks caused by synchronized offsets, as shown in Figure 88.

**Table 15. List of heating and cooling setpoint offset schedules applied in ResStock. Offset schedules are assumed to be 9 a.m. to 5 p.m. for daytime, and 10 p.m. to 7 a.m. for nighttime offsets. With this update, we introduced staggered versions of the schedules that are shifted forward or backward to prevent synchronized schedules across a large set of homes.**

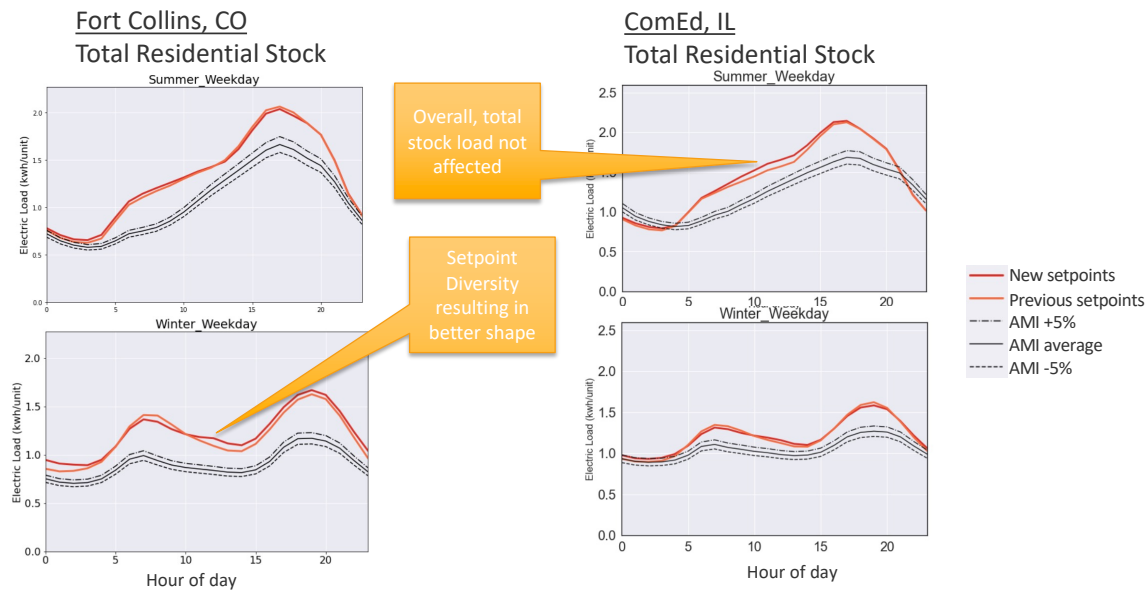
Heating Setpoint Offset Period Options	Day Setback Night Setback Day & Night Setback No Offset
Cooling Setpoint Offset Period Options	Day Setup Night Setup Night Setback Day & Night Setup Day & Night Setback Day Setup & Night Setback No Offset



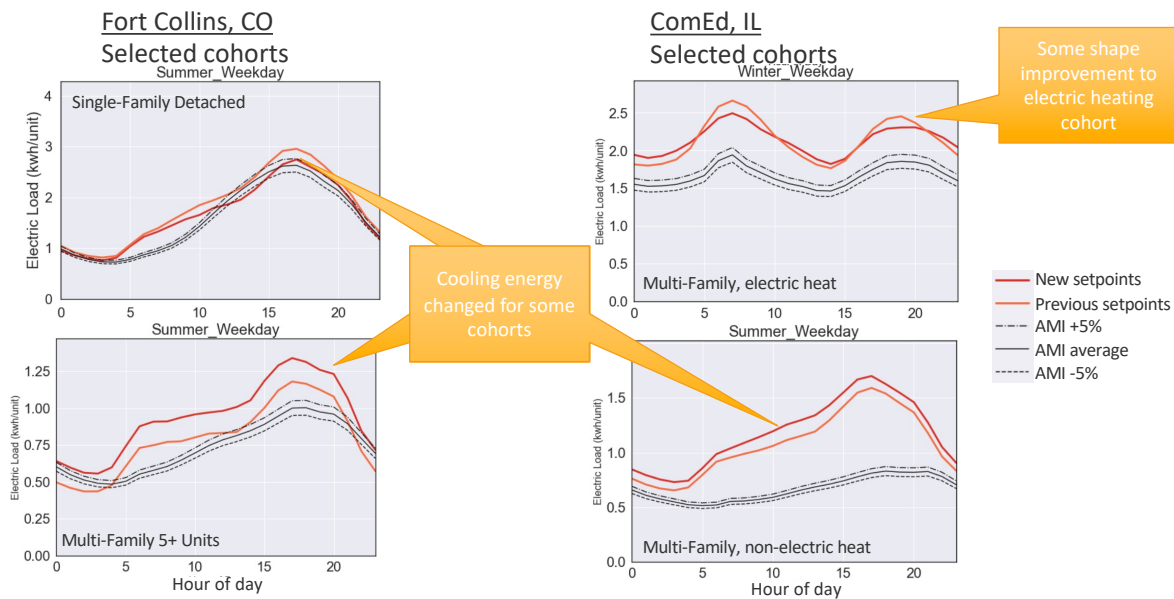
**Figure 88. Effects of increased setpoint schedule diversity ( $\pm 2$  hours) to the total residential building stocks of ComEd utility territory in Illinois. Note: final comparison is shown in Section 4.**

After introducing the thermostat schedule shift to generate more realistic profiles, we continued to observe higher than expected peaks driven by heating and cooling in other regions. To address this, we further increased the schedule diversity by staggering the offset schedules up to  $\pm 5$  hours, expanding the setpoint schedule options to range from -5 hours to +5 hours at increments of 1 hour. Increasing the setpoint schedule diversity further helped flatten peaks caused by synchronized offsets, mostly affecting individual cohorts of housing type and HVAC fuel type, as show in Figures 89 and 90.





**Figure 89. Effects of setpoint schedule diversity ( $\pm 5$  hours) and setpoint distribution changes to the total residential building stocks of Fort Collins, CO, and the ComEd utility territory in Illinois. In general, impacts to the shape of the loads can be attributed to changes in the setpoint schedule diversity, while changes in magnitude are likely a result of expanding the setpoint distributions to be based on climate zone humidity. Note: final comparison is shown in Section 4.**



**Figure 90. Effects of setpoint schedule diversity ( $\pm 5$  hours) and setpoint distribution changes to residential building stock cohorts in Fort Collins, CO, and the ComEd utility territory in Illinois. In general, impacts to the shape of the loads can be attributed to changes in the setpoint schedule diversity, while changes in magnitude are likely a result of expanding the setpoint distributions to be based on climate zone humidity. Note: final comparison is shown in Section 4.**

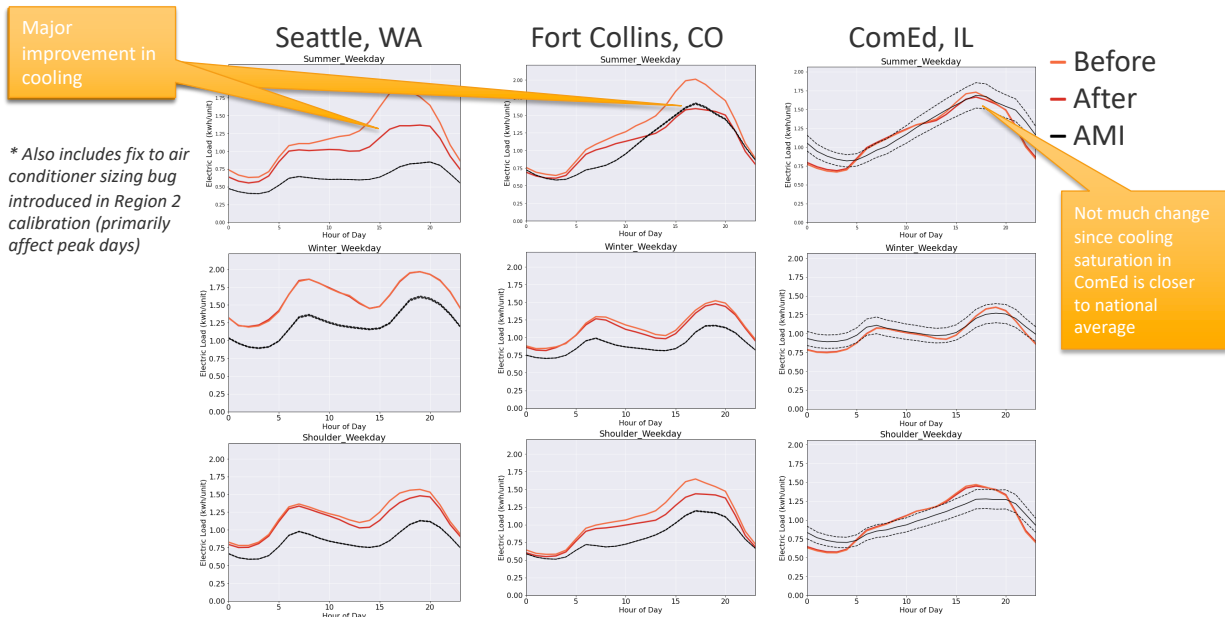
#### *Air Conditioner Saturation—Adding Climate Dependency*

The HVAC distribution restructure completed during Region 2 inadvertently removed the dependency of air-conditioner presence and type on location, giving all locations in the country the same distribution of homes with



central air conditioning or heat pump, room air conditioning, and no air conditioning. During Region 3 calibration, we added a climate dependency, using IECC climate zones, back into the air-conditioner type probability distributions to join the other dependencies of housing type, vintage, and heating type (ducts or not; heat pump or not).

With the addition of climate zones, the use of four dependencies can slice the 12,083 samples in RECS 2009 into segments with insufficient sample sizes, so we applied careful binning of response to ensure sufficient sample sizes ( $N > 10$ , based on EIA guidance for using RECS microdata (EIA 2017)) for all combinations.<sup>10</sup> The AC saturation update made a very significant impact on summer loads in Seattle and Fort Collins; AC saturation in ComEd was close to the national average, so the change was minimal (Figure 91).



**Figure 91. Impact of AC saturation climate dependency on AMI comparisons for Seattle, Fort Collins, and ComEd. The AC saturation update made a very significant impact on summer loads in Seattle and Fort Collins; AC saturation in ComEd was close to the national average, so the change was minimal. These comparisons include the impact of an AC sizing fix, but that fix only affected peak days. Note: final comparison is shown in Section 4.**

### Air-Conditioner Sizing Fix

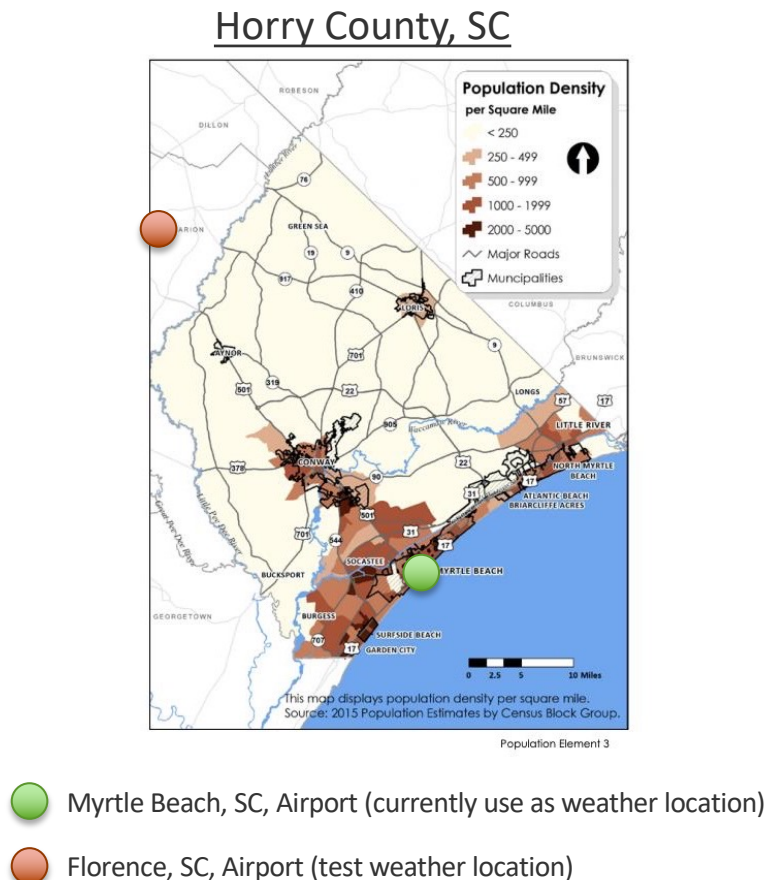
The stochastic occupancy feature added during Region 2 inadvertently increased the magnitude of internal gains used for the design cooling load calculation for air-conditioner sizing. This did not significantly affect annual energy use, only peak demand (about 1% of hours). After this discovery, we implemented automated regression testing (before vs. after checks) on heating and cooling capacities and other output variables such as unmet hours for heating and cooling setpoints. The impact of this change can be seen in the top 10 summer peak day QOI plots for ComEd, Fort Collins, and Seattle, included in Section 4, Results. The design cooling load calculation bug was introduced with the stochastic occupancy model in run number 16 (at the start of Region 2) and was fixed in run number 29 (at the start of Region 3).

### Sensitivity of Coastal vs. Non-Coastal Weather Stations

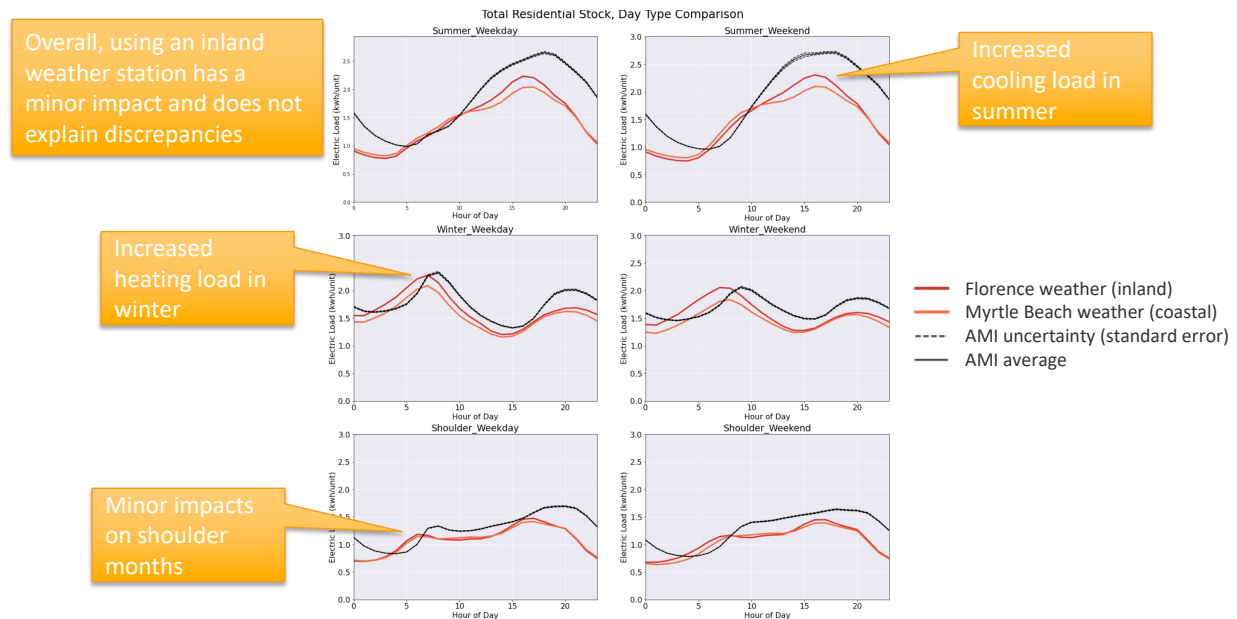
During Region 4 calibration, we investigated reasons for underestimation of summer cooling load for the Horry Electric Cooperative AMI dataset. One theory was that the main weather station in the county is at Myrtle Beach Airport,

<sup>10</sup>Due to low sample sizes for some Heating Types, Heating Type data for Non-Ducted Heating and None were grouped. Due to low sample sizes for some Housing Types, Housing Type data were grouped into: (1) Single-Family Detached and Single-Family Attached, (2) Multifamily 2–4 units and Multifamily 5+ units, and (3) Mobile Homes. Due to low sample sizes for some Vintages, Vintage ACS (20-year bins) were used instead of the typical 10-year bins used for RECS data. If a sample has both Central AC and Room AC, we assume it has Central AC only. If a sample indicates using a heat pump for AC but does not indicate using a heat pump for heating, then we either assign it a heat pump for heating (if electric heating was indicated), or we assign it Central AC (if non-electric heating was indicated).

which is located within 2 miles of the Atlantic coast. Horry Electric Cooperative serves much of the population of Horry County. Although the population density is highest along the coast, some of the population lives further inland (Figure 92). We investigated the sensitivity of summer load to weather station by simulating the Horry County housing stock using weather data from a station further inland, in Florence, SC (Figure 92). As shown in Figure 93, using an inland weather station had a minor impact on cooling and heating load, and does not fully explain the discrepancy.



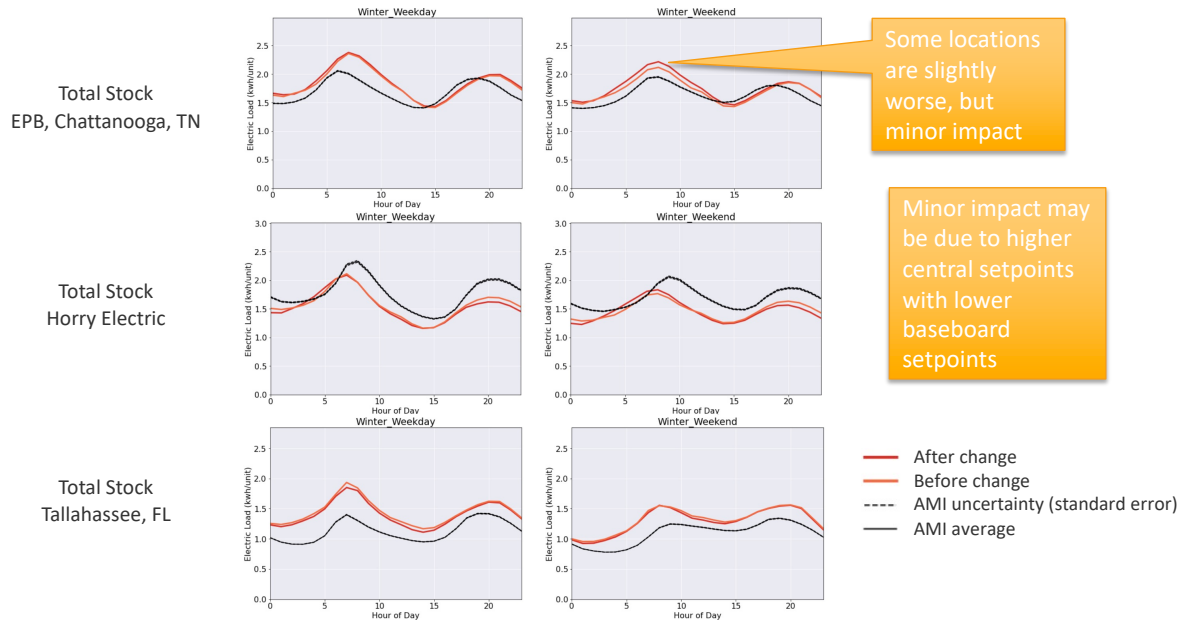
**Figure 92. Population density in Horry County, SC, is highest along the coast, but some of the population lives further inland, so we investigated the sensitivity of summer load to weather station by simulating Horry County housing stock using weather data from an inland station in Florence, SC. Source: Horry County (2018)**



**Figure 93. Using an inland weather station had a minor impact on cooling and heating load, and does not fully explain the discrepancy**

### *Zonal Electric Heating Setpoints*

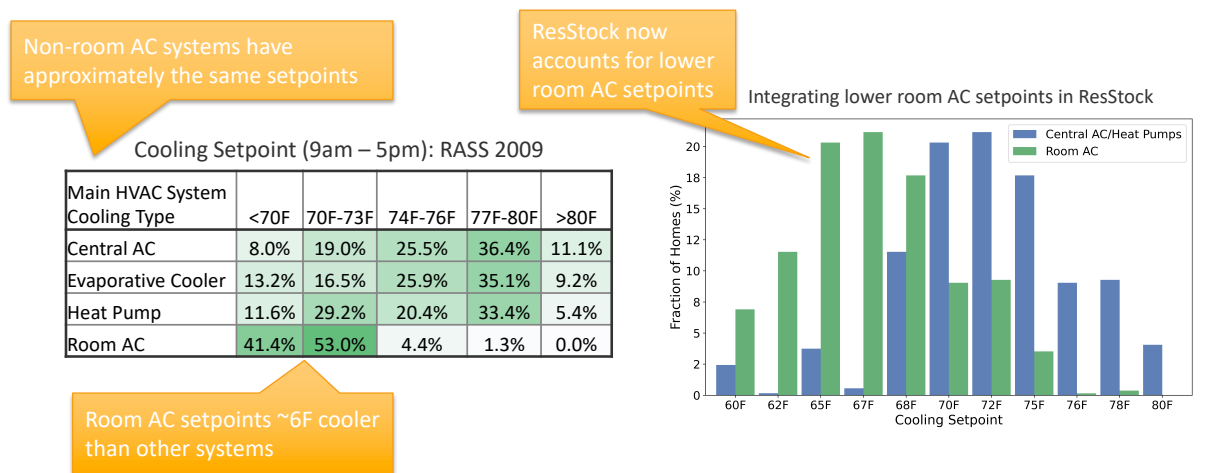
The NEEA 2011 Residential Building Stock Assessment (NEEA 2021c) provided evidence that homes with baseboard or plug-in electric heaters use less heating energy than homes with electric furnaces (RTF 2013). As documented in RTF (2013), modeling of those homes with the SEEM tool overestimated baseboard and plug-in heating, suggesting that zonal temperature control in different rooms could be causing inaccuracies. We found that RECS 2009 (EIA 2013) data on heating setpoints for baseboard and portable electric heaters are lower on average than ducted electric furnaces and heat pumps. By introducing zonal electric equipment as a dependency to heating setpoint, we hoped to better represent regional heating loads in areas with electric baseboard heating. This change was motivated by electric heating in the Southeast, but we found that this change had a greater impact in colder climates, because colder climates tend to have higher saturation of electric baseboard heating. Figure 94 shows how this change had minimal impact for the three AMI comparisons in the Southeast.



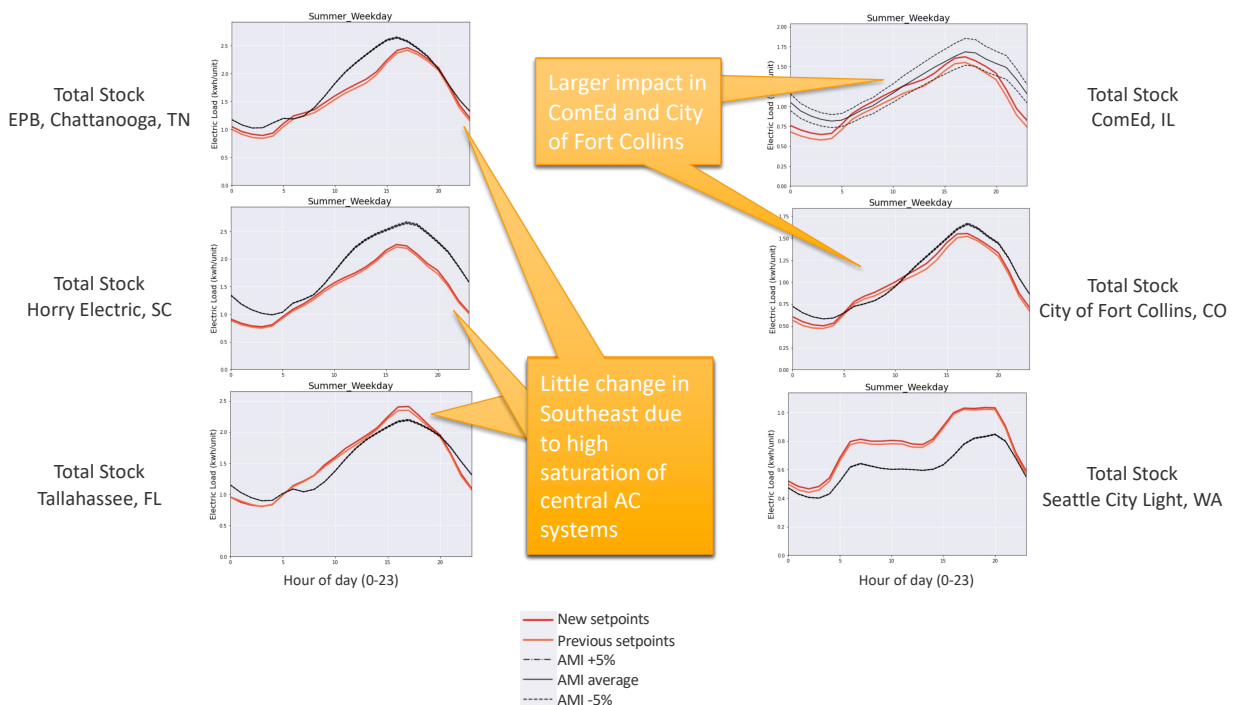
**Figure 94. Impact of zonal electric heating setpoint update on AMI comparisons for Chattanooga, TN, Horry, SC, and Tallahassee, FL. Note: final comparison is shown in Section 4.**

#### *Update Room Air-Conditioner Setpoints*

RECS 2009 and 2015 collected data on occupant-reported summer temperature inside the home (when someone is home during the day, when no one is home during the day, and at night) but only for homes with central air conditioners (EIA 2013). California's 2009 Residential Appliance Saturation Study (RASS) provides data for cooling setpoints broken out by HVAC cooling type, and indicates that homes with room air conditioners maintain a lower setpoint than homes with other cooling systems, such as central air conditioners and heat pumps (KEMA 2010). To capture this in our modeling, we introduced cooling system type as a dependency to cooling setpoint in ResStock. Cooling setpoint data previously relied solely on RECS 2009 and was based on climate zone and housing type. We continue to use the RECS 2009 cooling setpoint distributions, but shift the distribution of room air-conditioner setpoints by 6°F colder for all climate zones and housing types to align with RASS 2009, as shown in Figure 95. Figure 96 shows hourly impacts of the setpoint changes, which primarily affect colder regions with higher saturations of room air conditioners.



**Figure 95. RASS 2009 distribution of cooling setpoints based on cooling system type (left) and the distributions of room air-conditioning setpoints based on cooling type in ResStock after introducing RASS 2009 data (right).**

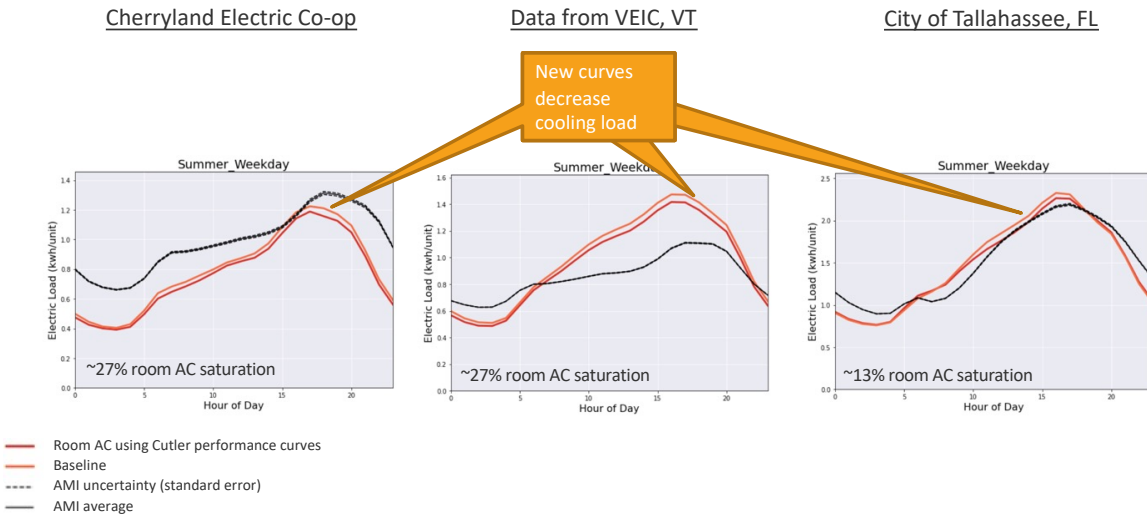


**Figure 96. Impact of room air-conditioner setpoint update on six summer weekday AMI comparisons. Influences are less prominent in warmer regions with a low saturation of room air conditioners. Note: final comparison is shown in Section 4.**

### Updating Room Air-Conditioner Performance Curves

Performance curves are used in our EnergyPlus modeling to define power draw and efficiency of air conditioners and heat pumps at a range of outdoor temperatures. The original implementation for modeling window/room air-conditioning units in ResStock relied on performance curves derived from a small sample of tested units (Winkler et al. 2012) which were later determined to likely be limited in their representativeness. We switched to standardized air-conditioner performance curves (Cutler et al. 2013), which are derived with data from numerous air-conditioner

manufacturers tested across multiple locations, and are now also used in the OpenStudio-HPXML-based workflows (Horowitz et al. 2019). As shown in Figure 97, the new performance curves reduce cooling loads on average, and impact regions with high a saturation of room air conditioners, such as the Northeast.

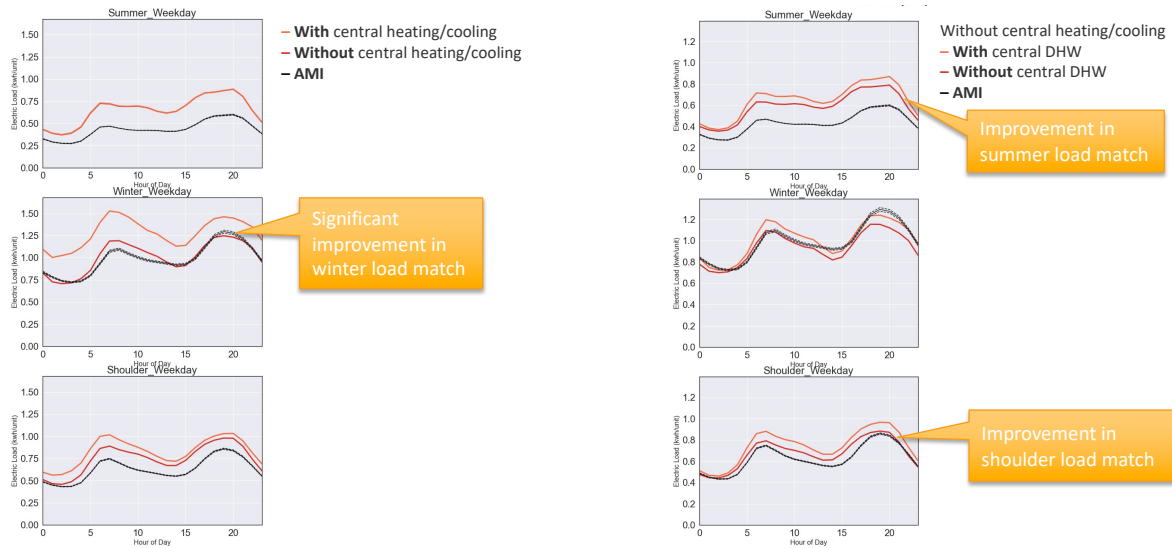


**Figure 97. Impact of standardized air-conditioner performance curves for room air-conditioner modeling on AMI comparisons for the Cherryland, MI, VEIC, and Tallahassee, FL data areas. Note: final comparison is shown in Section 4.**

#### *Removing Multifamily Central Heating and Water Heating Electricity*

During Region 3 of calibration, the overestimation of electric heating in multifamily buildings in Seattle led us to investigate whether building-level meters for centrally metered HVAC and domestic hot water (DHW) are included in the Seattle residential AMI data. We learned from Seattle City Light that individual units typically have a residential rate code, whereas multifamily building common areas and central meters are typically given a commercial rate code. Based on a conversation with EIA staff, we expect this assumption to hold true for most utilities. We removed central system HVAC and DHW from ResStock results for Seattle to see how this affects the comparison (Figure 98). To accomplish this, we used data from RBSA (Oregon and Washington) and RECS (remainder of the United States) on the prevalence of central HVAC and DHW, and introduced these distributions to ResStock. Note that central DHW is still modeled as in-unit water heaters, but tagging results facilitates the updated AMI comparisons. This modification is included in the final AMI comparisons presented in Section 4, Results, for Seattle City Light only. We expect this effect to be much smaller or non-existent in other utility AMI comparisons where multifamily building-level *electric* heating and water heating is much less common.





**Figure 98. Removing building-level central space heating electricity (left) and both central space heating and DHW electricity (right) from ResStock results dramatically improves the estimation of winter and shoulder load for Seattle aggregate multifamily meter AMI data. Note: final comparison is shown in Section 4.**

### 3.2.8 Thermal Envelope Updates

#### Updating Window-to-Wall Ratios

Prior to EULP calibration, ResStock used an assumption of 15% window-to-wall ratio (WWR) for all housing units. We know from past experience that there is diversity in actual WWR, and WWR can be much higher than 15% for some multifamily buildings. Data on WWR are not commonly available. We use data from NEEA’s 2016–2017 Residential Building Stock Assessment II (RBSA II), which includes measurement of window and wall areas for a representative sample of buildings in Washington, Oregon, Idaho, and Montana (NEEA 2021c). We fit exponential Weibull distributions to the WWR data in RBSA for different housing types, including low-rise (1–3 story) and mid-rise (4–7 story) multifamily buildings, and use these fits to derive the input distributions for ResStock (Figure 99). Sample size for high-rise (8+ story) multifamily buildings in RBSA II was low, so we assume all units in 8+ story buildings have 30% WWR.

Dependency=Geometry Building Type Height	6% WWR	9% WWR	12% WWR	15% WWR	18% WWR	30% WWR	sampling_ probability
Single-Family Detached	9%	29%	27%	17%	15%	4%	63%
Single-Family Attached	9%	29%	27%	17%	15%	4%	6%
Mobile Home	9%	29%	27%	17%	15%	4%	6%
Multifamily with 2-4 Units	10%	28%	26%	16%	15%	4%	8%
Multifamily with 5+ units, 1-3 stories	2%	13%	20%	19%	29%	17%	17%
Multifamily with 5+ units, 4-7 stories	1%	6%	10%	12%	27%	43%	1%
Multifamily with 5+ units, 8+ stories	0%	0%	0%	0%	0%	100%	0%

**Figure 99. Distributions of window-to-wall ratios for different residential building types were derived from the 2016–2017 RBSA II (NEEA 2021c). Sampling probability indicates the upstream probability of each row getting sampled by ResStock.**

#### Updating Air Infiltration Rates and Diversity

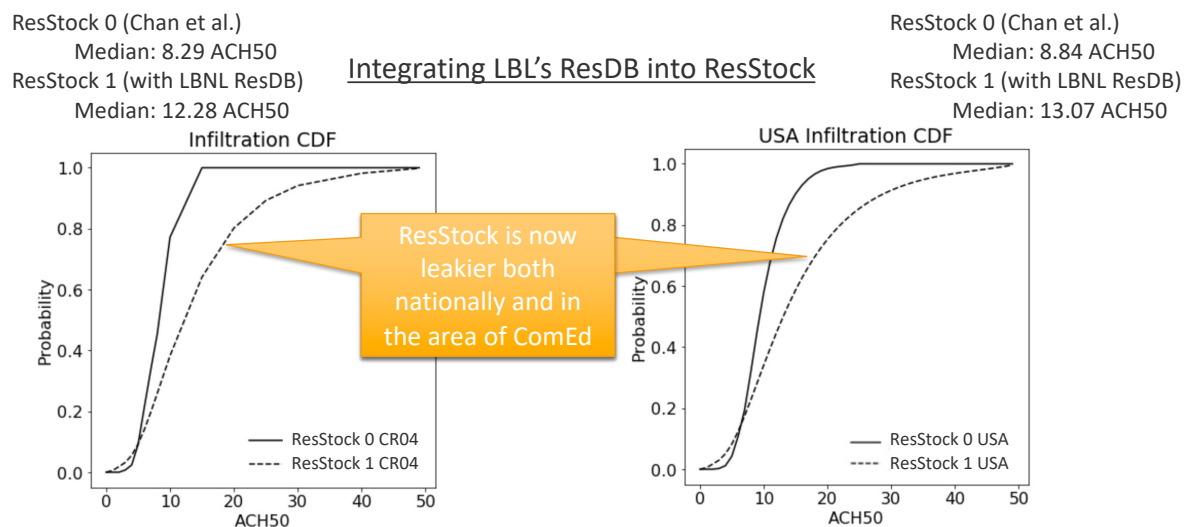
Prior to EULP calibration, air leakage (infiltration) for homes modeled in ResStock was based on a regression equation published by Chan, Joh, and Sherman (2012), which was derived from blower door air leakage measurements from 134,000 single-family detached homes, collected by LBNL researchers in the Residential Diagnostics Database (ResDB) (Walker, Chan, and Spears 2020). We had simplified the regression equation to use a reduced set of inputs:

climate region, vintage, floor area, and number of stories. The primary disadvantage of using the regression equation was that it only provided a single infiltration value for a given combination of dependency inputs. In other words, 100% of two-story 1,500–2,499 ft<sup>2</sup> homes built in the 1970s in the Southeast would be assigned a leakage value of 15 air changes per hour at 50 Pa (ACH<sub>50</sub>)—there was no diversity in infiltration rates for that segment of homes.

This update replaces the regression-based air infiltration inputs with probability distributions derived from the same database of air leakage measurements. We fit lognormal distributions to the cumulative distribution functions for ACH<sub>50</sub> published on the ResDB website (Walker, Chan, and Spears 2020). The fitted lognormal distributions were then used to assign probabilities to the discrete infiltration rate bins used by ResStock.

We made several assumptions in deriving and applying these infiltration distributions. The cumulative distribution functions are for single-family detached homes only, but we apply them to all housing types, reducing infiltration for units in multifamily and single-family attached buildings based on the fraction of their walls, ceilings, and floors that are exposed to exterior conditions (as opposed to adjacent units or corridors). Data from certain IECC climate zones were copied to other climate zones. IECC climate zone 2A was copied to 1A; IECC climate zone 6A was copied to 7A; and IECC climate zone 6B was copied to 7B. The ResStock vintage bins are more refined than the bins in ResDB. As a result, certain decade bins had to be copied to other bins. Cumulative distribution functions from the “Before 1960” ResDB category were used for the <1940, 1940s, 1950s vintages in ResStock. Cumulative distribution functions from the “After 2000” ResDB category were used for the 2000s and 2010s vintages in ResStock. Homes are assumed to not be Weatherization Assistance Program qualified and not ENERGY STAR certified.

The effect of this update on the distribution of infiltration rates can be seen in Figure 100. Despite being derived from the same source data, infiltration rates increased by about 50% with the new approach. We suspect this is because the regression equation was not able to capture the tail of the lognormal distributions, and there may have been other differences in assumptions between the two approaches. The increased diversity can be seen particularly in how the Great Lakes distribution was smoothed out. We evaluated the impact of this update on ComEd AMI comparisons, but the impact was obscured by a weather data bug that was present at the time.



**Figure 100. Distribution of ResStock infiltration rates before and after update for the Great Lakes (CR04) region (left) and nationally (right). Overall, infiltration rates increased by about 50%. The increased diversity can be seen particularly in how the Great Lakes distribution was smoothed out.**

### Roof Material Distributions

Prior to EULP calibration, all roofs were modeled as medium color asphalt shingles. We updated the roof material distributions based on data from RECS 2009 (example shown in Figure 101). Although this update was motivated by a different project and not an observed error, we wanted to evaluate the impact on calibration progress. Because this change was made during Region 3, we evaluated the impact on Seattle AMI comparisons and found that the change was negligible (Figure 102).

Dependency= Geometry Building Type RECS	Dependency= Location Region	Option= None	Option= Asphalt, Medium	Option= Composition Shingles	Option= Metal, Dark Slate	Option= Tile, Clay or Ceramic	Option= Tile, Concrete	Option= Wood Shingles
Mobile Home	CR06 (WA, OR)	0%	0%	49%	45%	0%	0%	7%
Single-Family Attached	CR06 (WA, OR)	0%	9%	74%	0%	4%	0%	12%
Single-Family Detached	CR06 (WA, OR)	0%	5%	84%	4%	0%	1%	6%

Figure 101. Example of updated roof material distributions for CR06 (Washington and Oregon), based on data from RECS 2009

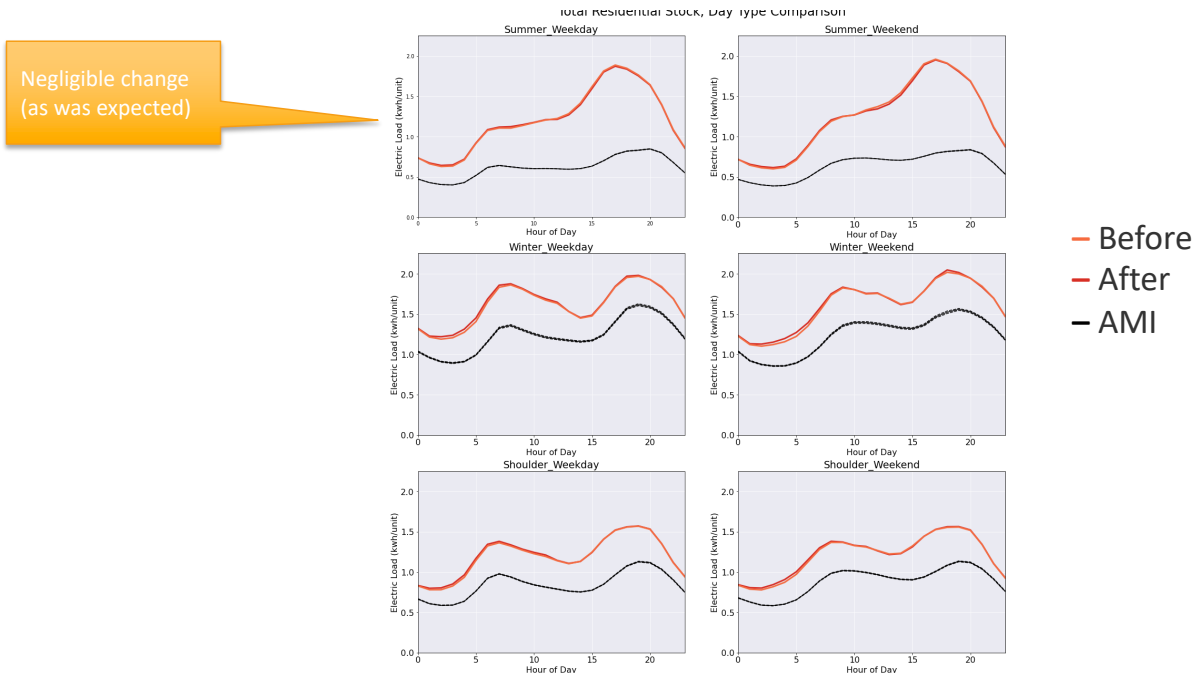


Figure 102. Impact of updated roof material distributions on Seattle AMI comparison—the change was negligible. Although this update was motivated by a different project and not an observed error, we wanted to evaluate the impact on calibration progress. Note: final comparison is shown in Section 4.

### Foundation Type Distributions

Prior to EULP calibration, ResStock’s distribution of foundation types was drawn from ORNL’s “Building Foundation Design Handbook” (Labs et al. 1988). In Region 3 of residential calibration, we updated the foundation types to draw from RECS 2009 data. RECS responses can indicate that a home has more than one foundation type, which we do not model in ResStock, so we divide the sample weight of those responses equally among the indicated foundation types. We also added dependencies on IECC climate zones and on vintage in addition to the existing dependency on housing type. A comparison of these distributions before and after the update is shown in Figure 103. Figure 104 shows the modest impact of this change on three residential AMI comparisons.

### Before:

Depends on state (1988 source)

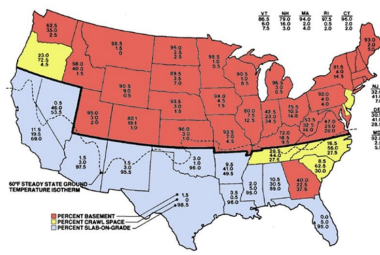
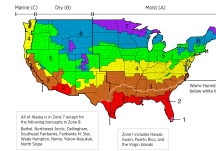


Figure 3. Share of residential foundations by state (Labs, et al., 1988)<sup>1</sup>  
From Building Foundation Design Handbook.  
ORNL/Sub-86-72143/1, Oak Ridge National Laboratory/US Dept. of Energy.

### After:

Depends on  
IECC Climate Zone,  
building type, and  
vintage



For example:

Dependence	Option=	Option=	Option=	Option=
CV=	Heated	Pier and	Slab	Unheated
ASHRAE	Bsmt	Beam		Bsmt
IECC				
Climate Zone	2004 Building Type	ReCS	Option=	Option=
			Crawl	Slab
4C	Single-Family Detached	<1940	55%	13%
4C	Single-Family Detached	1940-59	39%	2%
4C	Single-Family Detached	1960-79	55%	0%
4C	Single-Family Detached	1980-99	68%	2%
4C	Single-Family Detached	2000-09	64%	0%
4C	Single-Family Detached	2010s	64%	0%

#### Assumptions:

- All mobile homes have Pier and Beam foundations.
- Multi-family buildings cannot have Pier and Beam and Heated Basements
- Single-family attached buildings cannot have Pier and Beam foundations

Figure 103. Foundation type distributions before and after ResStock update

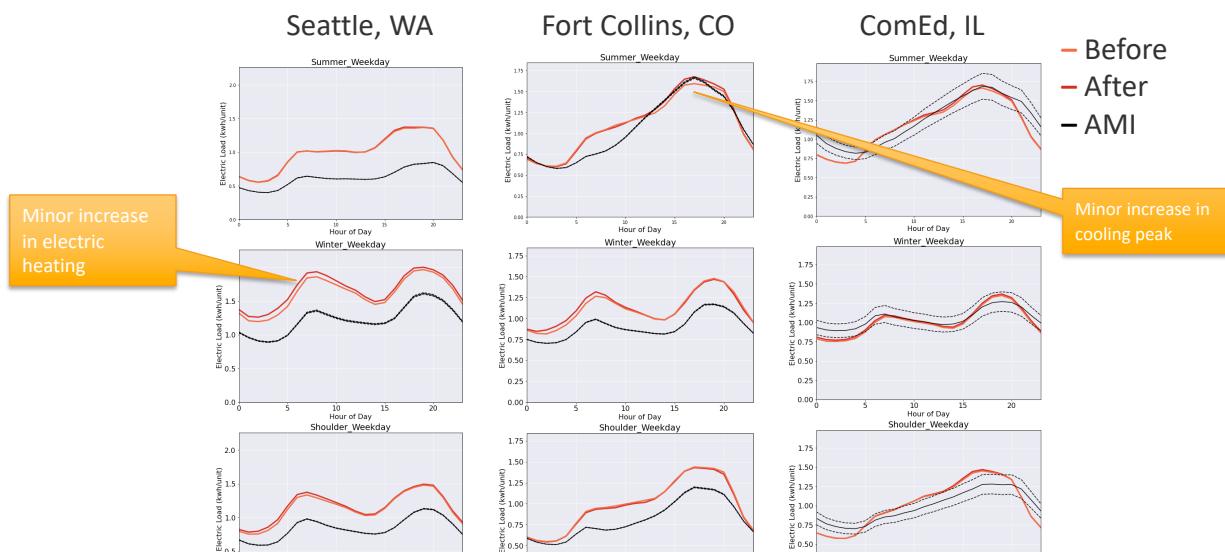
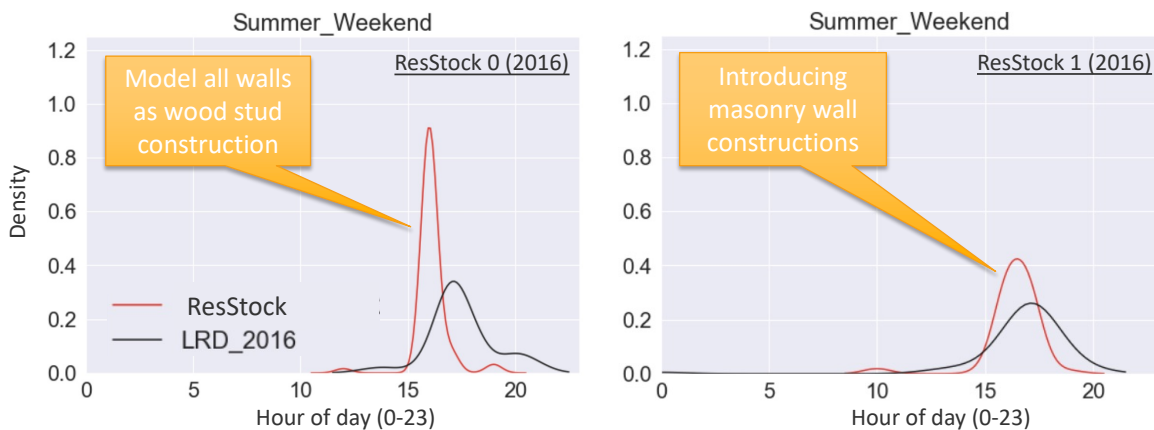


Figure 104. Impact of foundation distribution update on three utility service territory model results. Note: final comparison is shown in Section 4.

### Adding Modeling of Masonry Walls

At the start of the EULP project, all exterior walls were modeled as wood-framed walls in ResStock. We had input distributions for masonry vs. wood-framed walls in order to estimate which homes were eligible for drill-and-fill cavity insulation retrofits, but we had not yet developed the masonry wall assembly specifications necessary to model the masonry walls in EnergyPlus. To remedy this, we added options for multi-wythe brick walls and concrete block walls, and connected them to the previously defined input distributions from RECS 2009. These input distributions were subsequently refined, as discussed in the following sections. Because masonry walls have much more thermal mass than wood-framed walls, they slow the rate of heat transfer. For this reason, the modeling change had a significant effect on the timing of peak demand, as shown in Figure 105.



**Figure 105. Impact of modeling masonry wall constructions on the frequency of the daily peak occurring in a given hour for ComEd LRD and ResStock output for the predominant customer class of single-family gas-heated homes. Before the change (left) ResStock peaks were concentrated around 16:00 on most days. After the change (right) the ResStock peaks were more distributed over the afternoon hours, although still peaking earlier than the LRD.**

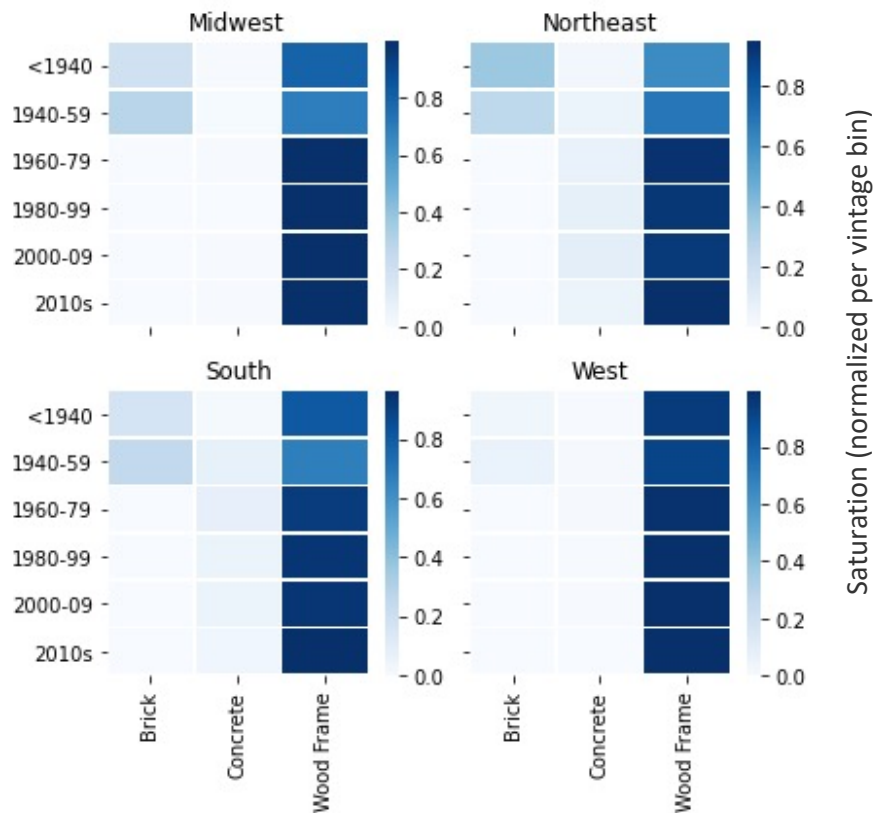
#### *Updating Masonry Wall Distributions*

Previously, the wall type distribution in ResStock came from the “WALLTYPE” field in RECS 2009, which is defined as the “major outside wall material” (EIA 2013). We interpret this as the material that can be observed from the outside. Therefore, it is ambiguous whether the “brick” option from RECS means a multi-wythe brick masonry wall or a wood-framed wall with 4-inch face brick. Additionally, because 2009 RECS has relatively few samples (12,083), we encountered insufficient sample sizes when querying wall type as a function of region, housing type, and vintage.

To update the description of wall construction type, we used the Homeland Infrastructure Foundation-Level Data (HIFLD) national parcel database (not available to the public), which consists of tax assessor data from millions of residential buildings in the United States (DHS 2021). The HIFLD parcel data include a field for “CONSTRUCTION\_CODE = Code indicating the type of construction (e.g., Brick / Concrete etc.)” (N=43M non-null values), which we can directly use as the wall construction type instead of needing to infer it from the exterior wall material. We queried this wall type as a function of housing type, state, and vintage. The primary effect of wall type is that brick masonry and concrete block wall assemblies have higher thermal mass than wood-framed walls. The mass can change the timing and magnitude of peak heating and cooling loads. As shown in Figure 106, the wall type is predominantly wood frame and most prominent in newer vintages.

#### *Updating Wall Exterior Finish Distributions*

The parcel data also include a field for “EXTERIOR\_WALL\_TYPE = Code indicating the type and/or finish of the exterior walls (e.g., Vinyl Siding, Brick Veneer)” (N=28M non-null values), so we used that to introduce distributions for exterior wall finish material. Previously all wall exterior finish material was modeled as light vinyl siding. We queried the distributions by wall type, vintage, and state as they show variation along these dimensions. For example, as shown in Figure 107, brick as an exterior finish (whether that is the surface of a multi-wythe brick masonry wall or a wood-framed wall with 4-inch face brick) is dominant in the Midwestern and Southern states but became less popular over time in the Northeast. Vinyl and wood are popular in the Northeast and West, in addition to stucco in the West. The primary effect of exterior finish is that 4-inch face brick adds thermal capacitance to wood-framed walls. A secondary effect is the impact of the reflectivity and absorbance of the material (e.g., light colors vs. dark colors) on the heat transfer through the wall. The parcel data do not include material colors, so we make assumptions based on the material type—an improvement over assuming all walls are finished with light vinyl siding.



**Figure 106. Wall type saturation by vintage bin and census region**

### Impact: Wall Type and Exterior Finish

As a result of the update to wall type, we saw the overall saturation of masonry walls drastically decrease and be replaced by wood-frame walls, which increased from 55% to 90% (Figure 106). Generally, a decrease in masonry wall type relative to wood frames leads to a lower building average thermal mass, and vice versa. Thermal mass provides thermal stability against outdoor temperature swings. The higher the thermal mass a building has, the slower its thermal response is to the outside elements. The exterior finish update introduced siding of darker colors, such as medium-dark wood and brick, which slightly increased the average surface absorptance of the buildings. These changes in thermal mass and surface absorptance caused minor increases or decreases in peaks for different utilities, as shown in Figure 108.

Although not directly relevant for this calibration effort, wall type is also significant because it determines the type of insulation retrofits that can be applied. For example, only wood-framed walls are eligible for drill-and-fill wall cavity insulation. Conversely, masonry walls can better support the weight of retrofit wall panel systems (Egerter and Campbell 2020). Similarly, wall exterior finish can determine the applicability and cost of different types of wall retrofits. When wood, metal, or vinyl siding needs to be replaced, continuous rigid insulation can be added under the new siding, which may have lower incremental cost than adding continuous rigid insulation over brick.



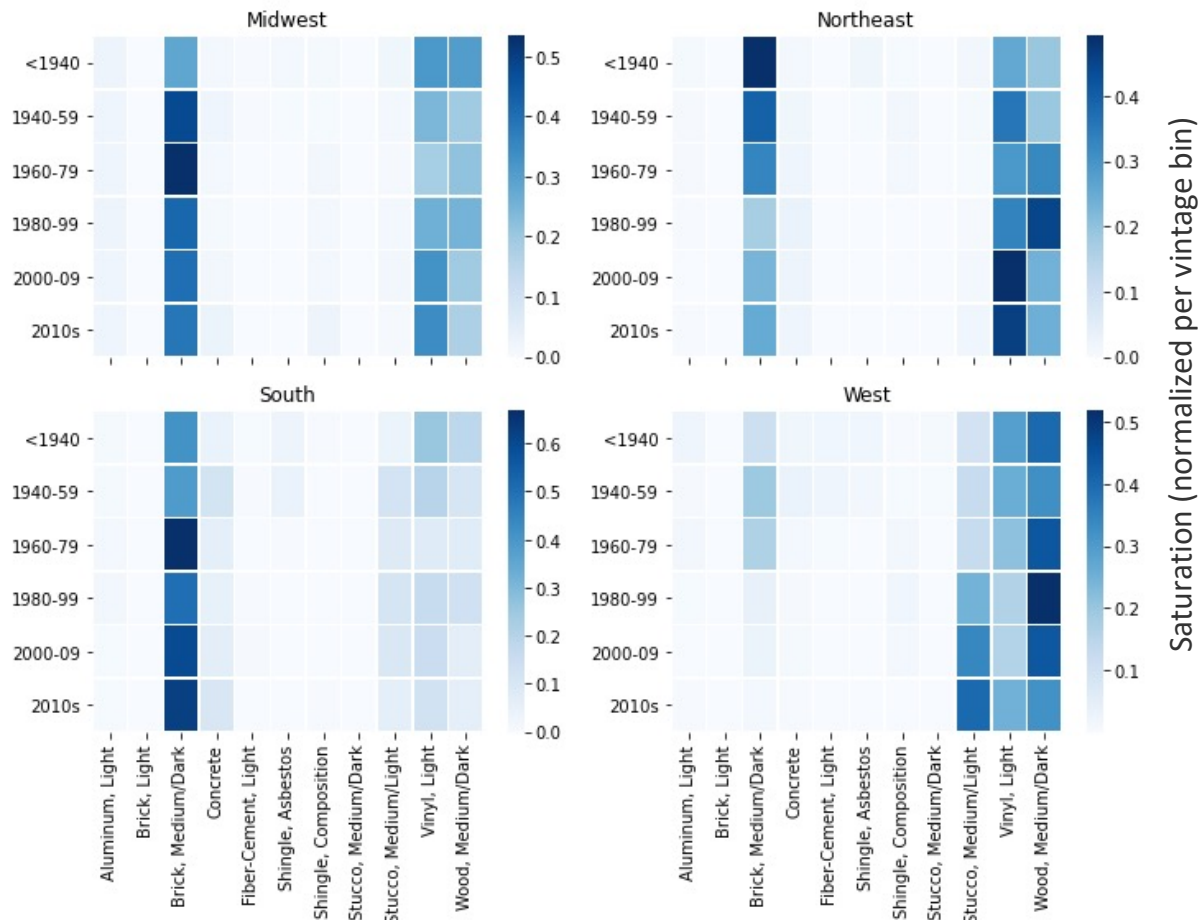
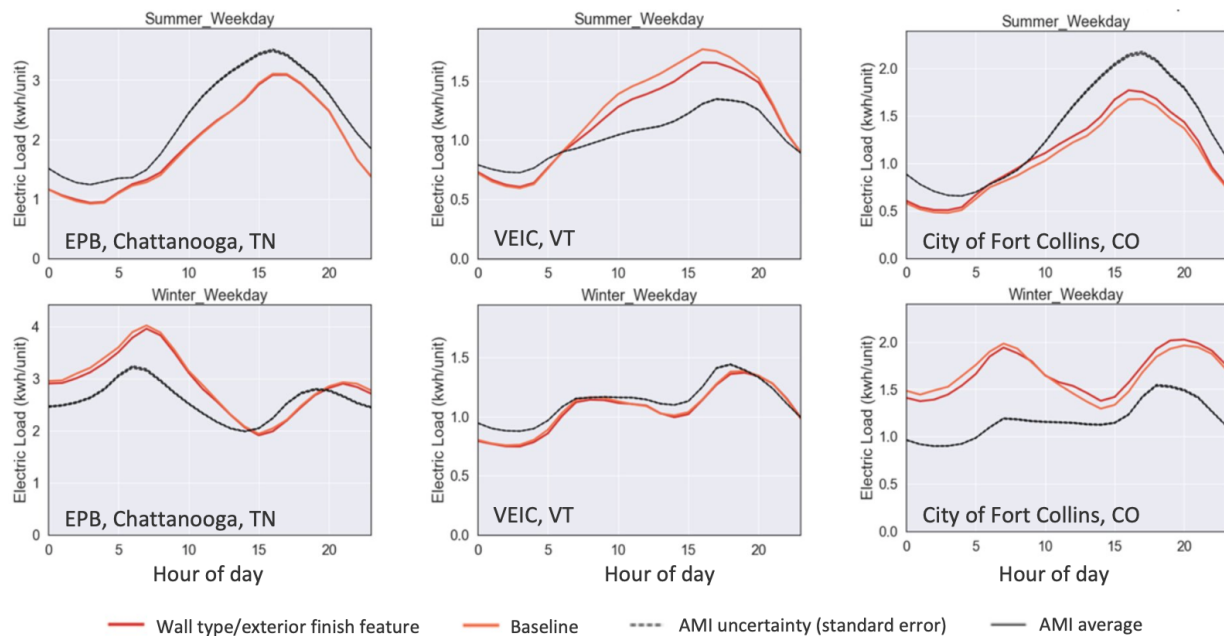


Figure 107. Wall exterior finish saturation by vintage bin and census region



**Figure 108. The impact of updated wall type and exterior finish distributions was relatively minor, shown here for comparisons between simulated and AMI average 24-hour electric load (per housing unit) for the top 10 summer and winter weekdays. Note: final comparison is shown in Section 4.**

### *Updating Window Types and Distributions*

At the start of EULP calibration, ResStock window type distributions were based on RECS 2009 data (EIA 2013), with three window type options: 1 pane, 2+ pane, and no windows.<sup>11</sup> RECS 2015 has more detailed response options for windows, including window frame material. The ResStock distributions were updated using these data.

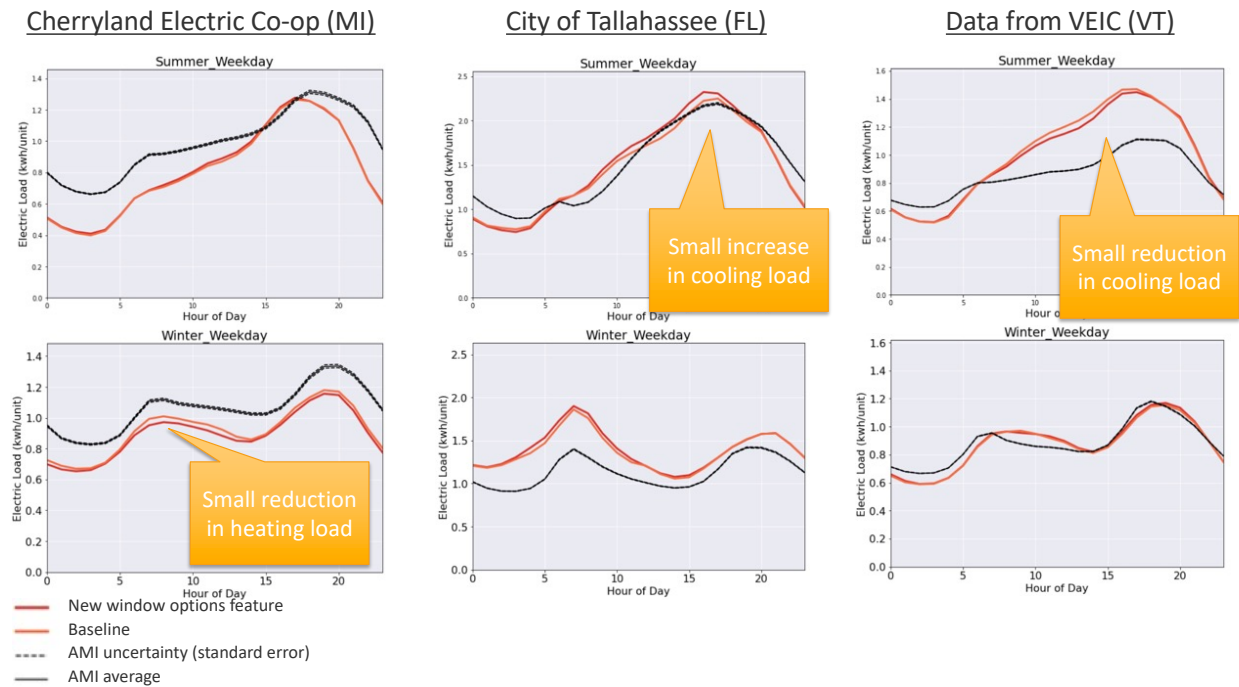
Due to high interest from stakeholders, the ResStock window options were further expanded to include options for external storm windows and low-E glass based on the best available data. For external storm windows, the stock distributions were drawn from Bickel, Phan-Gruber, and Christie (2013) and the physical and material properties were assumed to be the same as a window with one additional clear glass pane. For windows with low-E glass, the stock distributions were drawn from Ducker Worldwide studies of the U.S. Market for Windows, Doors, and Skylights (cite) and the physical properties were drawn from Culp, Widder, and Cort (2015).

This resulted in 10 different window type options in ResStock's input distributions, each with their own physical properties and material characteristics:

- Single pane, clear glass, metal frame
- Single pane, clear glass, metal frame, with exterior clear storm
- Single pane, clear glass, non-metal frame
- Single pane, clear glass, non-metal frame, with exterior clear storm
- Double pane, clear glass, metal frame, air filled
- Double pane, clear glass, metal frame, air filled, with exterior clear storm
- Double pane, clear glass, non-metal frame, air filled
- Double pane, clear glass, non-metal frame, air filled, with exterior clear storm
- Double pane, Low-E glass, non-metal frame, air filled, medium-gain
- Triple pane, Low-E glass, non-metal frame, air filled, low-gain.

The impact of these new window type options and their distributions is shown in Figure 109. The difference in total energy consumption is not large, and goes in different directions for different utilities.

<sup>11</sup> A very small percentage of RECS responses indicate that no windows are present.



**Figure 109.** This figure shows the impact of new window type options and window type distributions on average dwelling unit electric loads for summer and winter weekdays in the Cherryland, Tallahassee, and VEIC data areas. The differences are small, and do not go consistently in one direction. Note: final comparison is shown in Section 4.

### 3.2.9 On-Site Solar Photovoltaic Generation

For data on rooftop photovoltaic (PV) installations, we referenced the 2020 Tracking the Sun (TTS) report by LBNL and the 2020 solar market insight report from the Solar Energy Industries Association and Wood Mackenzie Power & Renewables (formerly Greentech Media Research) ((SEIA) 2020). For these data sources, we extracted probabilistic distributions of PV ownership, PV system size, and PV orientation. TTS contains detailed zip code and county-level PV installation project information, but the coverage is not consistent across counties and states. There is also a discrepancy in the number of installations and total installed capacity between TTS and the Wood Mackenzie report, which tracks the data by sector and state.

To amalgamate the data sources, we took the TTS county-level PV installations and added to them the difference in state-level residential installation count from the Wood Mackenzie report across all counties weighed by their population. In other words, we assumed the “missing” installations in TTS to be distributed proportionally to the county-level population. For the system size and array orientation, we used TTS exclusively and downselected the data to rooftop-only installations. To remove outliers, we capped the system sizes at 14 kW<sub>DC</sub>, which is the 95th percentile of all rooftop installations. Figures 110 and 111 show the rooftop PV saturation by county, system size distribution by state, and orientation distribution.

The impact of the new PV distributions is largest in California (see Figure 112), because the saturation of PV systems is the largest in California compared to other states. The decrease in the load is largest during the middle of the day. During the evening and morning hours the net site electricity (net site electricity = total site electricity - PV generation) is much closer to the total site electricity. During nighttime, there is no difference between the total site electricity and the net site electricity.

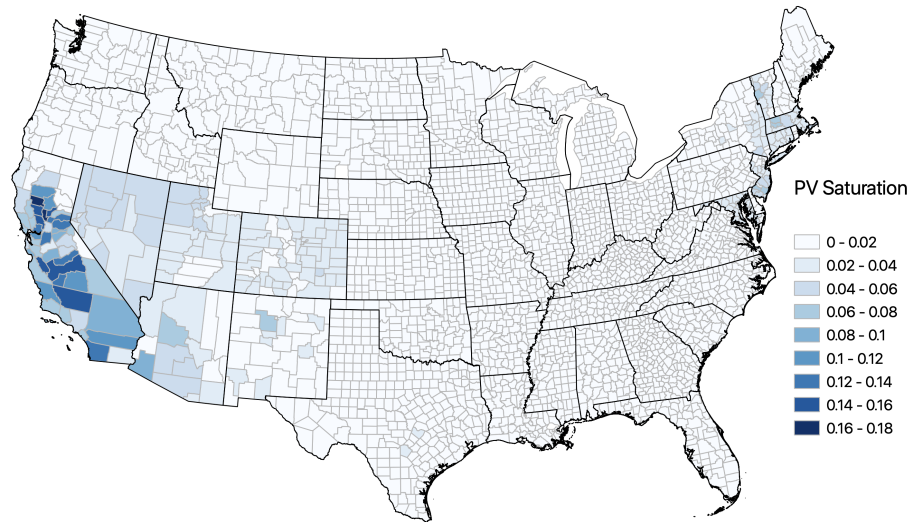


Figure 110. Fraction of dwelling units with rooftop PV, by county

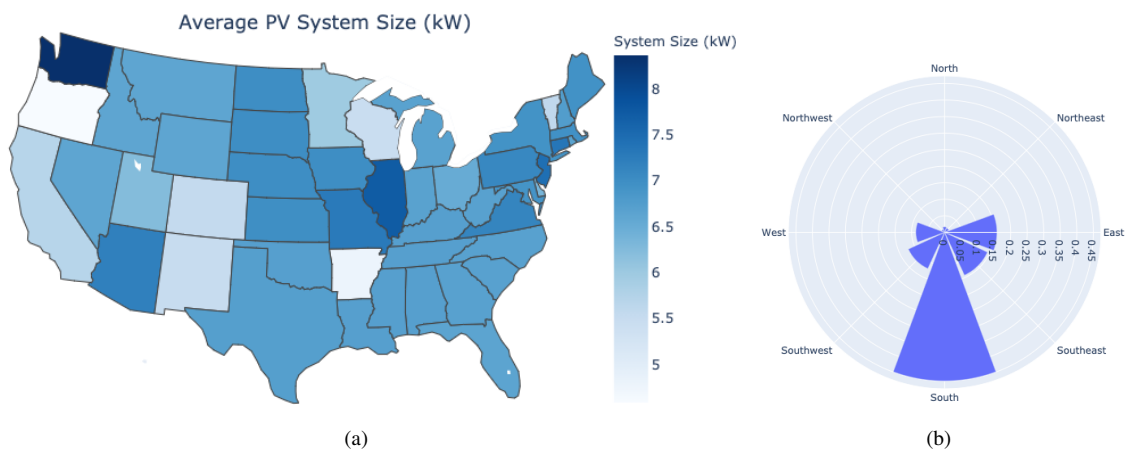
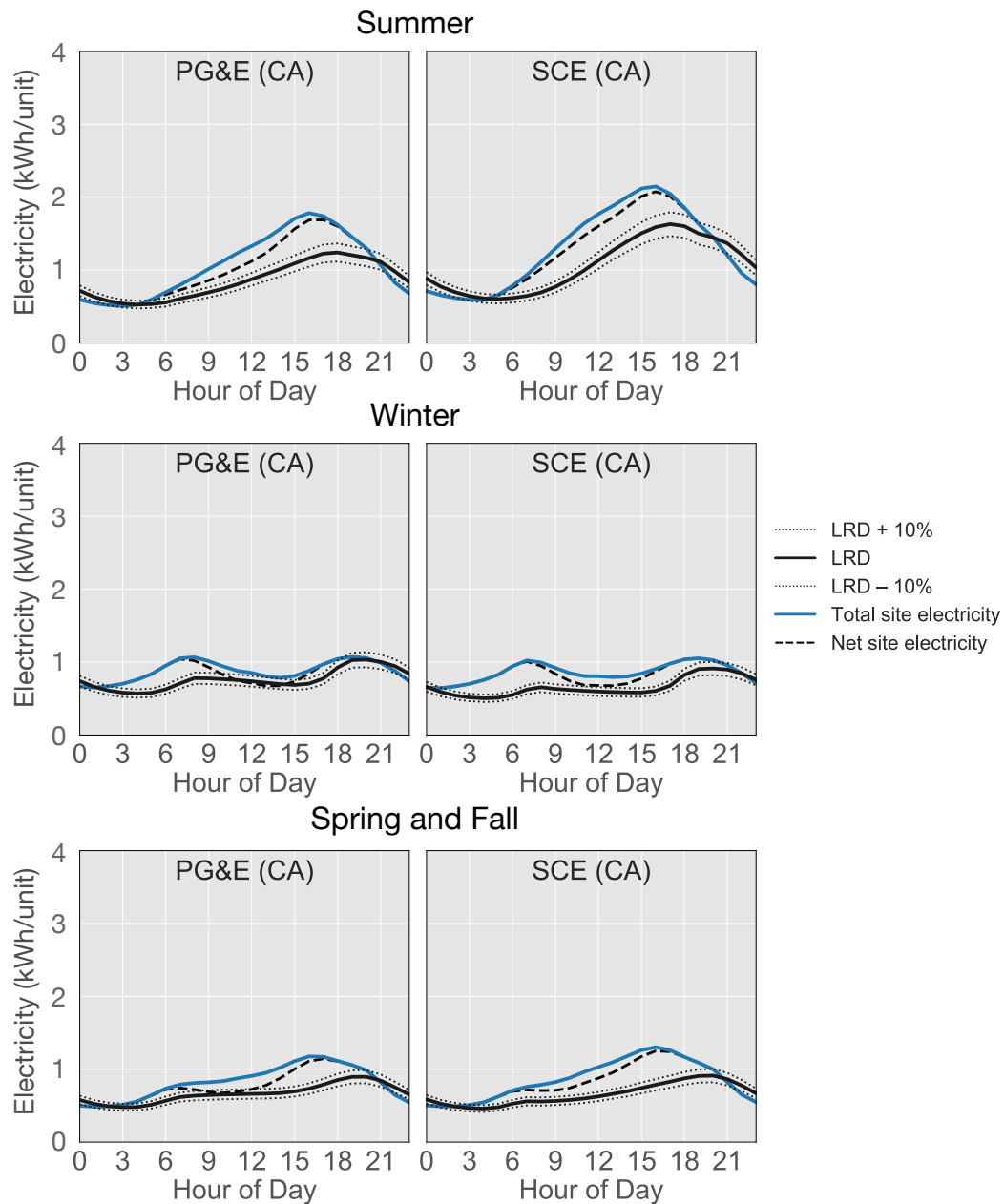


Figure 111. Rooftop PV system size by state (left) and system orientation (right)



**Figure 112.** This figure shows the impact of the new PV saturation, system size, and orientation distributions for the SCE and PG&E service territories. The net site electricity is reduced during the daytime hours due to PV-generated load. Note: final comparison is shown in Section 4.



### 3.2.10 Residential Output Correction Model

As described in Section 2.3.4, Our Calibration Philosophy, we prefer to make changes to model inputs—informed by data—to improve model fit. However, with this approach there will always be remaining model error with unknown cause. We suspect that some of the remaining error is because we do not have data on how heating and cooling setpoints or use of window air conditioners and space heaters might change seasonally or during more extreme hot or cold periods. Rather than trying to tune model inputs without supporting data, we determined, with input from our TAG, that it would be beneficial for users of the EULP dataset if we developed a method to judiciously apply an output correction model driven by consumption data available nationwide. Informed by our previous work on a correction model for a Los Angeles version of ResStock and correction approaches in the literature, we strove to adhere to the following principles in developing a new ResStock output correction model:

1. **Use data that are universally available** to avoid needing to extrapolate a correction model from one region to another. In our case, this was monthly electricity sales by state from EIA-861M.
2. **Only correct the heating and cooling end uses**, because remaining model error appears to be correlated to weather, and applying correction factors to other end uses is too much of an underdetermined problem.
3. **Do not attempt to correct load shapes at hourly resolutions**; only use monthly or daily multipliers that scale heating/cooling loads up or down. This is partially driven by data availability, but also because correcting at an hourly resolution is overly complicated and has less basis in the suspected behavioral causes of error described above.
4. **Make the correction optional and transparent**, so users can decide whether they wish to use it. Analysts are also welcome to apply their own correction models.

The result was a set of correction factors for heating and cooling end uses for each day of the year and state, which was published alongside the dataset<sup>12</sup>. The factors have not been applied to the profiles published in version (1.0) of the EULP dataset. Users wishing to use the factors should multiply all spaces heating and cooling end uses by their corresponding factor for each day of the year and state.

#### *Implementation*

Because we used EIA-861M data to guide our correction model, it is worth understanding how the EIA-861M data are collected. EIA 861M gives monthly energy consumption data for various utilities and states. However, different utilities report the monthly consumption data differently. According to private communication with EIA staff (EIA 2021a), utilities use two reporting methods:

1. **Calendar month reporting:**

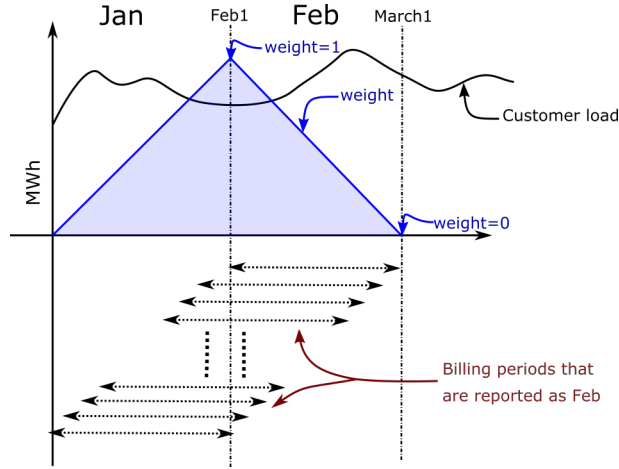
In this reporting method, the utility accurately sums up the energy consumed in a calendar month and submits this sum as the reported energy consumption for the month.

2. **Billing month reporting:**

In this reporting method, the utility sums up the monthly energy consumption of all customers whose billing period *ends* in a calendar month and reports that as the energy consumed in that month. For example, for the month of February, the utility would sum all energy consumption for customers whose billing period ends in February. This means a customer's billing period that starts on January 2 and ends on February 1 would be counted in the February monthly energy consumption. Assuming that the billing periods are uniformly distributed, the implication is that the days closest to February 1 will be included in almost all billing periods and days furthest away from February 1 will be included in the least number of billing periods. So, the reported February energy consumption will have the highest contribution from days closest to the February 1 and least contribution from days furthest away. This reported February energy consumption can be computed from the actual consumption by using a triangular weight function as shown in Figure 113. We multiply the consumption data in January and February by the triangular weights and sum them up to get the reported February consumption.

Because EIA-861M is composed of reported data that can follow either of these two formats (or a mixture of these), we can transform ResStock outputs into these reporting formats to achieve a more like-for-like comparison. We illustrate the process of doing so next.

<sup>12</sup>Available at <https://www.nrel.gov/buildings/end-use-load-profiles.html>



**Figure 113. The total load for February when using billing-month reporting is calculated by a triangular-weighted summation of January and February load**

For each month, we define two sets of weights, the rising triangle weight (*RTW*) and falling triangle weight (*FTW*):

$$RTW_{m,d} = \frac{d-1}{D_m-1} \text{ for } d = 1 : D_m - 1 \quad (3.6)$$

$$FTW_{m,d} = 1 - \frac{d-1}{D_m-1} \text{ for } d = 1 : D_m - 1 \quad (3.7)$$

where,

$m$  is the month (1 to 12)

$d$  is the day of the month, and

$D_m$  is the total number of days in the month.

We represent the *RTW* vector by  $\triangleleft$  and *FTW* vector by  $\triangleright$ , with the assumption that they are for the same month to which they are applied.

Let  $A_{m,d}$  represent the actual daily loads for month  $m$  and day  $d$ , and  $R_m$  represent the reported monthly load for the month  $m$ . Then, if the utility is using calendar month reporting:

$$R_m = \sum_{d=1}^{D_m} A_{m,d} = A_m = \vec{1} \cdot \vec{A}_m \quad (3.8)$$

where,

$\vec{1} = [1, 1, \dots]$  of dimension  $\dim(D_m)$

$\vec{A}_m = [A_{m,1}, A_{m,2}, \dots, A_{m,D_m}]$

If the utility is using billing month reporting:

$$R_m = \sum_{d=1}^{D_{m-1}} A_{m-1,d} RTW_{m-1,d} + \sum_{d=1}^{D_m} A_{m-1,d} FTW_{m,d} \quad (3.9)$$

or

$$R_m = \triangleleft \cdot \vec{A}_{m-1} + \triangleright \cdot \vec{A}_m \quad (3.10)$$

EIA-861M provides the reported monthly data for the whole state, which is composed of multiple utilities. Some of those utilities could be using calendar month reporting and some could be using billing month (in fact, even different customers within the same utility could have this discrepancy). So, let's assume that  $\alpha$  is the fraction of the load that is reported using calendar months, and  $(1 - \alpha)$  is the fraction of load reported using billing month. Then, we can find the mixed load reporting for a month,  $M_m$ , for a state using,

$$M_m = \alpha \cdot \text{calendar\_month\_aggregation} + (1 - \alpha) \cdot \text{billing\_month\_aggregation} \quad (3.11)$$

or

$$M_m = \alpha A_m + (1 - \alpha)(\Delta \cdot \overrightarrow{A_{m-1}} + \Delta \cdot \overrightarrow{A_m}) \quad (3.12)$$

or

$$M_m = (1 - \alpha)\Delta \cdot \overrightarrow{A_{m-1}} + (\alpha \vec{1} + (1 - \alpha)\Delta) \cdot \overrightarrow{A_m} \quad (3.13)$$

If we assume that the energy consumption for the month of December last year is the same as this year, that will let us assume that if  $m = 1$ , then  $m - 1 = 12$  (wrap around to December). Then the mixed load reporting for each month of the year can be written with the following equations:

$$M_1 = (1 - \alpha)\Delta \cdot \overrightarrow{A_{12}} + (\alpha \vec{1} + (1 - \alpha)\Delta) \cdot \overrightarrow{A_1} \quad (3.14)$$

$$M_1 = (1 - \alpha)\Delta \cdot \overrightarrow{A_1} + (\alpha \vec{1} + (1 - \alpha)\Delta) \cdot \overrightarrow{A_2} \quad (3.15)$$

$$M_1 = (1 - \alpha)\Delta \cdot \overrightarrow{A_2} + (\alpha \vec{1} + (1 - \alpha)\Delta) \cdot \overrightarrow{A_3} \quad (3.16)$$

$$M_1 = (1 - \alpha)\Delta \cdot \overrightarrow{A_3} + (\alpha \vec{1} + (1 - \alpha)\Delta) \cdot \overrightarrow{A_4} \quad (3.17)$$

$$M_1 = (1 - \alpha)\Delta \cdot \overrightarrow{A_4} + (\alpha \vec{1} + (1 - \alpha)\Delta) \cdot \overrightarrow{A_5} \quad (3.18)$$

$$M_1 = (1 - \alpha)\Delta \cdot \overrightarrow{A_5} + (\alpha \vec{1} + (1 - \alpha)\Delta) \cdot \overrightarrow{A_6} \quad (3.19)$$

$$M_1 = (1 - \alpha)\Delta \cdot \overrightarrow{A_6} + (\alpha \vec{1} + (1 - \alpha)\Delta) \cdot \overrightarrow{A_7} \quad (3.20)$$

$$M_1 = (1 - \alpha)\Delta \cdot \overrightarrow{A_7} + (\alpha \vec{1} + (1 - \alpha)\Delta) \cdot \overrightarrow{A_8} \quad (3.21)$$

$$M_1 = (1 - \alpha)\Delta \cdot \overrightarrow{A_8} + (\alpha \vec{1} + (1 - \alpha)\Delta) \cdot \overrightarrow{A_9} \quad (3.22)$$

$$M_1 = (1 - \alpha)\Delta \cdot \overrightarrow{A_9} + (\alpha \vec{1} + (1 - \alpha)\Delta) \cdot \overrightarrow{A_{10}} \quad (3.23)$$

$$M_1 = (1 - \alpha)\Delta \cdot \overrightarrow{A_{10}} + (\alpha \vec{1} + (1 - \alpha)\Delta) \cdot \overrightarrow{A_{11}} \quad (3.24)$$

$$M_1 = (1 - \alpha)\Delta \cdot \overrightarrow{A_{11}} + (\alpha \vec{1} + (1 - \alpha)\Delta) \cdot \overrightarrow{A_{12}} \quad (3.25)$$

$\alpha$  can be solved for each state as part of a multi-dimensional optimization that fits a degree day regression model to the state's average temperature and electricity consumption. EIA performed this optimization and gave us the alpha values for the states (see discussion on the EIA developed degree day model on page 125).

Keeping the above point in mind, we explored the following different output correction models that scale the output kWh such that the monthly state total kWh matches with the EIA-861M data.<sup>13</sup>

Type 1: We uniformly scale all the end uses such that the total statewide load matches with the statewide load from EIA-861M.

Type 2: We scale only the HVAC load such that the total statewide load matches with the statewide load from EIA-861M. When HVAC loads are scaled, we can choose to scale only the heating ( $\theta_0$ ), only the cooling load ( $\theta_1$ ), or both heating and cooling loads ( $\theta_{0\_5}$ ).

<sup>13</sup>Types 3 and 4 were similar to Type 5 and Type 6, respectively, but used national average CDD and HDD instead of state average CDD and HDD. They did not perform well compared to Type 5 and 6, and were omitted from the comparisons.

Type 5: Similar to Type 2, we only scale HVAC load such that the total load matches with EIA-861M. However, we scale some days more than others. In Type 5, we compare the cooling degree days (CDD) and heating degree days (HDD) in each county for each day to the state average CDD and HDD for the whole year, and scale extreme days more than milder days for both heating and cooling.

Type 6: Like Type 5 but scale milder days more than extreme days.

Type 7: Like Type 5 but scale extreme days more for heating and milder days more for cooling.

Type 8: Like Type 5 but scale milder days more for heating and extreme days more for cooling (inverse of Type 7).

In order to use the above correction models, we would need the EIA-861M data for the year for which the simulation is run. In order to make the correction model more general, we also explored models that rely on a more general source.

EIA developed a degree day model that provides a relationship between the state average daily temperature and the state average kWh per customer per day. The model has six parameters:

**Alpha:** The proportion of loads reported in billing month format

**CDD Tbase:** Cooling degree day base temperature - the average temperature associated with the start of cooling

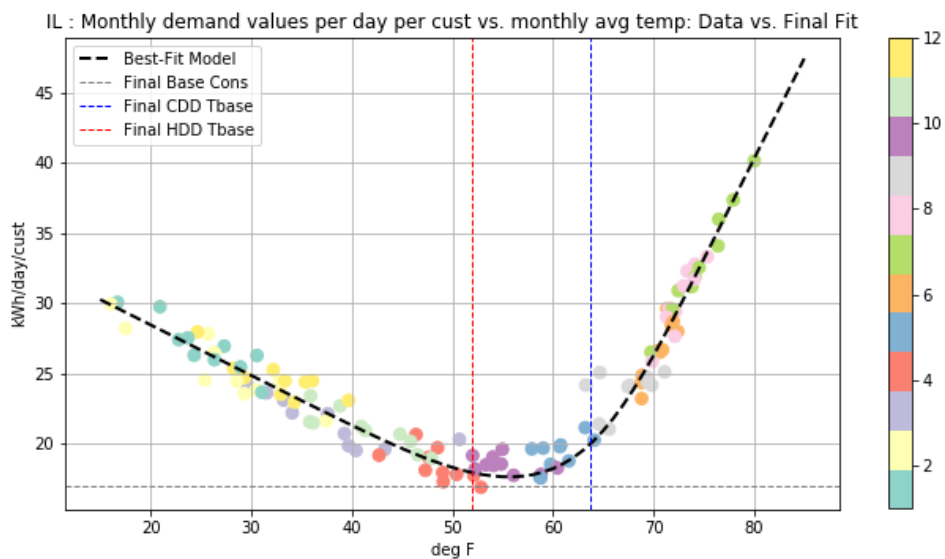
**HDD Tbase:** Heating degree day base temperature - the average temperature associated with the start of heating

**Base consumption:** The baseload (kWh/cust/day)

**Cooling slope:** The slope of the cooling region (kWh/cust/day)

**Heating slope:** The slope of the heating region

The parameters are minimized by optimizing the fit over 10 years of EIA-861M data. An illustration of the fit is shown in Figure 114<sup>14</sup>.



**Figure 114. Example of the EIA degree day model fit for the state of Illinois. Source: Derived from EIA Form 861M and Climate Prediction Center Population-Weighted Daily Degree Days.**

<sup>14</sup>Models developed in collaboration with Greg Lawson, U.S. Energy Information Administration. Modeling approach is still evolving. Model parameters and results are not final.

The following four correction models scale the load to get kWh/customer/day to the same level estimated by the degree day model discussed above.

Type 9: Scale the daily state-level HVAC load so that the total load per customer per day matches the value estimated by the degree day model obtained by using population-weighted state average temperature.

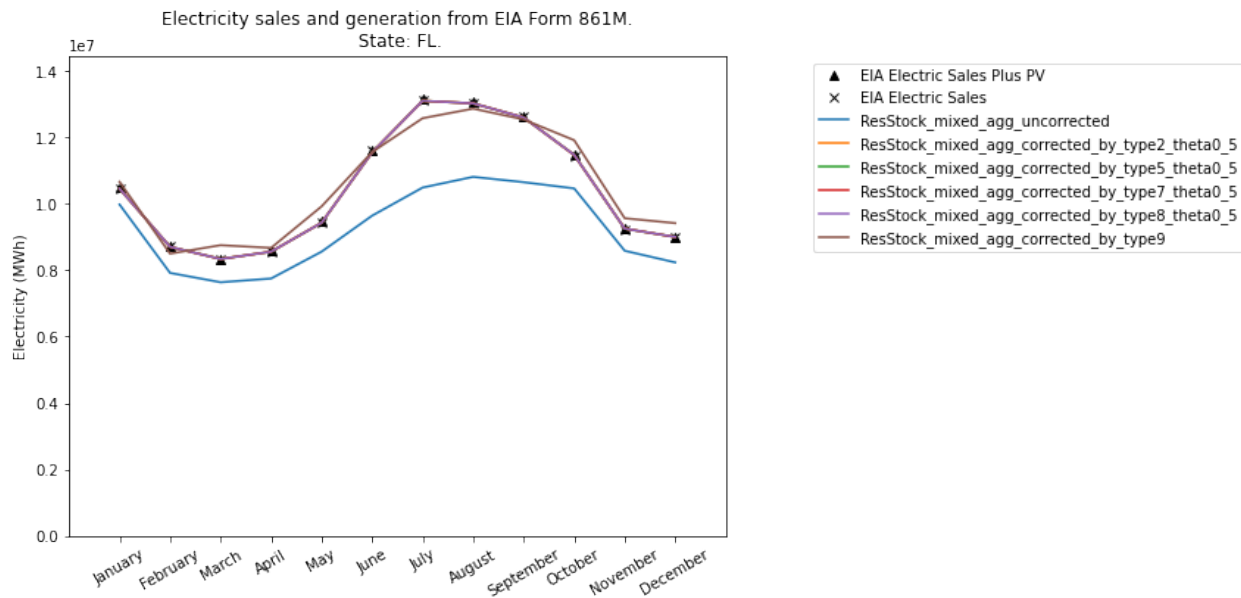
Type 10: Like 9, but don't scale the baseload; only make the heating and cooling slope match the degree day model.

Type 11: Scale the daily county-level HVAC load so that the total load per customer per day matches the value estimated by the degree day model (for the state) obtained by using county average temperature.

Type 12: Like 11, but don't scale the baseload; only make the heating and cooling slope match the degree day model.

## Performance

### Performance in Regard to State Average



**Figure 115. Correction model performance example for Florida. In this example, Types 2, 5, 7, and 8 result in almost identical monthly values.**

Figure 115 shows an example of how the state average load is shifted after the application of various correction models. We can see that most correction models adjust the state average load such that it becomes very close to the EIA-861M level. It should be noted that monthly aggregation of ResStock load is done using the mixed aggregation technique discussed above using equation 3.14 through equation 3.25 and using the alpha value for the state so that it can be compared with the EIA-861M data.

Figure 116 shows the coefficient of variation of the root mean square error (CVRMSE) between the corrected version of ResStock's result for 2018 and the state's 2018 EIA-861M data. Hawaii and Alaska are not currently modeled in ResStock. EIA-861M does not have data for Washington, D.C., so the correction model cannot be applied and thus is not included in the table.

For correction models on the left section (which are calibrated using 2018 EIA-861M), models that can scale both heating and cooling loads (i.e., models with  $\theta_{0.5}$ ) have very low CVRMSE for most states. Since there is a restriction on how much the HVAC load can be scaled for correction ( $0.5 \leq \text{factor} \leq 2.0$ ), for states where ResStock was especially off, there is still remaining discrepancy, and this results in substantial CVRMSE. Models that can only scale either the cooling load or heating load ( $\theta_0$  or  $\theta_1$ ) have poorer performance because they can't correct for all of the offset from EIA-861M.

Correction models on the right section (which are calibrated using the multi-year EIA-861M degree day model) have substantially lower CVRMSE than the uncorrected version of ResStock, but some discrepancy with 2018 EIA-861M still exists. The remaining discrepancy is because the degree day model was based on the last 10 years of EIA-861M data, and the actual load in 2018 varied from the degree day model fit.

### ***Performance in Regard to Various Utility Regions***

Figure 117 shows an example of monthly average kWh fit improved for the VEIC utility region (a utility in Vermont) by a correction model (Type 9); this is compared side by side to the improvement in the state-level fit for Vermont. Because ResStock overestimates the energy consumption at both state and the utility region level, when we make the statewide adjustment, the fit for the utility region also improves.

Figure 118 shows an example of monthly average kWh fit deteriorated for a Tallahassee utility (a utility in Florida) region by a correction model (Type 9); this is compared side by side to the improvement in the state-level fit. Because ResStock underestimates the energy consumption at the state level, but is mostly already correct at the utility region level, when we make the statewide adjustment as per the correction model, while the state wide fit improves, the fit for the utility region deteriorates.

A summary of performance of different correction models for different utility regions is shown in Figure 119. The CVRMSE in the table is obtained by averaging the CVRMSE based on an hourly, monthly, and top 100 hours comparison between the AMI/LRD data for the utility and corrected ResStock result aggregated over the same region. We can see that all correction models achieve similar CVRMSE when averaged across the regions, and none of the correction models improves AMI/LRD fit consistently across all regions. Typically, for each AMI/LRD data region, there is a specific correction model that achieves the best fit, and that model is different for different regions.

Considering all these performance data, we pick the Type 9 correction model as the final correction model because of two reasons. First, it belongs to the class of correction model that is trained in multi-year EIA-861M degree day model instead of specific year's EIA-861M data. This means, this model can be applied to any year's ResStock run, including ResStock run using TMY weather data. This makes the model more generalizable. The second reason is that it performs the best in its class when considering both the utility region fit (Figure 119) and state total fit (Figure 116). Additional plots comparing modeled load to AMI data (example weeks and daily sums) with and without the correction model can be found in Appendix C.



End-Use Load Profiles for the U.S. Building Stock:  
Methodology and Results of Model Calibration, Validation, and Uncertainty Quantification

	Average monthly CVMSE with EIA861M for 2018																	
State	Calibrated using corresponding 2018 EIA861M												Calibrated using mulit-year EIA861M degree day model					
Correction Model Type	type1	type2_theta1	type2_theta0	type2_theta0_5	type5_theta0_5	type6_theta0_5	type5_theta0	type6_theta0	type5_theta1	type6_theta1	type7_theta0_5	type8_theta0_5	type9	type10	type11	type12	Uncorrected	
AVG	0.01	0.29	0.37	0.05	0.08	0.06	0.38	0.37	0.42	0.29	0.07	0.08	0.24	0.35	0.28	0.38	0.48	
AL	0.01	0.04	0.42	0.01	0.01	0.01	0.42	0.42	0.05	0.04	0.01	0.01	0.16	0.15	0.20	0.17	0.42	
AR	0.02	0.21	0.59	0.02	0.02	0.02	0.29	0.54	0.21	0.21	0.02	0.02	0.16	0.22	0.16	0.21	0.29	
AZ	0.00	0.05	0.48	0.01	0.05	0.01	0.48	0.48	0.48	0.05	0.01	0.05	0.59	0.51	0.63	0.46	0.48	
CA	0.02	0.37	0.45	0.28	0.30	0.29	0.45	0.45	0.49	0.38	0.30	0.30	0.38	0.65	0.42	0.63	0.49	
CO	0.00	0.54	0.50	0.05	0.11	0.09	0.50	0.50	0.75	0.54	0.09	0.12	0.14	0.30	0.14	0.21	0.75	
CT	0.01	0.30	0.21	0.06	0.08	0.07	0.38	0.20	0.38	0.30	0.07	0.07	0.25	0.41	0.26	0.43	0.38	
DE	0.01	0.13	0.29	0.01	0.01	0.01	0.32	0.29	0.32	0.13	0.01	0.01	0.15	0.22	0.15	0.19	0.32	
FL	0.01	0.01	0.46	0.01	0.01	0.01	0.48	0.46	0.01	0.01	0.01	0.01	0.11	0.12	0.19	0.15	0.48	
GA	0.01	0.04	0.44	0.01	0.01	0.01	0.44	0.44	0.06	0.04	0.01	0.01	0.16	0.16	0.16	0.21	0.44	
IA	0.00	0.39	0.26	0.01	0.02	0.02	0.26	0.26	0.48	0.39	0.02	0.02	0.24	0.27	0.25	0.27	0.48	
ID	0.00	0.19	0.47	0.05	0.15	0.06	0.49	0.47	0.53	0.19	0.10	0.17	0.18	0.32	0.50	0.53	0.53	
IL	0.00	0.27	0.22	0.01	0.01	0.01	0.23	0.22	0.36	0.27	0.01	0.01	0.22	0.35	0.27	0.71	0.36	
IN	0.00	0.10	0.42	0.01	0.01	0.01	0.42	0.42	0.44	0.10	0.01	0.01	0.22	0.12	0.24	0.27	0.44	
KS	0.00	0.43	0.09	0.04	0.07	0.06	0.11	0.10	0.45	0.43	0.07	0.06	0.18	0.45	0.19	0.45	0.45	
KY	0.01	0.37	0.37	0.01	0.01	0.01	0.53	0.37	0.53	0.37	0.01	0.01	0.19	0.29	0.20	0.33	0.53	
LA	0.02	0.41	0.59	0.03	0.06	0.05	0.65	0.59	0.40	0.41	0.05	0.07	0.14	0.18	0.14	0.18	0.65	
MA	0.01	0.74	0.32	0.17	0.24	0.20	0.33	0.34	0.81	0.74	0.19	0.25	0.21	0.53	0.22	0.42	0.81	
MD	0.00	0.32	0.12	0.01	0.01	0.01	0.13	0.12	0.36	0.32	0.01	0.01	0.28	0.37	0.29	0.46	0.36	
ME	0.00	0.32	0.33	0.10	0.15	0.11	0.34	0.33	0.46	0.32	0.12	0.14	0.22	0.32	0.23	0.28	0.46	
MI	0.00	0.29	0.12	0.01	0.02	0.02	0.13	0.13	0.33	0.29	0.02	0.02	0.31	0.45	0.33	0.32	0.33	
MN	0.00	0.53	0.33	0.01	0.04	0.04	0.33	0.33	0.63	0.53	0.02	0.05	0.23	0.14	0.25	0.19	0.63	
MO	0.00	0.18	0.23	0.01	0.01	0.01	0.30	0.23	0.18	0.18	0.01	0.01	0.26	0.14	0.26	0.25	0.30	
MS	0.03	0.23	0.48	0.02	0.02	0.12	0.40	0.51	0.40	0.23	0.04	0.03	0.14	0.16	0.17	0.18	0.40	
MT	0.01	0.26	0.28	0.04	0.09	0.05	0.30	0.28	0.39	0.27	0.09	0.10	0.18	0.27	0.18	0.31	0.39	
NC	0.01	0.27	0.35	0.01	0.02	0.01	0.36	0.35	0.46	0.27	0.02	0.02	0.16	0.35	0.18	0.32	0.46	
ND	0.00	0.47	0.45	0.14	0.25	0.15	0.46	0.45	0.66	0.48	0.19	0.24	0.22	0.48	0.47	0.60	0.66	
NE	0.01	0.65	0.27	0.01	0.10	0.01	0.30	0.27	0.65	0.65	0.11	0.01	0.16	0.19	0.42	0.62	0.70	
NH	0.01	0.37	0.34	0.16	0.26	0.17	0.35	0.35	0.46	0.37	0.16	0.23	0.23	0.47	0.24	0.33	0.46	
NJ	0.00	0.29	0.15	0.01	0.06	0.02	0.15	0.15	0.35	0.29	0.02	0.06	0.30	0.59	0.31	0.61	0.35	
NM	0.01	0.47	0.41	0.16	0.29	0.18	0.41	0.41	0.62	0.47	0.18	0.29	0.33	0.81	0.39	0.82	0.62	
NV	0.00	0.07	0.56	0.01	0.02	0.01	0.56	0.56	0.09	0.07	0.01	0.03	0.57	0.85	0.65	0.90	0.56	
NY	0.01	0.69	0.52	0.21	0.33	0.23	0.76	0.52	0.86	0.69	0.25	0.28	0.70	0.87	0.75	1.12	0.86	
OH	0.00	0.14	0.33	0.01	0.01	0.01	0.37	0.33	0.37	0.14	0.01	0.01	0.27	0.20	0.30	0.41	0.37	
OK	0.00	0.14	0.25	0.01	0.01	0.01	0.29	0.25	0.29	0.14	0.01	0.01	0.21	0.21	0.21	0.27	0.29	
OR	0.00	0.14	0.32	0.02	0.11	0.05	0.32	0.32	0.39	0.15	0.07	0.11	0.24	0.22	0.23	0.19	0.39	
PA	0.01	0.21	0.08	0.01	0.01	0.01	0.08	0.08	0.25	0.21	0.01	0.01	0.30	0.31	0.32	0.31	0.25	
RI	0.00	0.75	0.36	0.26	0.30	0.30	0.81	0.38	0.81	0.75	0.28	0.32	0.25	0.62	0.26	0.75	0.81	
SC	0.01	0.15	0.44	0.01	0.02	0.01	0.47	0.44	0.16	0.15	0.02	0.01	0.18	0.28	0.18	0.29	0.47	
SD	0.01	0.27	0.51	0.08	0.17	0.08	0.51	0.51	0.59	0.27	0.10	0.19	0.20	0.56	0.20	0.50	0.59	
TN	0.02	0.30	0.94	0.02	0.02	0.04	0.52	0.91	0.52	0.31	0.02	0.02	0.24	0.22	0.24	0.22	0.52	
TX	0.01	0.11	0.41	0.02	0.02	0.02	0.42	0.41	0.12	0.11	0.02	0.02	0.13	0.16	0.41	0.41	0.42	
UT	0.00	0.16	0.67	0.06	0.10	0.07	0.67	0.67	0.70	0.16	0.07	0.10	0.22	0.24	0.22	0.21	0.70	
VA	0.00	0.26	0.34	0.01	0.01	0.01	0.34	0.34	0.27	0.26	0.01	0.01	0.28	0.56	0.30	0.49	0.44	
VT	0.00	0.24	0.44	0.10	0.15	0.11	0.44	0.44	0.50	0.24	0.13	0.14	0.23	0.51	0.24	0.38	0.50	
WA	0.01	0.26	0.28	0.11	0.19	0.10	0.29	0.28	0.39	0.26	0.14	0.20	0.25	0.31	0.25	0.24	0.39	
WI	0.00	0.30	0.32	0.01	0.02	0.01	0.32	0.32	0.45	0.30	0.02	0.02	0.11	0.12	0.13	0.13	0.45	
WV	0.00	0.19	0.19	0.01	0.01	0.01	0.20	0.19	0.29	0.19	0.01	0.01	0.24	0.27	0.26	0.28	0.29	
WY	0.01	0.19	0.21	0.01	0.03	0.01	0.22	0.21	0.31	0.19	0.01	0.03	0.24	0.25	0.25	0.25	0.31	

Figure 116. Correction model performance matrix by state

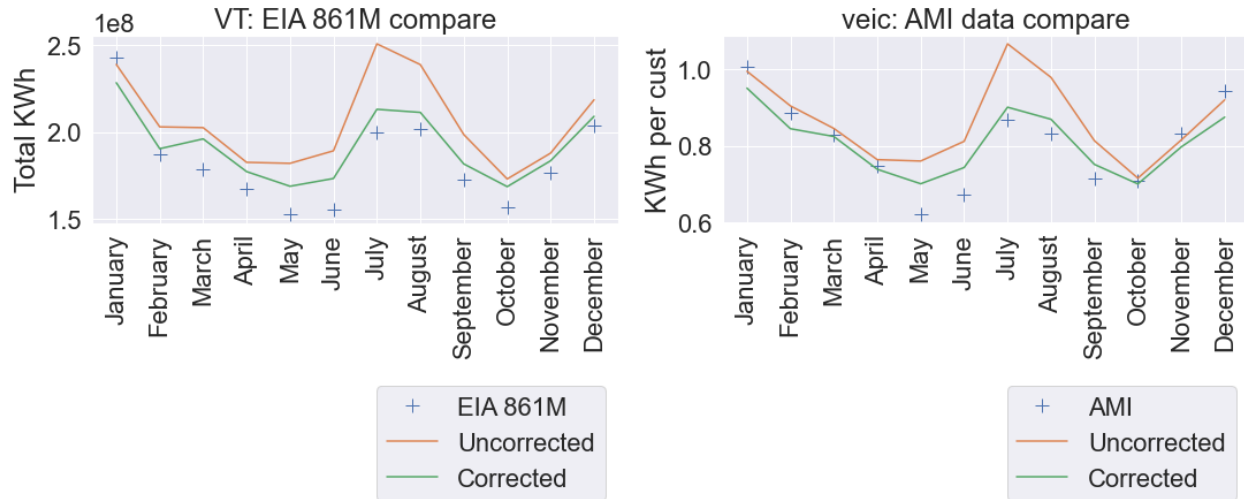


Figure 117. Correction model performance example for VEIC

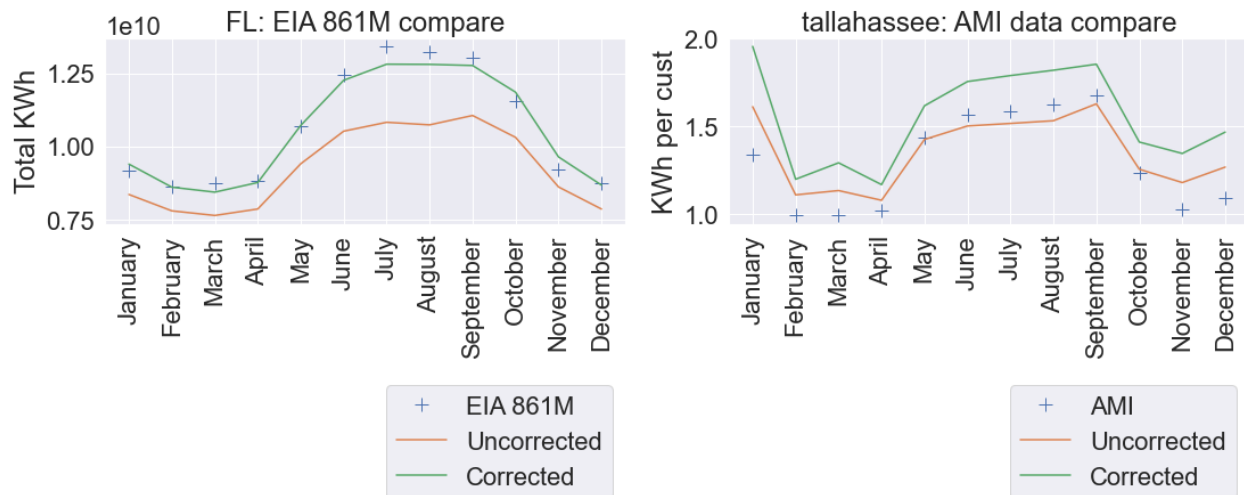


Figure 118. Correction model performance example for Tallahassee

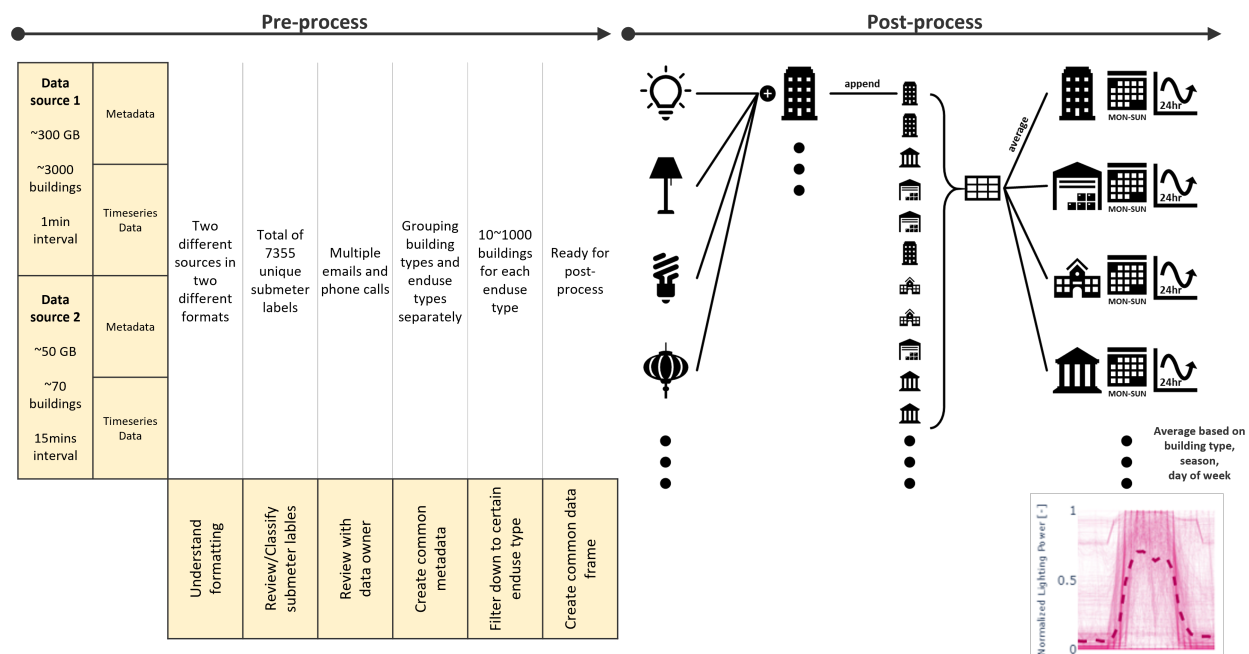
Average of CVRMSE																		
Utilities		Calibrated to corresponding year's EIA 861M												Calibrated using multi-year EIA861M degree day model				
		type1	type2_theta0	type2_theta0_5	type2_theta1	type5_theta0	type5_theta0_5	type5_theta1	type6_theta0	type6_theta0_5	type6_theta1	type7_theta0_5	type8_theta0_5	type9	type10	type11	type12	Without correction
AMI cherryland	Correction Model Type	0.24	0.25	0.24	0.19	0.25	0.24	0.20	0.25	0.24	0.19	0.24	0.24	0.19	0.18	0.21	0.23	0.20
AMI epb		0.17	0.54	0.17	0.11	0.12	0.20	0.12	0.54	0.20	0.13	0.15	0.25	0.17	0.15	0.17	0.14	0.12
AMI fort_collins		0.11	0.15	0.13	0.25	0.15	0.14	0.23	0.12	0.14	0.24	0.12	0.16	0.16	0.18	0.14	0.10	0.23
AMI horry		0.11	0.16	0.11	0.12	0.17	0.11	0.12	0.16	0.12	0.12	0.12	0.12	0.15	0.11	0.15	0.13	0.17
AMI seattle		0.28	0.27	0.25	0.26	0.42	0.22	0.24	0.25	0.24	0.27	0.30	0.19	0.20	0.19	0.40	0.39	0.24
AMI tallahassee		0.25	0.42	0.30	0.22	0.14	0.40	0.22	0.33	0.26	0.22	0.32	0.28	0.38	0.33	0.39	0.30	0.14
AMI veic		0.09	0.19	0.07	0.06	0.19	0.08	0.19	0.19	0.08	0.06	0.08	0.08	0.07	0.13	0.08	0.11	0.19
LRD AEP (OH)		0.07	0.09	0.09	0.11	0.11	0.12	0.11	0.08	0.08	0.10	0.07	0.12	0.15	0.12	0.13	0.14	0.11
LRD Ameren (MO)		0.28	0.25	0.28	0.36	0.35	0.28	0.36	0.27	0.30	0.36	0.27	0.30	0.28	0.26	0.29	0.33	0.35
LRD Appalachian (VA)		0.10	0.11	0.09	0.12	0.11	0.09	0.12	0.12	0.10	0.12	0.09	0.10	0.11	0.09	0.19	0.14	0.15
LRD BGE (MD)		0.19	0.16	0.18	0.29	0.16	0.18	0.29	0.20	0.21	0.29	0.17	0.21	0.23	0.25	0.26	0.32	0.29
LRD Cleveland (OH)		0.25	0.15	0.28	0.29	0.17	0.33	0.17	0.16	0.24	0.25	0.24	0.33	0.35	0.31	0.49	0.51	0.17
LRD ComEd (IL)		0.17	0.21	0.16	0.14	0.21	0.15	0.18	0.21	0.16	0.14	0.17	0.15	0.13	0.14	0.12	0.17	0.18
LRD ERCOT		0.13	0.15	0.14	0.13	0.10	0.16	0.14	0.14	0.13	0.11	0.14	0.14	0.10	0.14	0.10	0.10	0.10
LRD MetEd (PA)		0.14	0.13	0.13	0.16	0.12	0.13	0.16	0.14	0.14	0.16	0.12	0.14	0.08	0.09	0.13	0.13	0.16
LRD OhioEd (OH)		0.15	0.12	0.17	0.18	0.12	0.22	0.12	0.12	0.15	0.15	0.15	0.21	0.23	0.19	0.28	0.30	0.12
LRD PECO (PA)		0.08	0.08	0.08	0.10	0.08	0.10	0.09	0.08	0.08	0.09	0.08	0.10	0.11	0.10	0.10	0.11	0.09
LRD Penelec (PA)		0.14	0.13	0.13	0.16	0.12	0.14	0.16	0.14	0.14	0.16	0.13	0.15	0.13	0.13	0.16	0.16	0.16
LRD PG&E (CA)		0.25	0.16	0.23	0.22	0.16	0.22	0.16	0.16	0.23	0.23	0.23	0.22	0.12	0.23	0.11	0.19	0.16
LRD PP (PA)		0.30	0.31	0.29	0.27	0.31	0.29	0.28	0.30	0.29	0.28	0.30	0.28	0.25	0.25	0.17	0.17	0.28
LRD SCE (CA)		0.19	0.17	0.19	0.20	0.17	0.21	0.17	0.17	0.19	0.19	0.19	0.20	0.15	0.21	0.17	0.20	0.17
LRD ToledoEd (OH)		0.22	0.16	0.25	0.27	0.18	0.31	0.18	0.16	0.21	0.23	0.21	0.31	0.33	0.29	0.38	0.41	0.18
LRD WPP (PA)		0.25	0.25	0.24	0.21	0.26	0.24	0.22	0.25	0.24	0.22	0.25	0.23	0.22	0.21	0.17	0.17	0.22
Grand Total		0.18	0.20	0.18	0.19	0.18	0.20	0.18	0.20	0.18	0.19	0.18	0.20	0.19	0.19	0.21	0.22	0.18

Figure 119. Correction model performance matrix for various utility regions

### 3.3 Commercial Model Updates

#### 3.3.1 Interior Lighting Schedules

One of the end uses that contributes significantly to energy consumption in commercial buildings is the interior lighting system (as discussed in *ComStock Sensitivity Results*). To improve the representation of subhourly time-dependent behavior of the interior lighting in commercial buildings, end-use data were used to derive average daily profiles (hereafter referred to as schedules) of interior lighting usage for different building types. Figure 120 shows the process of how two different end-use datasets were processed to extract insights. Because these datasets started in different formats, several steps were necessary to standardize the format: (1) reviewing the end-use labels with the data provider, (2) reviewing and grouping major end-use categories (e.g., interior lighting, plug load, refrigeration) based on specific end-use labels (e.g., office 1 light, coffee machine, walk-in freezer), and (3) combining and standardizing the data to extract insights. The end-use data processing shown in this figure provided a way to disaggregate measured data by building type, region, day of week, and month of year. The combined dataset created during this step was used later in other parts of the commercial model updates.



**Figure 120. Data processing workflow for commercial end-use data**

The standardized end-use data generated by the process shown in Figure 120 include 3,723 unique buildings across the United States. Not all buildings included the full set of end uses, so it is necessary to understand the coverage of the data that were used for each end-use model update. Figure 121 shows the portion of the sample that includes interior lighting measurements. This set of 1,089 buildings spans several regions (and climate zones) and includes several building types: mercantile, food service, education, warehouse, office, etc. Data collected in these buildings include a continuous 12 months of measurements at either 1-minute or 15-minute intervals, as shown in Figure 121. Although the lighting data include 1,089 unique buildings, the count varies significantly by building type, with 787 of the 1,089 buildings belonging to the mercantile type. Because some building types do not have enough samples to reflect the average and typical behavior, and others (e.g., service, other) are not included in the ComStock model, only four building types were chosen for interior lighting schedule updates: mercantile, food service, education, and office.

Figure 122 presents the average interior lighting schedules derived for each day of the week from the end-use data for four building types (mercantile, food service, education, and office). The table next to the figure shows which ComStock building types use these schedules and how many end-use data samples informed these schedules. As shown in Figure 122, on average, the interior lighting in retail and strip mall buildings is assumed to be turned off almost completely during unoccupied hours and assumed to stay on until 9 p.m. Full-service restaurant lighting op-

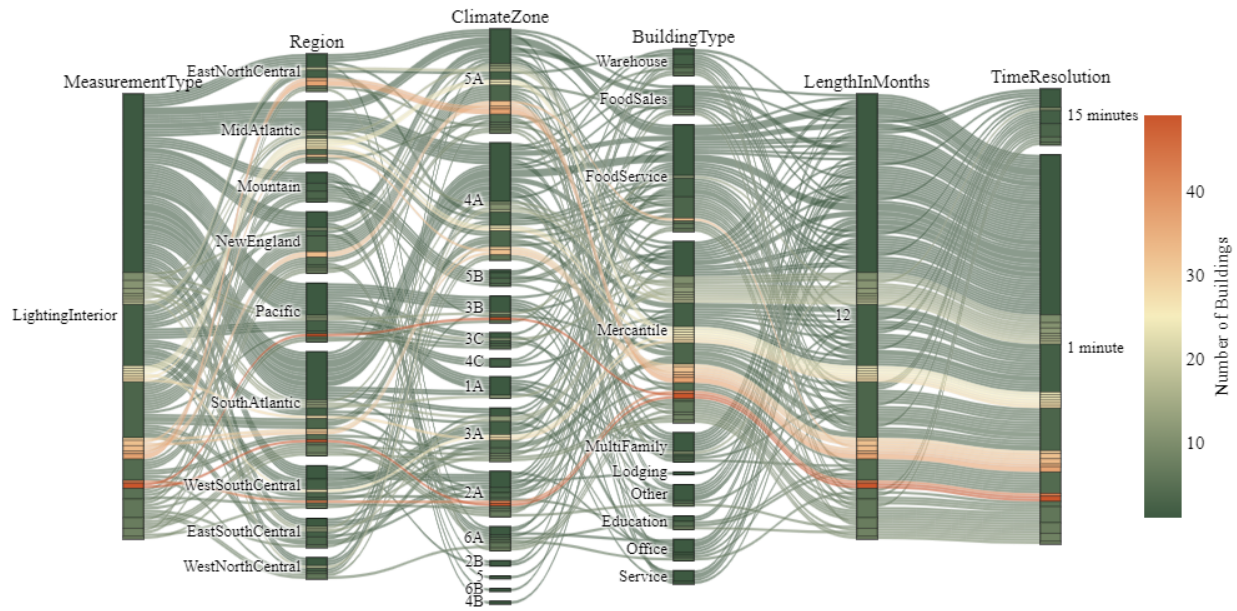


Figure 121. Sample space covered by the commercial building end-use data: Interior lighting

erates more (at 30% of the daily peak value) during the unoccupied period and also stays on later at night. Because schools can have different energy usage behavior between academic period and summer break period, the end-use data collected from the schools were analyzed to differentiate the usage between academic and summer break periods. Based on this differentiation, Figure 122 shows two separate schedules. As expected for schools and offices, the schedules for these building types also differ between weekdays and weekends. Unfortunately, the dataset used to derive these schedules did not include information on the presence of automated lighting controls (occupancy or photosensors).

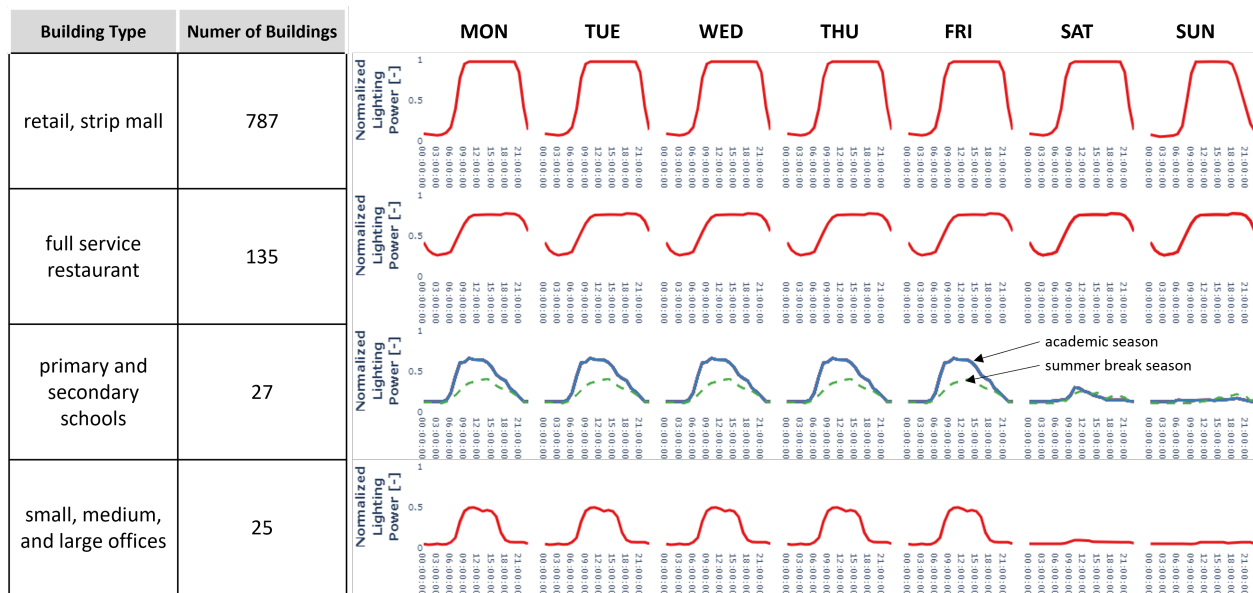
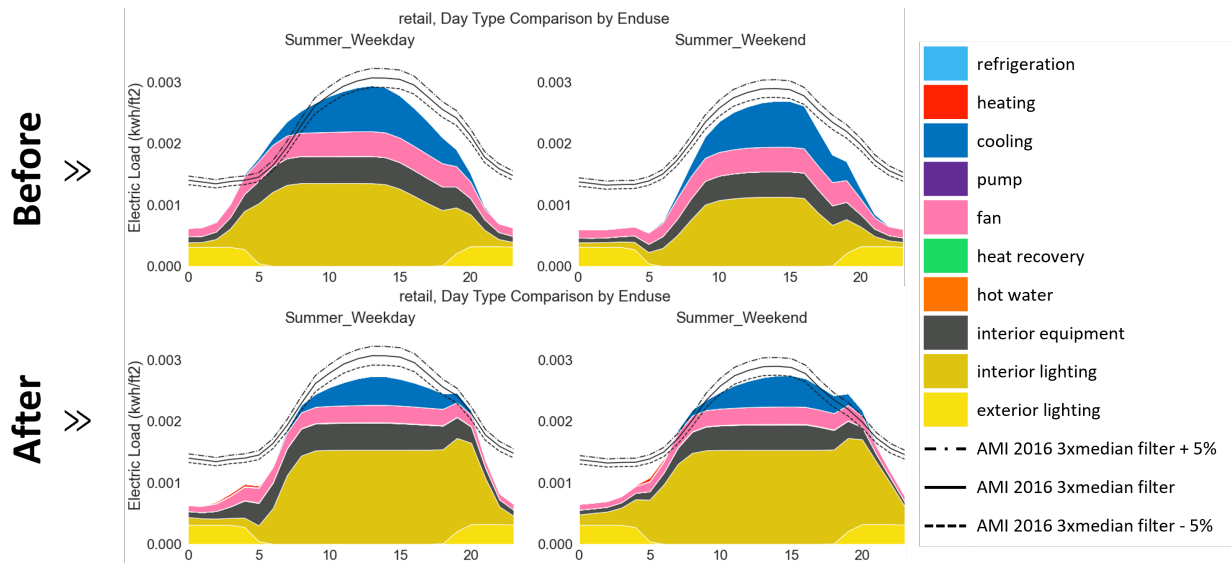


Figure 122. Median schedules derived from the commercial end-use data: Interior lighting

The schedules shown in Figure 122 are only used as the starting point schedules in the ComStock workflow. They are modified later in the workflow (before the energy simulation) to reflect more reality by adding more variability



(described in Section 3.3.4 and 3.3.5 in detail). Figure 123 shows an example of the impact of updating the lighting schedule in retail building models. It shows the average daily summer profiles with electric load intensity (in kWh/ft<sup>2</sup>) for the retail building type and shows the average of both the measured building stock of retail buildings (from AMI data) and the simulated building stock of retail buildings (from ComStock). The schedule update mainly affected the magnitude and/or timing of the end-use component and, as shown in the figure, the update to the retail building type shifted the peak and duration timing of the interior lighting usage later in the day.



**Figure 123. Calibration results before and after the update: Interior lighting schedule**

### 3.3.2 Interior Equipment Schedules

Another major end use that contributes significantly to the energy consumption (as discussed in *ComStock Sensitivity Results*) is the electric interior equipment (i.e., plug load) usage in commercial buildings. Similar to the interior lighting schedule update, average schedules of interior equipment usage for different building types were extracted from the standardized dataset created by the process shown in Figure 120. Figure 124 presents the sample space covered by a total of 1,157 buildings that collected interior equipment end-use measurements. The interior equipment covered in this update includes not only plug loads such as office equipment, but also cooking or kitchen equipment such as electric ovens in restaurants. 697 of the 1,157 buildings were food service buildings. Because some building types did not have enough samples to reflect the average and typical behavior, and others (e.g., service, other) are not included in the ComStock model, only four building types were chosen for interior equipment schedule updates: food service, mercantile, education, and office.

Figure 125 presents the average interior equipment schedules derived for each day of week from the end-use data for the four selected building types. The table next to the figure shows which ComStock building types use these schedules and how many end-use data samples informed these schedules. As shown in the figure, the average interior equipment usage for restaurants on Friday, Saturday, and Sunday is slightly higher than on other days, as peak customer demands are concentrated in these three days. Retail and strip mall interior equipment usage is concentrated in the afternoon, with a high base load during unoccupied hours. Interior equipment schedules in schools are differentiated between academic and summer break periods, and the schedules derived for the office are applied in all small, medium, and large office models in ComStock. As mentioned in Section 3.3.1, schedules shown in Figure 125 are also used as a starting point and are modified further before the simulations to add more variability (described in Section 3.3.4 and 3.3.5 in detail).

Figure 126 shows the impact of the schedule update on primary school building model results. The figure shows the average daily summer profiles with electric load intensity (in kWh/ft<sup>2</sup>) for the primary school building model. It also shows the average of both the measured building stock of retail buildings (from AMI data) and the simulated building stock of primary school buildings (from ComStock). The figure shows the impact of the lighting schedule update



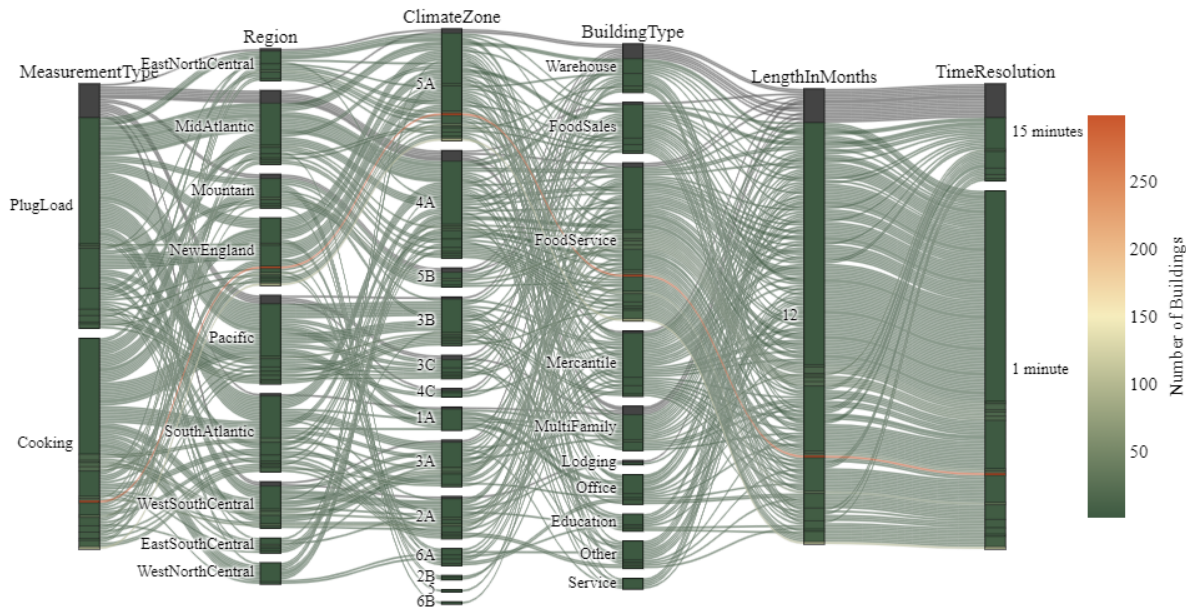


Figure 124. Sample space covered by the end-use data: Interior equipment

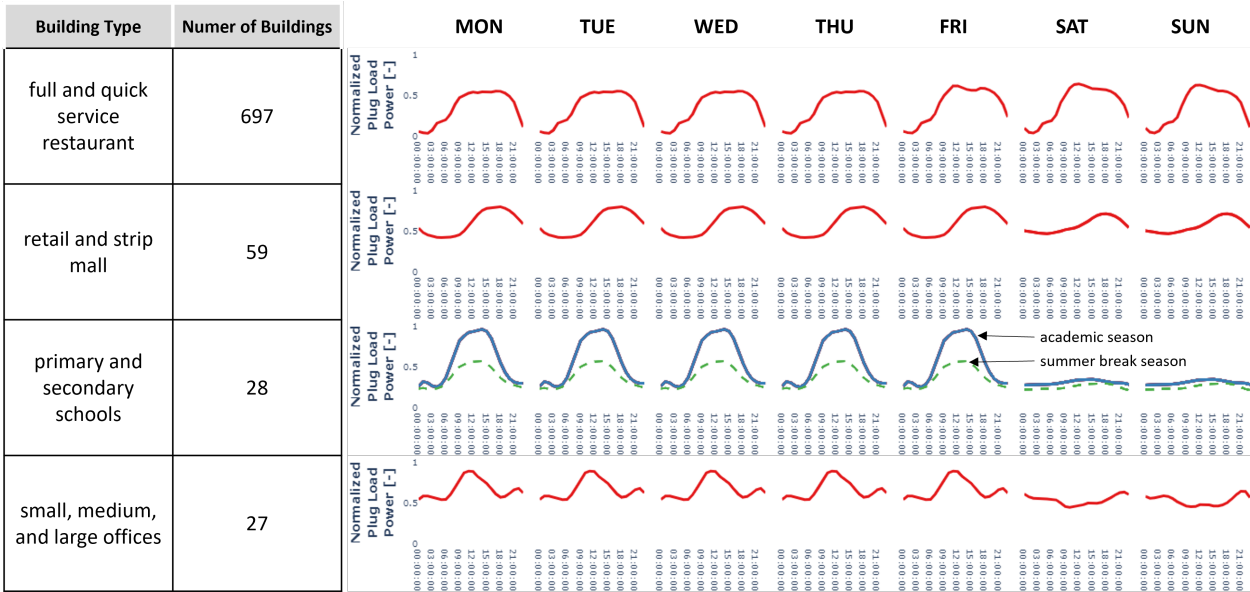


Figure 125. Median schedules derived from the commercial end-use data: Interior equipment

described in Section 3.3.1 as well as the impact of interior equipment schedule update, where both the magnitude and the overall profile changed significantly.

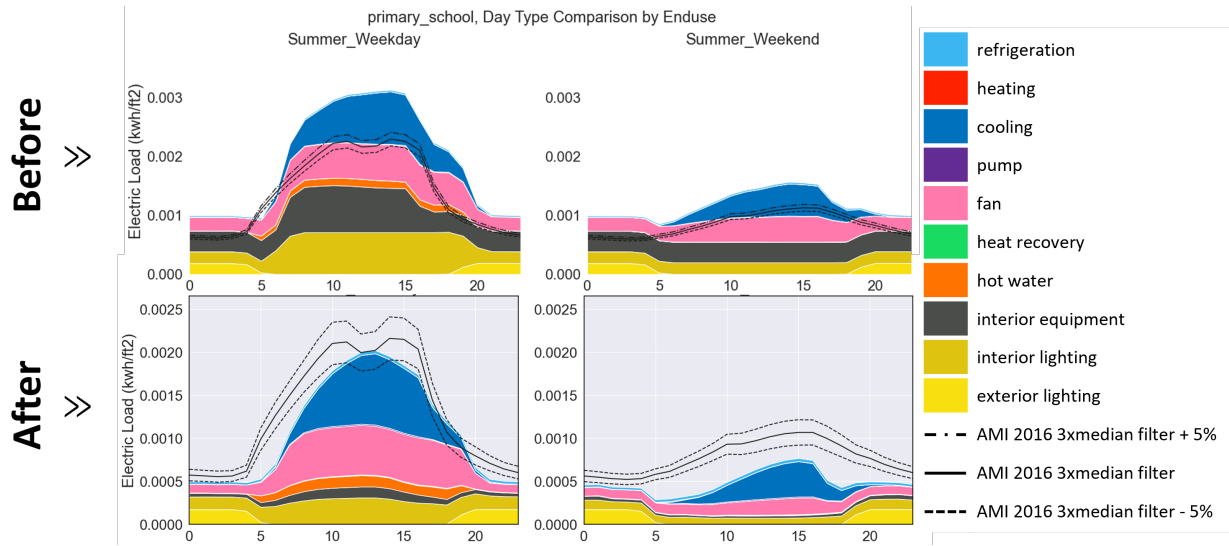


Figure 126. Calibration results before and after the update: Interior equipment schedule

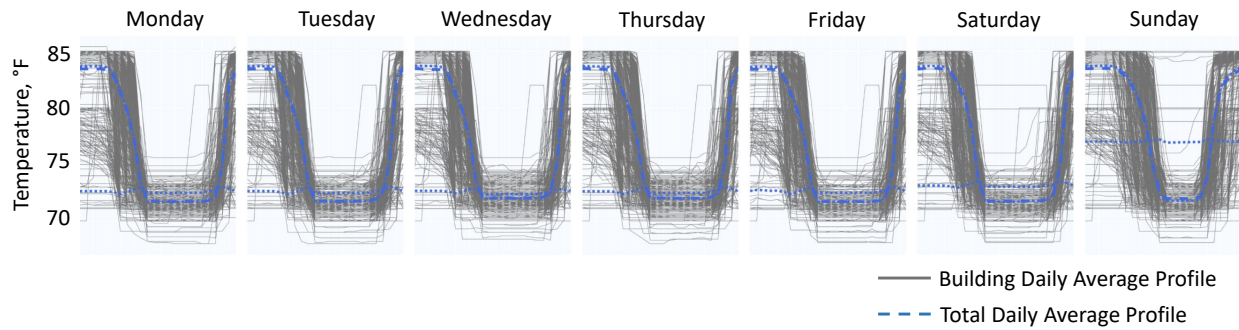
### 3.3.3 Thermostat Setpoints

The sensitivity analysis described in *ComStock Sensitivity Results* shows that thermostat heating and cooling setpoints have a large impact on commercial building energy consumption. To improve our model assumptions, we extracted average thermostat setpoint profiles from three BAS datasets. Altogether, we processed the thermostat setpoints for 3,732 buildings for this analysis, with the counts per building type shown in Table 16. The restaurant and retail building types each contained large sample sizes exceeding 1,500 buildings. The office and school building types contained marginal sample sizes of 31 and 16 buildings, respectively. We assumed that the office data would be suitable for use despite the marginal sample size because it represented a range of building owners across several states. The school data, however, was deemed insufficient due to the small sample size under 25 buildings, and because the buildings in that sample are predominately from a single school district, making it less likely to be representative. All other building types with an equivalent ComStock building type were determined to have counts too small for use and were not changed as part of this effort.

Table 16. Building Counts With Thermostat Data by Building Type

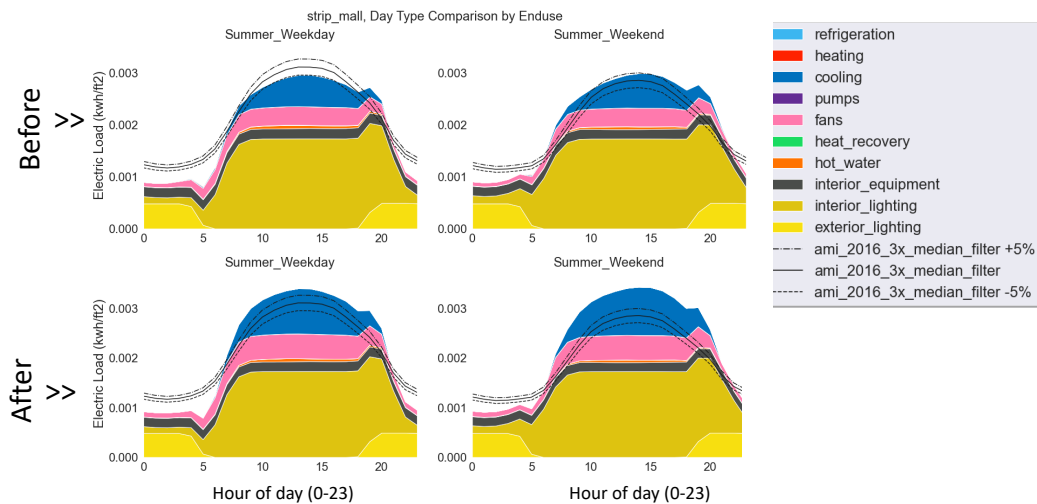
Building Type	Building Count
Food Service/Restaurant	1,817
Mercantile Retail	1,692
Food Sales/Grocery	164
Office	31
School	16
Warehouse	4
Hotel	4
Hospital	2
Outpatient	2

To create average thermostat profiles, we first combined the three data sources into a single dataset by matching corresponding variables (e.g., “htspt” & “heating setpoint”). We then grouped the thermostat variables by individual building and day of the week using a building-count-weighted mean to form average daily profiles per building for both heating and cooling, which were then sorted by building type. Figure 127 illustrates this process for restaurants, where the grey lines represent the average day-of-week profiles for each building sample, and the blue thick dashed lines represent the average profile of the group. Day types with similar profiles were grouped together when appropriate. For most building types, this resulted in separate profiles for weekdays, Saturday, and Sunday.



**Figure 127. Sample averaged cooling thermostat setpoints for retail buildings from end-use data**

Figure 128 illustrates the impact of the thermostat schedule update for retail buildings in Region 1, where an increase in cooling consumption can be seen as a result of the new schedules.



**Figure 128. Impact of thermostat setpoint update for strip malls in Region 1**

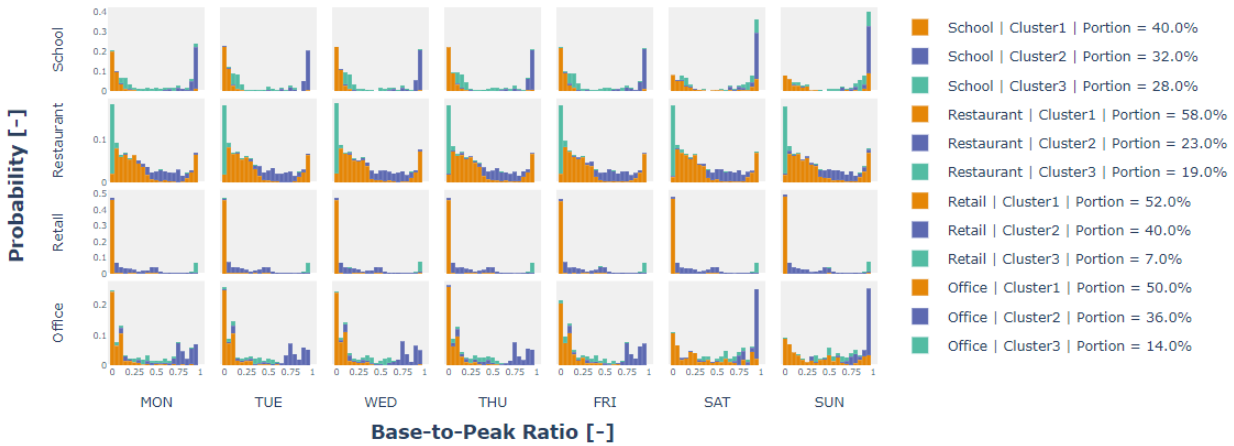
### 3.3.4 Interior Lighting Schedule Magnitude Variability

The average daily usage profile (or schedule) of interior lighting in different building types was derived from the end-use data as described in Section 3.3.1. Modeling all buildings using the same average schedule is not realistic, especially when the focus is narrowed to a certain building type, region, and end use. For example, varying characteristics such as different operating hours (i.e., when a building starts and ends its business operation), different peak timing (i.e., when a building has the highest demand during the day), and different levels of base load (i.e., how much electricity is used during unoccupied hours) should be captured to reflect the realistic performance of buildings in the stock. To reflect this variation in building operation (hereafter referred to as variability) in the models for interior lighting usage in commercial buildings, we used the standardized end-use data from the process shown in Figure 120 to derive a distribution of building schedule characteristics.

Before the update described in this section, ComStock applied a distribution of operating hours derived from whole-building metered data (Bianchi et al. 2020). As shown in Figure 122, while this approach added variability in the “horizontal” dimension by stretching, shifting, or compressing the start and stop time, it did not modify the “vertical” dimension of variability to represent diversity in operation during unoccupied hours. In this update, a distribution of the base-to-peak ratio (BPR) metric was derived from measured data and then used as an input to the ComStock workflow. The BPR, described in Li et al. (2021b), is calculated by dividing the daily minimum (base) load by the

daily maximum (peak) load. The data shown in Figure 121 were used to calculate BPR distributions for different building types.

Figure 129 shows the distributions of BPR generated for the four building types (school, restaurant, retail, and office), divided into three clusters, one for each building type and day of the week. The stacked distribution of the three clusters shows the final distribution of BPR for each building type. It is possible to create one BPR distribution without clustering; however, the clustering analysis was conducted to understand whether characteristics of these groups of buildings (within each cluster) are representative of real buildings. For example (and as shown in the figure), in 52% (cluster 1) of 787 retail buildings, the lighting power dropped to nearly zero during unoccupied hours, while the other groups of buildings either did not have a single typical reduction in lighting power (cluster 2, 40% of 787 buildings), or kept their lights operating even during unoccupied hours (cluster 3, 7% of 787 buildings). Similar to the average schedules, the usage characteristics represented by BPR are also different between weekdays and weekends in school and office building types. For these buildings, the weekend BPRs are highly concentrated close to 1, meaning there is almost no reduction in lighting power during the unoccupied hours. This is because their peak lighting power during the weekend is already low (compared to weekdays) and remains low for the entire weekend, not because schools and offices keep their lights on during the weekend.



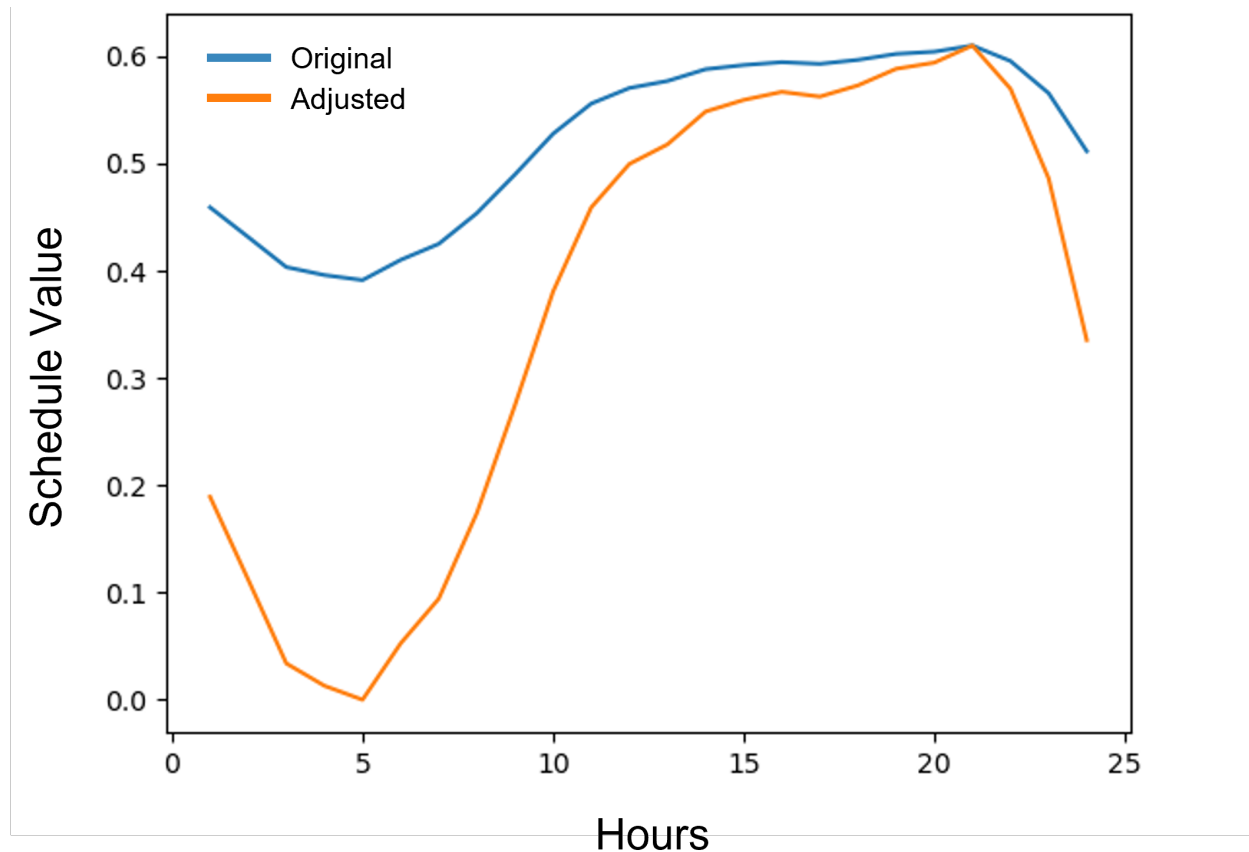
**Figure 129. Base-to-peak ratio distribution: Interior lighting**

The BPR distributions are added into the model workflow through an OpenStudio measure. The BPR distribution for offices in Figure 129 is applied to small, medium, and large office models; the BPR distribution for retail is applied to retail and strip mall models; the BPR distribution defined for schools is applied to primary and secondary school models; and the BPR distribution defined for restaurants is applied to quick and full-service restaurant models. Although no BPR variability is applied to the remaining ComStock building types due to a lack of data, this workflow can be used if supporting data (i.e., BPR distributions for other building types) become available. The measure adjusts interior lighting schedules for individual models such that the distribution of BPR in population matches the distribution shown in Figure 129. Weekday and weekend schedules are modified separately because there are unique BPR distributions for weekdays and weekends. The measure methodology maintained the peak value of each schedule and modified all other points in the schedule so that the ratio of the base value to peak matched the specified BPR for that model. The measure went through each lighting schedule and looped through every timestep,  $t$ , in the schedule and applied Equation 3.26 to calculate the new value at each timestep. In Equation 3.26,  $val$  is the existing schedule value at timestep  $t$ ,  $newVal$  is the new schedule value at  $t$ ,  $bpr$  is the specified BPR,  $peak$  is the peak value in the existing schedule, and  $base$  is the base value in the existing schedule.

$$newVal = \frac{bpr \times peak - base}{peak - base} \times (peak - val) + val \quad (3.26)$$

Because the weekday and weekend BPR values were determined independently of one another, there were some instances in which the adjusted weekday schedules had a lower base value than the adjusted weekend schedules. This is not representative of most commercial buildings because it is unlikely that the building will have a higher baseload

on the weekends. To address this issue, we added a check to the measure after both schedules were adjusted. If the adjusted weekend base value was higher than the adjusted weekday base, we modified the adjusted weekend schedule once again to have the same base value as the adjusted weekday schedule. Figure 130 shows a single lighting schedule before and after modification by the BPR measure. The peak value remains the same, and the rest of the schedule is scaled to meet the specified BPR (0.0 in this case).



**Figure 130. Example schedule adjustment by the BPR measure**

Adding lighting BPR variability to ComStock produced a small change in the lighting energy use at the stock scale, but added more realistic variability within the stock. Figure 131 shows an end-use breakdown of retail electric load in the summer. Between the “before” and “after” plots, there is a noticeable increase in base interior lighting electric load (labeled “interior lighting”). This change brought the whole baseload closer to the AMI data, accounting for a portion of the missing nighttime load.



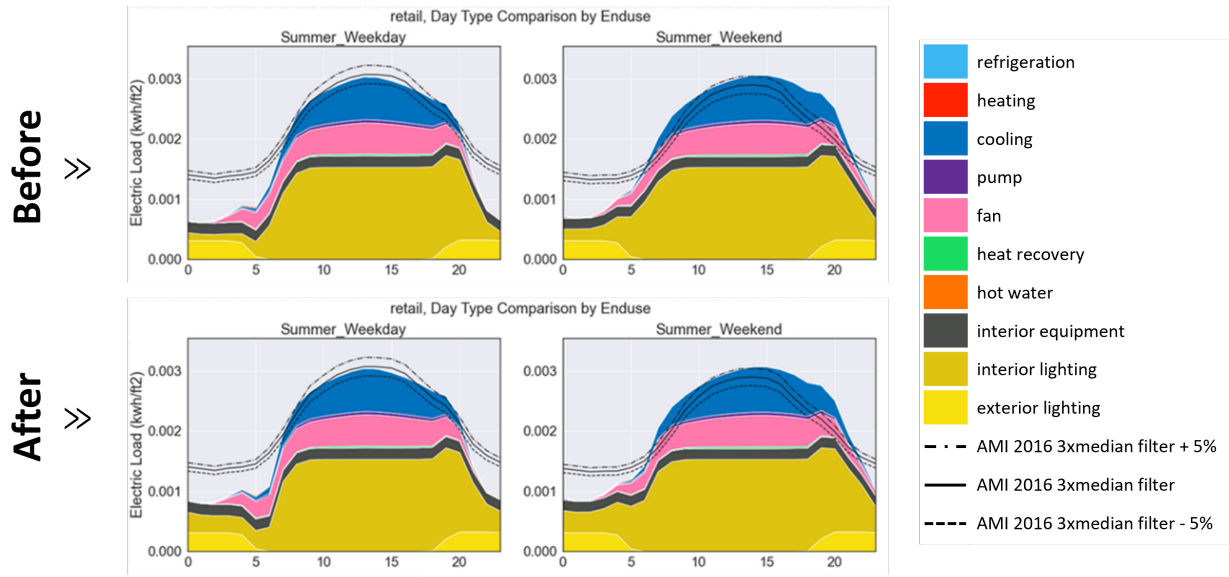


Figure 131. Calibration results before and after the update: Interior lighting BPR variability

### 3.3.5 Interior Equipment Schedule Magnitude Variability

In order to add more realistic variability in interior equipment usage for different building types, we applied the methodology described in Section 3.3.4 to data covering the sample space shown in Figure 124 to calculate BPR distributions. Figure 132 shows the distributions of BPR generated for the four building types (school, restaurant, retail, and office) grouped into three clusters for each building type by day of the week. The distributions of the three clusters for each building type are stacked to show the final distribution of BPR for each building type and for each day of the week. As shown in the figure, in 66% (clusters 1 and 3) of 697 food service buildings, the interior equipment power dropped to less than 50% of the peak during unoccupied hours. In the other 24% (cluster 2), interior equipment base load varied without strict control of equipment between occupied and unoccupied hours. Similar to the average schedules, the usage characteristics represented by BPR vary between weekdays and weekends in school and office building types. For office and school buildings, the weekend BPRs are highly concentrated close to 1, meaning there is almost no usage reduction during unoccupied hours. This is because their peak power on the weekend is already low (compared to weekdays) and it remains low for the entire weekend, not because schools and offices keep all of their interior equipment operating.

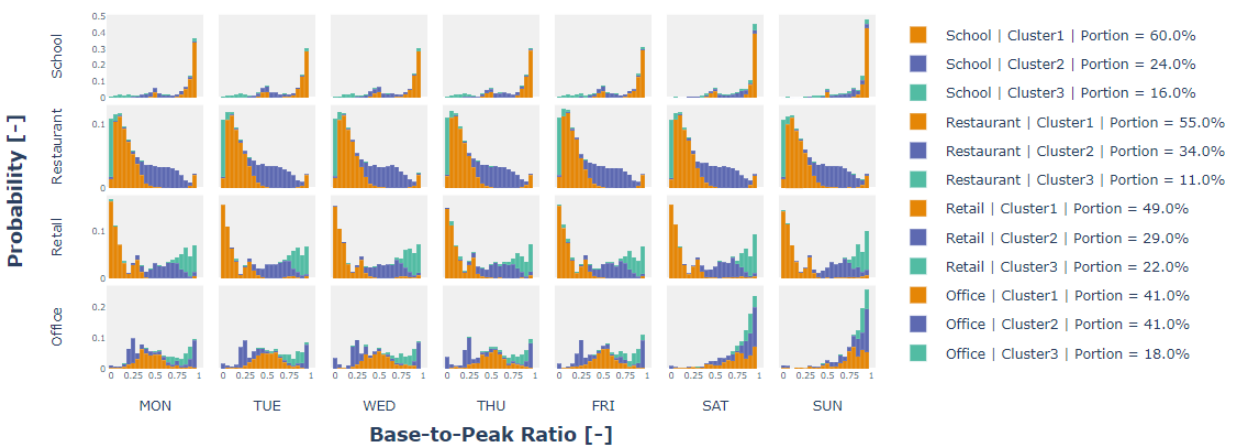
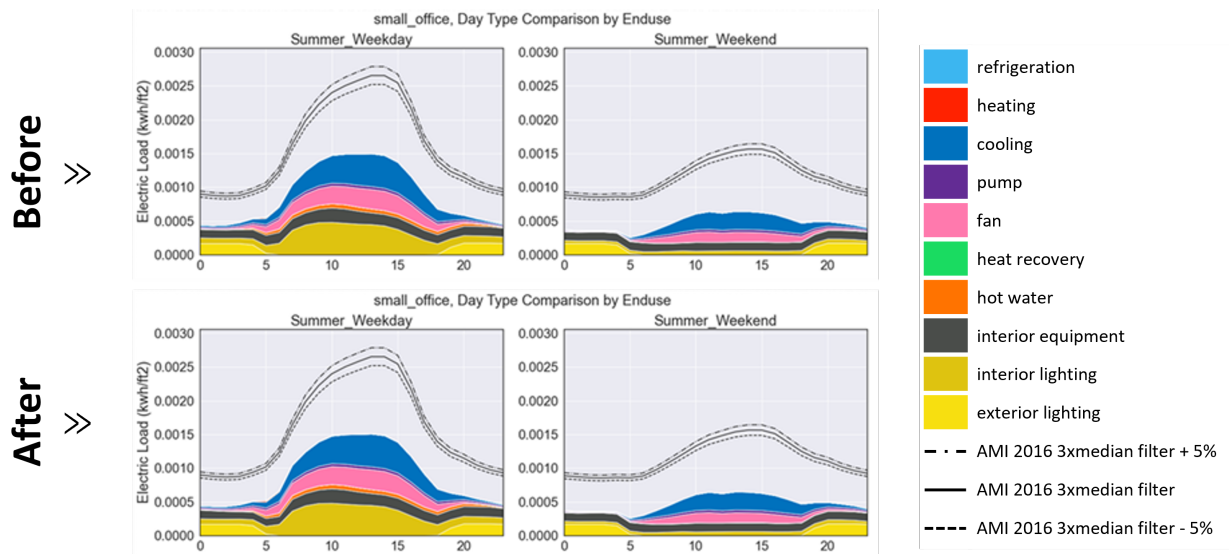


Figure 132. Base-to-peak ratio distribution: Interior equipment



The OpenStudio measure workflow described in Section 3.3.4 is used to apply the interior equipment BPR variability in the models. The measure excluded vertical transport such as elevators. Weekday and weekend schedules were modified separately as there are unique BPR values for weekdays and weekends. The measure methodology maintained the peak value of each schedule and modified all other points so that the ratio of the base value to peak matched the specified BPR. Details of schedule modification methodology can be found in Section 3.3.4. Adding plug load BPR variability to ComStock did not produce a major change in plug load energy use at the stock total level. Figure 133 shows an end-use breakdown of small office electric load in the summer. There is an almost imperceptible difference between the plug load profile (labeled “interior\_equipment”) in the “before” and “after” plots. A slight difference can be observed in the morning baseload. Although the aggregate change is not significant, it does add realistic diversity within the stock.



**Figure 133. Calibration results before and after the update: Interior equipment BPR variability**

### 3.3.6 Exterior Lighting Power

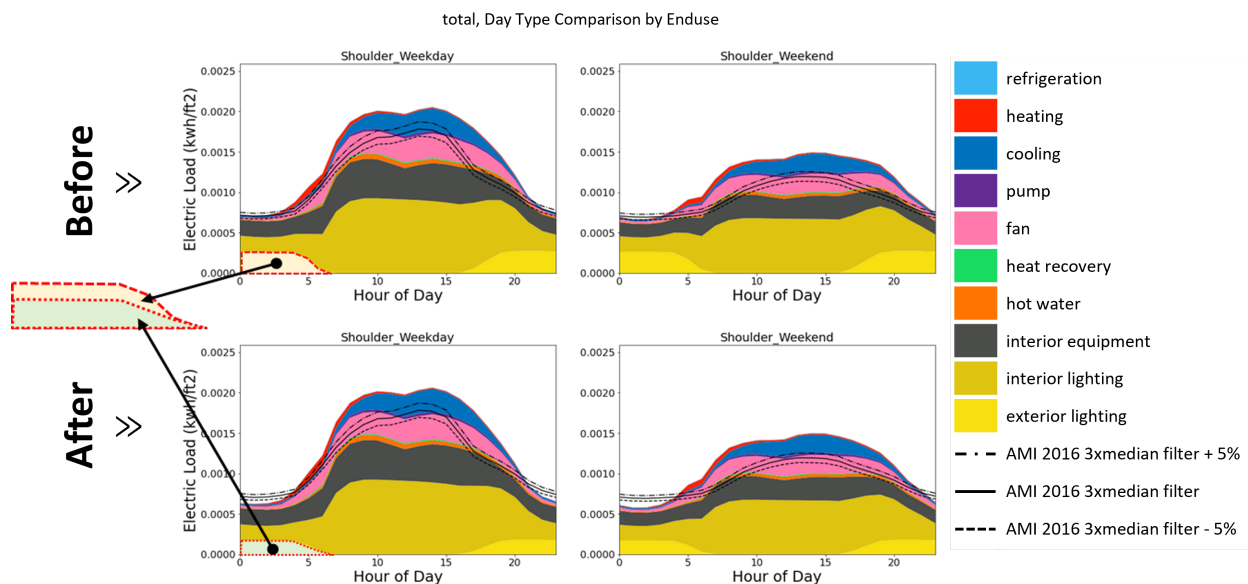
In addition to the significant model parameters that were emphasized in *ComStock Sensitivity Results*, additional efforts were made to review other building model parameters based on available data. One of the parameters that was reviewed was exterior lighting energy usage. While the operation behavior (or profile) of the exterior lighting is typically relatively simple (e.g., lights illuminating only when it is dark outside) compared to interior lighting, the variety of exterior lighting types (e.g., parking area, walkway, main entry, loading dock) and the power density of each lighting type can vary. Thus, in this update, the 2015 U.S. Lighting Market Characterization Study (Buccitelli et al. 2017) was used to derive the exterior lighting power density, with an emphasis on parking lot lights, whose lights make up a majority of the exterior lighting energy consumption.

The study provides “summary estimates of the installed stock, energy use, and lumen production of all general illumination lighting products operating in the U.S.,” where the estimates are based on “both public and confidential sources of information including building lighting audits, industry surveys, national lighting product shipment data, and interviews with lighting professionals and subject matter experts.” Based on the level of granularity the estimates provide, it is possible to extract the average wattage of parking lighting system (216 W/system) and average number of parking lot spots (13 lot-spots/system) that can be used to derive the average wattage used in each parking lot spot (16.6 W/lot-spot). Based on the average square footage of each lot spot (405 ft<sup>2</sup>/lot-spot) derived from Richman et al. (2008), the lighting power density of the parking lot light that represents the average of nationwide consumption can be calculated as 0.041 W/ft<sup>2</sup>-nationwide. Table 17 presents the revised parking lot lighting power densities for each exterior lighting standard category. The previous lighting power densities (from the energy code) and the portion from the national building samples in each exterior lighting standard category were used to weight the national average (0.041 W/ft<sup>2</sup>) for each category. Note that the Title 24 energy code is applicable to buildings in California, and the ASHRAE 90.1 series of energy codes is used for buildings in other states.

**Table 17. Exterior parking lot lighting power density derived from the data source**

Exterior Lighting Standard	Portion From National Samples	Revised Parking Lot LPD [W/ft <sup>2</sup> ]
DOE Ref 1980–2004	34.0%	0.050
ASHRAE 90.1-2004	12.6%	0.042
ASHRAE 90.1-2007	27.5%	0.042
ASHRAE 90.1-2010	8.22%	0.028
ASHRAE 90.1-2013	5.87%	0.028
Title 24/DEER Pre-1975	0.00%	0.037
Title 24/DEER 1985	0.12%	0.037
Title 24/DEER 1996	0.91%	0.037
Title 24/DEER 2003	1.67%	0.037
Title 24/DEER 2007	2.81%	0.037
Title 24/DEER 2011	2.35%	0.037
Title 24/DEER 2014	0.81%	0.018
Title 24/DEER 2015	1.55%	0.018
Title 24/DEER 2017	1.58%	0.018

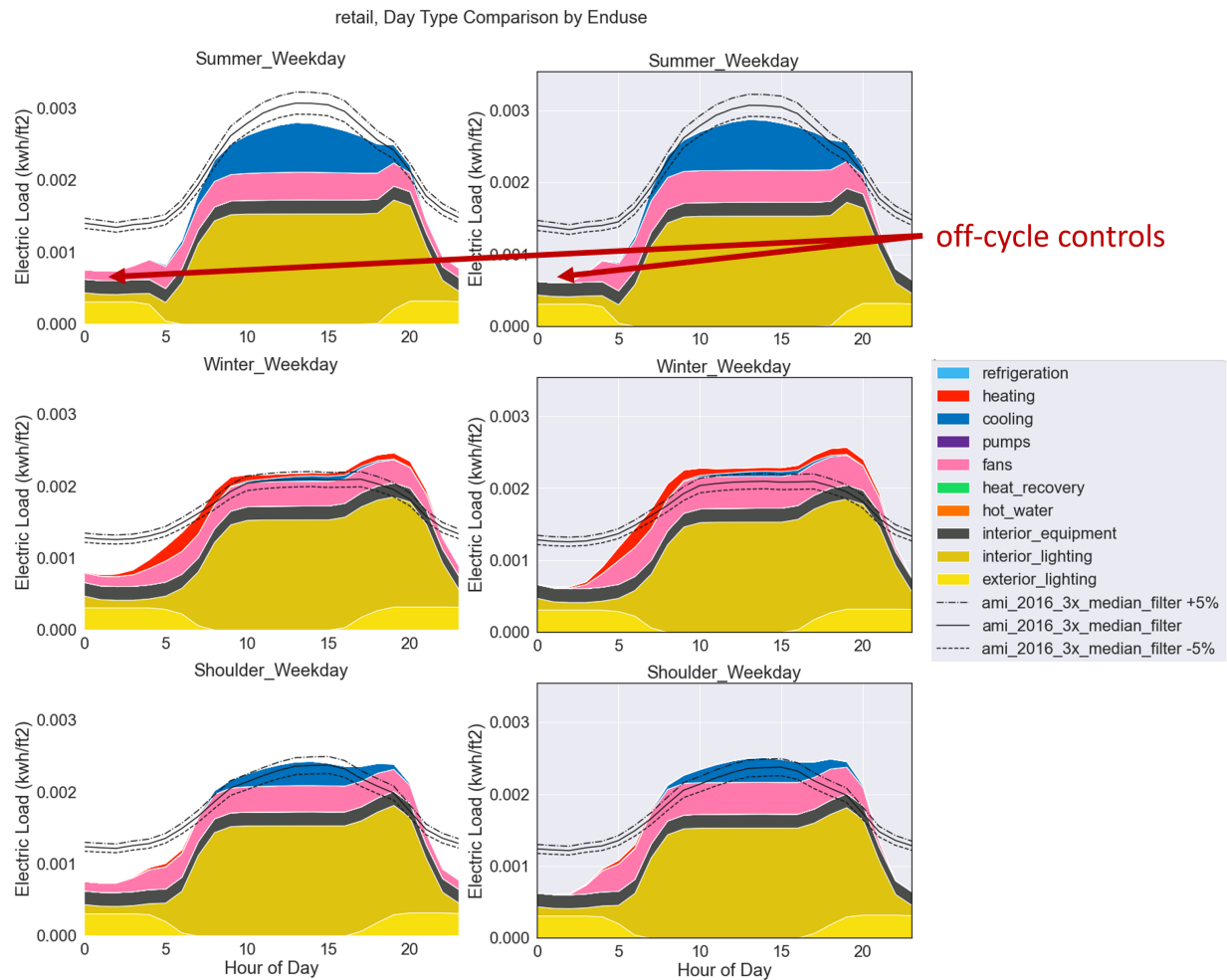
Figure 134 is an example of the impact of the exterior lighting power density update on the total load profile. It shows a reduction in the average exterior lighting load in the simulated building stock (from ComStock). While this update appears to have increased the difference between the modeled and measured load profiles, the change was kept because we believe it accurately reflects the best available data for exterior lighting in the stock.



**Figure 134. Calibration results before and after the update to exterior lighting (parking lots) power density**

### 3.3.7 Off-Cycle Controls for Packaged Single-Zone Systems

The previous behavior for packaged single-zone HVAC systems was to stay fully on (fans and space conditioning) if they provide ventilation, in order to meet ventilation minimums, including during nighttime operation. This resulted in unrealistic fan energy use at night. Packaged single-zone system controls were updated to default to the zone occupancy schedule if the system provides ventilation, and cycling operation at all other times or if the unit does not provide ventilation. Packaged single-zone systems that are affected include PTAC, PTHP, PSZ-AC, PSZ-HP, four pipe fan coil, VRF terminal units, and water-to-air heat pumps. If paired with a ventilation system, as is typically the case for four pipe fan coil, VRF, and water-to-air heat pump systems, the units follow cycling operation to meet zone load.



**Figure 135. Example of off-cycle control impact on retail buildings. Previously, fans ran continuously to provide ventilation air during unoccupied times. Off-cycle controls changed this to cycling operation, which results in little to no fan energy use in this region because of aggressive thermostat setbacks. Later changes reduce the thermostat setback severity, resulting in more nighttime fan energy use than shown for this region.**

### 3.3.8 Demand Control Ventilation and Energy Recovery

Newer versions of ASHRAE 90.1 require demand control ventilation (DCV) and/or heat/energy recovery for HVAC systems under certain conditions. The DCV control was being negated by a software programming mistake that was converting all ventilation requirements to a per-area value before DCV controls were applied, causing the DCV in the models to never operate. Also, in certain circumstances, energy recovery was not being added when it should have if percentage of outdoor air fell within a certain range. These programming mistakes were located and fixed in the OpenStudio-Standards library. Both DCV and heat recovery are rare, as most buildings either use systems that don't require it, or were built before these code requirements.

### 3.3.9 Nighttime HVAC System Operation

The default operation for HVAC equipment in ComStock was originally to schedule fans and ventilation to shut off at night and only cycle on when needed to maintain thermostat setpoints. In reality, the nighttime behavior of HVAC equipment in the commercial building stock varies. Some rooftop units (RTUs) follow this behavior, while others operate their fan 24/7, regardless of the hours of operation of the building. Additionally, RTUs that cycle on and off at night may or may not control their intake of outdoor ventilation during unoccupied hours. Fan operation and

ventilation both have a significant impact on energy consumption, so we wanted to better align the behavior of our models to those of the building stock.

We used one of the procured commercial building BAS datasets to better understand the probability distributions for the different unoccupied fan and ventilation behaviors. Overall, 843 building sites containing 5,706 RTUs were analyzed. The counts per building type are shown in Figure 18. Building types with counts less than 25 utilize the average of all building types in this analysis. The timeseries behavior of three BAS datapoints from this dataset was analyzed: (1) *OccupiedHours*, which specifies if the RTU was scheduled to be in an occupied mode for the timestep, (2) *FanCallHrs*, which specifies if the RTU fan was active for the timestep, and (3) *VentModeHrs*, which specifies if the RTU outdoor air damper was open for the timestep. These three variables were used to categorize RTUs into 1 of 3 nighttime operation modes, summarized in Table 19, by observing the behavior between the hours of 1 a.m. and 4 a.m., which could be considered the most reliable hours for when a typical RTU would experience nighttime setbacks. Note that Hotels and Hospitals were omitted from this analysis due to their unique 24/7 operation hours.

**Table 18. Building and RTU counts per building type from commercial BAS data source used for nighttime HVAC analysis**

Building Type	Site Count	Unit Count
All	843	5,706
Retail	541	3,300
Unknown	164	1,391
Office	43	466
Restaurant	39	155
Grocery	35	212
Hotel	6	46
Education	6	29
Warehouse	5	94
Healthcare	4	13

**Table 19. Nighttime RTU operation modes from BAS data variables. Note that RTUs with Modes='off' would fall into either Mode 2 or Mode 3 should the RTU need to cycle on to meet zone thermostat requirements.**

	Description (Night)	Night Occupied Status	Night Fan Status	Night Ventilation Status
<b>Mode 1</b>	Scheduled on, running	Active	Active	Active
<b>Mode 2</b>	Scheduled off, fan cycles with ventilation	Inactive	Active	Active
<b>Mode 3</b>	Scheduled off, fan cycles without ventilation	Inactive	Active	Inactive
<b>Off</b>	Scheduled off, fan off	Inactive	Inactive	Inactive

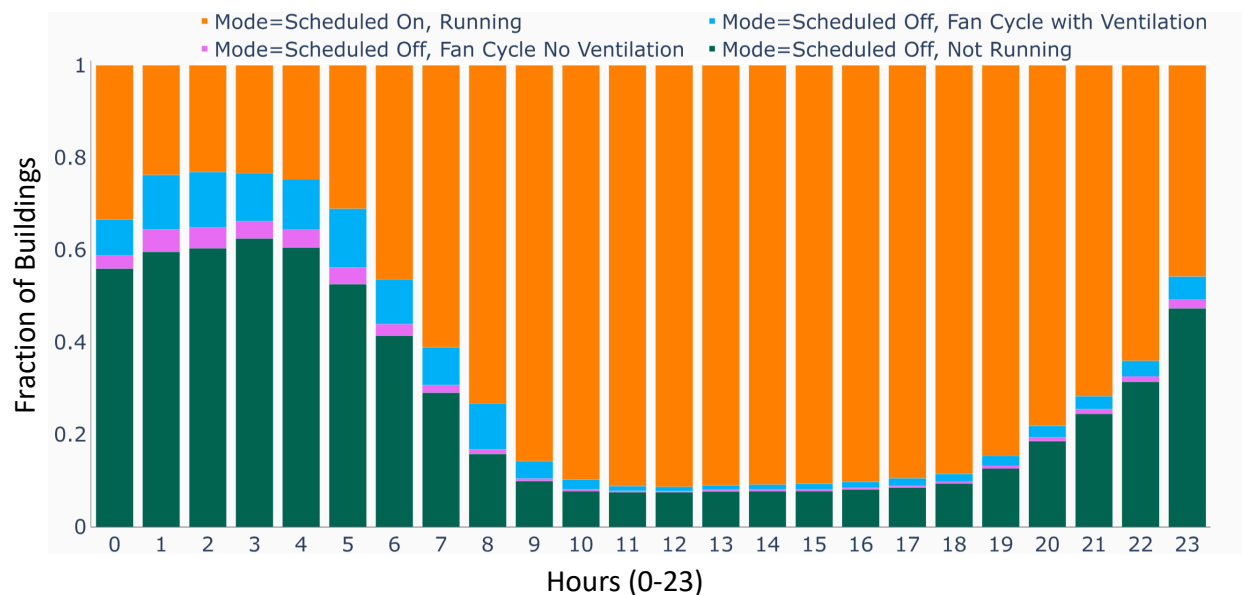
Figure 136 illustrates the average hourly fraction of the buildings in the BAS dataset across all building types that utilize the different nighttime operation modes. For ComStock implementation, probability distributions were created for these operation modes per building type, to then be applied to models with applicable RTU-based HVAC system types.

Mode 1, shown in orange, indicates that the RTU's fans and outdoor air ventilation are scheduled to be fully active through the night, with no scheduled unoccupied mode. This would be the most energy consuming operation mode due to the additional fan operation hours and outdoor ventilation air that requires conditioning. The data suggest that this operation mode is used in nighttime operation in more than 20% of buildings based on the minimum fraction of buildings experiencing this mode for any given hour, generally between 1 a.m. and 4 a.m. We chose the minimum fraction as our indicator to minimize the inclusion of units that may have very late end times or very early start times

when deriving the fraction of units that do not turn down at night. Note that variation is prevalent between building types. Offices, for example, showed more than 50% of units operating in this mode.

Modes 2 and 3 are both variations of RTUs that are scheduled to be off at night; combined, they make up the remaining portion of the probability distribution apart from Mode 1. Mode 2 and Mode 3 both schedule systems to shut down at night, but will cycle on when needed to maintain zone thermostat setpoints. The difference between them is that when Mode 2 cycles on its fans, the outdoor air (ventilation) dampers are scheduled to be open, whereas when Mode 3 cycles its fans, the outdoor air (ventilation) dampers are scheduled to be closed. This would generally make Mode 3 use less energy because it would not require conditioning of outdoor air. Note that for this analysis we ignore data points where a unit is both scheduled to be off and is not running, as indicated in green. It is not possible to differentiate between Mode 1 or Mode 2 when the system is not running. Additionally, some of the Mode 2 datapoints could reflect the operation of an economizer, which brings in extra cool outdoor air during times when the unconditioned outdoor air is cool enough to provide some cooling. There was no economizer status indicator in the dataset.

To determine the probability distribution of different operation schemes, we analyzed the ratio between Mode 2 and Mode 3 at hours most likely to be good indicators of nighttime behavior, between 1 a.m. and 4 a.m. We based the distribution on the hour that had the lowest ratio of Mode 2 to Mode 3 to minimize the influence of economizers on the distribution, which could inflate the prevalence of Mode 2. Furthermore, this analysis was also repeated separately for each season, and the minimum ratio of Mode 2 to Mode 3 across seasons was used to further mitigate economizer influence. It can be seen in Figure 136 that Mode 2, in blue, is more prevalent during the nighttime hours between 1 a.m. and 4 a.m. than Mode 3, in purple. Utilizing the multiseason analysis approach, 73% of units with a scheduled unoccupied mode fell under Mode 2, while the remaining 23% fell under Mode 3. Normalizing between all modes yielded an overall stock-wide probability distribution of 27% for Mode 1, 50% for Mode 2, and 23% for Mode 3. In ComStock, building types with less than 25 samples used the stock-wide distribution, while building types with more than 25 samples used the building-type-specific distribution.

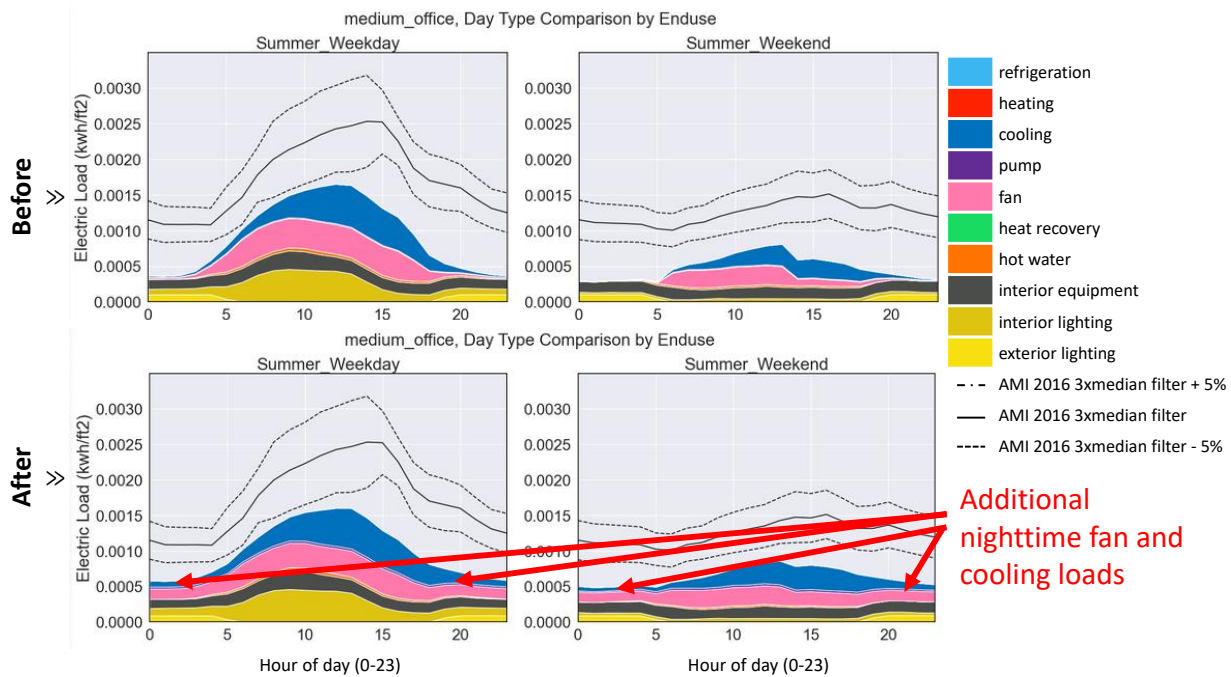


**Figure 136. Average hourly RTU operation mode fractions from BAS datasource; data shown are for 843 buildings containing 5,706 RTUs**

An OpenStudio measure script was written to implement the three operation schemes; models with applicable RTU HVAC system types are assigned these schemes based on the probability distributions derived from the BAS data. This implementation requires modifications to outdoor air schedules, air loop operation schedules, and fan operation schedules in the models. An example of the impact of this analysis can be seen in Figure 137 for medium office buildings for calibration Region 2b, Portland, OR. Increases in fan energy and cooling energy during the early morn-



ing and late evening hours can be observed as a result of adding RTUs with constant nighttime fan and ventilation operation to a fraction of buildings in the stock.



**Figure 137. Impact of adding HVAC nighttime behavior variability on Region 2b—Portland, OR. Higher nighttime fan and cooling loads can be observed.**

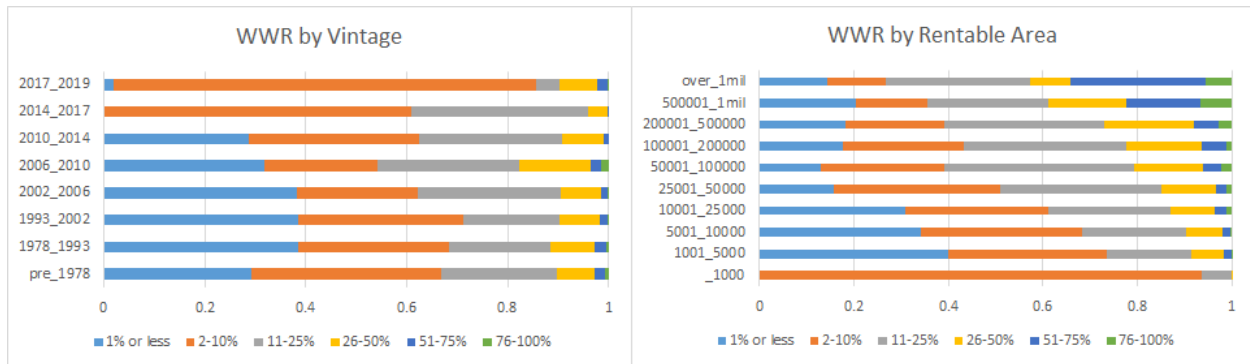
### 3.3.10 Window-to-Wall Ratio

Window-to-wall ratio (WWR) is defined as the fraction of above-grade wall area that is covered by fenestration. Previously, ComStock used the WWR from the DOE prototype building models. While each building type had a different WWR, there was no variability within each building type, which is not representative of the building stock. To address this issue, we referenced the NFRRC Commercial Fenestration Market Study, conducted by Guidehouse (Barbour et al., Expected publication 2021). The study intended to characterize the national commercial window stock through data collection and analysis. Six primary data sources representing all regions of the United States were used in the study—a 2020 Guidehouse survey, NEEA CBSA, DOE Code Study, CAEUS, CBECS, and RECS (multifamily). A variety of window properties were collected, including WWR, number of panes, frame material, glazing type, low-E coating, gas fill, and many others. In total, the database contained approximately 16,000 samples, each with an appropriate weighting factor based on coverage, completeness, and fidelity of each data source. We incorporated the WWR results from this study into the ComStock model and may incorporate other fields in the future to further refine our window modeling methodology.

From the Guidehouse data, we developed a WWR distribution for each combination of building type, floor area, and vintage. We first analyzed the WWRs separately by building type, floor area, geographic location, and vintage to determine which filters were appropriate to use for the final distributions. Geographic location did not have a significant impact on WWR, so it was left out of the final distributions. As can be seen in Figure 138 below, we see a noticeable change in the WWR in buildings built after 2014, indicating that new buildings are trending toward larger windows. Similarly, we see a distinct trend in the WWR as a function of floor area, where larger buildings have more windows. While the previous methodology only varied WWR by building type, these new distributions introduce more WWR variability through the consideration of vintage and floor area.

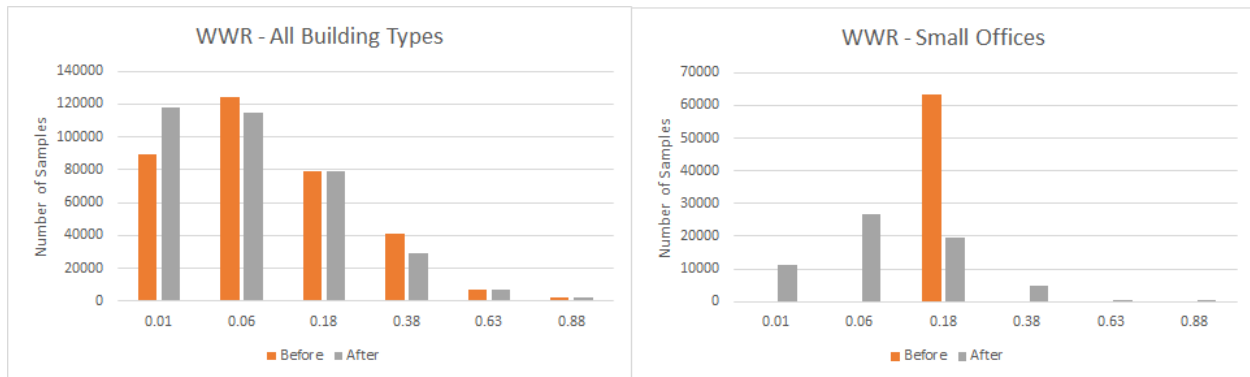
The before and after WWR distributions are shown in Figure 139. The distinct bins in the graph are a result of the way WWR is binned in the CBECS Show Card: 0–1% WWR is binned to 0.0, 2–10% to 0.06, 11–25% to 0.18, 26–50% to 0.38, 51–75% to 0.63, and 76–100% to 0.88. The final distributions do not change the stock total energy





**Figure 138. Window-to-wall ratio breakdowns by vintage and rentable floor area**

consumption significantly, but add realistic variability within building types, as can be seen in Figure 139. Previously, all small offices had the same WWR of 0.15, whereas using the new distributions, small office WWRs vary from 0.01 to 0.88.



**Figure 139. Window-to-wall ratio distributions before and after incorporating NFRC data**

### 3.3.11 Restaurants in Strip Malls

In the case of strip malls, we recognized a need to add a restaurant space type, as food services are commonly found in shopping plazas. Previously, strip mall models were only made up of retail space types, which resulted in our strip mall models having low internal loads and low variability within this building type. In reality, however, every strip mall has a different combination of businesses, some of which may be restaurants with much higher cooking loads than retail spaces.

In order to determine what fraction of strip mall models should be the restaurant space type, we surveyed strip malls in the Denver area, recording the number of restaurant and non-restaurant businesses in each plaza. We surveyed approximately 20 strip malls, a total of 189 businesses, 40 of which were restaurants. The average restaurant fraction was 21%, median 20%, minimum 5%, and maximum 50%. With such a large range of restaurant fractions, we decided to implement a distribution rather than a hard-coded restaurant fraction to better capture the variability in real strip malls. We centered the distribution around 20%, the median of our survey results, and chose a minimum of 0% and maximum of 40%. Our final distribution of restaurant fractions is shown in Figure 140.

As a result of adding 0–40% restaurant space types in strip mall models, EUIs increased considerably due to added kitchen internal loads. In addition, there is more variation in EUIs now that the strip mall models contain different combinations of space types. Figure 141 shows strip mall electric EUIs before and after restaurants were added.



Figure 140. Distribution of restaurant fraction in strip malls

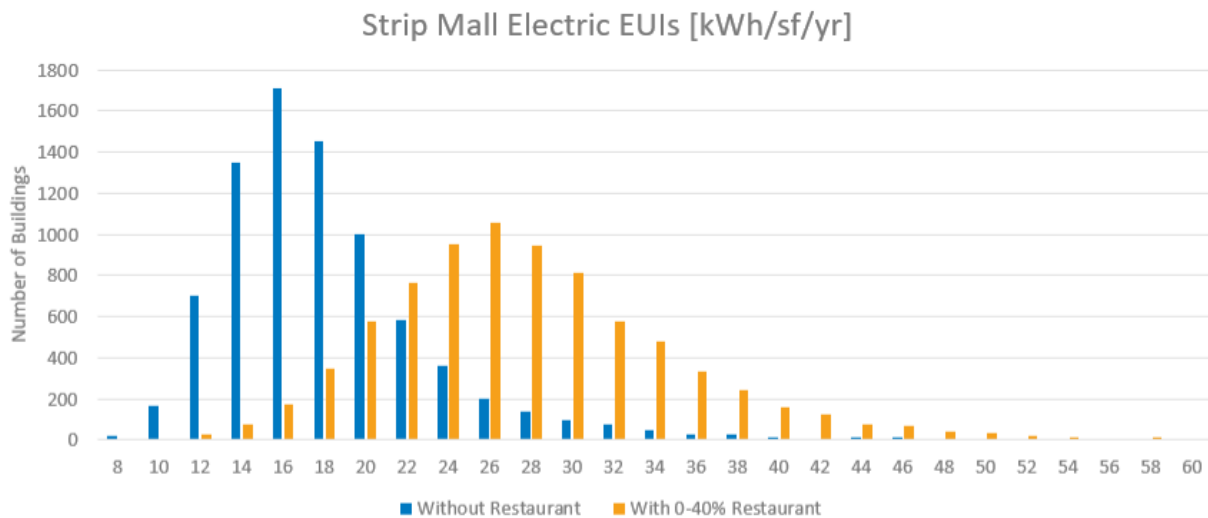


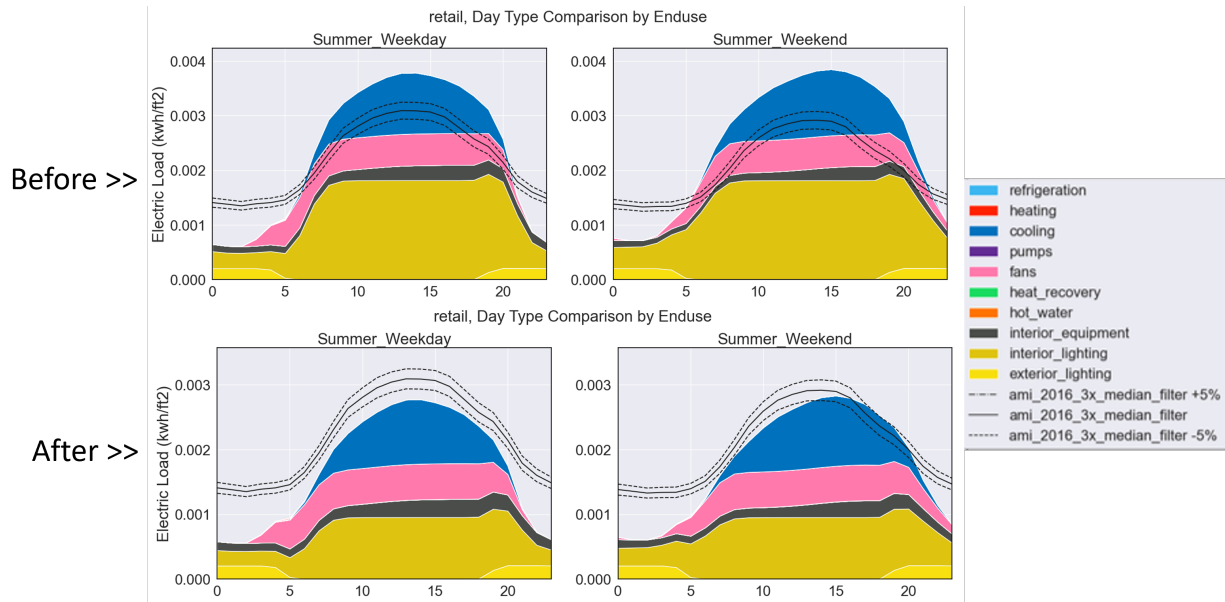
Figure 141. Strip mall electric EUIs with and without restaurant space type

### 3.3.12 Lighting Power Density

Lighting is the largest commercial end use and building system most frequently retrofitted to save energy. Most buildings install far less lighting power than the ASHRAE 90.1 allowance with lower lighting levels and better lighting technology. The 2019 NEEA Commercial Building Stock Assessment (CBSA) and 2015 DOE U.S. Lighting Market Characterization Report (LMC) include lighting power density estimates by building type, shown in Table 20. The CBSA report shows an average lighting power density of around 1.0 W/ft<sup>2</sup> in 2014, and 0.75 W/ft<sup>2</sup> in 2019. The LMC report shows an average of around 0.7 W/ft<sup>2</sup> from 2015. The ASHRAE 90.1-2019 prescriptive lighting power density allowances are the ComStock template version that most closely represents the stock installed lighting power density around 2018. Figure 142 shows a comparison of retail lighting before and after the update. As Table 20 shows, this update brings installed lighting power densities closer to typical buildings for all building types. With the update, educational and warehouse buildings in ComStock have slightly lower lighting power densities than show in the CBSA or LMC reports, likely reflecting longer retrofit timelines for these buildings types.

**Table 20. Commercial Lighting Power Density Comparison by Building Type**

<b>ComStock Building Type</b>	<b>CBSA Building Type</b>	<b>CBSA 2014 (W/ft<sup>2</sup>)</b>	<b>CBSA 2019 (W/ft<sup>2</sup>)</b>	<b>LMC Building Type</b>	<b>LMC (W/ft<sup>2</sup>)</b>	<b>ComStock Previous Avg. (W/ft<sup>2</sup>)</b>	<b>ComStock ASHRAE 90.1-2019 pre- scriptive (W/ft<sup>2</sup>)</b>
Full-service restaurant	Restaurant	1.2	0.7	Food service	0.7	1.76	0.77
Quick-service restaurant	Restaurant	1.2	0.7	Food service	0.7	1.41	0.85
Outpatient	Hospital	NA	0.9	Health care—outpatient	0.5	1.22	0.87
Hospital	Hospital	NA	0.9	Health care— inpatient	0.6	1.56	0.98
Strip mall	Retail/service	1.2	0.8	Retail—mall & non-mall	0.8	1.48	0.8
Retail	Retail/service	1.2	0.8	Retail—mall & non-mall	0.8	1.89	0.98
Small hotel	Lodging	0.9	0.6	Lodging	0.6	1.41	0.71
Large hotel	Lodging	0.9	0.6	Lodging	0.6	1.11	0.44
Small office	Office	1	0.8	Offices (non- medical)	0.6	1.18	0.67
Medium office	Office	1	0.8	Offices (non- medical)	0.6	1.11	0.67
Large office	Office	1	0.8	Offices (non- medical)	0.6	1.13	0.67
Primary school	School	1	0.9	Education	1.4	1.3	0.69
Secondary school	School	1	0.9	Education	1.4	1.22	0.71
Warehouse	Warehouse	0.6	0.5	Warehouse and storage	0.8	0.77	0.45



**Figure 142. Comparison of retail lighting profile after the lighting power density update to reflect ASHRAE 90.1-2019 prescriptive allowances**

### 3.3.13 Energy Code Adoption

In ComStock, the assumption is made that major building systems are replaced periodically over the lifespan of the building. Ideally, when these systems are replaced, the efficiency level of the new system meets or exceeds the requirements of the energy code in force, as determined by the location of the building and the time of replacement. In reality, however, some buildings will lag (not comply with the energy code) and other buildings will over-perform (lead the energy code).

#### Energy Code Adoption

The adoption history of energy codes in the United States is complex. Some states have a statewide code, while others have codes determined at a city or county level. For ComStock, the adoption of energy code is assumed to be a function of year and state. For states with no statewide code, to the best of our ability, the code covering the biggest cities was selected. Where a state code was not a derivative of the ASHRAE 90.1 series of codes, the most similar versions of ASHRAE 90.1 was used for that state. The exception to this is California, where the Title 24 series of codes, as represented in DEER (*Database for Energy Efficient Resources* 2021), was used because this series of codes was known to be significantly different from ASHRAE 90.1. Most of the information used to develop the code adoption history was taken from the Buildings Codes Assistance Project (*Code Status Maps: Commercial Energy Code Adoption*).

#### Energy Code Compliance

For this discussion, energy code compliance is defined as the extent to which a building constructed to comply with a certain energy code meets the requirements of that code. For example, a building built to comply with ASHRAE 90.1-2010 may meet all envelope requirements but fail to meet some HVAC control requirements. Unfortunately, there is little information available on commercial energy code compliance at a national level, and what information does exist is not detailed. The status of this information is described in a detailed report (*Residential and Commercial Sector Energy Code Adoption and Compliance Rates* 2017) generated for EIA's NEMS modeling effort. What we do know, both from this information and anecdotally, is that commercial energy compliance is imperfect. Some buildings exceed code and others lag behind.

In ComStock, a framework was put in place to allow for a distribution of buildings lagging, meeting, or leading the energy code on a building-system level. However, because of the data limitations, we assumed that all building systems met the requirements of the energy code that was in force in their location, both when built and as replaced

over time. In the future, we hope to move away from this code-compliance-based framework toward a model driven by distributions of known building characteristics, but if more data about code compliance do become available, these can be incorporated into ComStock.

### **3.3.14 Building System Turnover**

As noted above, in ComStock, the assumption is made that major building systems are replaced periodically over the lifespan of the building. The reason for these replacements could be equipment failure, building remodeling, energy efficiency upgrades, etc. To model the turnover of building systems, it is necessary to understand how often these building systems are replaced, which reflects how long they last in the building stock.

#### *Effective Useful Life*

The metric commonly used by the energy efficiency community to describe the lifespan of a measure is effective useful life (EUL). In *Energy Efficiency Policy Manual Version 6 For Post-2018 Programs* (2020), the California Public Utility Commission defines EUL as “An estimate of the median number of years that the measures installed under the program are still in place and operable.” In the reliability community (*Reliability terminology* 2021), EUL is typically referred to as “median time to failure,” whereas ASHRAE Abramson 2006 uses the term “median service life.”

For ComStock, the primary source of EULs is the California Public Utilities Commission (CPUC) Database of Energy Efficiency Resources (DEER) (*Database for Energy Efficient Resources* 2021). While previous work on EULs indicates that there is a wide range in the quality of EUL data found in various sources nationally, it also indicates that the studies performed in DEER are generally the best available (Skumatz 2012). The values in DEER were cross-referenced against the lifetimes used in the EIA NEMS Commercial Demand Module (*Commercial Demand Module of the National Energy Modeling System: Model Documentation* 2017) and the ASHRAE Service Life and Maintenance Cost Database (ASHRAE 2021). Table 21 shows the EULs assumed for different building systems in ComStock.

**Table 21. Effective Useful Life of Major Commercial Building Systems**

Major Building System	EUL (Years)	Notes
Envelope—Wall Insulation	200	This value was based on engineering judgement. DEER EULs are capped at 20 years per CPUC policy. NEMS does not appear to model wall turnover separate from whole-building replacement.
Envelope—Roof Insulation	200	This value was based on engineering judgement. DEER EULs are capped at 20 years per CPUC policy. NEMS does not appear to model roof turnover separate from whole-building replacement.
Envelope—Windows	70	Based on a reliability analysis of windows from the 2014 Commercial Building Stock Assessment from the Pacific Northwest. DEER uses an EUL of 20 years for window replacement as the DEER EULs are capped at 20 years per CPUC policy.
Exterior Lighting	15	This closely matches the highest EUL in DEER for outdoor lighting of 16 years. NEMS does not break out exterior lighting, but all NEMS commercial lighting technology types have a 10-year EUL.
Interior Lighting	10	This is in line with the EULs in DEER for interior lighting, and matches the 10-year EUL for all commercial lighting technologies in NEMS.
HVAC	20	The highest EUL in DEER for HVAC is 20 years. For rooftop air conditioners, which serve by far the largest portion of the building stock, NEMS uses a 21-year EUL. NEMS HVAC EULs range from 9.5 years for window AC units up to 30 years for some boilers. ASHRAE (2021) includes 33 packaged DX rooftop units with a mean lifetime of 21 years, and appears to be the source of some NEMS HVAC lifetimes.
Service Water Heating (SWH)	15	The highest SWH EUL in DEER is 20 years for a tankless water heater. Most tank-based SWH equipment in DEER has an EUL of 15 years or less. NEMS non-solar SWH equipment EULs range from 10 to 15 years. ASHRAE (2021) includes 5 gas-fired water heaters with a mean lifetime of 15 years, and 36 electric water heaters with a mean lifetime of 10 years.
Interior Equipment (Plug and Process Loads)	15	This value was based on engineering judgement and is meant to represent an average over all types of plug and process loads. If plug and process loads are addressed in future iterations, splitting the plug and process loads into information technology (IT) equipment and other equipment will be investigated, as IT equipment typically has a higher turnover rate than other process loads such as hospital equipment or commercial kitchen equipment. NEMS includes commercial kitchen equipment with an EUL of 12 years, commercial ice machines with an EUL of 8 years, commercial vending machines with an EUL of 13.5 years, and commercial refrigeration equipment with an EUL of 10 years.



### *Building Envelope*

For the building envelope (windows, wall insulation, and roof insulation), the DEER database was not informative because the maximum EUL is capped at 20 years per CPUC policy. Based on this, other sources of envelope lifetime information were sought out.

### **Windows**

As part of the DOE-funded Advanced Building Construction initiative, a team collected information on windows from a variety of commercial building surveys, including a new survey of buildings built since roughly 2010. Unfortunately, because windows are long-lived, while most of the surveys did ask about windows, only one survey had enough information to perform a reliability analysis. This survey was the 2014 Commercial Building Stock Analysis (Navigant 2014), which covers the Pacific Northwest. This survey included information on the age of the building, whether the windows had ever been replaced, an estimate of the year replacement if replaced, and a weighting factor to describe how each sample fits into the whole building population. From these data, a reliability analysis was performed. Figure 143 shows the estimated survival curve. As indicated by the black 'X' on the figure, the EUL estimate for windows is 70 years. Also of note is that the maximum lifespan extends to more than 400 years. In practice, what this indicates is that windows on some buildings will never be replaced, lasting the entire lifespan of the building.



**Figure 143. Reliability analysis for windows in commercial buildings, from 2014 CBSA (Navigant 2014)**

### **Walls and Roofs**

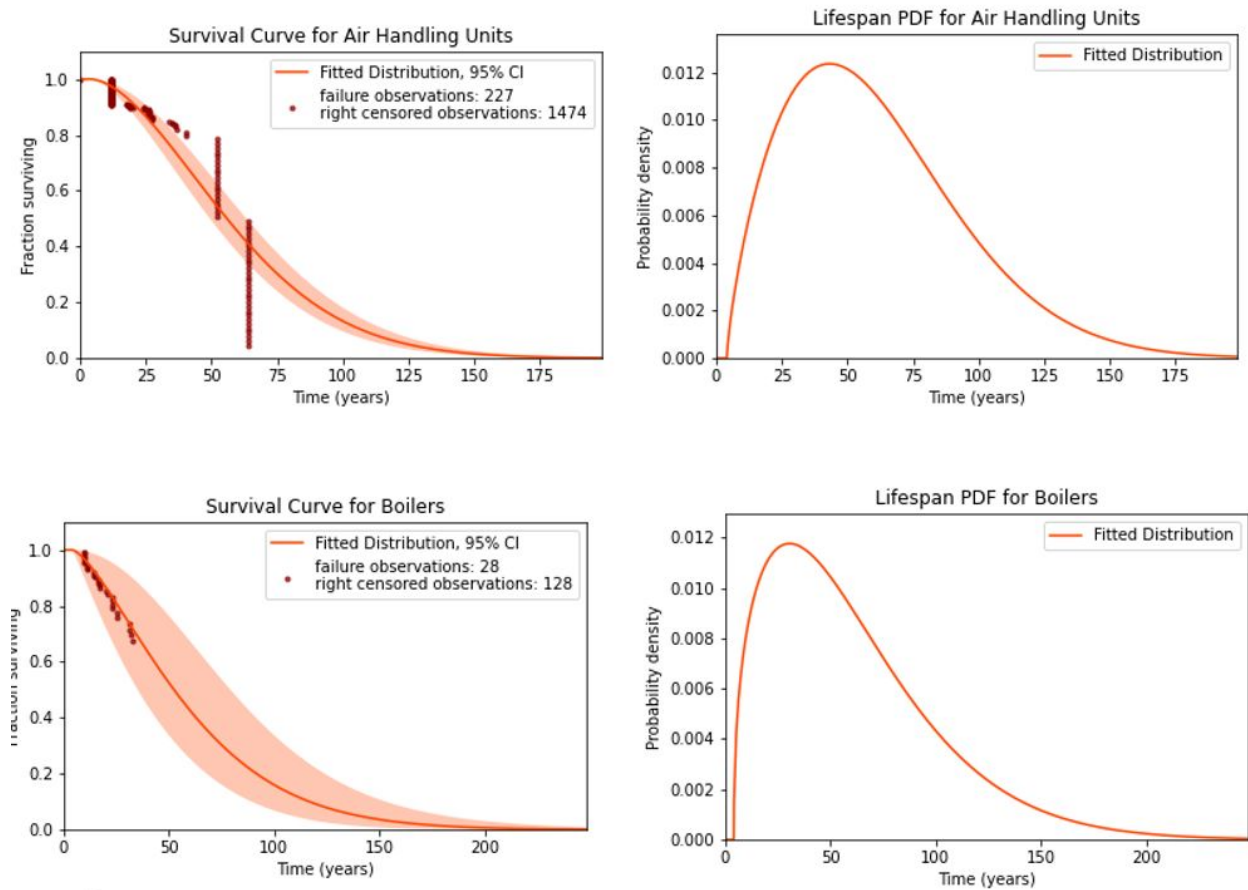
None of the data sources identified included information on EULs for walls and roofs, or more specifically the insulation on these surfaces. DEER EULs are capped at 20 years per CPUC policy. NEMS does not appear to model wall turnover separate from whole-building replacement. Based on engineering judgement, an EUL of 200 years was selected to indicate that for most buildings, the wall and roof insulation will not be replaced before the building is demolished.

### *Distribution of Lifespans*

The EUL estimates above represent the median lifespan for a given building system. However, not all systems will fail and be replaced at exactly that length of time; some will be replaced earlier, and others will be replaced later. To represent this diversity of failure rates, a distribution is used.

The simplest approach would be to use a normal distribution centered on the EUL. However, studies of reliability data show that this is not a good assumption, and often use a Weibull distribution to represent lifetimes. To check whether a Weibull distribution appeared to be an appropriate form for representing the lifetimes of building equipment, a reliability analysis was performed on data from the ASHRAE Service Life and Maintenance Cost Database (ASHRAE 2021). This analysis was performed following the methodology described in an ASHRAE journal article

(Hiller 2000), and was implemented using the Reliability package (Reid 2020) in Python. Four categories of equipment with a reasonable number of entries were investigated: Air Handling Units, Boilers, Chillers, and Air Source DX Equipment (all types of each available in the database).



**Figure 144. Survival curves and derived lifespan probability density functions for commercial HVAC equipment**

As shown in Figure 143, Weibull distributions are a good fit for several categories of HVAC equipment failure data. Although the ASHRAE database includes data for many different types of HVAC equipment, it was not selected as the primary source for deriving EULs for ComStock due to the limitations and biases in the database described by its creators (Abramson 2006). Instead, the decision was made to use the EUL sources described in the previous section and develop Weibull curve parameters around these EULs. The selected parameters are shown in Table 22. For the 70-year EUL, the parameters came from the window reliability analysis. For the 10, 15, and 20-year EULs, the only constraint during was to match the EUL definition: 50% of the equipment would still be operable at the EUL. A minimum lifespan of 60% of the EUL was selected with the assumption that although individual components of a system might fail, it is unlikely that products exist on the market that routinely fail at a whole-building scale in only a few years. The 200-year EUL parameters were selected to represent no failure for the life of the building.

**Table 22. Commercial Equipment Lifetime Weibull Distribution Parameters**

EUL	Shape (beta)	Scale (alpha)	Shift (gamma)
10	1.6	EUL/2 = 5	6/6/00
15	1.6	2/7/05	6/9/00
20	1.6	EUL/2 = 10	6/12/00
70	1.3	91	0
200	1	1	200

### 3.3.15 Heating and Service Water Heating Fuel Determination

The distribution of space heating and service water heating fuels—particularly the difference between using electricity and direct combustion of fossil-fuels for space heating—can have a large impact on the energy consumption of the load profiles of the buildings stock.

The CBECS 2012 microdata (EIA 2012) has information on whether each building is heated and what heating fuel(s) are used. This dataset has a geographic resolution at the census division level, and each census division includes several states. Observations of several data sources, including maps of natural gas utility service territories, utility load research data from various subregions of ERCOT (Texas), and the residential American Community Survey (ACS; Ruggles et al. 2021) show that within these census divisions, the geographic distribution is not uniform; natural gas is more prevalent in urban areas than rural areas, and certain areas use more fuel oil and propane for historical reasons. A process to capture this geographic diversity was proposed based on the hypothesis that the distribution of fuel availability, particularly of natural gas, would be similar between commercial and residential buildings in a given region. To test this hypothesis, residential data for single-family detached homes from ACS was aggregated to the census division and compared to the distributions from CBECS. As shown in Figure 145, this hypothesis was broadly correct in regard to the prevalence electricity and natural gas.

However, some residential fuel saturation levels, particularly the natural gas saturation in the Pacific division and the fuel oil saturation in the Middle Atlantic and New England divisions, were deemed too different from CBECS to be used directly. Instead, a decision was made to preserve the geographic granularity of the residential data at a county level for the distribution of electricity, natural gas, fuel oil, and propane, while uniformly scaling the distributions for all counties within each census division to match the distributions of fuel types in CBECS for these fuels. Because the residential data does not include information on district heating or unheated buildings, the census-division level distributions for these two “fuels” were used for all counties in a given census division.

The result is a geographically diverse fuel breakdown that matches CBECS. Figures 146–151 show how the saturation levels vary by county. Note that the color scales on each figure are different, and that the total saturation for a given county sums to 1 across all fuel types.

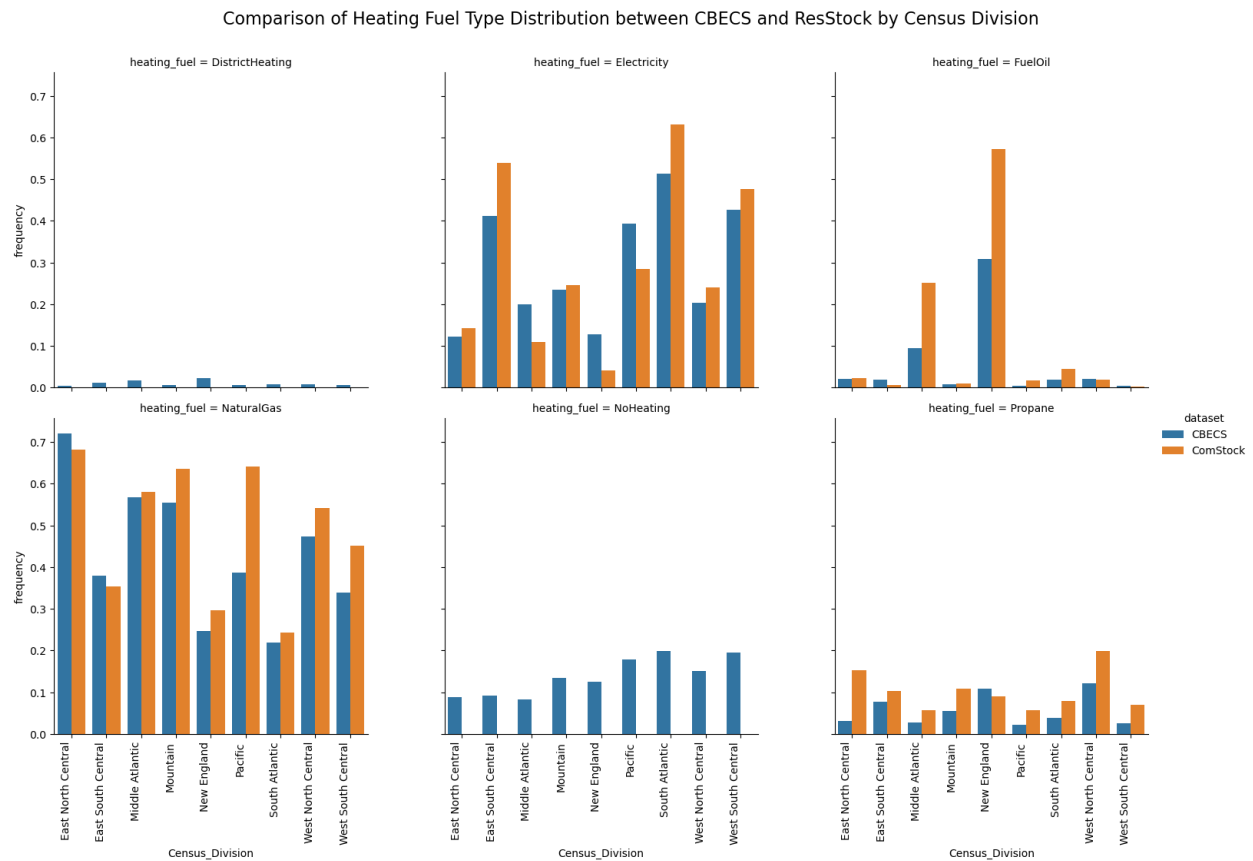
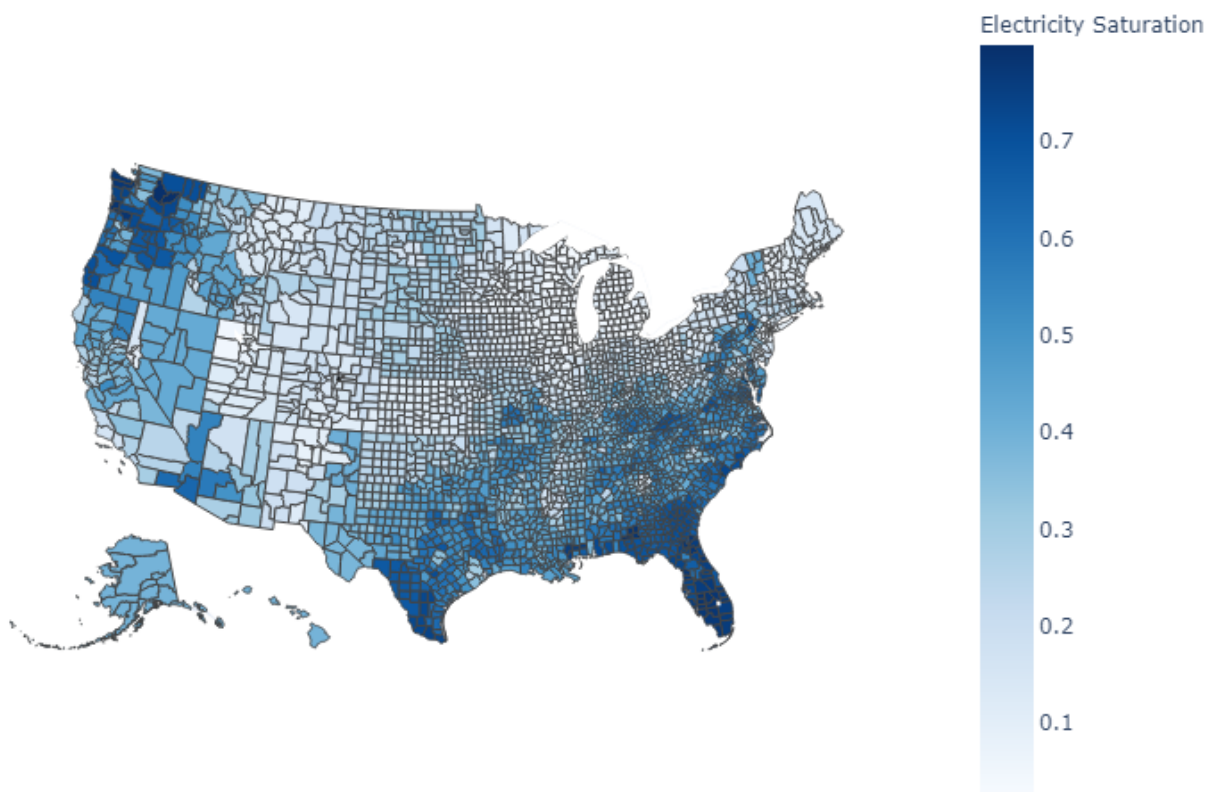


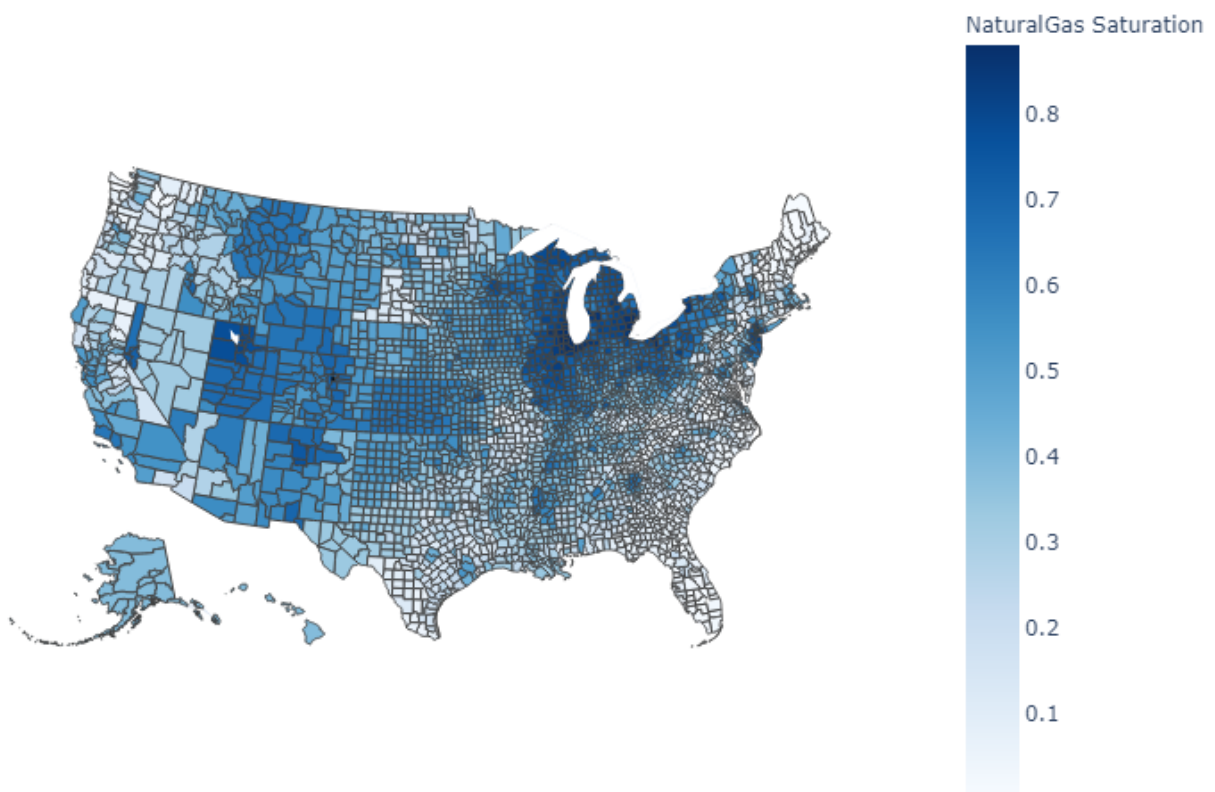
Figure 145. Comparison of heating fuel distribution between CBECS and ACS by census division

### Electricity Heating Fuel by County



**Figure 146.** Map of the revised saturation of natural gas heating in commercial buildings, by county

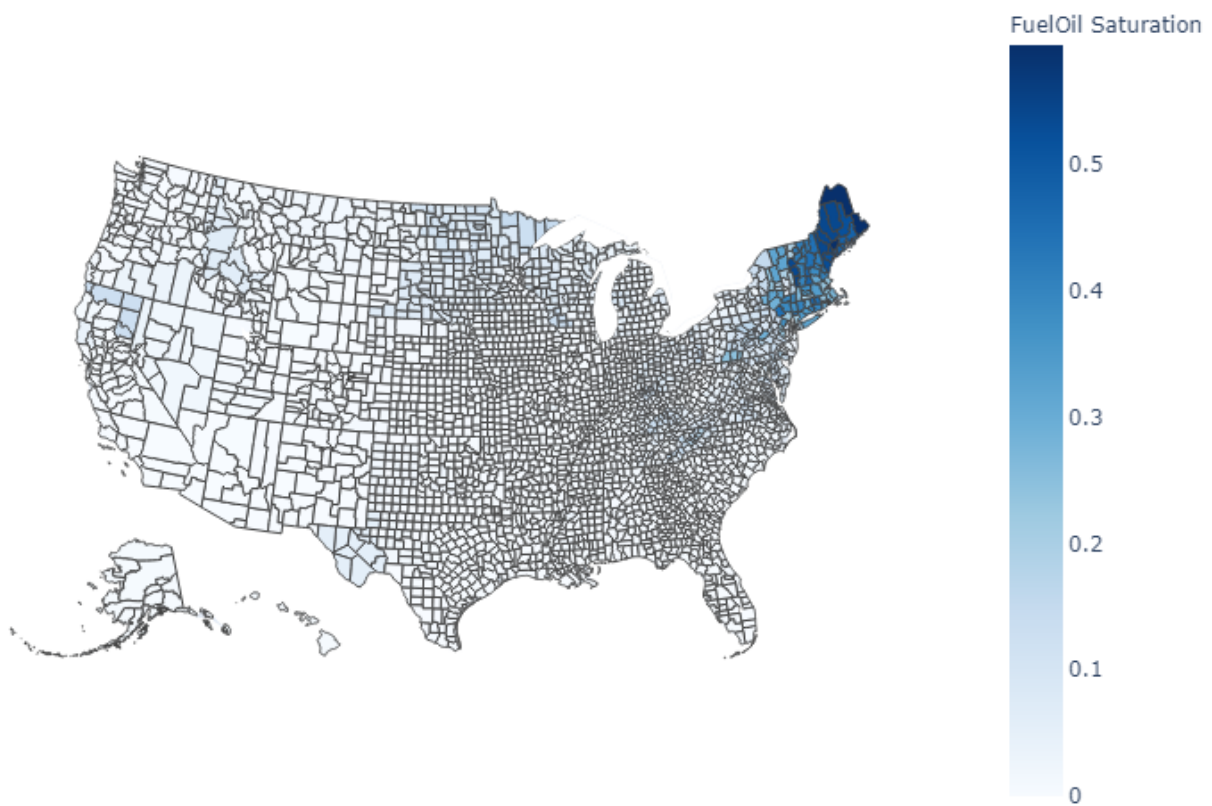
### NaturalGas Heating Fuel by County



**Figure 147. Map of the revised saturation of electric heating in commercial buildings, by county**

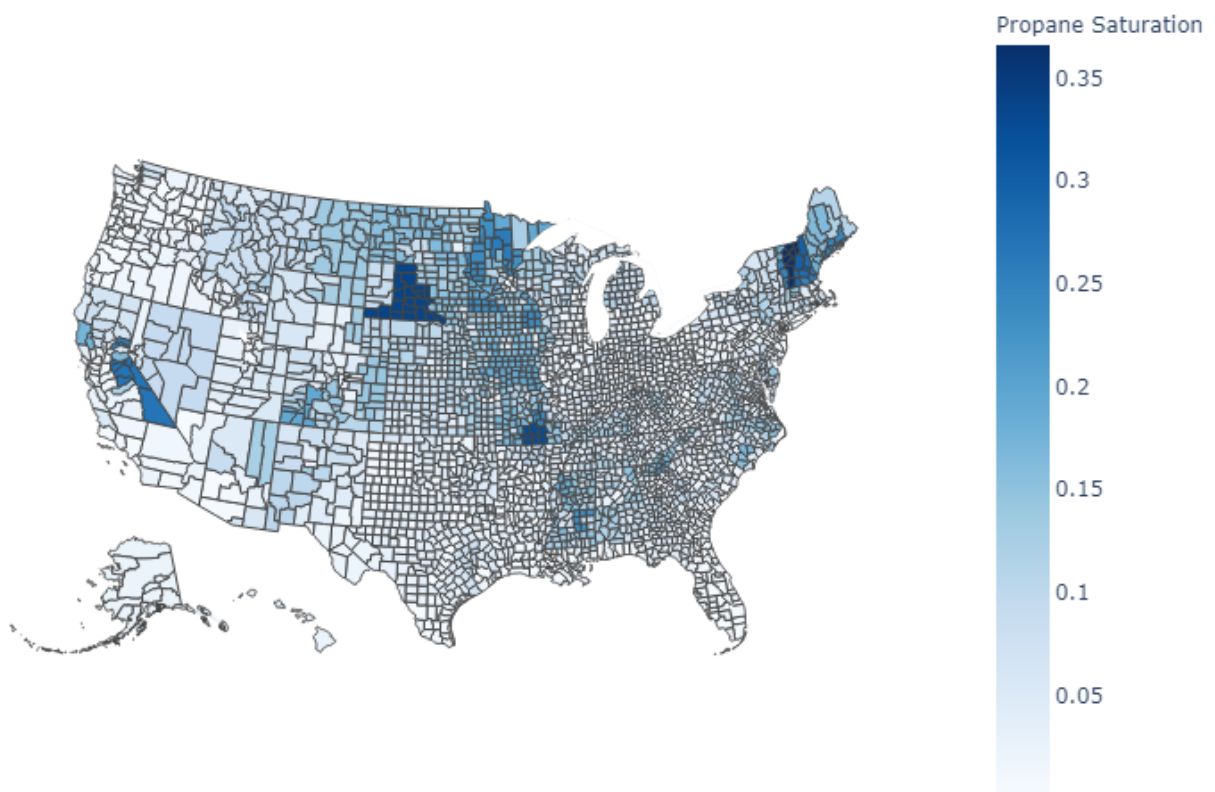


### FuelOil Heating Fuel by County



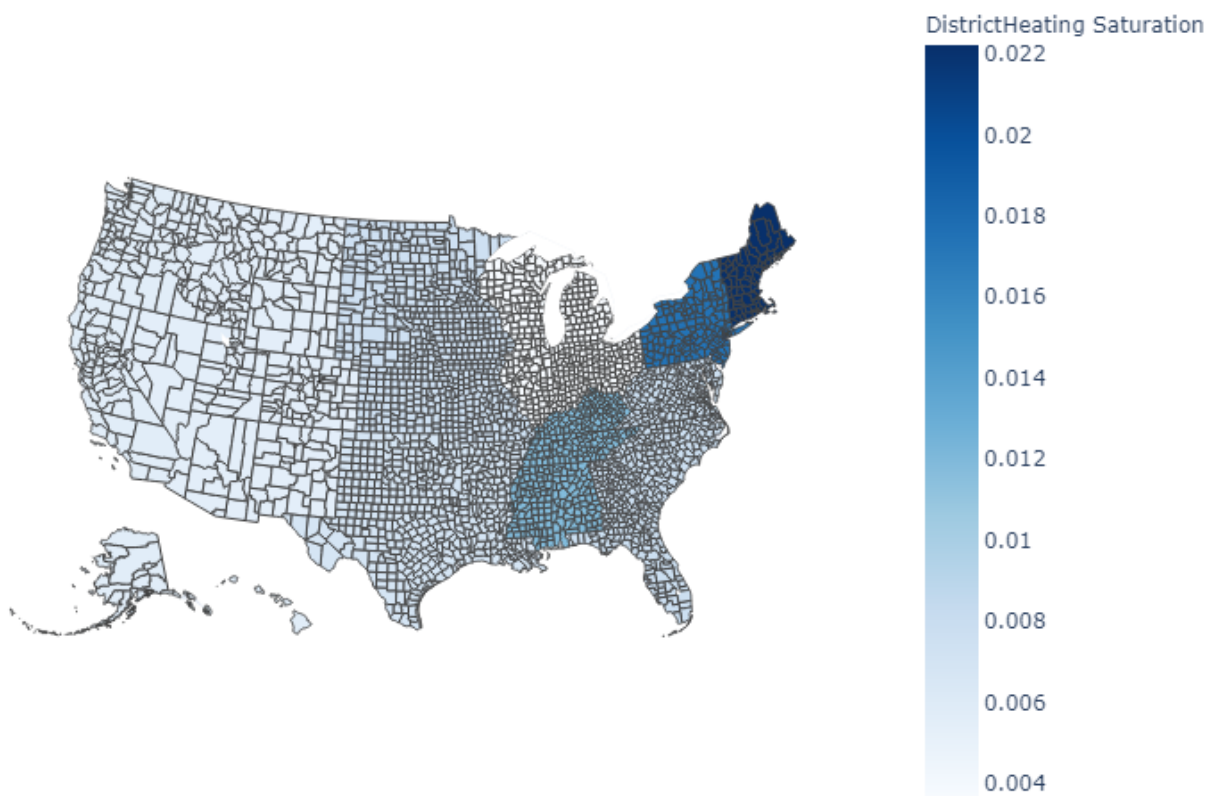
**Figure 148. Map of the revised saturation of fuel oil heating in commercial buildings, by county**

### Propane Heating Fuel by County



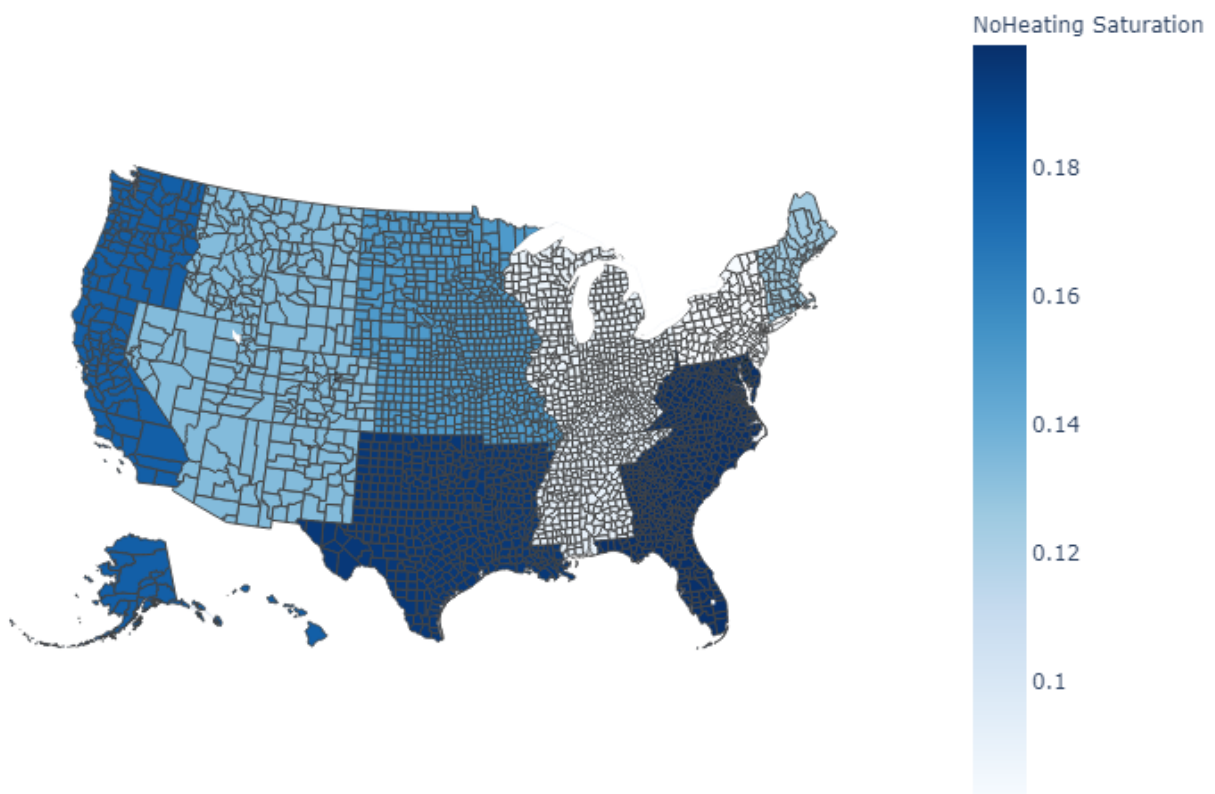
**Figure 149. Map of the revised saturation of propane heating in commercial buildings, by county**

### DistrictHeating Heating Fuel by County



**Figure 150. Map of the revised saturation of district heating in commercial buildings, by county**

### NoHeating Heating Fuel by County

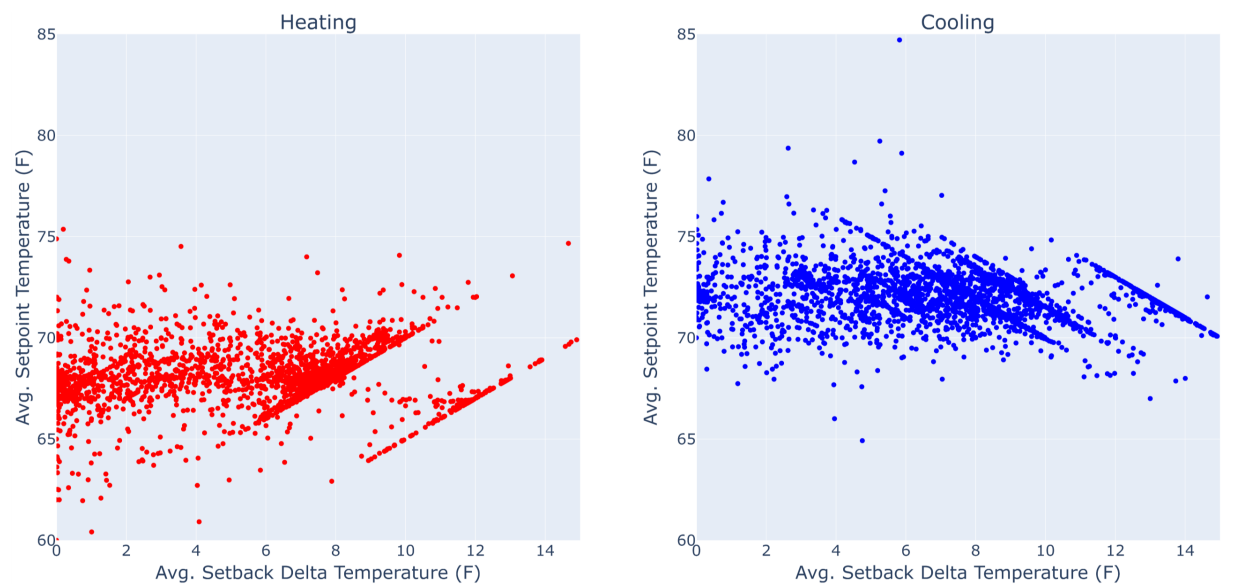


**Figure 151. Map of the revised saturation of unheated commercial buildings by county**

### 3.3.16 Thermostat Variability

Section 3.3.3 introduces our methodology for implementing data-informed thermostat schedules. However, this method applies identical thermostat setpoints and setbacks to every building based on building type. In reality, thermostat setpoints and setbacks will vary between buildings to some extent, and this variability can provide a more realistic distribution of energy behavior in ComStock.

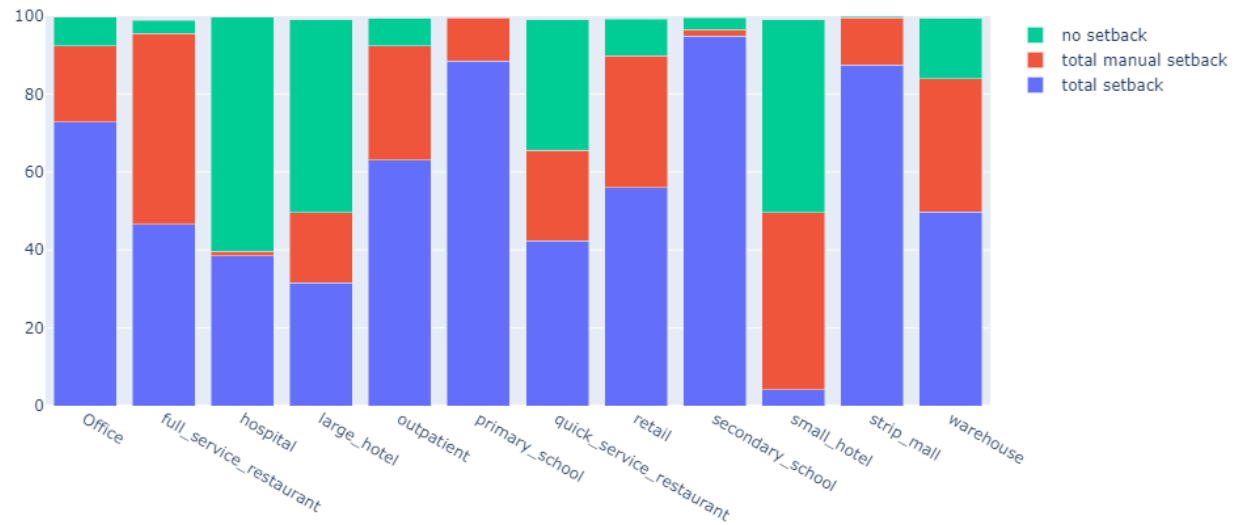
The BAS data sources used in Section 3.3.3 were used for this effort, and are summarized by building type in Table 16. The goal was to determine probability distributions, per building type, for heating setpoint, heating setback, cooling setpoint, and cooling setback. Figure 152 illustrates the relationship between the different heating and cooling setpoints, and their relative setbacks, across all building types in the BAS dataset used to create these setpoint-setback distributions. For building types with fewer than 25 samples, we used the overall distribution across all building types in the data. Note that healthcare and hotel building types were not modified as part of this effort due to their typical 24/7 operation schedules and unique indoor air requirements.



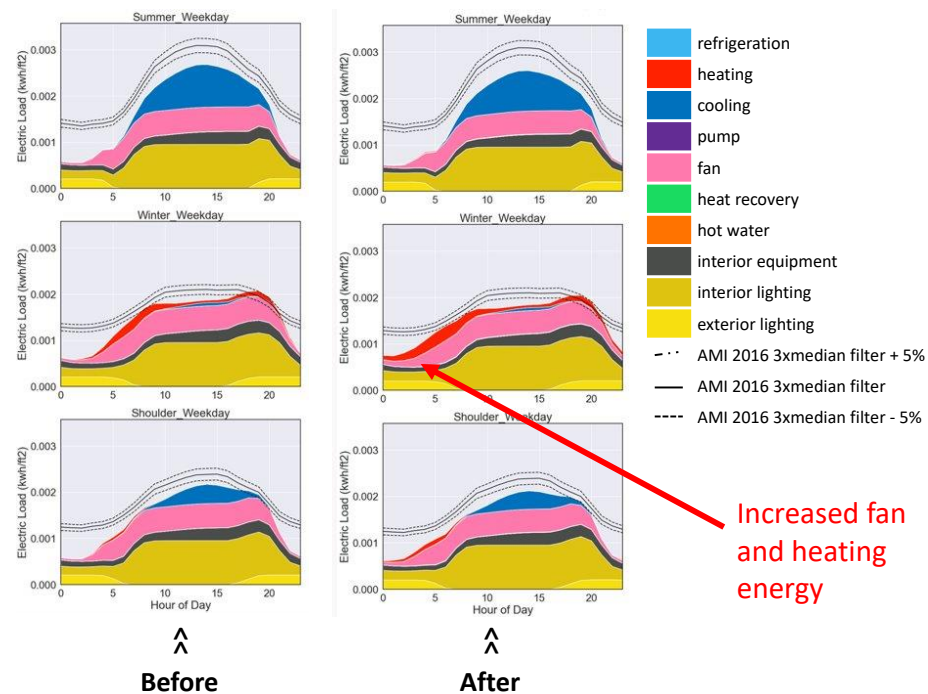
**Figure 152. Average heating and cooling thermostat setpoint-setback correlation from BAS datasources**

Due to the smaller sample sizes for many building types in the BAS dataset, we used CBECS 2012 microdata (EIA 2012), which has much larger sample sizes across building types, to help inform this effort. CBECS does not provide data on setpoint or setback temperatures, but it does provide survey responses as to whether heating and cooling setbacks are used, and if these setbacks are manual or automatic. Figure 153 illustrates these percentage breakdowns by building type. For the purposes of energy consumption, it does not matter if a setback is manual or automatic. However, we are skeptical that all CBECS survey respondents who claim to implement manual nighttime setbacks actually do so regularly. With no additional data sources to investigate this question, we made the conservative assumption that only 20% of buildings that claim to use manual thermostat setbacks actually do so on a nightly basis.

For ComStock implementation, the heating and cooling setpoint temperature probability distributions were generated by building type using the BAS data sources described in this section. The setback temperature deltas also used these same data sources; however, the probability distribution of whether a thermostat setback exists, which would be equivalent to a thermostat setback of 0°F, was derived using the CBECS data. An OpenStudio measure was developed to implement these setpoints and setbacks in the models at the frequencies informed by the probability distributions. Figure 154 shows the impact of this change for retail buildings in calibration Region 1. A small increase in fan and heating energy can be seen due to the added prevalence of buildings without thermostat setbacks. Figure 155 illustrates how adding thermostat variability caused the ComStock distribution of electric EUIs to flatten slightly.

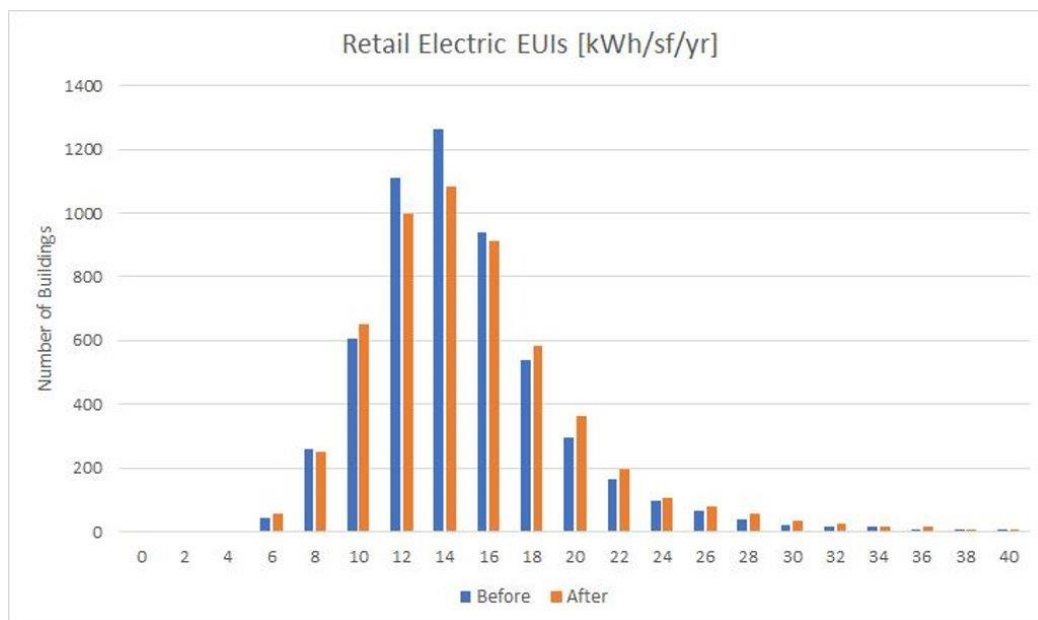


**Figure 153. Average percentage of building stock utilizing heating and cooling thermostat setbacks, manual or automatic, from weighted CBECS**



**Figure 154. Impact of thermostat variability on retail buildings in calibration Region 1—Fort Collins, CO. An increase in heating and fan energy can be observed as a result of the added prevalence of buildings without thermostat setbacks.**



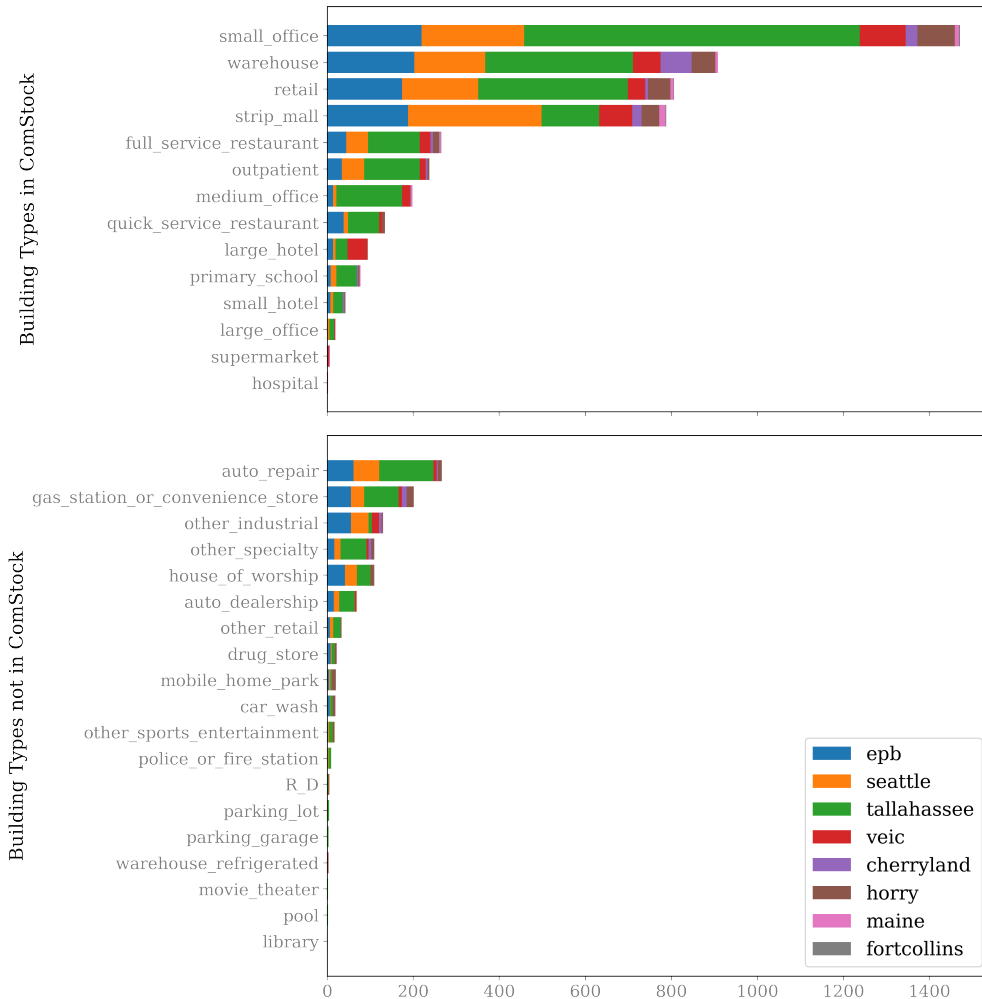


**Figure 155. Impact of thermostat variability on the retail building electric EUI distribution in calibration Region 1—Fort Collins, CO. The distribution becomes flatter as a result of this added variability.**

### 3.3.17 Hours of Operation Schedules

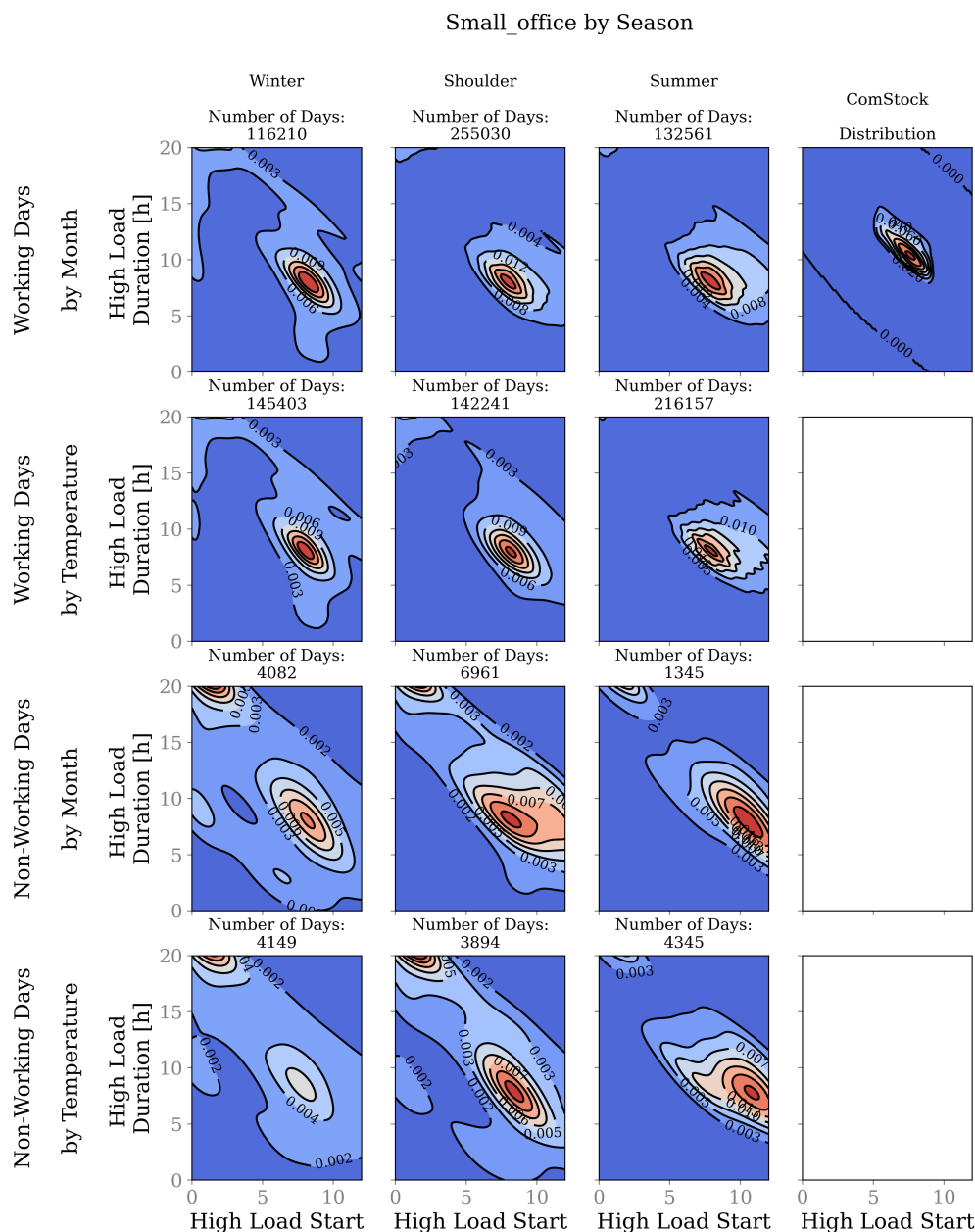
In this section, we applied the method introduced in Bianchi et al. (2020) to analyze 1-year of AMI data from 6,070 buildings spread across eight utilities (the “commercial schedules AMI dataset”). We first extracted the two-dimensional distribution of *High Load Start Time* and *High Load Duration* (as introduced in Section 2.3.10 and defined in Table 11) from this AMI dataset, as an approximation of the schedule of hours of operations for each building type, and then compared this distribution with the inputs of ComStock at the start of EULP calibration.

Figure 156 lists the number of buildings for each building type from each utility. Among the 15 building types considered in ComStock, 14 can be found in the commercial schedules AMI dataset. The only exception is secondary schools, because all schools were grouped together in the AMI data.



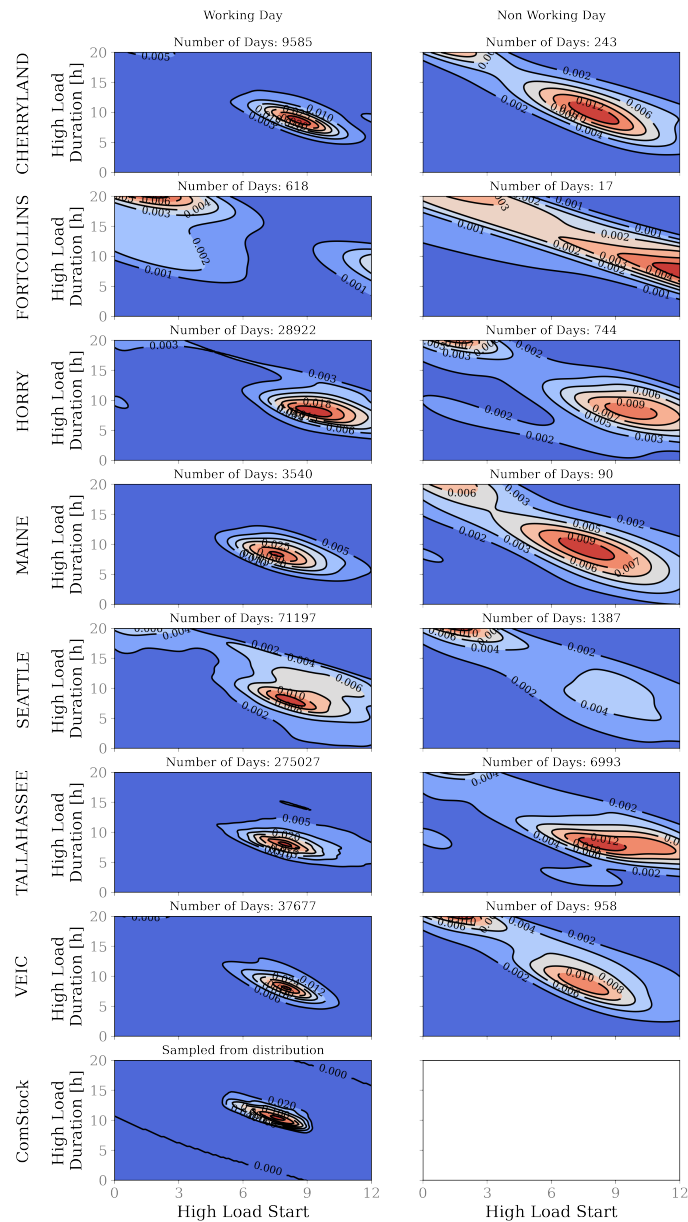
**Figure 156. Number of samples by building type and utility in the commercial AMI dataset used to derive hours of operation schedules**

We compared the distribution extracted from the commercial schedules AMI dataset with the inputs of ComStock at the start of EULP calibration. The results of small office are presented to illustrate the process, because it has the largest sample size in the AMI dataset. The result is shown in Figure 157. We considered two day types: working day vs. non-working day, and two season definitions: defined by month and defined by daily average outdoor air temperature. The distribution of hours of operations is more diffuse in the AMI dataset compared with the inputs to ComStock at the start of EULP calibration. Also, the duration of high load is smaller in real AMI data than in the previous ComStock assumptions.



**Figure 157. Distribution of small office hours of operations extracted from AMI data (from seven utilities), by day type and season, and compared to ComStock before updates. This figure shows how the hours of operation (start time and duration of the high load period) are influenced by season (all utilities are combined in this plot).**

We explored whether and how the hours of operation are influenced by season in Figures 157 and by utility in 158. Some differences can be observed but considering the modeling complexity, the desire to create a nationally applicable approach, and the desire to avoid overfitting the national model to a specific utility region, we combined the AMI data across seasons and utilities together to generate a distribution of hours of operation for each building type. These new distributions were applied to ComStock in place of the existing distributions.



**Figure 158. Distribution of small office hours of operations extracted from AMI data (from seven utilities), by utility and day type, and compared to ComStock before updates. This figure shows how the hours of operation (start time and duration of the high load period) are influenced by utility region (all seasons are combined in this plot).**

## 4 Results

### 4.1 Final Calibration and Validation Results—Residential

#### 4.1.1 Residential EIA Data Validation

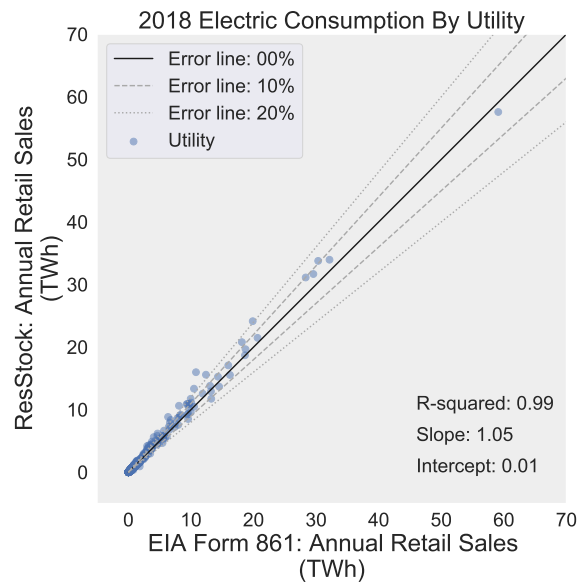
The data from EIA used in the residential sector validation is from (1) EIA-861 (EIA 2021b), (2) EIA-861M (EIA 2021c), and (3) EIA-176 (EIA 2021e), and (4) RECS (EIA 2021f). These data sources are described in Section 2.3.6. The first three surveys allow for ResStock to be compared to utility annual retail sales, monthly state electric sales, and monthly state natural gas energy. RECS 2009 and RECS 2015 report modeled end-use breakouts of collected utility bill data, along with housing metadata. The RECS data can be compared to ResStock for various housing stock cohorts.

##### *Residential EIA-861/176*

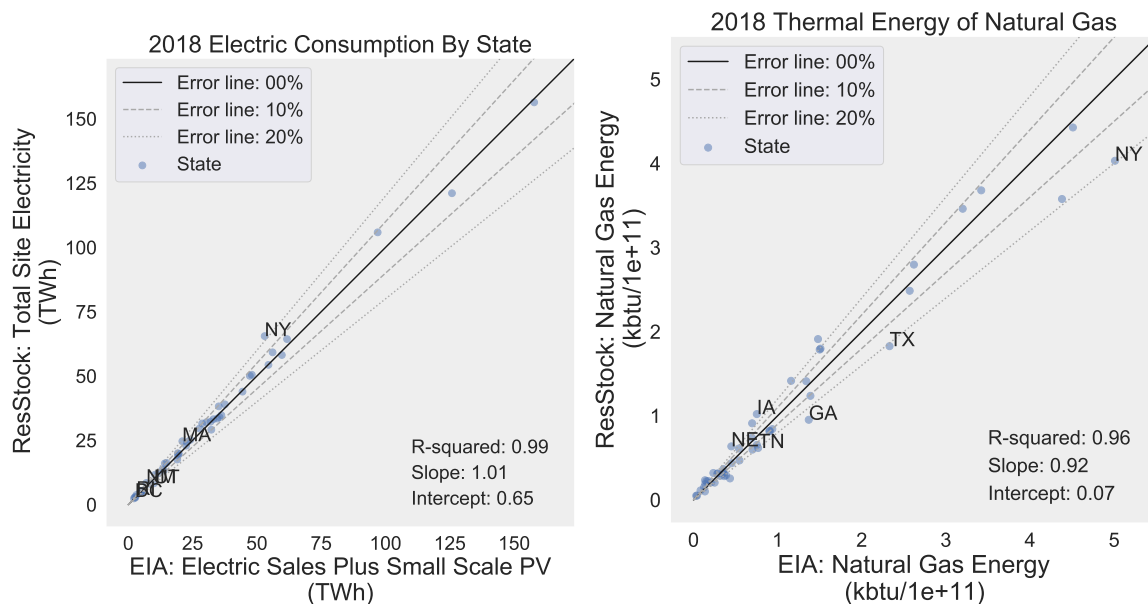
Figure 159 shows a comparison between ResStock (including output correction, Section 3.2.10) and the annual retail sales in EIA-861 in 2018. This comparison includes 1,769 utilities of the 2,125 utilities in EIA-861, 2018. ResStock uses the reported service territory (counties served) to map results from ResStock to utilities. The utilities missing in the comparison do not have service territories reported in EIA-861, 2018. ResStock shows good agreement with the utility level annual retail sales with a regression slope of 1.05, an intercept of 0.01 (TWh) and an  $R^2$  of 0.99. EIA provides a state annual retail sales adjustment for non-respondents and for missing PV generation that ranges between -0.4% and 9.6% of the retail sales for a given state. This adjustment could indicate some errors in reporting by some utilities and result in some of the scatter seen in Figure 159. As discussed in Section 5.1.2, the calibration model improvements resulted in a significant improvement in this comparison, both a reduction in the overestimation (slope) and reduction in scatter ( $R^2$ ).

The same data at the utility level was aggregated to the state level to compare ResStock state retail sales to EIA-861 (Figure 160). ResStock outputs show good agreement between the state wide annual retail electric sales, with a slope of 1.01, an intercept of 0.65 (TWh), and an  $R^2$  of 0.99. As discussed in Section 5.1.2, the calibration model improvements resulted in a significant reduction in scatter ( $R^2$ ) and a minor improvement in slope for this comparison.

Also shown in Figure 160 is a comparison between ResStock and natural gas energy by state from EIA-176. The data in EIA-176 was converted from delivered volume of natural gas to natural gas energy using natural gas heat content factors provided by EIA (EIA 2021d). Calibration to natural gas data was less of a focus of the calibration effort, and as a result the more scatter can be seen in the figure. However, most states are still within 20% of the reported natural gas energy from EIA. The natural gas annual energy scatterplot regression has a slope of 0.92, an intercept of 7e9 kBtu, and an  $R^2$  of 0.96. As discussed in Section 5.1.2, the calibration model improvements resulted in a significant improvement in this comparison, reducing the underestimation (slope), although there is still a fair degree of scatter ( $R^2$ ) with more than 20% error for several states.



**Figure 159. ResStock modeled annual retail sales compared to 2018 EIA-861 annual retail sales for approximately 1,700 utilities (including heating and cooling correction based on EIA-861M; see Section 3.2.10)**



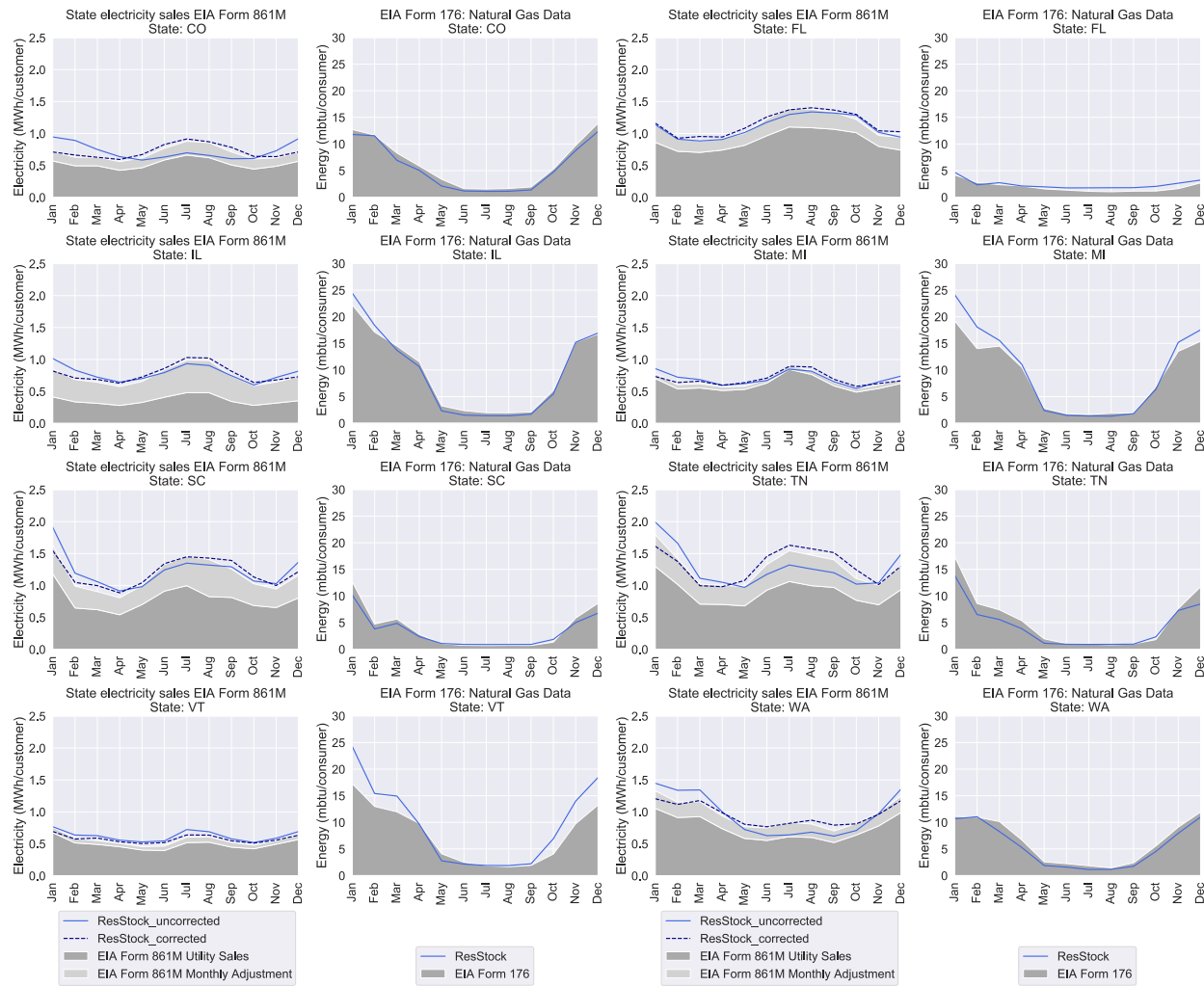
**Figure 160. ResStock modeled annual total site electricity (including heating and cooling correction based on EIA-861M; see Section 3.2.10) compared to 2018 EIA-861 state electric sales plus small scale solar (left), and ResStock modeled natural gas energy compared to 2018 EIA-176 state natural gas energy (right)**

### Residential EIA-861M

While EIA-861 is an annual survey, EIA-861M is monthly survey of a subset of utilities. EIA imputes monthly sales, revenue, and customer counts for the utilities that are not surveyed, which are provided as state-level monthly adjustment factors for the state total sales, number of customers, and revenue. Monthly state-level natural gas residential sales is provided by EIA-176. Figure 161 shows the monthly data from EIA-861 and EIA-176 for the eight states where residential AMI data were obtained for this project. Also shown is the ResStock estimated natural gas energy



## End-Use Load Profiles for the U.S. Building Stock: Methodology and Results of Model Calibration, Validation, and Uncertainty Quantification



**Figure 161. ResStock monthly electric sales with/without correction and natural gas energy compared to 2018 electric sales and natural gas data reported in EIA-861M and EIA-176 for the states that are part of the residential calibration regions. For a complete set of state monthly comparisons to EIA-861M and EIA-176 see Appendix F.**

and the state retail sales before and after the correction model is applied (Section 3.2.10). For a complete set of state monthly comparisons to EIA-861M and EIA-176, see Appendix F.

For the states shown in Figure 161, ResStock is able to accurately estimate the low level of natural gas monthly energy during the summer. During these warmer months the natural gas energy is expected to be mostly water heating, with some gas for cooking, clothes dryers, and pool heating. The largest summer discrepancy in the states shown is in FL. The colder months have higher natural gas heating energy consumption. ResStock is overestimating the natural gas energy during colder months for IL, MI, and VT. However, the regression slope of 0.92 in Figure 160 suggests that across all states, ResStock slightly underestimates annual natural gas energy.

The monthly electricity retail sales comparisons show that the corrected version of ResStock performs well compared to the EIA-861M data. The correction model only adjusts the HVAC related end-use loads. Generally, the correction model for ResStock increases the cooling loads and decreases the heating load in most states. Some states see more correction than others. Of the states shown in Figure 161, CO, SC, TN, and WA see the largest correction from the ResStock correction model.

## EIA RECS

The RECS surveys from 2009 (EIA 2013) and 2015 (EIA 2018b) collected utility bills of individual households and characteristics about these households. The utility bills are then disaggregated to end uses through various end-use models (EIA 2018a). These annual end-use loads are compared to 2018 ResStock results (including heating and cooling correction based on EIA-861M; see Section 3.2.10). More end uses are included in RECS 2015 compared to RECS 2009. The method of disaggregating end uses changed between 2009 and 2015 RECS. Since the end uses are modeled in both RECS and ResStock, we expect a certain amount of scatter in the visualizations and look for general trends rather than quantifying exact errors. However, the unique advantage of using RECS data for the end-use energy comparisons is the ability to slice the end-use energy by various housing characteristics. RECS surveys only occupied dwelling units and ResStock models both occupied and unoccupied dwelling units, so the results from ResStock are downselected to only occupied dwelling units.

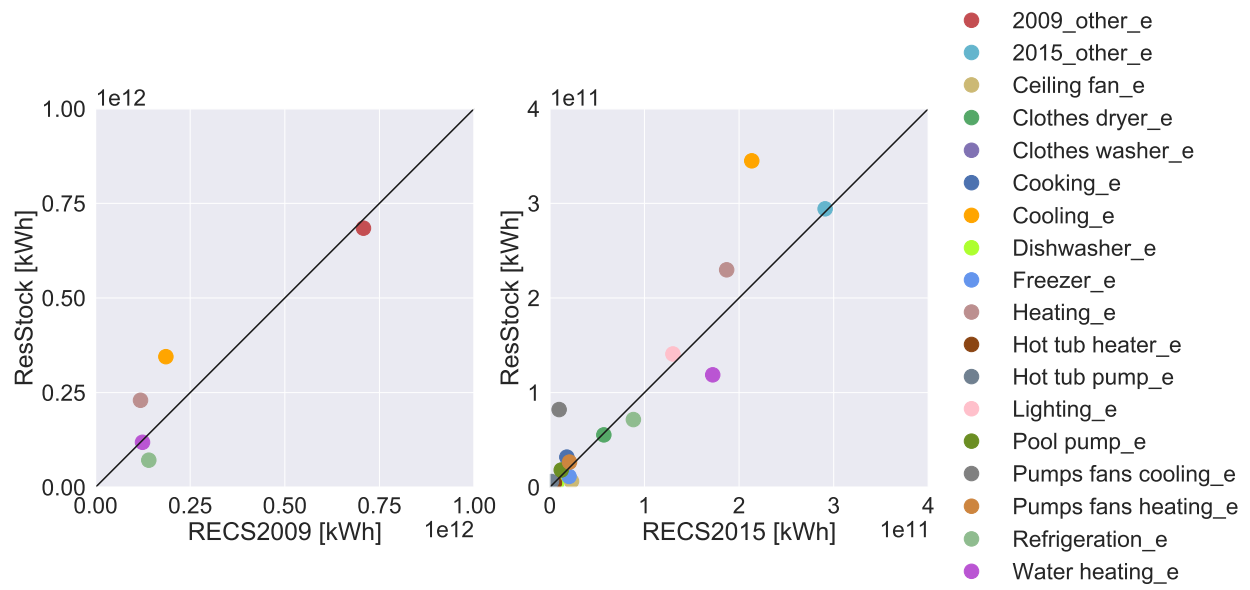
The comparisons to 2009 and 2015 RECS data performed in this project are electric end-use totals of (1) the national stock (Figure 162), (2) the national stock by building type (Figure 163)<sup>15</sup>, and (3) each Building America climate zone (Figure 164). Looking at all of these figures a few common trends can be identified. Across many of the end uses ResStock annual end-use totals show relatively good agreement to the RECS survey data; however, there are two exceptions.

Cooling appears to be consistently overestimated in 2018 ResStock (corrected) compared to RECS 2015 and 2009. 2018 was a warmer year than both 2015 (6% fewer cooling degree days) and 2009 (20% fewer cooling degree days) (NOAA 2021), so the observed difference may be due to this difference in weather. Weather differences aside, the cooling overestimation is consistent with the findings from comparison to HEMS (Figure 176), but is contrary to the comparison of ResStock to EIA-861M, which drives the heating and cooling correction model. While comparisons to AMI and LRD datasets in Figures 165, 166, and 171 are seasonal averages and therefore are approximate, they show overestimation of cooling in some regions (Fort Collins, Seattle, Chattanooga, Tallahassee, PG&E, SCE), underestimation in others (ComEd, Horry, Cherryland, AEP, OhioEd, ToledoEd), and approximate magnitude agreement in others (VEIC, Ameren Missouri, Appalachian Power, BGE, MetEd, PECO, ERCOT). When looking at the EIA-861M state monthly electric retail sales (Figure 161), ResStock (corrected) estimates the summer load relatively well, and in fact tended to underestimate summer electricity before the heating/cooling output correction model was applied.

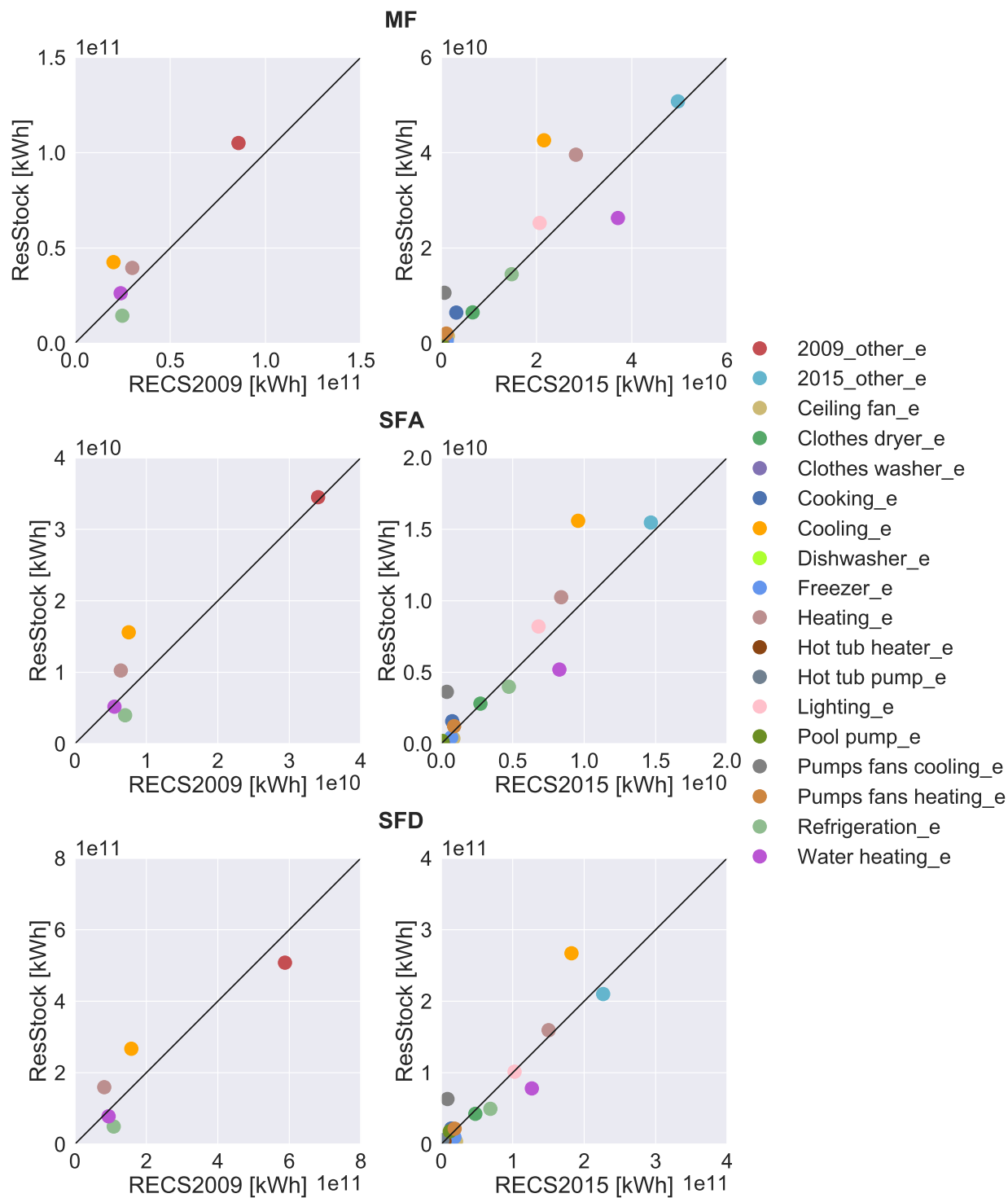
Aside from the weather differences, one other possible explanation for overestimating cooling is that the correction model is incorrectly increasing cooling energy when it is some other end use that should be increased in the summer. However, the fact that the EIA-861M errors are so well correlated to degree days makes us doubt that this would be the case, although it is possible that non-AC cooling alternatives, such as fans or evaporative coolers, that are not categorized as air conditioner cooling by RECS, are part of the explanation. Another possible explanation is that the cooling load fractions between occupied and vacant dwelling units is incorrect. ResStock could be over approximating cooling energy use in occupied homes and underestimating cooling energy use in vacant homes. Part of this discrepancy could be due to seasonally occupied dwelling units. It is unclear if RECS included some surveys sent to homes that are seasonally occupied (resulting in lower utility bills), and ResStock does not model seasonally occupied homes.

Electric water heating appears to be underestimated compared to RECS 2015, though it estimates well compared to RECS 2009. A dwelling unit's water heater fuel type is closely correlated with its space heating fuel type. ResStock uses RECS 2009 to describe this relationship because RECS 2015 does not include enough samples to query with the necessary geographic granularity. The geographic granularity of the data we queried from RECS 2009 was multiple states (N=10 "custom regions" that are aggregations of RECS reportable domains). The dwelling unit's main heating fuel type is based on PUMS data which has PUMA geographic resolution (N=2,400). Because of the geographic granularity discrepancy between these two data sources we could be under counting the number of electric water heaters. There could also be a different share of electric water heating in the homes sampled in RECS 2015 compared with RECS 2009. Finally, the end-use energy disaggregation method used for RECS changed between 2009 and 2015, which could contribute to the differences between the two years.

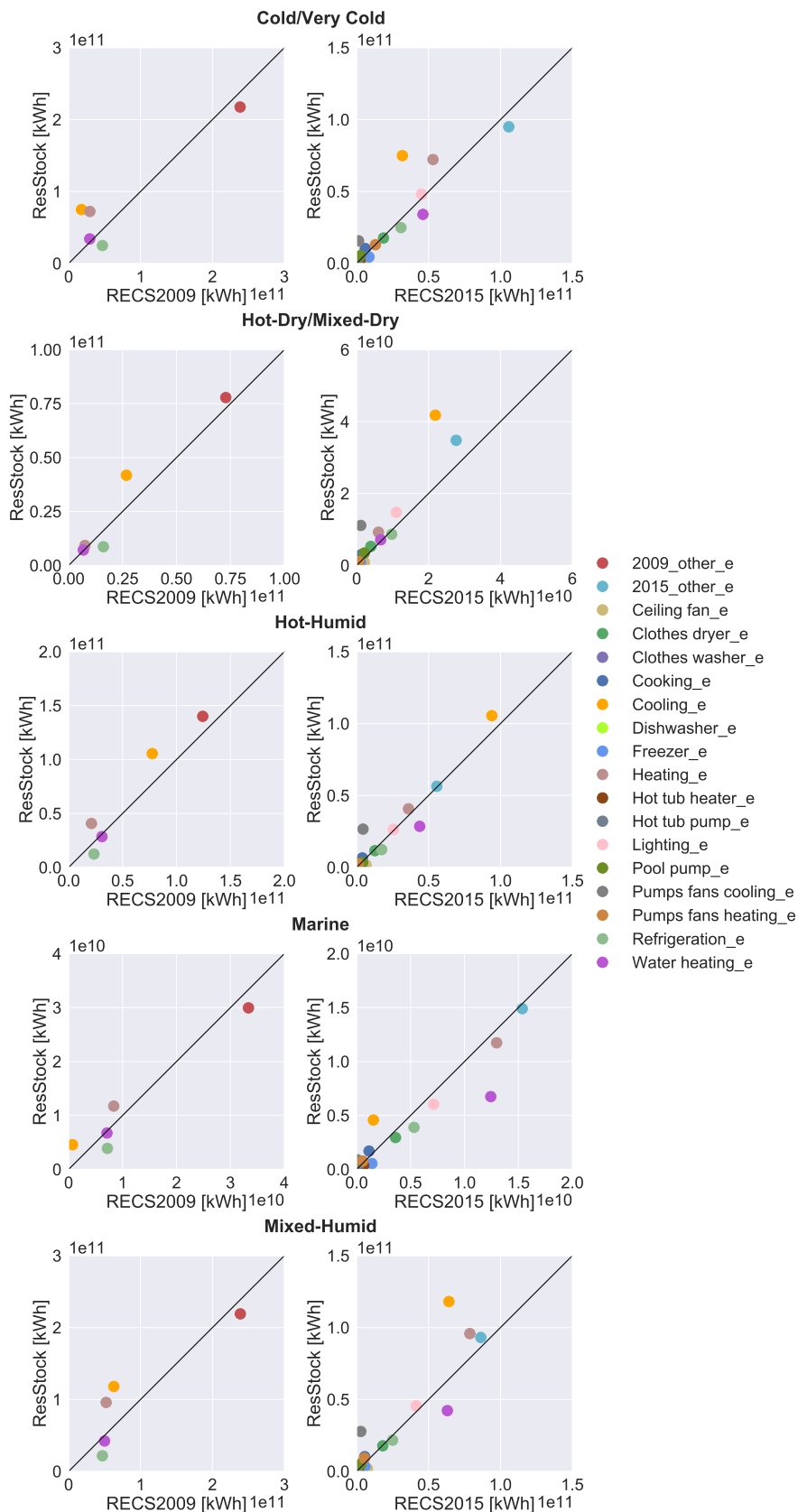
<sup>15</sup>Mobile homes are included in single-family detached and all multifamily units are combined into a single category.



**Figure 162. ResStock (2018 weather and 2018 stock) occupied dwelling unit electric end-use annual total (including heating and cooling correction based on EIA-861M; see Section 3.2.10) compared to RECS 2009 and RECS 2015.**



**Figure 163. ResStock (2018 weather and 2018 stock) occupied dwelling unit electric end-use annual total (including heating and cooling correction based on EIA-861M; see Section 3.2.10) compared to RECS 2009 and RECS 2015 by building type**



**Figure 164. ResStock (2018 weather and 2018 stock) occupied dwelling unit electric end-use annual total (including heating and cooling correction based on EIA-861M; see Section 3.2.10) compared to RECS 2009 and RECS 2015 by Building America Climate Zone**

**Table 23. Residential AMI Data Season Definitions**

Region	Region name	Data year	Summer months	Winter months	Shoulder months
Region 1	ComEd, IL	2018	6, 7, 8	1, 2, 3, 4, 10, 11, 12	5, 9
Region 2	Fort Collins, CO	2018	6, 7, 8	1, 2, 3, 4, 10, 11, 12	5, 9
Region 3	Seattle, WA	2019	6, 7, 8	1, 2, 3, 10, 11, 12	4, 5, 9
Region 4a	Chattanooga, TN	2019	5, 6, 7, 8, 9	1, 2, 3, 11, 12	4
Region 4b	Tallahassee, FL	2019	5, 6, 7, 8, 9, 10	1, 2, 11, 12	3, 4
Region 4c	Horry, SC	2018	5, 6, 7, 8, 9	1, 2, 3, 11, 12	4, 10
Region 5a	Cherryland, MI	2018	6, 7, 8	1, 2, 3, 4, 10, 11, 12	5, 9
Region 5b	State of VT (VEIC)	2019	7, 8	1, 2, 3, 4, 10, 11, 12	5, 6, 9

#### 4.1.2 Residential AMI Data Validation

In this section, we show the final comparisons to the eight residential utility load datasets collected for the five residential calibration focus regions. Details of these datasets are discussed in Section 2.3.6 and the map of the data sets can be seen in Figure 7. ResStock stacked end-use loads are compared to the AMI meter data for weekdays and weekends for three seasons.<sup>16</sup> The seasons are summer (cooling dominated months), winter (heating dominated months), shoulder (low load months), as determined by the procedure in *Definition of Seasons in QOIs*. As listed in Table 23, the seasons are different for each of the AMI calibration regions to best isolate the load shape effects of cooling loads, heating loads, and base loads.

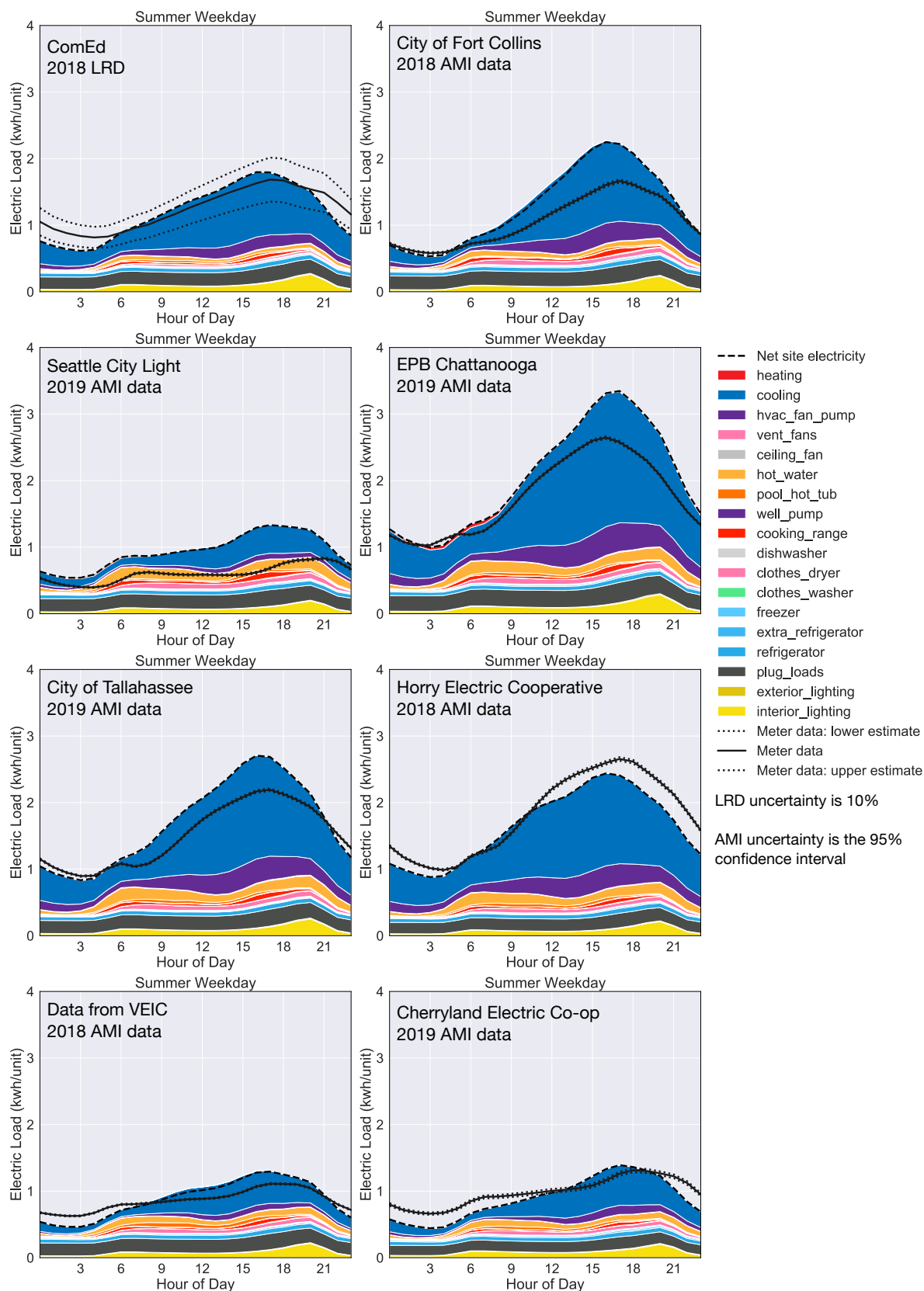
Figures 165 and 166 show ResStock (corrected) end-use loads compared to the regional calibration AMI datasets for summer. In some regions ResStock overestimates the summer load: Fort Collins, CO; Chattanooga, TN; and Tallahassee, FL. ResStock also seems to be overestimating the baseload in Seattle, WA as the majority of the load compared to the AMI data is taken up by non-HVAC loads. The minimum daily load for the seasons seem to be low for ComEd, IL; State of VT (VEIC); and Cherryland, MI. For Horry, SC the ResStock load is a bit low in the afternoon and evening. There are no major differences between the weekday and weekend comparisons.

Figures 167 and 168 show ResStock (corrected) end-use loads compared to the regional calibration AMI datasets for winter. The largest issue in winter is in Tallahassee, FL. ResStock substantially overestimates the heating load and heating peak. However, when looking at the EIA monthly retail sales for Florida in Figure 161, ResStock estimates the winter electricity use for Florida overall quite well, suggesting that the discrepancy is specific to Tallahassee. The season average load is about 10% low for ComEd, IL, but LRD are used to compare ResStock results. The sampling uncertainty of this LRD is estimated to be around 10%, but is hard to quantify exactly how uncertain these load data are. For Horry County, SC, ResStock is underestimating the heating load. ResStock is likely missing some load from vacant or seasonal use dwelling units. Horry County, SC has a large fraction of vacant dwelling units especially in multifamily dwelling units, which is 66% vacant.

Figures 169 and 170 show ResStock end-use loads compared to the regional calibration AMI datasets for the shoulder season. For most of the regions in these figures, the heating and cooling loads are likely the reason for discrepancy between ResStock and the meter data. In Fort Collins, CO; Chattanooga, TN; and Tallahassee, FL the ResStock cooling load is too high during the shoulder months. In Horry County, SC and Cherryland, MI, both heating and cooling loads could be increased to improve the ResStock load shape. In Seattle, WA the ResStock baseloads are likely overestimated as the baseloads take up the entire area under the AMI meter data.

<sup>16</sup>PV generation is included in these comparisons, although for the eight AMI datasets, the PV generated load is negligible. The dashed line representing ResStock net site electricity should be directly compared to the meter data, not just the top of the stacked end uses.





**Figure 165. ResStock modeled stacked end-use loads (including heating and cooling correction based on EIA-861M; see Section 3.2.10) compared to summer weekday meter data for each of the regional calibration datasets**

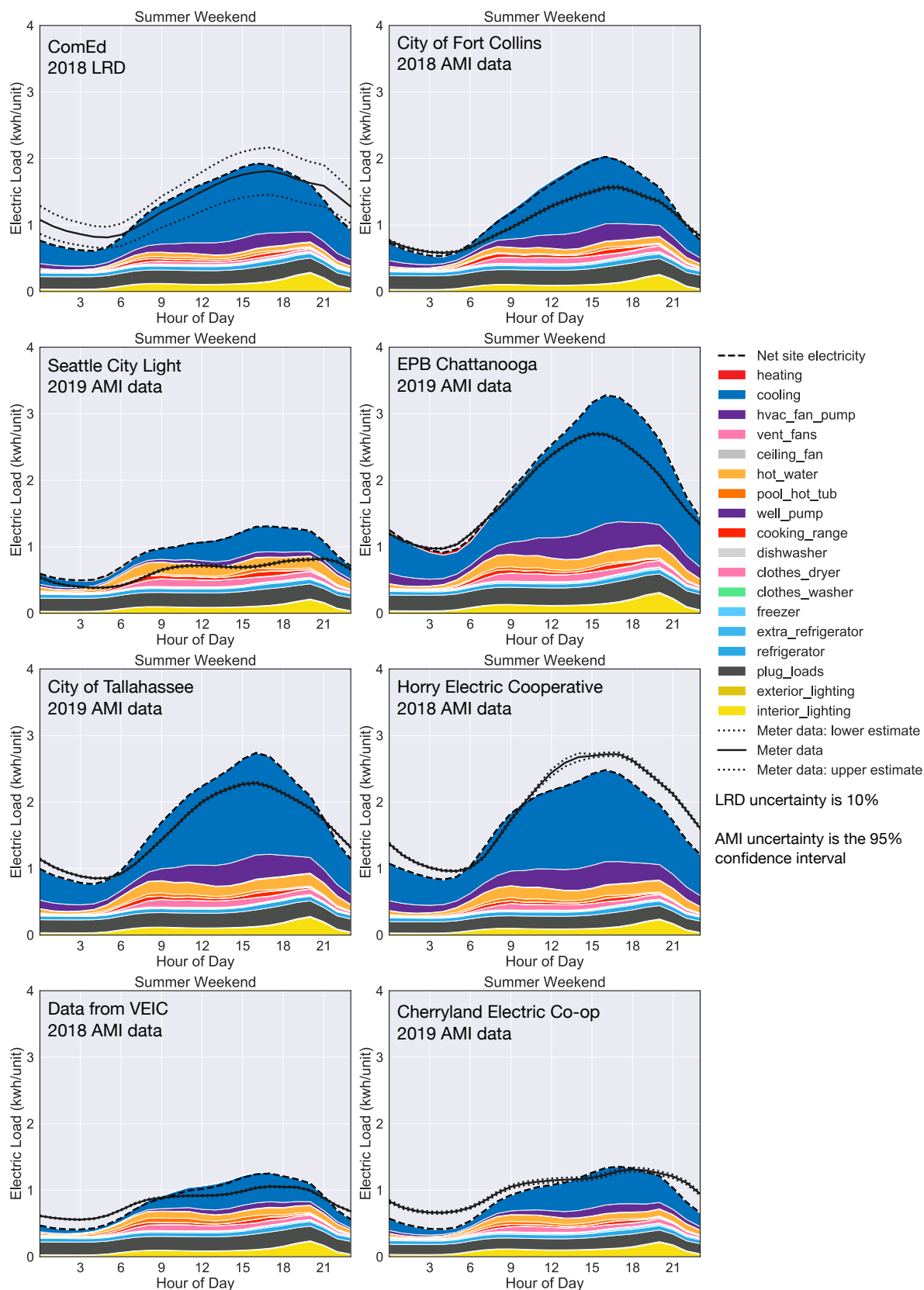
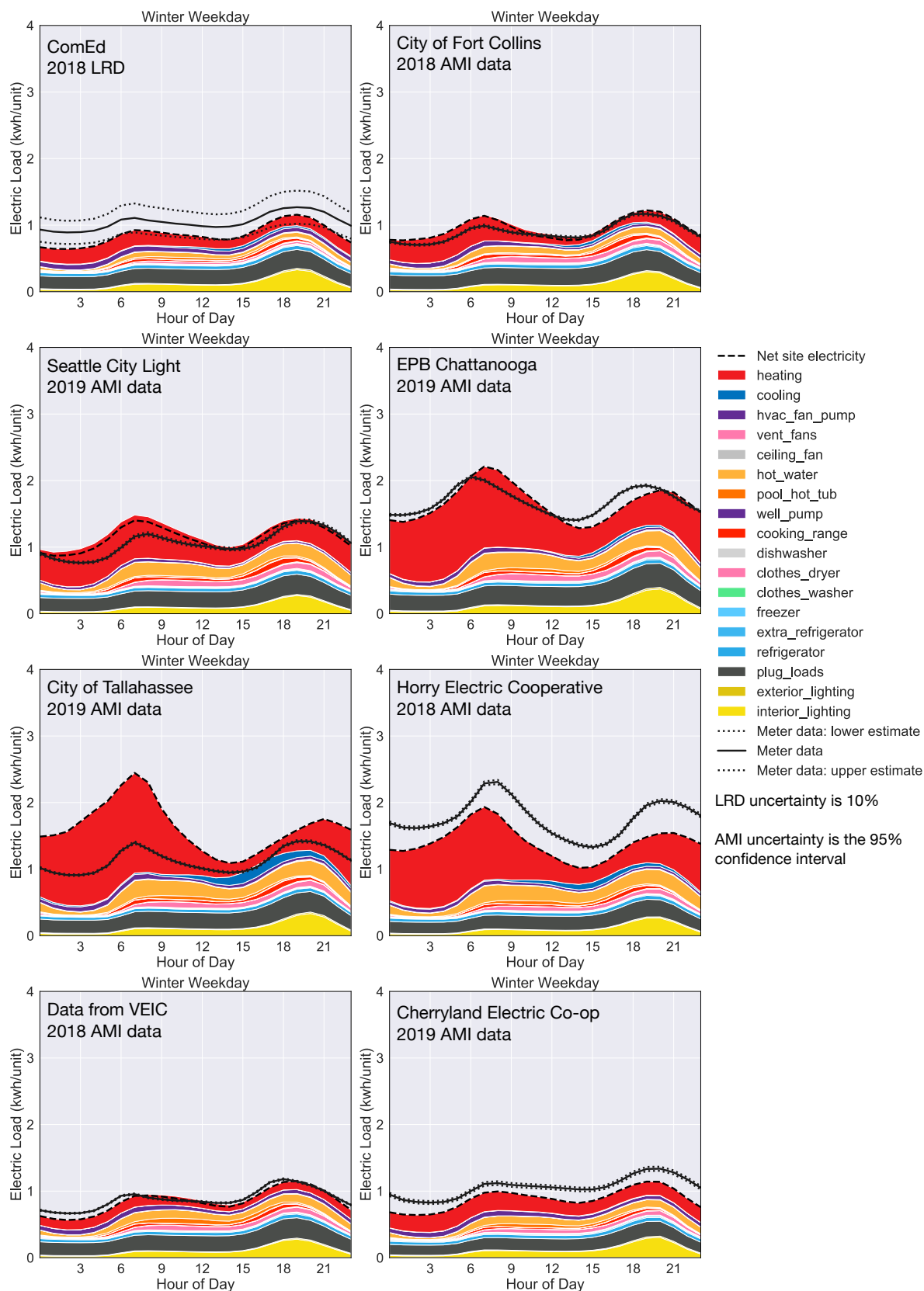
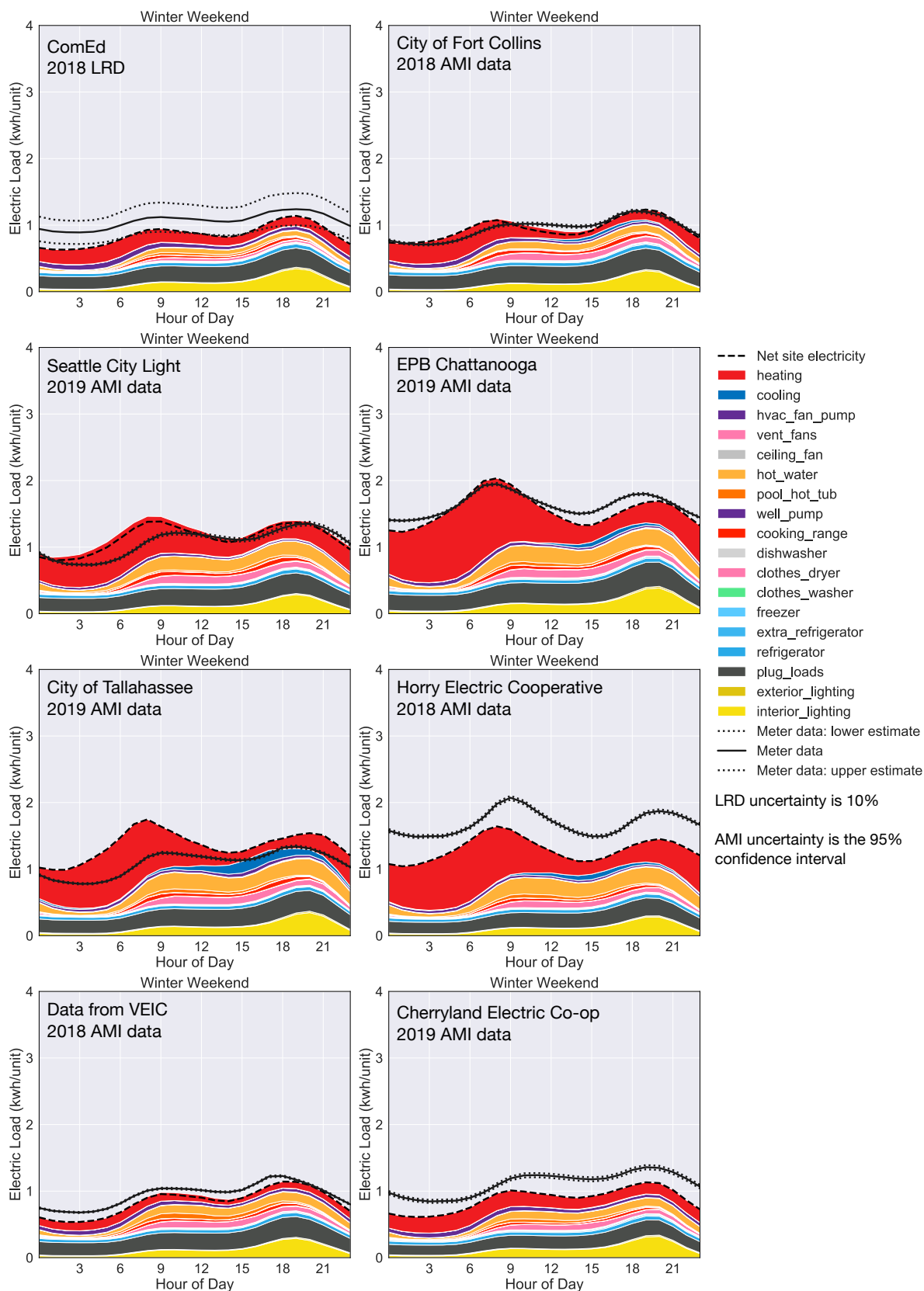


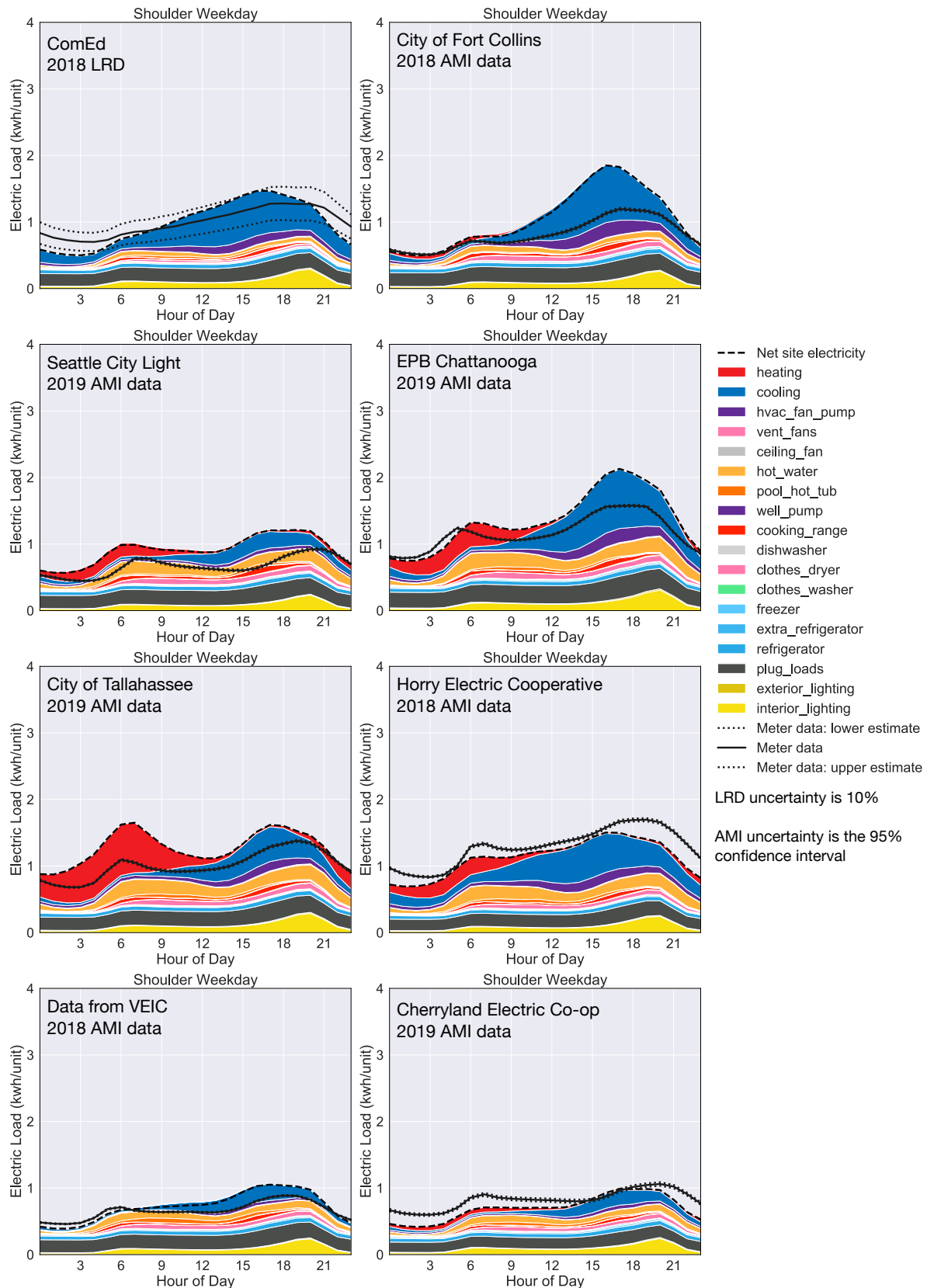
Figure 166. ResStock modeled stacked end-use loads (including heating and cooling correction based on EIA-861M; see Section 3.2.10) compared to summer weekend meter data for each of the regional calibration datasets



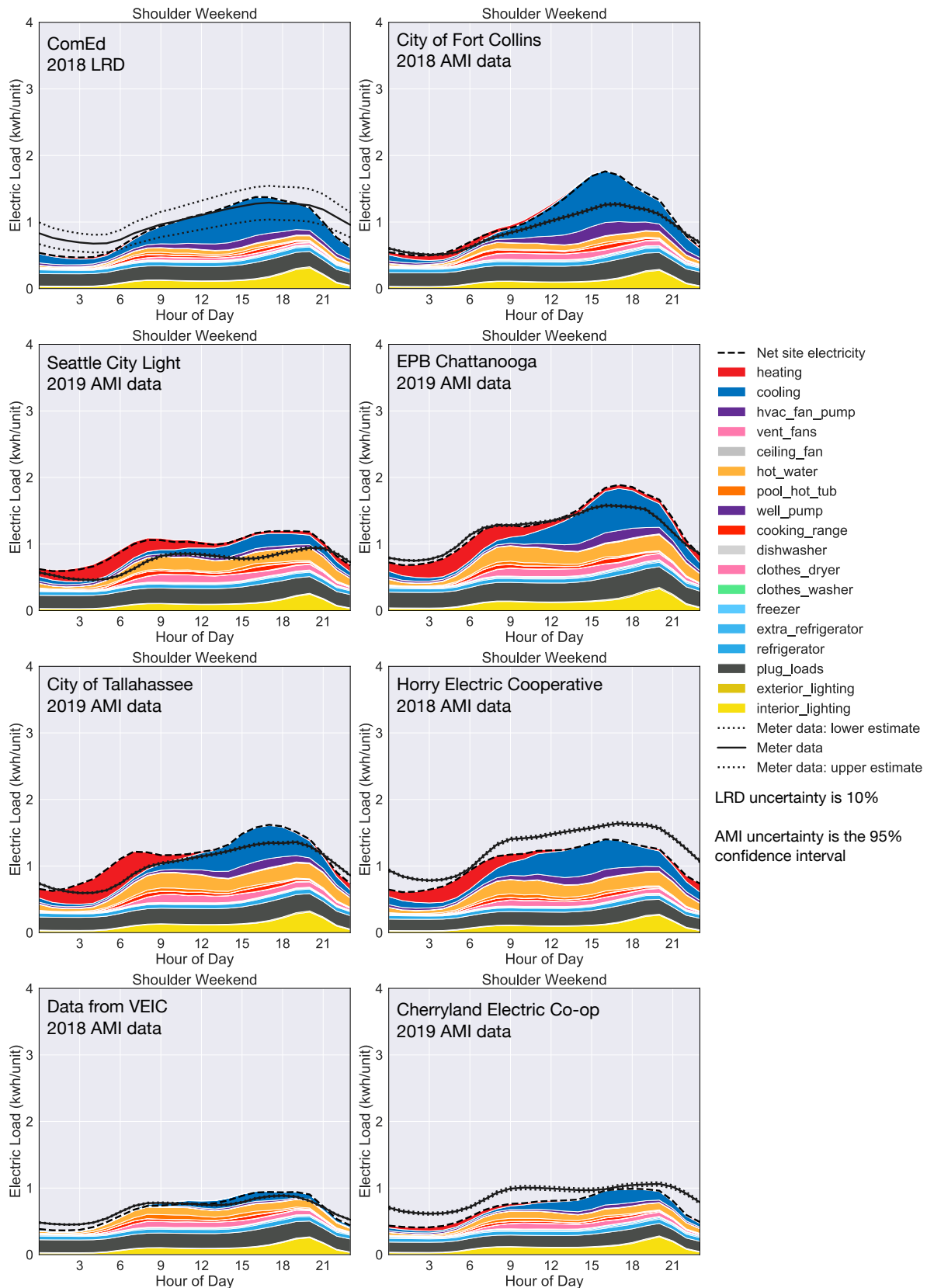
**Figure 167. ResStock modeled stacked end-use loads (including heating and cooling correction based on EIA-861M; see Section 3.2.10) compared to winter weekday meter data for each of the regional calibration datasets**



**Figure 168. ResStock modeled stacked end-use loads (including heating and cooling correction based on EIA-861M; see Section 3.2.10) compared to winter weekend meter data for each of the regional calibration datasets**



**Figure 169. ResStock modeled stacked end-use loads (including heating and cooling correction based on EIA-861M; see Section 3.2.10) compared to shoulder weekday meter data for each of the regional calibration datasets**



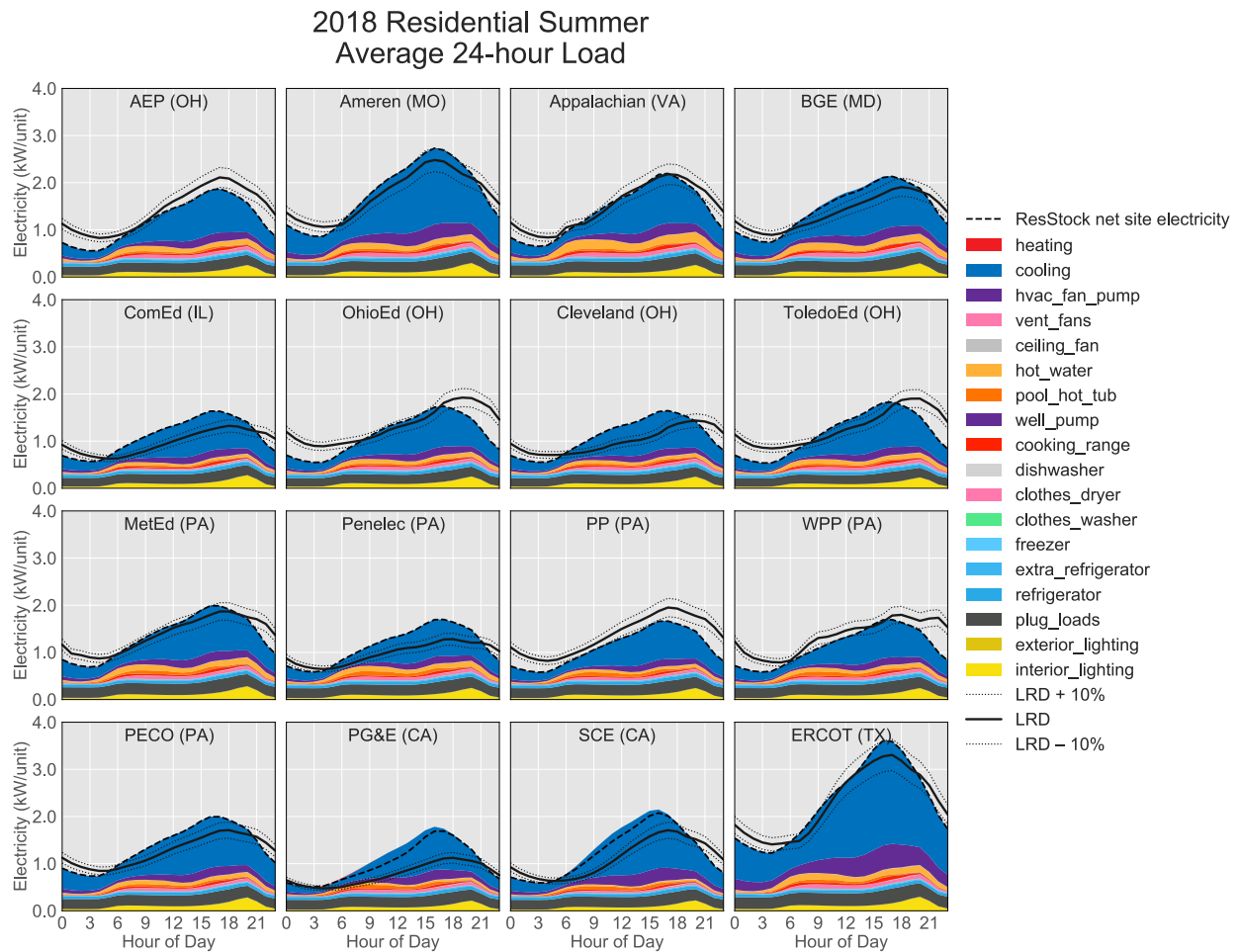
**Figure 170. ResStock modeled stacked end-use loads (including heating and cooling correction based on EIA-861M; see Section 3.2.10) compared to shoulder weekend meter data for each of the regional calibration datasets**



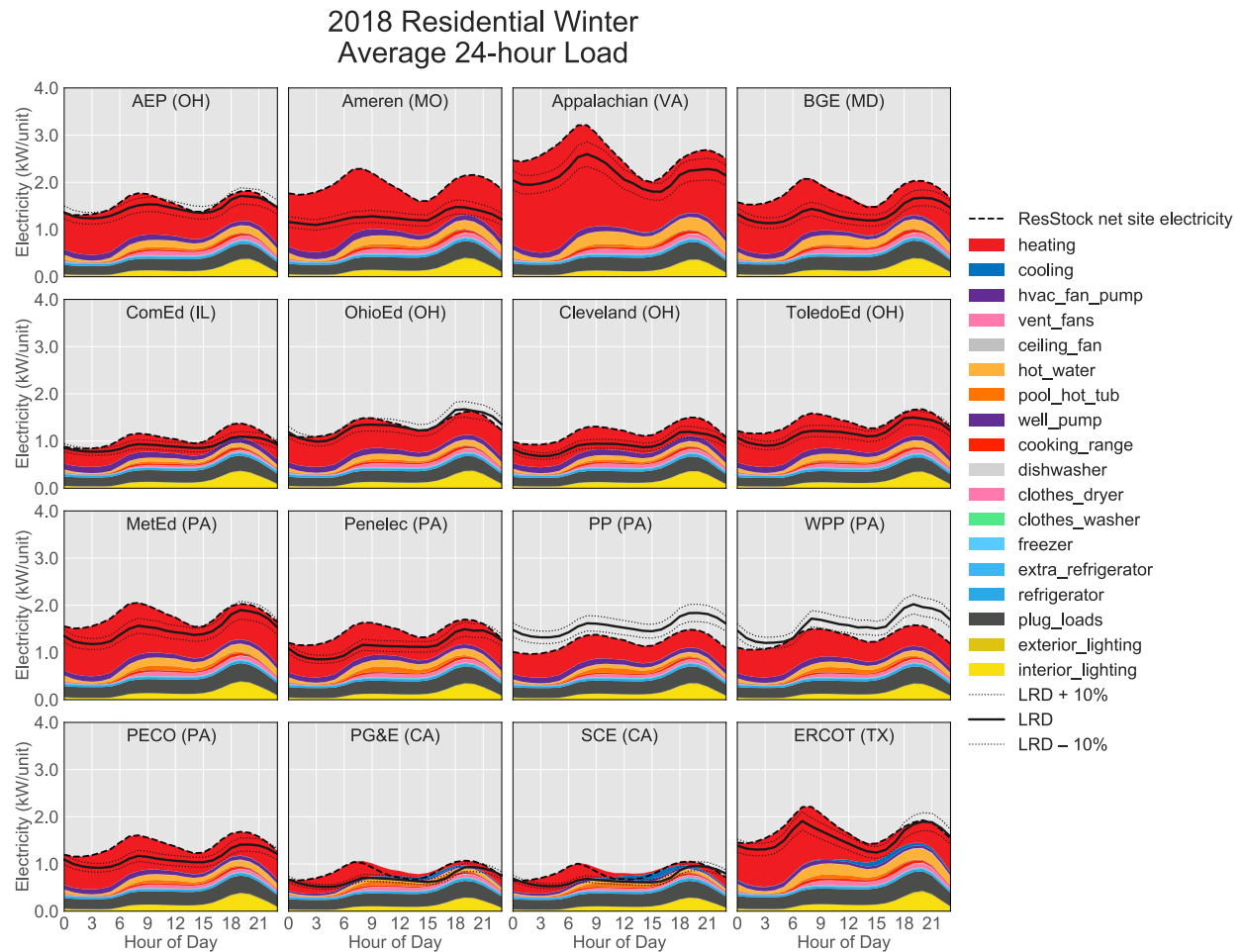
### 4.1.3 Residential LRD Validation

For this project a set of 16 LRD datasets were collected for comparison to ResStock. A description of these data is in Section 2.3.6. The LRD covers large sections of California, Texas, Ohio, Pennsylvania, but also smaller sections of Illinois, Missouri, Maryland, Virginia, and West Virginia. A map of the counties that are served by these utilities in 2018 can be seen in Figure 8. Figure 171, Figure 172, and Figure 173 show ResStock stacked end uses compared to the LRD for summer months (June, July, and August), winter months (December, January, and February), and spring and fall months (March, April, May, September, October, November). These data comparisons were created relatively late in the calibration process, so these comparisons did not inform most calibration decisions and resulting changes to ResStock. The calibration strategy has been to use the AMI datasets to inform nationally relevant changes to ResStock. However, because the AMI data comparisons informed nationally relevant modeling changes and did not focus on only updating the AMI region, the LRD comparisons improved throughout the calibration process.<sup>16</sup>

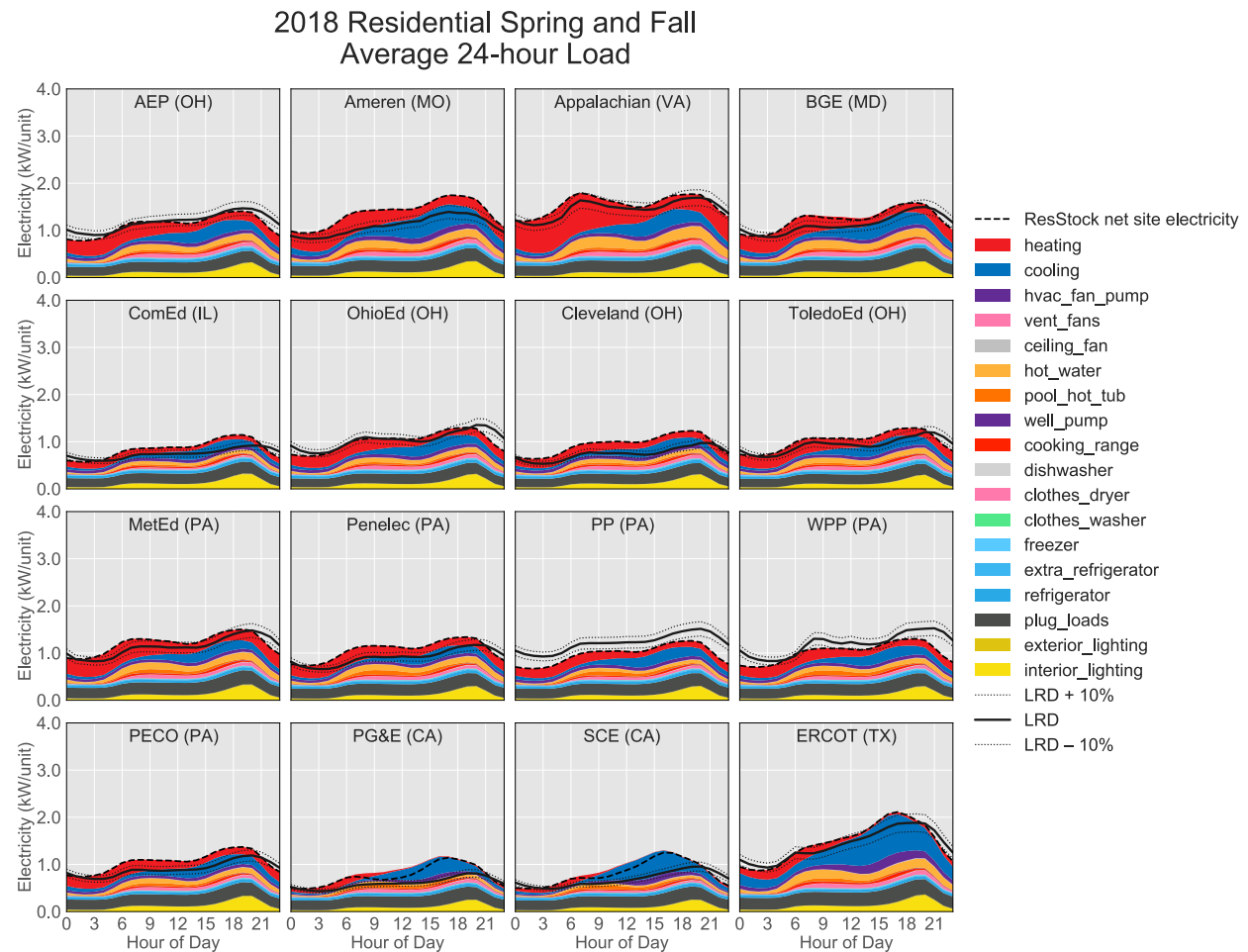
Although we are generally pleased with these comparisons, there are some areas for improvement. In summer, Figure 171, PG&E (CA), SCE (CA), and Penelec (PA) have high loads compared to the LRD. For the Ohio LRD, ResStock is missing some load on the evenings and early morning hours. In winter, Figure 172, ResStock is overestimating the load in Ameren (MO), and to a lesser degree in Appalachian (VA) and MetEd (PA). For spring and fall, Figure 173, ResStock is likely overestimating the cooling load in PG&E (CA) and SCE (CA). For Ameren (MO), ResStock is likely overestimating the heating load.



**Figure 171. ResStock modeled stacked end-use loads (including heating and cooling correction based on EIA-861M; see Section 3.2.10) compared to 2018 summer LRD from various utilities (see map in Figure 8)**

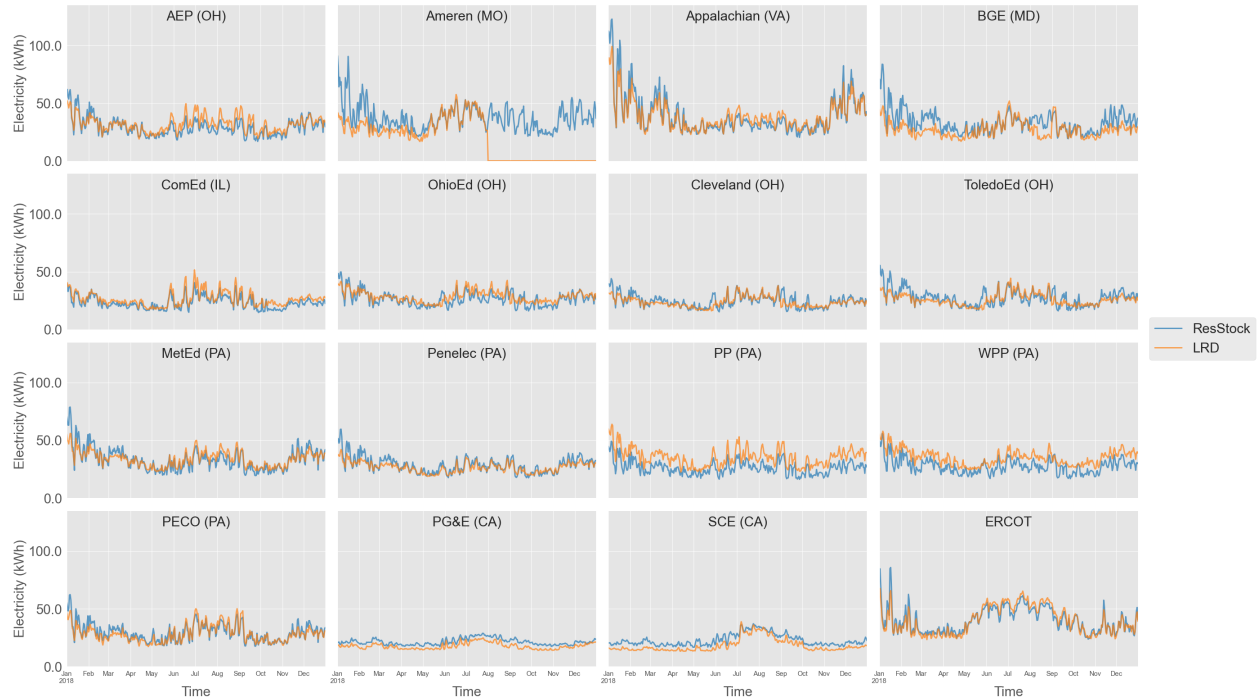


**Figure 172. ResStock modeled stacked end-use loads (including heating and cooling correction based on EIA-861M; see Section 3.2.10) compared to 2018 winter LRD from various utilities (see map in Figure 8)**



**Figure 173. ResStock modeled stacked end-use loads (including heating and cooling correction based on EIA-861M; see Section 3.2.10) compared to 2018 spring and fall LRD from various utilities (see map in Figure 8)**

Another way to compare modeled and measured load profile data is by looking at daily sums of electricity use per unit. Figure 174 shows this comparison using residential LRD datasets. These plots do not show diurnal dynamics, but they do show seasonal dynamics, including how the models' weather response compares to LRD. Comparisons of modeled daily sums to residential AMI data can be found in Appendix C.



**Figure 174. ResStock modeled daily sums of electricity per unit (kWh/unit/day) (including heating and cooling correction based on EIA-861M; see Section 3.2.10) compared to 2018 LRD. Ameren (MO) data were missing for half of 2018; only the non-zero portions were used for the previous season-day comparisons.**

#### 4.1.4 Residential End-Use Data Validation

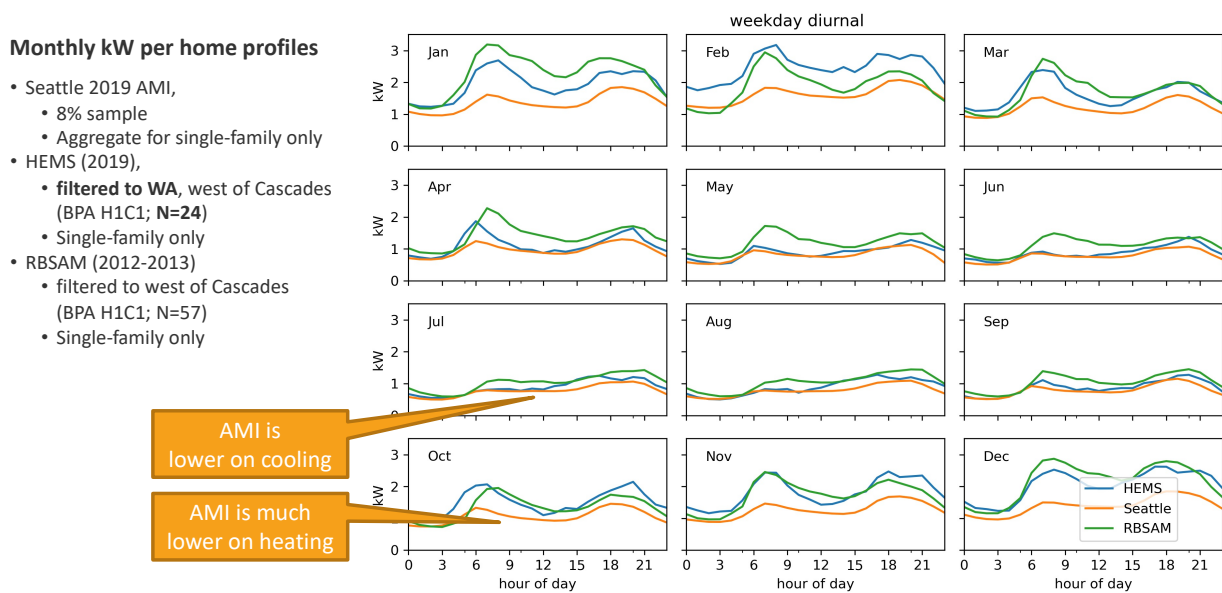
Our validation of heating and cooling end uses was accomplished through the AMI comparisons, because the peaks and seasonal variation can easily be identified with weather-driven heating and cooling loads. Validation of non-weather-dependent end uses data is primarily covered by the comparison of the new residential stochastic occupant behavior model against the RBSAM dataset (Section 3.2.1) and complemented with the residential end-use transferability study (Section 2.3.9). However, we were able to compare ResStock end use outputs to submetered end use data in one instance, where we had access to both AMI data and end-use data in the same region: Seattle.

First, as shown in Figure 175, we compared whole-home load profiles to see what differences there were between the AMI data for Seattle and submetered end-use data from NEEA's ongoing HEMS submetering study (filtered to 24 homes in western Washington; NEEA 2021b) and 2012–13 data from NEEA's RBSA Metering study (RBSAM; filtered to 57 homes in western WA and OR; Larson et al. 2014). The AMI data average is lower than HEMS and RBSAM during summer months and much lower during winter months. Some of the difference between the older RBSAM data and the 2019 AMI and HEMS data may be explained by different weather and by changes in the stock characteristics, such as increasing saturation of LED lighting. However, the differences between the AMI data and HEMS suggest that the characteristics of sampled homes (such as vintage, floor area, or heating fuel) differ between the datasets.

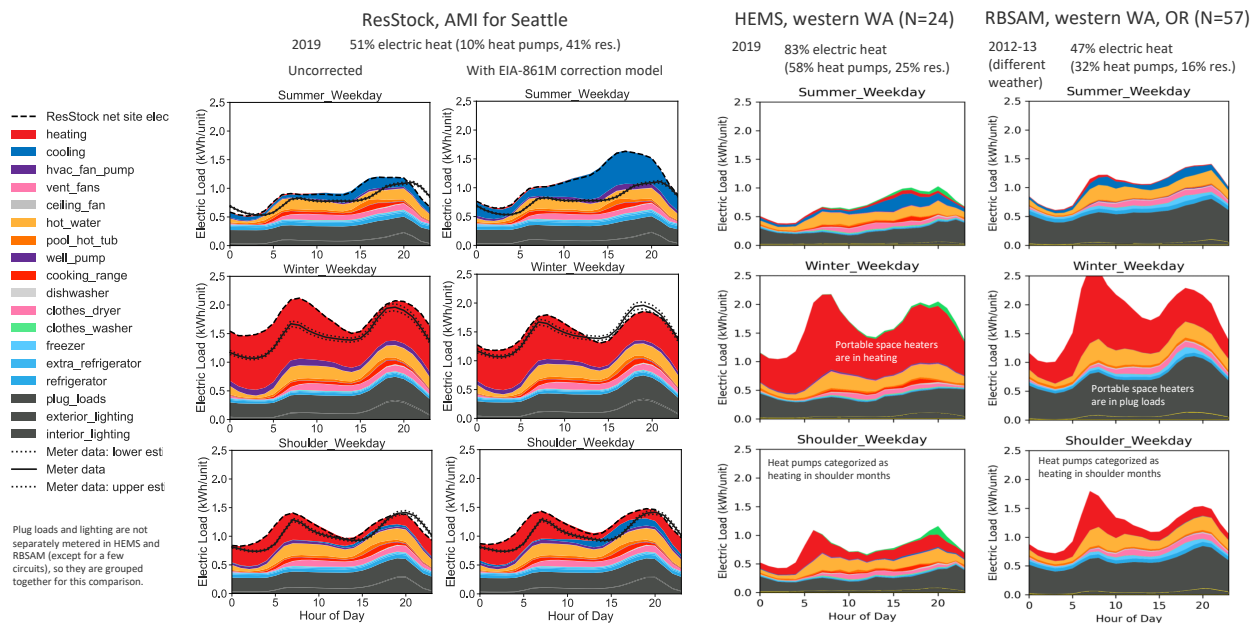
Next, we compared seasonal end-use profiles across the three datasets (Figure 176) and ResStock 2019 outputs (corrected). As discussed in Section 2.3.9, it was a substantial effort to align end-use categories across the datasets. As noted in the figure caption, in most cases lighting could not be metered separately from plug load circuits, so we grouped plug loads and lighting for this comparison. One can observe that the magnitude of ResStock non-HVAC loads appears to be greater than that of HEMS and less than that of RBSAM.

It appears that ResStock (uncorrected) estimates summer cooling energy relatively well, compared to HEMS and RBSAM, and relative to the Seattle AMI data. However, the EIA-861M correction model (as shown in Figure 161) increased summer cooling electricity for Washington state, because ResStock was about 25% low compared to EIA-861M data for Washington. Thus, the correction model is responsible for the summer overestimation seen in the second column of Figure 176. This is an example of a case where discrepancies between two empirical data sources mean that it is challenging for the model to compare well against both sources.

The ResStock (uncorrected) estimate of winter load is high compared to Seattle AMI data, and the EIA-861M correction model appears to correct this overestimation quite well, although it results in a winter load that is low compared to HEMS and RBSAM. We found that the subset of HEMS homes in western Washington (N=24) has a higher percentage of homes with electric heat compared to the data used in ResStock for Seattle (Ruggles et al. 2021). The filtered RBSAM dataset has a similar share of heating that is electric, but a higher share of heat pumps compared to electric resistance heating, and is for a different weather year. These observations highlight some of the challenges of comparing directly to end use submetering datasets with relatively low sample sizes.



**Figure 175. Comparison of 2019 AMI data from Seattle to 2019 HEMS data from western Washington and 2012–13 RBSAM data from western Washington and Oregon, by month (single-family homes only). The AMI data average is lower during summer months and much lower during winter months, suggesting that there may be some sampling differences between the datasets.**



**Figure 176. Comparison of 2019 ResStock outputs (with and without the heating and cooling correction based on EIA-861M; see Section 3.2.10) and Seattle AMI data to 2019 HEMS data from western Washington and 2012–13 RBSAM data from western Washington and Oregon, by season.**

Notes:

1. For these comparisons, seasons are defined using calendar months (summer = June–Aug., winter = Dec.–Feb.).
2. The share of heating with electric resistance and heat pumps differs between HEMS, RBSAM, and the data we have for Seattle in ResStock (percent electric heating is specific to Seattle and is from PUMS (Ruggles et al. 2021); percent heat pump is for IECC zone 4C and is from RECS 2009 (EIA 2013)).
3. In HEMS and RBSAM, because they are often on the same circuit, plug loads and lighting were not be metered separately for most homes, so plug loads and lighting were grouped (dark grey) for this comparison.
4. In the RBSAM data, portable space heaters show up as plug loads, whereas in HEMS, circuits with portable space heaters were labeled as such and we were able to categorize them as heating.
5. In both HEMS and RBSAM, electricity used by heat pumps could not be distinguished as heating or cooling, so for these graphs, we assume they are in cooling mode during summer and in heating mode during the winter and shoulder seasons. The ResStock results for shoulder season suggest that there is likely a mix of both heating and cooling during the shoulder season months.



#### 4.1.5 Residential Quantities of Interest Over Time

For residential calibration, we were able to track the progress of the quantities of interest (QOI) over the course of all five calibration regions. Table 24 shows the 62 ResStock runs that were simulated over the five-region process and which model changes were included in each. Most of these changes were not targeted to change only the specific regions being calibrated, but were updates to the national housing stock characteristics or weather data. As a result, each subsequent calibration region benefited from the calibration effort of previous regions.

**Table 24. Description of Each ResStock Model Run Shown in the QOI Plots (most descriptors link to the appropriate section elsewhere in the report)**

ResStock EULP Run Number	Calibration Focus Region	Region Run	Description (with hyperlinks to relevant report sections)
1	1	ComEd Only	ResStock before EULP
2	1	ComEd Only	Daylight savings and ceiling fan bug fixes
3	1	ComEd Only	Introduce new weather data (see Section 2.4)
4	1	ComEd Only	<i>Adding Modeling of Masonry Walls</i>
5	1	ComEd Only	<i>Increased Geospatial Resolution</i>
6	1	ComEd Only	Increase interior thermal mass of dwelling units (sensitivity, not included in future runs)
7	1	ComEd Only	<i>Lighting Technology Saturation, Plug Load Schedule</i>
8	1	ComEd Only	<i>HVAC Setpoint Schedule Diversity, Preliminary test of AHS for AC ownership, Bug fixes on previous improvements</i>
9	1	ComEd Only	<i>Major Appliance Saturation, Increased Resolution of Floor Area Bins, Preliminary plug load equation adjustments, Refrigerator efficiency updates (see Section 3.2.7), Minor bug fixes</i>
10	1	ComEd Only	<i>Updating Air Infiltration Rates and Diversity, Updating Window-to-Wall Ratios</i>
11	1	National	National run with improvements thus far
12	1	National	Weather data comparisons (MERRA and NOAA ISD)
13	1	ComEd Only	Weather data comparisons (MERRA and NOAA ISD)
14	1	National	Weather precipitation issue fix
15	2	National	Weather precipitation issue fix (different weather year) <i>Updated Mapping of Dwelling Units to Utilities</i>
16	2	ComEd, Fort Collins, Horry	<i>Introduce New Residential Stochastic Occupant Behavior Model (select regions)</i>
17	2	National	<i>Introduce New Residential Stochastic Occupant Behavior Model (national)</i>
18	2	National	Changes to interior thermal mass
19	2	National	Natural ventilation sensitivity Run 1
20	2	National	Natural ventilation sensitivity Run 2
21	2	National	<i>Heating Fuel Type—Higher Geographic Resolution</i>
22	2	National	<i>Accounting for Climate Zone Humidity in HVAC Setpoint Distributions</i>
23	2	National	<i>Introduce Plug Load Equations by Housing Type</i>
24	2	National	New baseline incorporating 21, 22, and 23
25	2	National	<i>Introduce Unoccupied Housing Units</i>
26	2	National	Change in setpoint seasonality (not included in future runs)

*table continued on next page*

Table 24 – *continued from previous page*

<b>ResStock EULP Run Number</b>	<b>Calibration Focus Region</b>	<b>Region Run</b>	<b>Description</b> (with hyperlinks to relevant report sections)
27	2	National	Introduce <i>Improving the Structure of HVAC Characteristics, Updating Room Air-Conditioner Efficiency Distributions</i> , and HVAC efficiency distributions (see Section 3.2.7)
28	3	National	<i>Roof Material Distributions</i> , Revert interior thermal mass changes, bug fixes
29	3	National	<i>Foundation Type Distributions, Air-Conditioner Sizing Fix</i>
30	3	National	<i>Air Conditioner Saturation—Adding Climate Dependency</i>
31	3	National	New baseline incorporating improvements to date
32	3	National	<i>Increased Geographic Resolution of Weather Data</i>
33	3	National	<i>Faster Multifamily Unit Modeling</i>
34	3	National	<i>Household Size—Adding Geographic Dependency</i>
35	3	National	<i>Regional Variation in Lighting Efficiency, Regional Variation in Plug Load Usage</i>
36	3	National	<i>Water Heating Updates, Removing Multifamily Central Heating and Water Heating Electricity</i>
37	4	National	Plug load bug fixes
38	4	National	Additional work on <i>Increased Geographic Resolution of Weather Data</i>
39	4	National	Fix errors in <i>Faster Multifamily Unit Modeling</i>
40	4	National	Final ResStock before switching to single-unit multifamily implementation
41	4	National	More error fixes in <i>Faster Multifamily Unit Modeling</i>
42	4	National	<i>Refinements to Unoccupied Housing Units</i>
43	4	National	<i>Further Update Lighting Technology Saturation</i>
44	4	National	<i>Zonal Electric Heating Setpoints</i>
45	4	Horry County Only	<i>Sensitivity of Coastal vs. Non-Coastal Weather Stations</i>
46	4	National	<i>Occupant Schedule Shifting Based on Time Use Data</i>
47	4	National	Fix cooling setpoint errors
48	4	National	<i>Update Room Air-Conditioner Setpoints</i>
49	5	National	New baseline incorporating improvements to date
50	5	National	Correct bedroom size assignment for extreme house sizes
51	5	National	Fix garage sizing errors
52	5	National	<i>Adding Modeling of Dwelling Units in Mid-Rise and High-Rise Buildings</i>
53	5	National	ResStock back-end updates
54	5	National	<i>Updated Lighting Usage Equation</i>
55	5	National	<i>Updating Window Types and Distributions</i>
56	5	National	New baseline with improvements to date
57	5	National	Minor changes (e.g., structure of foundation insulation) to allow for future ResStock features
58	5	National	<i>Updating Masonry Wall Distributions, Updating Wall Exterior Finish Distributions</i>
59	5	National	<i>Updating Room Air-Conditioner Performance Curves</i> , fix room AC performance curve bug
60	5	National	<i>Residential Output Correction Model</i>

*table continued on next page*

Table 24 – *continued from previous page*

<b>ResStock EULP Run Number</b>	<b>Calibration Focus Region</b>	<b>Region Run</b>	<b>Description (with hyperlinks to relevant report sections)</b>
61	N/A	National	EULP final run, raw ResStock output, including <i>On-Site Solar Photovoltaic Generation</i>
62	N/A	National	EULP final run with <i>Residential Output Correction Model</i> applied, including <i>On-Site Solar Photovoltaic Generation</i>

### Annual Electricity Error Over Time

Table 24 provides a way to look up which model changes correspond to the changes in QOIs shown in the subsequent QOI plots. For annual electricity per dwelling unit for each calibration region (Figure 177) Seattle was the most improved, but at the end of Region 5, ResStock still was overestimating annual electric load per dwelling unit for Seattle, even with the EIA-861M correction model applied. Before correction, the discrepancy is largely due to overestimation of heating load. After correction it appears to be due to overestimation of cooling load. Fort Collins annual electricity per dwelling unit decreased throughout the calibration process. The improvements to Fort Collins are likely because of reducing heating and cooling loads steadily throughout calibration regions three and four. ResStock overestimates the annual electricity per dwelling unit for the City of Tallahassee. The overestimation is due to high heating loads, especially on cold days.

### ComEd, IL Timeseries QOIs Over Time

Figure 178 shows the timeseries magnitude and timing QOIs for ComEd, IL throughout the calibration process. There were a few changes that had a significant impact on the ComEd, IL QOIs. In the Region 1 calibration effort, using NOAA ISD weather data rather than MERRA weather data decreased the summer peak load significantly. At the end of region one, the residential stochastic load model was introduced into ResStock, causing the baseloads and peaks to significantly increase. This large increase continued to be present until the end of Region 2 when a bug was fixed in how the schedules were being interpreted by OpenStudio and EnergyPlus. Throughout the calibration process, the baseload (minimum daily load by season) continued to be underestimated by ResStock. It is likely that ResStock is missing heating and cooling in the late evenings and early morning hours. The increase of summer peaks and decrease of winter peaks at the end of calibration Region 5 is due to the correction model and adjusting the HVAC loads to better match the Illinois monthly electric retail sales according to EIA-861M.

### Fort Collins, CO Timeseries QOIs Over Time

Figure 178 shows the timeseries magnitude and timing QOIs for Fort Collins, CO throughout the calibration process. As with ComEd, the impacts of the residential stochastic load model being integrated and the bug being fixed can be seen at the beginning and end of calibration Region 2. The winter peak timing metric was difficult for ResStock to estimate because morning and late afternoon peaks are roughly equal in winter, so when averaging a set of peak times days together the timing metric is pulled to the middle of the day. As a result, the timing metric is not overly reliable for this region. The final few changes of the calibration process were the integration and application of the residential correction model. For Colorado, ResStock (uncorrected) underestimates the cooling load in summer months and overestimates the heating load in winter months. As a result there were significant changes to the QOIs in the last couple ResStock runs.

### Seattle, WA Timeseries QOIs Over Time

Figure 180 shows the timeseries magnitude and timing QOIs for Seattle, WA throughout the calibration process. The summer peak loads were the most improved aspect of the Seattle QOIs. As described in Section 3.2.7, the dependency of air conditioner saturation on climate was accidentally removed during Region 2 and was fixed in Region 3. As a result, all climate zones had the same air conditioner saturation. This bug is a particularly large problem in marine climates as the mild climate typically has fewer air conditioner systems than other climates. In a similar way to Fort Collins, the winter timing metric is a less reliable metric as ResStock estimates the winter

peaks in the morning and late afternoon to be about equal. This causes the average time of the peak to shift towards the middle of the day. The last couple runs of the calibration effort were the introduction of the correction model. ResStock underestimates the Washington load in summer months and as a result the load is increased during these months. The increase caused ResStock to overestimate the summer peaks. As Seattle is not a cooling dominated climate, the large percent error is largely because the load is relatively low in summer. Another explanation for the large overestimation of summer load is that the baseloads in Seattle may be high. The baseloads almost sum up to the AMI data.

#### **Chattanooga, TN Timeseries QOIs Over Time**

Figure 181 shows the timeseries magnitude and timing QOIs for Chattanooga, TN throughout the calibration process. Chattanooga, TN benefited from the nationally relevant changes leading up to the Region 4 phase. The starting point for the QOIs for Chattanooga, TN was relatively close to the AMI data compared to earlier regions. ResStock does estimate the summer peak time to be a bit later than the AMI data. The EIA-861M correction model shifted the peak time a little earlier and closer to the AMI data. The correction model shifted the average summer peak magnitude and the top 10 peak magnitude up, and as a result ResStock is now overestimating the summer peaks.

#### **Horry County, SC Timeseries QOIs Over Time**

Figure 182 shows the timeseries magnitude and timing QOIs for Horry County, SC throughout the calibration process. Although Horry County, SC was not part of a focused calibration region until Region 4, ResStock runs during Region 2 and Region 3 were queried and the impact of the changes can be seen in Figure 182. The nationally relevant changes in Region 2 and Region 3 affected the QOIs for Horry County. Horry County, SC is a unique area compared to the other AMI data locations because it has a large fraction of seasonally occupied dwelling units. This underestimation is likely due to assumptions around unoccupied or periodically unoccupied dwelling units. Seasonally operated dwelling units are currently modeled as fully unoccupied in ResStock. At the end of calibration, ResStock (uncorrected) was overestimating the winter load and underestimating the summer load for South Carolina. As a result, the correction model makes significant adjustments to the summer and winter peak QOIs at the end of the calibration effort.

#### **Tallahassee, FL Timeseries QOIs Over Time**

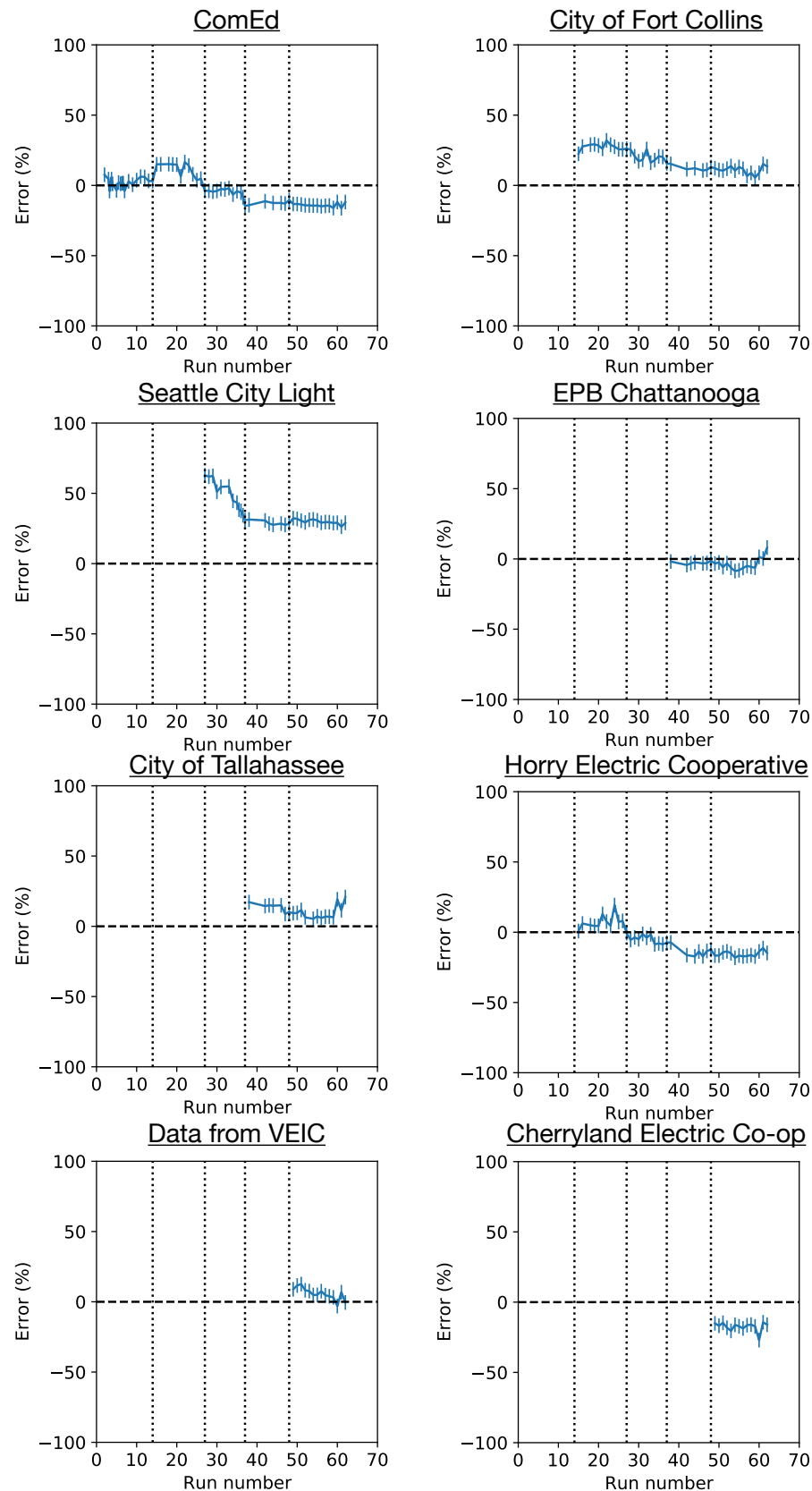
Figure 183 shows the timeseries magnitude and timing QOIs for Tallahassee, FL throughout the calibration process. The baseload and summer cooling load in Tallahassee, FL likely benefited from previous calibration regions as these QOI errors are small. The winter peak load and timing are the largest discrepancy between ResStock and all the AMI data. Before applying the EIA-861M correction model, ResStock was overestimating the heating load in Tallahassee, but underestimating the load for winter months in Florida overall (based on 2019 EIA-861M data). As a result, the correction model pushes the winter peak loads in Tallahassee even higher. As the winter peak is so much higher than the AMI data, the timing of the peak is also earlier than expected.

#### **State of VT (VEIC) Timeseries QOIs Over Time**

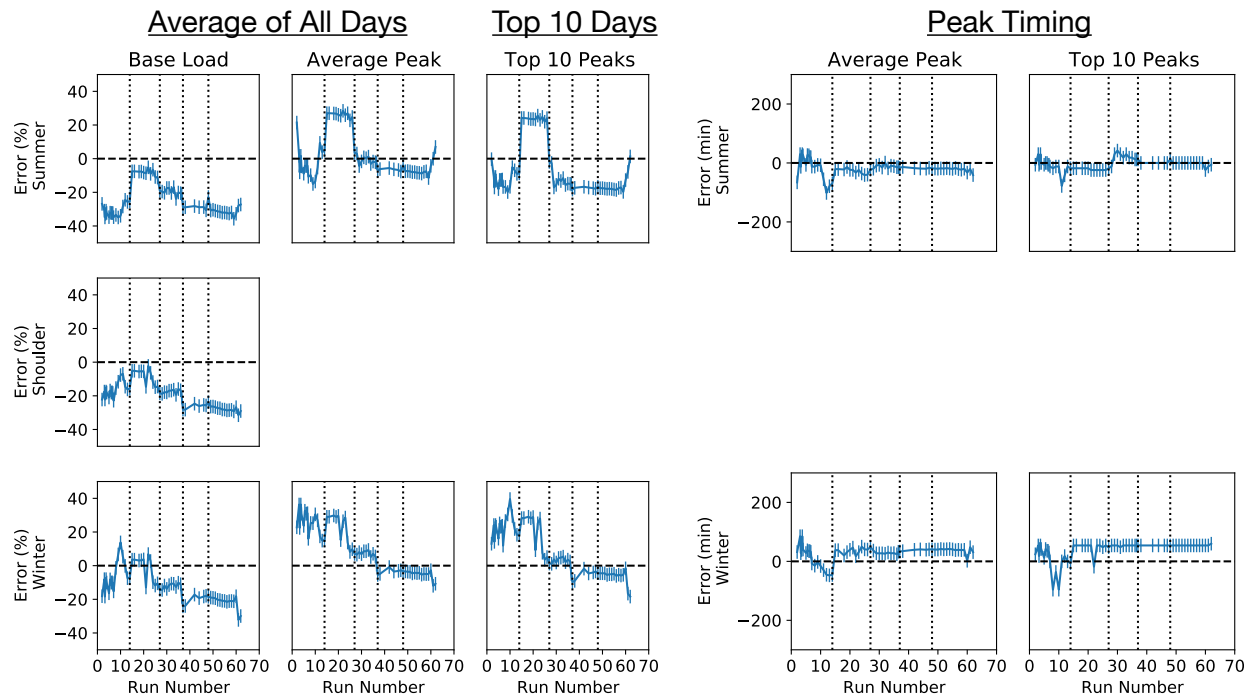
Figure 184 shows the timeseries magnitude and timing QOIs for the State of VT (VEIC) throughout the calibration process. VT likely benefited from the nationally relevant changes leading up to calibration Region 5. The starting point for the QOIs for VT was relatively close to the AMI data compared to earlier regions. Through the regional calibration the cooling loads were decreased making ResStock more accurate in estimating the summer peak loads.

#### **Cherryland, MI Timeseries QOIs Over Time**

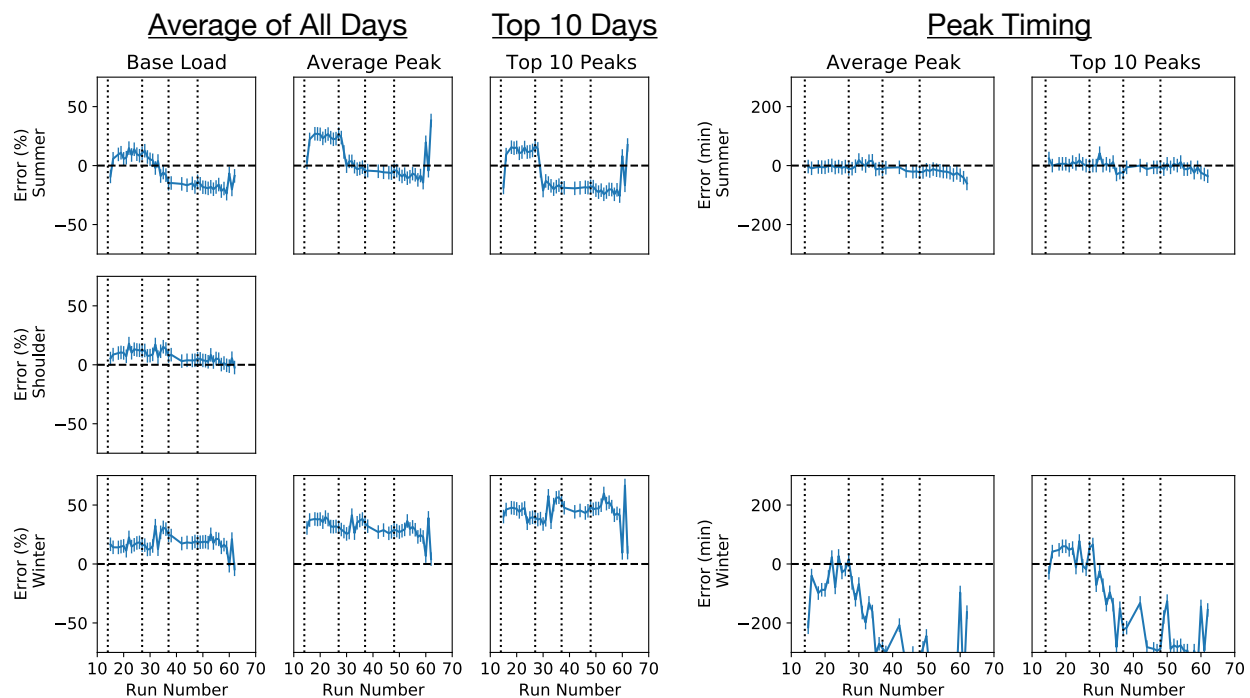
Figure 185 shows the timeseries magnitude and timing QOIs for Cherryland, MI throughout the calibration process. The Cherryland, MI comparisons likely benefited from the nationally relevant changes leading up to calibration Region 5. The starting point for the QOIs for Cherryland, MI was relatively close to the AMI data compared to earlier regions. The season average daily minimum load is underestimated by ResStock. As in other regions, we suspect that this missing load is HVAC related.



**Figure 177. ResStock annual electricity per dwelling unit for each calibration region for each model update during the calibration process compared to the region's LRD or AMI data. Table 24 lists the changes made in each run number. Vertical dotted lines indicate transitions between regional phases.**

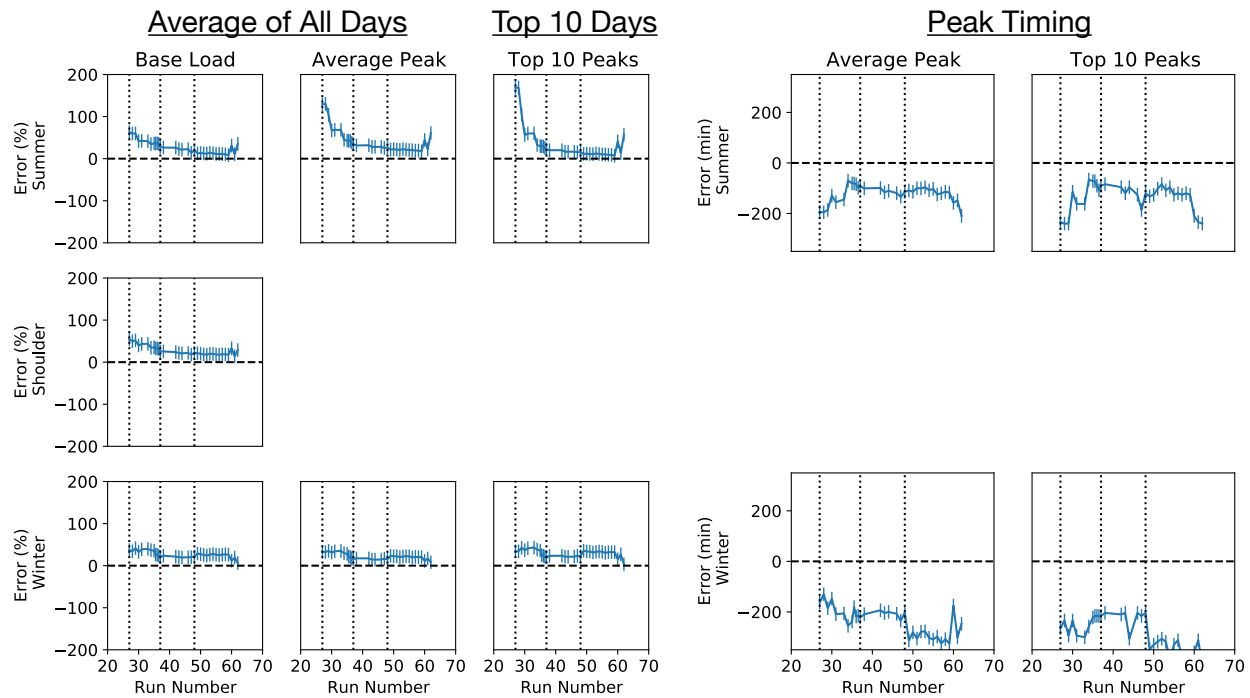


**Figure 178. ResStock timeseries QOIs for ComEd for each model update during the calibration process compared to the 2018 LRD. Table 24 lists the changes made in each run number. Vertical dotted lines indicate transitions between regional phases.**

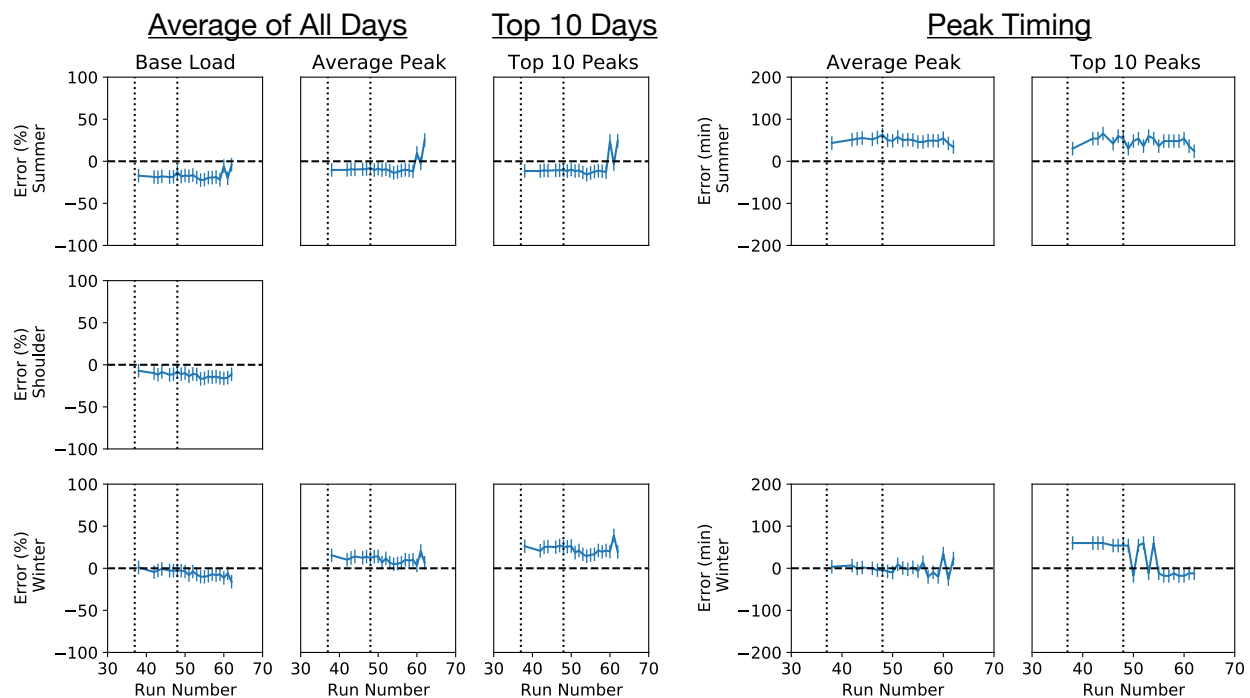


**Figure 179. ResStock timeseries QOIs for the City of Fort Collins for each model update during the calibration process compared to the 2018 AMI data. Table 24 lists the changes made in each run number. Vertical dotted lines indicate transitions between regional phases.**

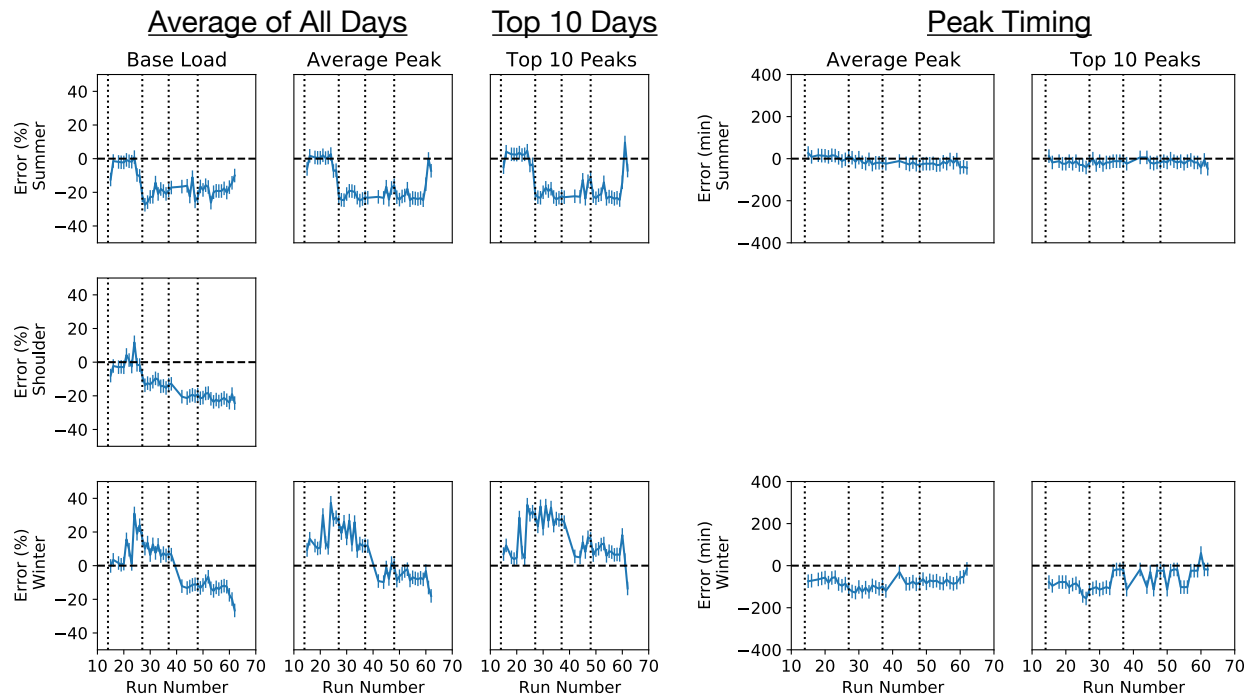




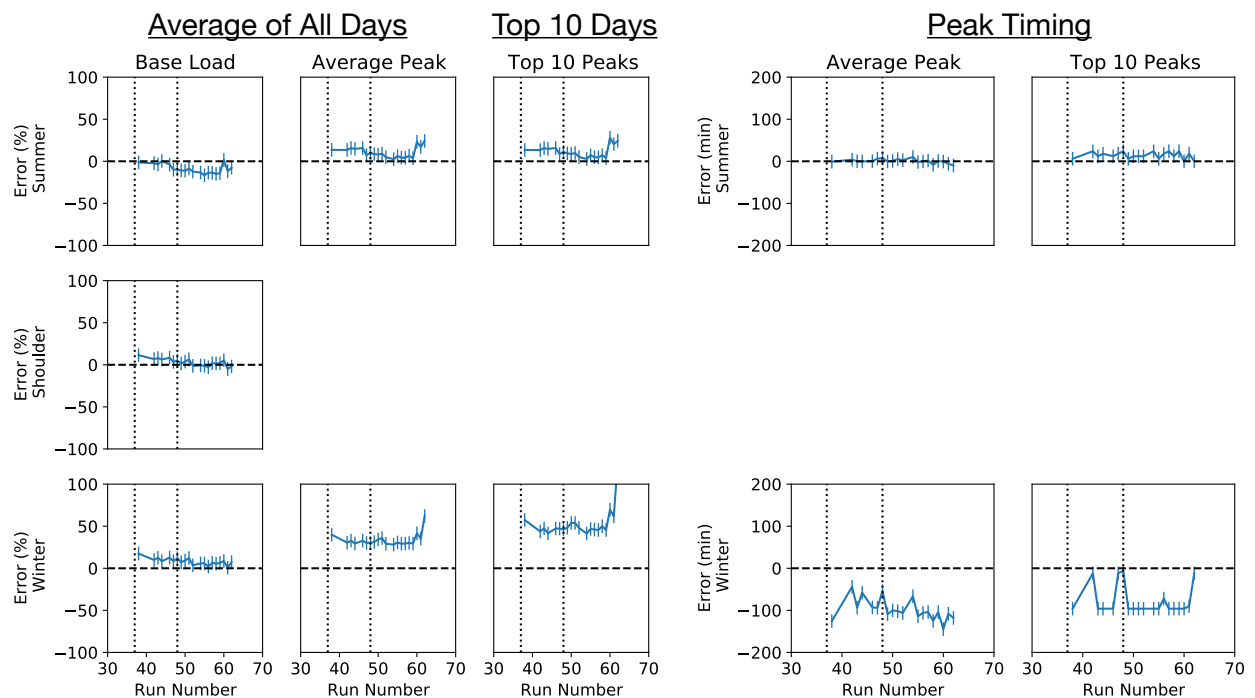
**Figure 180. ResStock timeseries QOIs for Seattle City Light for each model update during the calibration process compared to the 2019 AMI data. Table 24 lists the changes made in each run number. Vertical dotted lines indicate transitions between regional phases.**



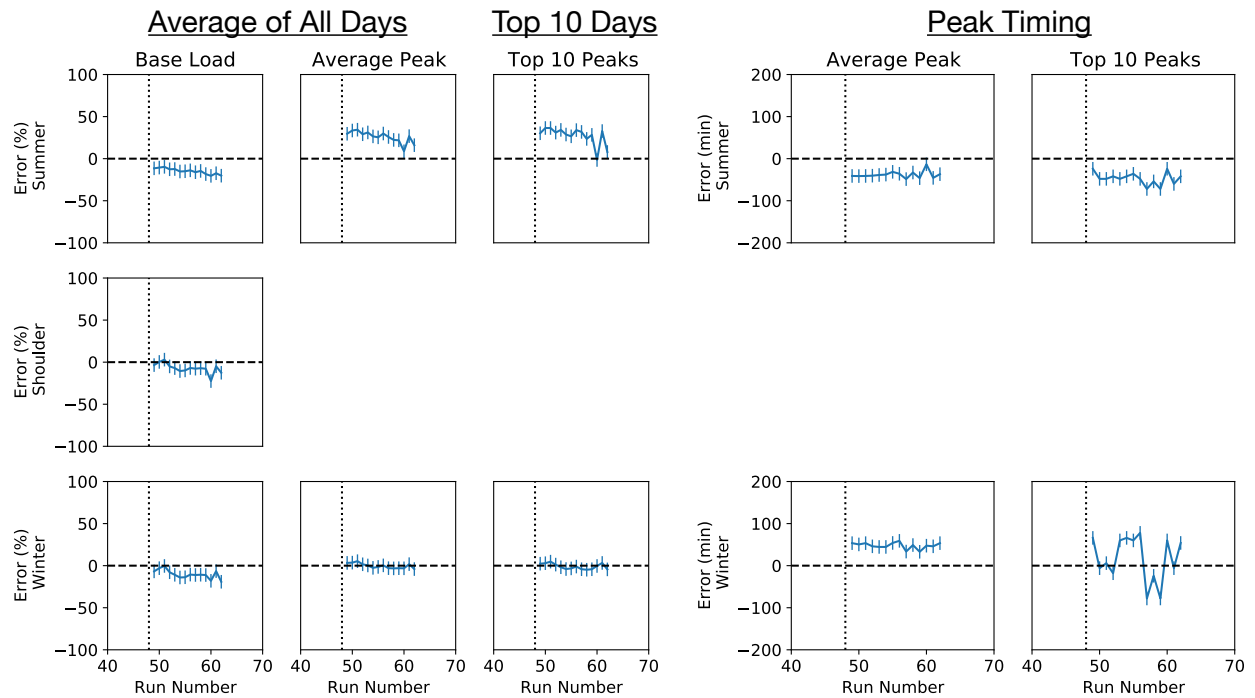
**Figure 181. ResStock timeseries QOIs for EPB Chattanooga for each model update during the calibration process compared to the 2019 AMI data. Table 24 lists the changes made in each run number. Vertical dotted lines indicate transitions between regional phases.**



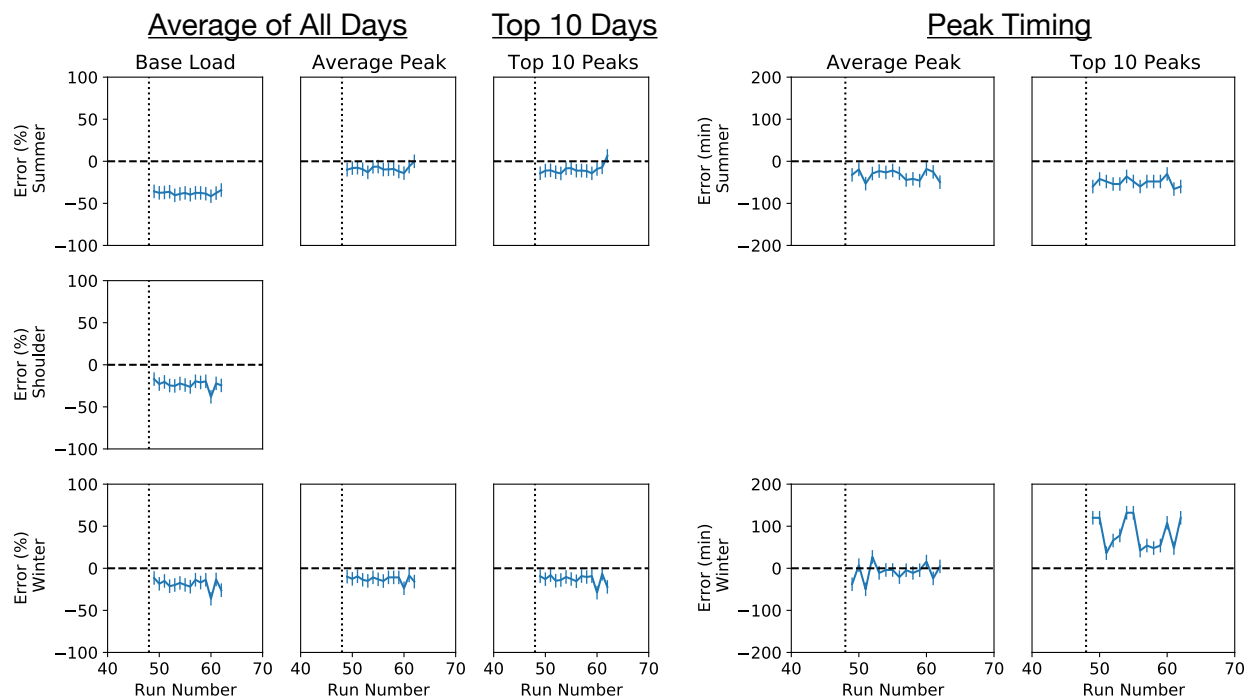
**Figure 182. ResStock timeseries QOIs for Horry Electric Cooperative for each model update during the calibration process compared to the 2018 AMI data. Table 24 lists the changes made in each run number. Vertical dotted lines indicate transitions between regional phases.**



**Figure 183. ResStock timeseries QOIs for the City of Tallahassee for each model update during the calibration process compared to the 2019 AMI data. Table 24 lists the changes made in each run number. Vertical dotted lines indicate transitions between regional phases.**



**Figure 184. ResStock timeseries QOIs for the data from VEIC for each model update during the calibration process compared to the 2018 AMI data. Table 24 lists the changes made in each run number. Vertical dotted lines indicate transitions between regional phases.**



**Figure 185. ResStock timeseries QOIs for Cherryland Electric Co-op for each model update during the calibration process compared to the 2019 AMI data. Table 24 lists the changes made in each run number. Vertical dotted lines indicate transitions between regional phases.**

### Residential QOI Summary

The final state of all the QOIs is presented in Figures 186, 187, and 188. The figures show both the magnitude QOIs and timing QOIs. For the magnitude QOIs, both the absolute discrepancy in kWh/unit and relative discrepancy in percent are shown. When interpreting the accuracy of ResStock compared to the QOIs, a few considerations should be kept in mind. For the shoulder season, the peak magnitude and peak timing metrics are less meaningful than summer and winter peak magnitude and timing. The shoulder seasons was specifically chosen for each region to minimize the load and extract as much baseload as possible. The season average daily baseloads are small in magnitude, so the relative percent discrepancies may be large in some regions, but when looking at the absolute discrepancy in kWh/unit the discrepancies are much smaller.

Accurately estimating the peak loads for summer and winter is more critical in regions where the heating and cooling loads are large. Estimating cooling peak loads is more critical in ComEd, Fort Collins, Chattanooga, Tallahassee, and Horry County. Estimating heating peak loads is more critical in Seattle, Chattanooga, Tallahassee, and Horry. These climates are not necessarily cold climates, but have a substantial number of dwelling unit with electric heating. The largest discrepancies are in winter for Tallahassee, FL. As was described above, before the EIA-861M correction model was applied, ResStock overestimated the winter peak load in Tallahassee, FL. Because ResStock somewhat underestimates the Florida state total load during the winter months, the correction model pushes the Tallahassee winter peaks in the wrong direction. Finally, the peak timing metrics do not necessarily give a good indication of the accuracy in areas like Fort Collins, Seattle, Vermont, and Cherryland as the morning and evening peaks are approximately equal. Averaging the peak times shifts the peak time towards the middle of the day and can cause large timing discrepancies.

	qoi_type	average_daily_base			average_daily_peak			top_10_average_daily_peak		
	season	shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region 1	ComEd, IL	-0.2	-0.2	-0.3	-0.0	0.1	-0.1	0.1	0.0	-0.3
region 2	Fort Collins, CO	-0.0	-0.1	-0.0	0.6	0.6	0.1	0.8	0.4	0.3
region 3	Seattle, WA	0.1	0.1	0.0	0.4	0.5	0.1	0.4	0.5	0.1
region 4a	Chattanooga, TN	-0.1	-0.0	-0.2	0.5	0.7	0.1	1.0	0.9	0.7
region 4b	Tallahassee, FL	-0.0	-0.1	0.1	0.6	0.5	1.0	1.6	0.7	3.1
region 4c	Horry County, SC	-0.2	-0.1	-0.3	-0.2	-0.2	-0.4	0.0	-0.3	-0.7
region 5a	Cherryland, MI	-0.2	-0.2	-0.2	-0.0	0.1	-0.2	0.3	0.1	-0.4
region 5b	State of VT	-0.1	-0.2	-0.1	0.2	0.2	-0.0	0.3	0.1	-0.1

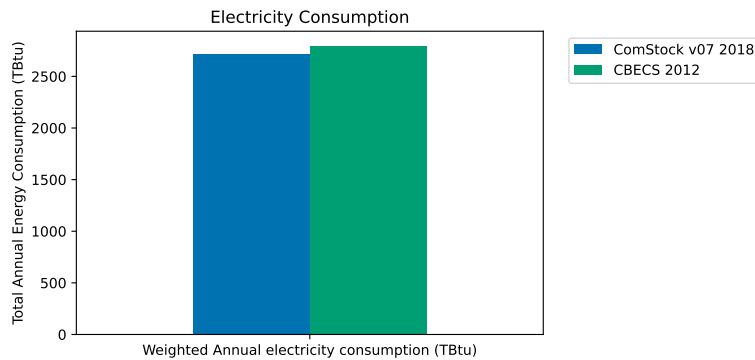
Figure 186. Magnitude QOI discrepancies for the total residential building stock in kW/unit

	qoi_type	average_daily_base			average_daily_peak			top_10_average_daily_peak		
	season	shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region 1	ComEd, IL	-29%	-27%	-30%	-1%	7%	-11%	4%	1%	-18%
region 2	Fort Collins, CO	-7%	-9%	-2%	51%	36%	8%	47%	18%	20%
region 3	Seattle, WA	29%	35%	6%	46%	61%	8%	35%	57%	3%
region 4a	Chattanooga, TN	-11%	-4%	-16%	32%	26%	7%	47%	24%	19%
region 4b	Tallahassee, FL	-2%	-8%	7%	41%	24%	62%	81%	24%	128%
region 4c	Horry County, SC	-24%	-12%	-28%	-10%	-8%	-17%	1%	-8%	-14%
region 5a	Cherryland, MI	-34%	-35%	-26%	-3%	6%	-15%	22%	6%	-21%
region 5b	State of VT	-16%	-27%	-17%	20%	18%	-2%	22%	8%	-4%

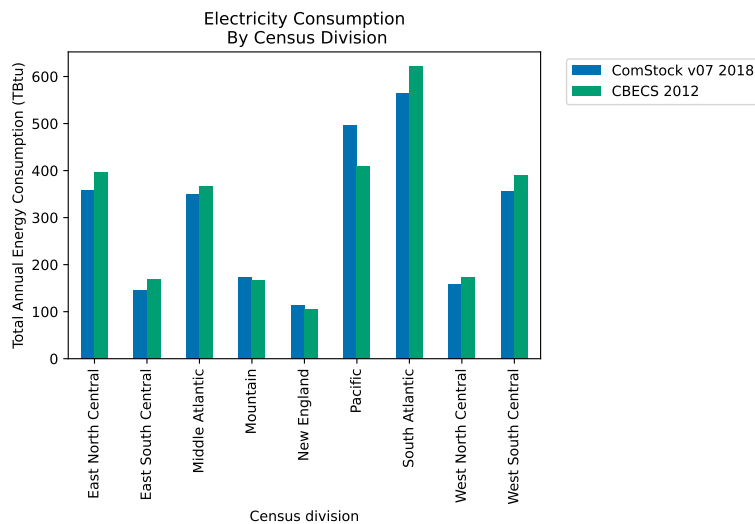
Figure 187. Magnitude QOI discrepancies for the total residential building stock relative to the meter data

	qoi_category	timing					
	qoi_type	average_daily_peak			top_10_average_daily_peak		
	season	shoulder	summer	winter	shoulder	summer	winter
region	region_name						
region 1	ComEd, IL	-0.2	-0.7	0.5	-0.5	-0.1	1.0
region 2	Fort Collins, CO	-1.5	-1.0	-3.3	-0.7	-0.8	-1.8
region 3	Seattle, WA	1.9	-3.5	-4.1	-2.3	-4.0	-6.8
region 4a	Chattanooga, TN	-0.2	0.6	0.4	2.2	0.4	-0.2
region 4b	Tallahassee, FL	-1.8	-0.2	-2.0	-4.9	0.0	-0.2
region 4c	Horry County, SC	-2.4	-0.6	-0.5	-2.5	-0.6	-0.3
region 5a	Cherryland, MI	-1.2	-0.7	-0.7	-1.9	-1.0	2.2
region 5b	State of VT	-0.9	-1.1	0.9	-1.8	-0.4	0.6

Figure 188. Timing QOI discrepancy (hours) for the total residential building stock



**Figure 189. ComStock compared to CBECS 2012 annual national electricity consumption**



**Figure 190. ComStock compared to CBECS 2012 annual electricity consumption by census division**

## 4.2 Final Calibration and Validation Results—Commercial

### 4.2.1 Commercial EIA Data Validation

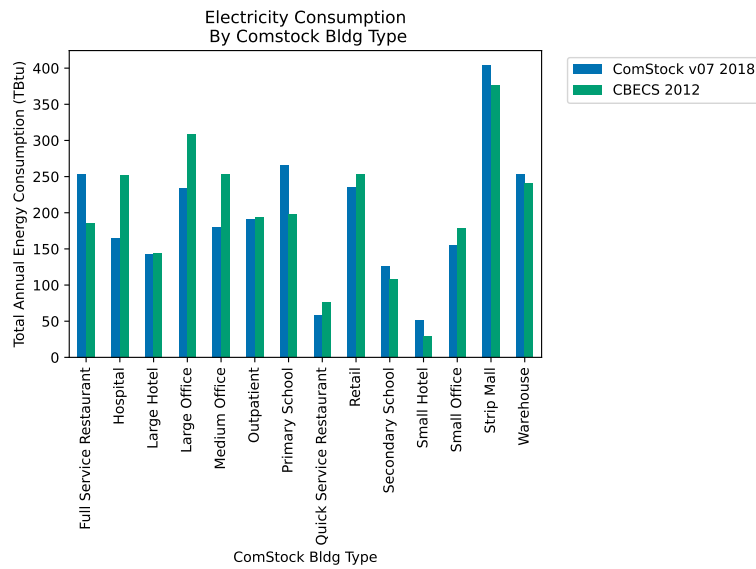
#### *Commercial EIA CBECS*

The comparison of ComStock to CBECS must start by stating the fact that the latest available CBECS energy consumption data is from the 2012 survey, and the ComStock model represents the building stock circa 2018. We know from looking at the 2014 and 2019 CBSE surveys of the building stock in the Pacific Northwest that it is possible for major change (roughly a 30% decrease in lighting power density) to happen in the lighting end use during this time span. We do not know if this change applies nationally, but in general we would expect the ComStock annual totals on a per-area basis to be lower than CBECS. Because ComStock uses CBECS to scale the total national square footage of each building type, by definition the floor areas between the two datasets are identical at the national level, so the following comparisons should be interpreted as differences primarily on a per-area basis. After the 2018 CBECS data are released, a new comparison should be performed.

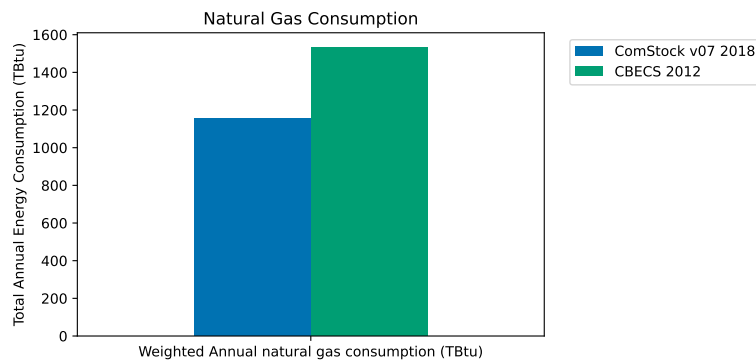
#### **Electricity**

Looking at the national electricity consumption total in Figure 189, we see that ComStock model results are close to CBECS. Looking at the breakdown in Figure 190, we see that ComStock overestimates most in the Pacific census division and underestimates most in the South Atlantic census division.





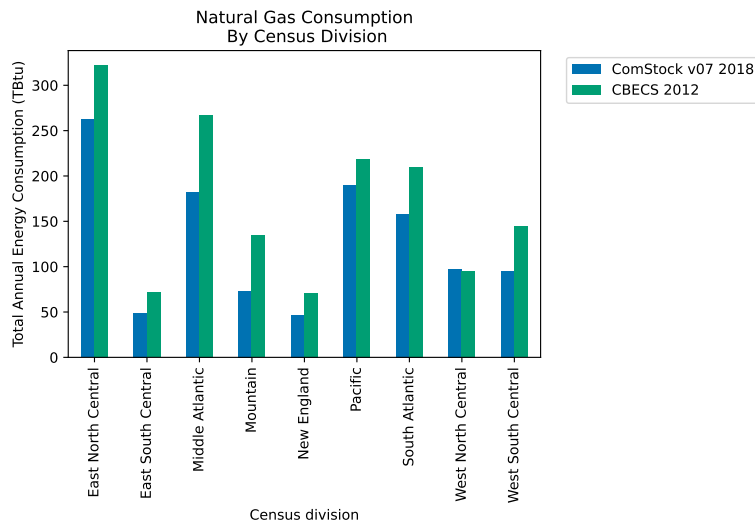
**Figure 191. ComStock compared to CBECS 2012 annual electricity consumption by building type**



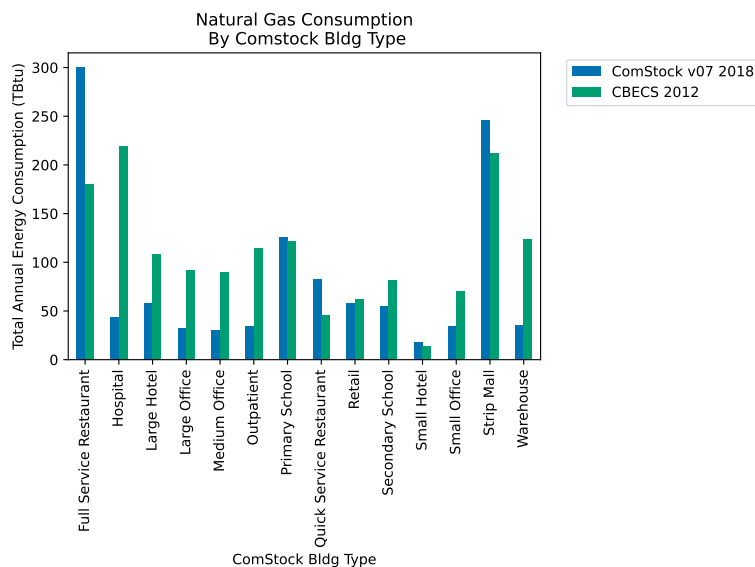
**Figure 192. ComStock compared to CBECS 2012 annual national natural gas consumption**

## Natural Gas

Looking at the national natural gas consumption total in Figure 192, we see that the ComStock is about 20% below the CBECS total. Looking at the breakdown in Figure 193, we see this underestimate reflected across all census divisions except West North Central. Perhaps the most illuminating view is the building-type comparison in Figure 194, which shows that the estimate is not uniform, but varies widely between building types. The overestimate in full-service restaurants and strip malls (which are modeled with a portion of full-service restaurant space), indicates a likely overestimate in either cooking or ventilation loads in those building types.



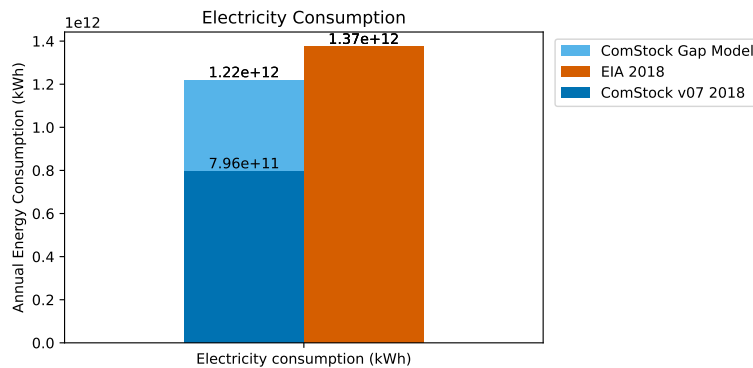
**Figure 193. ComStock compared to CBECS 2012 annual natural gas consumption by census division**



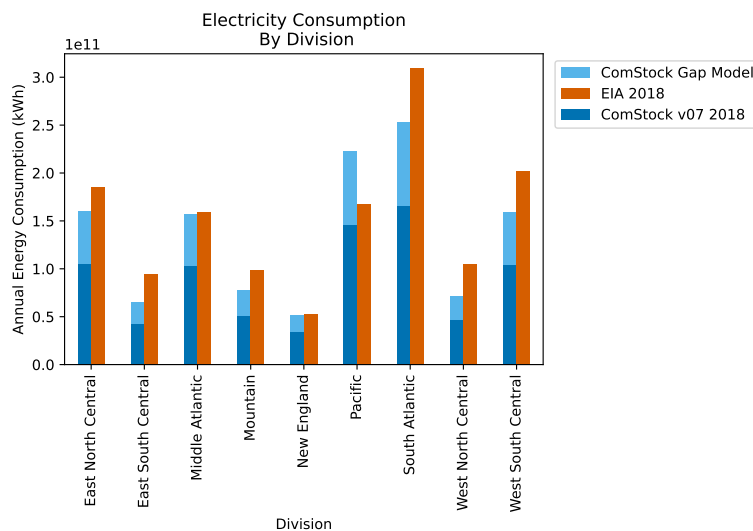
**Figure 194. ComStock compared to CBECS 2012 annual natural gas consumption by building type**

#### *Commercial Form EIA-861M and Form EIA-176*

These EIA datasets and the ComStock Gap Model are described in more detail in Section 2.3.7, but one important note is that these data are reported by utilities, who are responsible for making the distinction between commercial and industrial customers. Although EIA provides guidance to utilities, based on discussion with EIA staff, these datasets likely include energy consumption from buildings that we would consider industrial. Importantly, for this comparison, both the EIA datasets and the ComStock model are from 2018, so the time lag difference found in the CBECS-to-EIA comparisons are not found. However, because the national square footage scaling of ComStock is based on the 2012 CBECS, any new construction or demolition that occurred during the 2012–2018 period is not reflected, likely meaning that ComStock underestimates the square footage for some building types. After the 2018 CBECS data are released, a new comparison should be performed.



**Figure 195. ComStock compared to EIA-861M annual electricity consumption**



**Figure 196. ComStock compared to EIA-861M annual electricity consumption by census division**

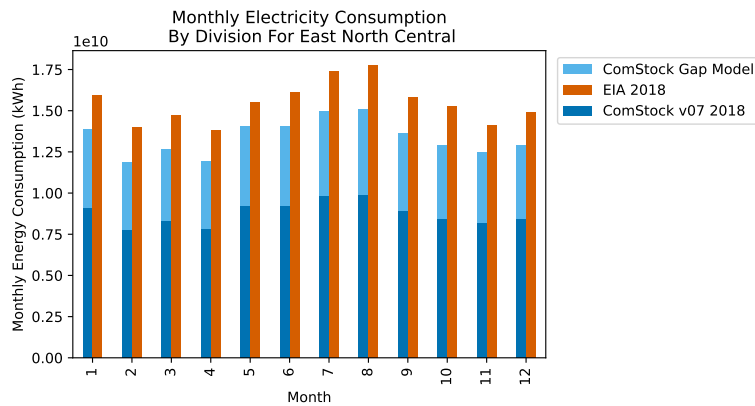
As described in Section 4.1.1, there is some time shift in the EIA data caused by the difference between utility billing systems and the calendar months used to report the data. In the residential comparisons, this shift was corrected for. In the commercial comparisons, the raw (uncorrected) monthly EIA data are shown.

### Electricity

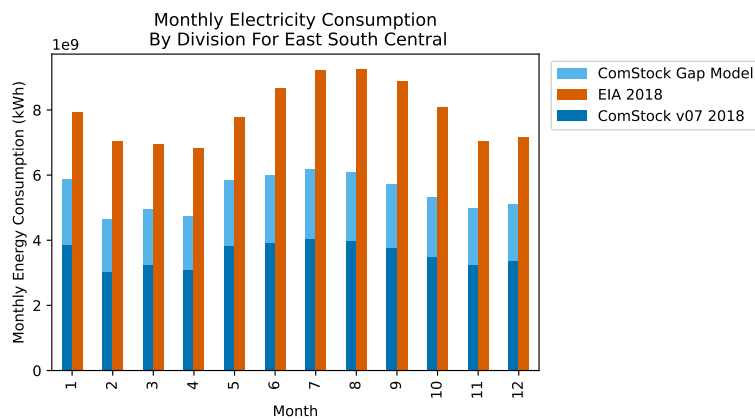
Looking at the national electricity consumption total in Figure 195, we see that ComStock plus the ComStock Gap Model slightly underestimates the annual consumption. Looking at the breakdown in Figure 196, we see that ComStock overestimates most in the Pacific census division and underestimates most in the South Atlantic census division, matching the pattern found in the CBECS comparison in Figure 190. Looking at the monthly consumption patterns in Figures 197 through 205, we see that ComStock generally follows the monthly contours of the EIA data well.

### Natural Gas

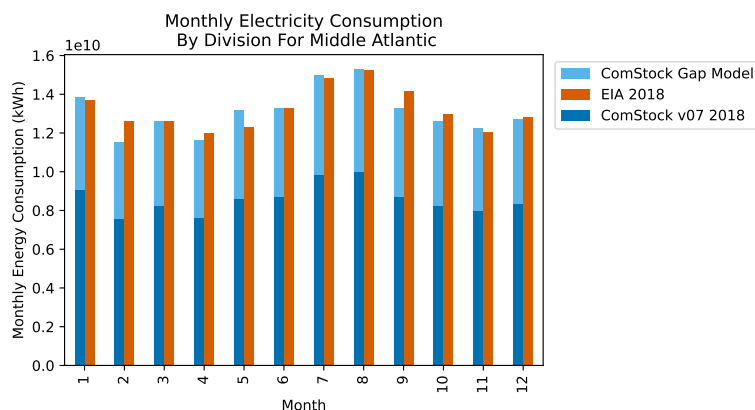
Looking at the national natural gas consumption total in Figure 206, we see that ComStock plus the ComStock Gap Model significantly underestimates annual consumption. Looking at the breakdown in Figure 207, we see that ComStock underestimates most significantly in the coldest climates, matching the pattern found in the CBECS comparison in Figure 193. Looking at the monthly consumption patterns in Figures 208 through 216, we see that ComStock underestimates in both the summer and winter months, indicating that the issue is underestimation in



**Figure 197. ComStock compared to EIA-861M monthly electricity consumption by division for East North Central**



**Figure 198. ComStock compared to EIA-861M monthly electricity consumption by division for East South Central**



**Figure 199. ComStock compared to EIA-861M monthly electricity consumption by division for Middle Atlantic**

both base load (primarily composed of water heating and cooking) and the weather-dependent heating load. Further analysis and refinement of these loads should be undertaken before analyses looking at measures that might impact natural gas consumption are evaluated, particularly any evaluation of load electrification measures.

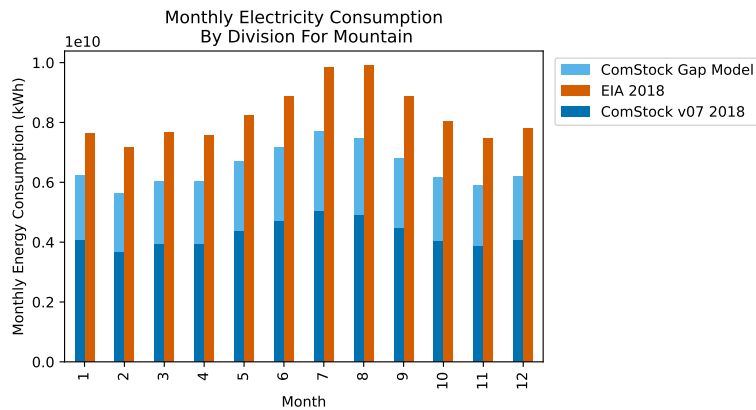


Figure 200. ComStock compared to EIA-861M monthly electricity consumption by division for Mountain

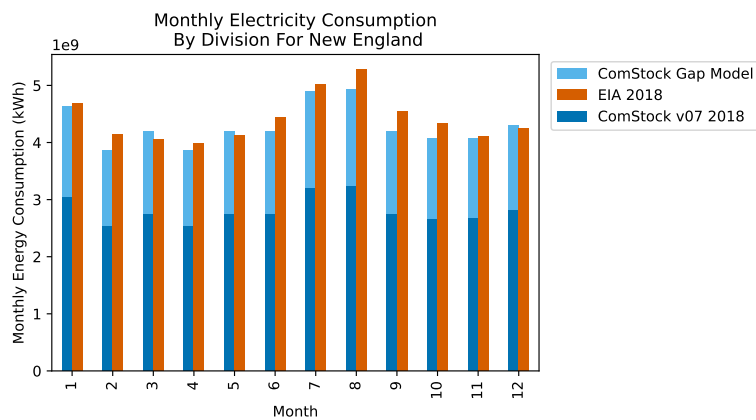


Figure 201. ComStock compared to EIA-861M monthly electricity consumption by division for New England

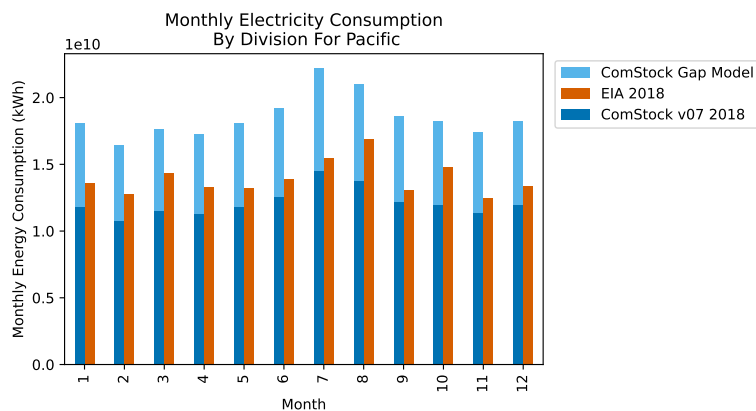


Figure 202. ComStock compared to EIA-861M monthly electricity consumption by division for Pacific

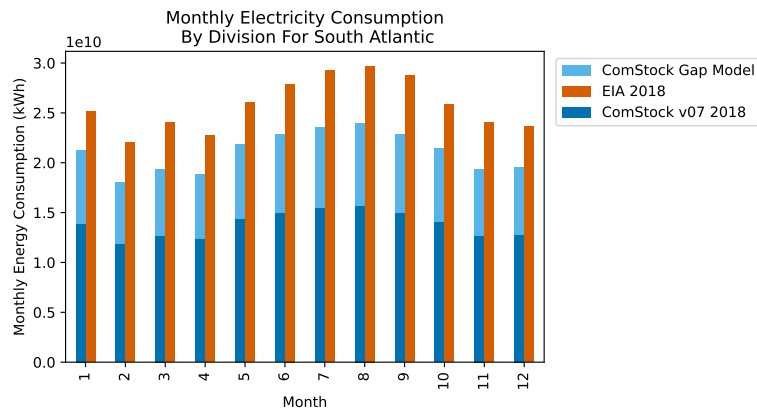


Figure 203. ComStock compared to EIA-861M monthly electricity consumption by division for South Atlantic

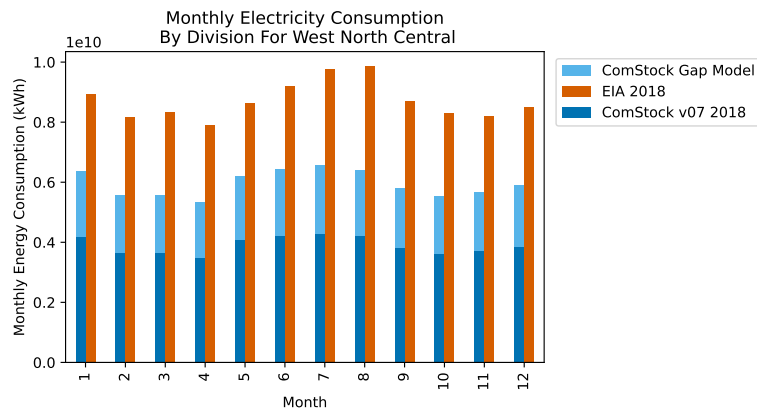


Figure 204. ComStock compared to EIA-861M monthly electricity consumption by division for West North Central

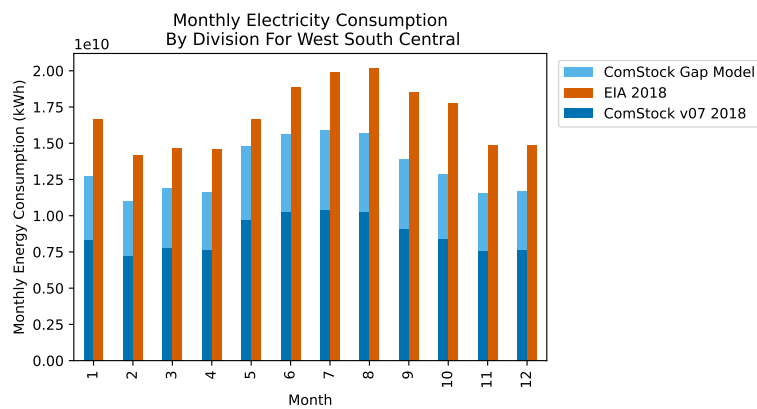
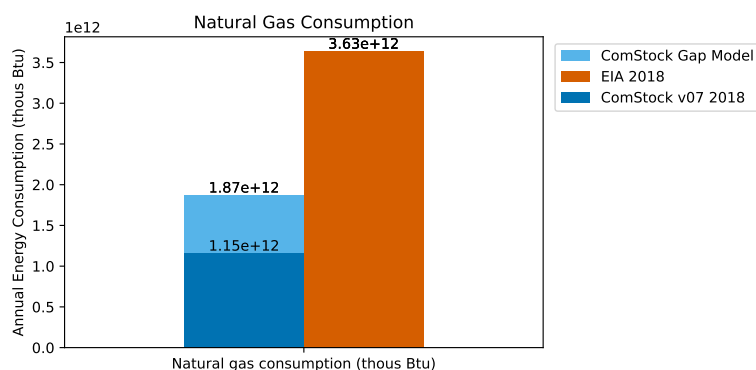
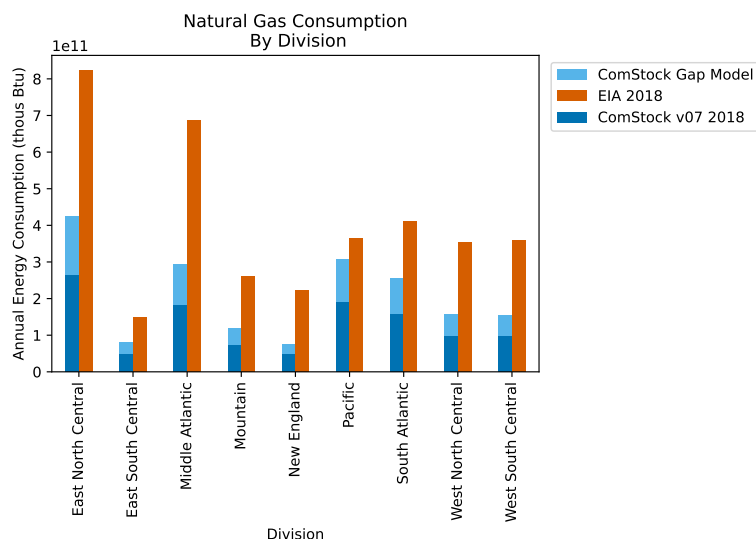


Figure 205. ComStock compared to EIA-861M monthly electricity consumption by division for West South Central

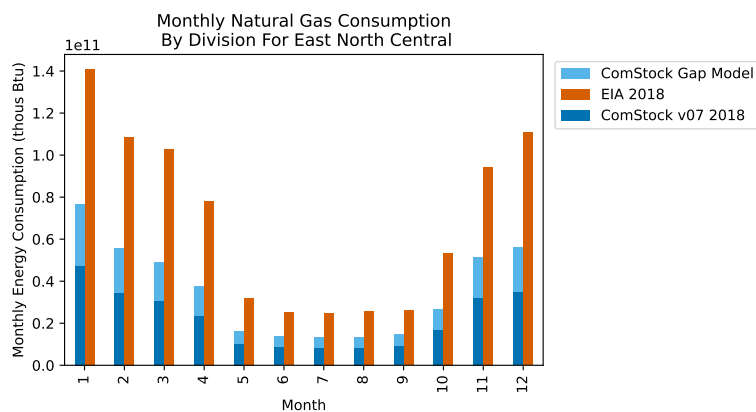




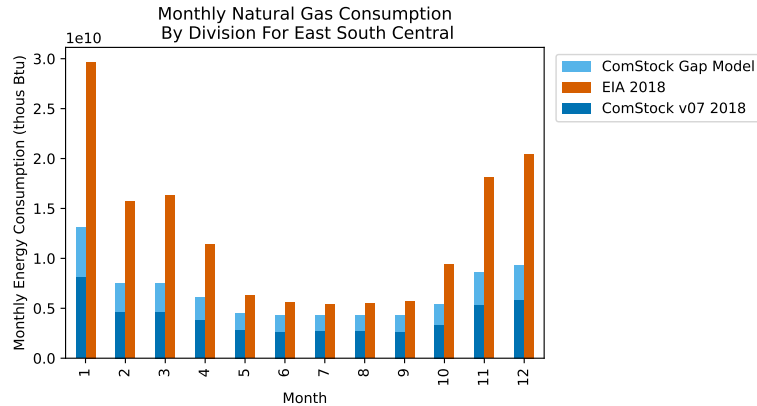
**Figure 206. ComStock compared to EIA-176 natural gas consumption**



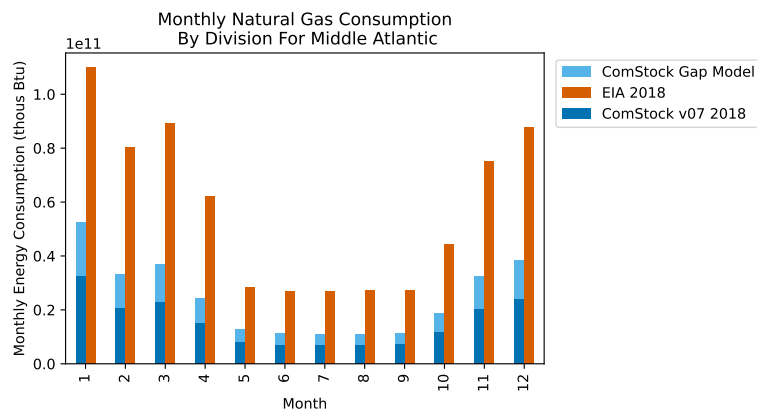
**Figure 207. ComStock compared to EIA-176 natural gas consumption by division**



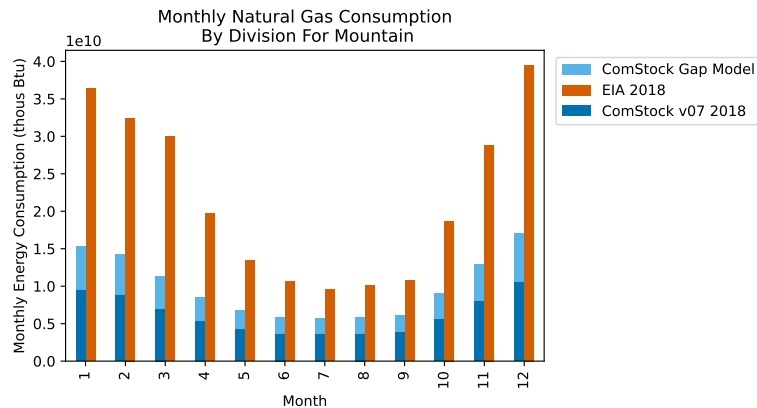
**Figure 208. ComStock compared to EIA-176 monthly natural gas consumption by division for East North Central**



**Figure 209. ComStock compared to EIA-176 monthly natural gas consumption by division for East South Central**



**Figure 210. ComStock compared to EIA-176 monthly natural gas consumption by division for Middle Atlantic**



**Figure 211. ComStock compared to EIA-176 monthly natural gas consumption by division for Mountain**

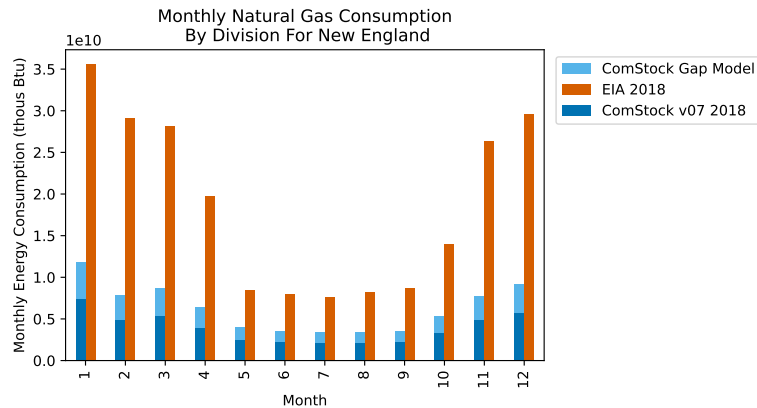


Figure 212. ComStock compared to EIA-176 monthly natural gas consumption by division for New England

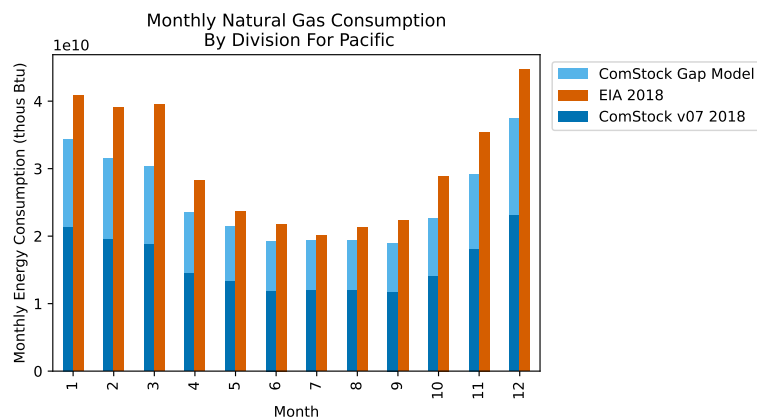


Figure 213. ComStock compared to EIA-176 monthly natural gas consumption by division for Pacific

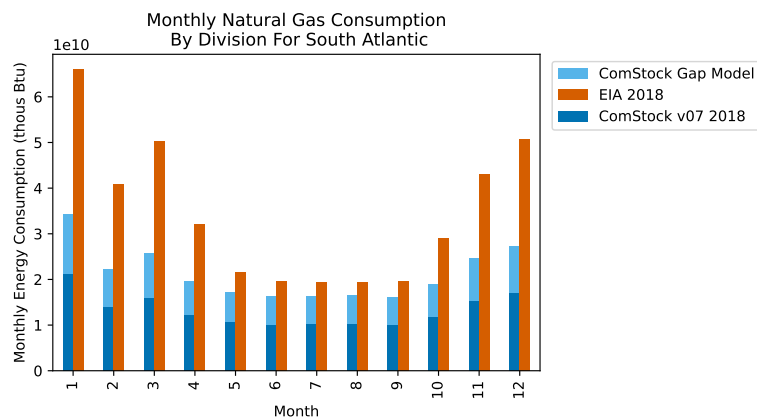


Figure 214. ComStock compared to EIA-176 monthly natural gas consumption by division for South Atlantic

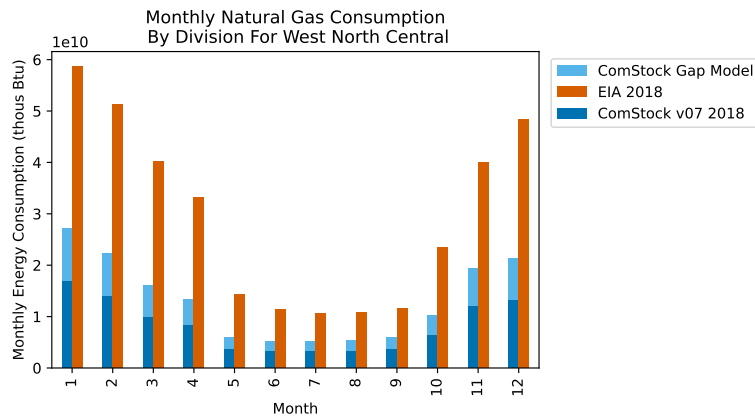


Figure 215. ComStock compared to EIA-176 monthly natural gas consumption by division for West North Central

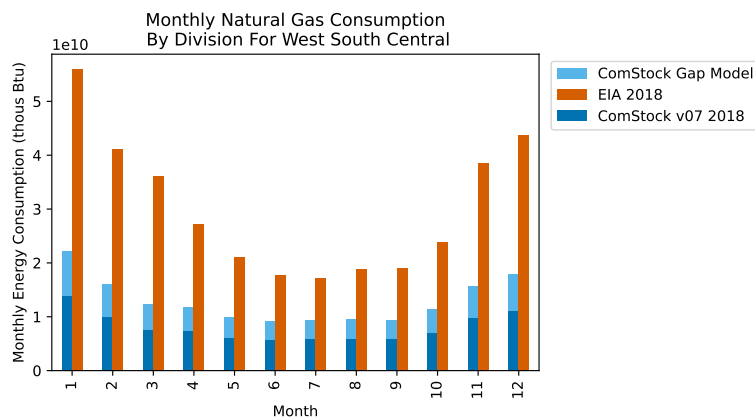


Figure 216. ComStock compared to EIA-176 monthly natural gas consumption by division for West South Central

#### 4.2.2 Commercial AMI Data Validation

Evaluating the difference between datasets requires an understanding of the details of the datasets. As described in Section 2.3.5, the commercial AMI data collection process was both complex and significantly impacted by COVID-19. Throughout the data collection and cleaning process, a variety of issues were encountered and many were resolved. The known limitations of the AMI are described below.

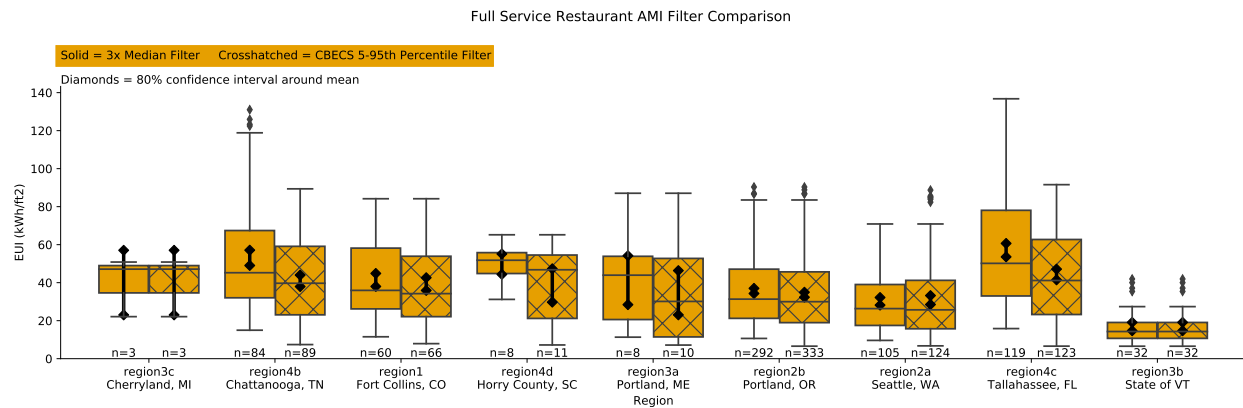
##### *Biases in Commercial AMI Data*

For electric utilities, the fundamental unit of record is an electricity meter. Other than a service address where the meter is physically located and the rate class associated with this meter, utilities have very little information about the load being served. Rate classes typically include a few residential rates, a few commercial and industrial rates based on peak power draw, and a few other miscellaneous rates for things such as street lighting and irrigation. For residential buildings, this presents less of a challenge because residential dwelling units, and in particular single-family units, are typically served by a single meter. As described in Section 2.3.5, in order to create a meaningful comparison for commercial buildings, it was necessary to assign all meters to a specific building, aggregate all the meters assigned to each building, and then normalize the energy consumption based on building characteristics. The potential biases added by each step in the process are discussed below, along with our associated recommendations for future studies.

*Building characteristics* such as building type and floor area are critical for evaluating the energy consumption of a commercial building. Because the utilities had no information about the buildings served by their meters, our first task was to find this information. As discussed in Section 2.3.5, our source of building characteristics was the CoStar commercial real-estate database. The main purpose of this database is to facilitate real-estate transactions and investment. As such, the building types assigned to buildings in the database are geared toward this purpose, not toward building energy. For many of the CoStar building types, there was an obvious mapping to a building-energy-focused building type (e.g., a school in CoStar is easily mapped to a school for building energy modeling). However, for other building types, the mapping was less definite (e.g., a restaurant in CoStar could represent either a quick-service or full-service establishment). The choice of mappings was based on a manual review of a sample of CoStar using information readily available online. However, there could be biases such as CoStar agent or regional differences that were not identified during this review. Additionally, the building usage may have had tenancy changes between when the data were last updated and the period represented by the AMI data. The building type distinctions we saw the most ambiguity in were full-service vs. quick-service restaurants, strip mall vs. retail, and warehouse. For future studies, it may be worthwhile to assign a building-energy-focused building type to each building.

*Meter matching* is the process of matching a utility meter with the building it serves. As discussed in Section 2.3.5, this was done by matching the meter service addresses to the building addresses in the CoStar database. The first requirement for a meter to be matched to a building was for the building to exist in the CoStar database. While we found the database to be broadly comprehensive, it did have some limitations. First, it only contained buildings that are typically bought and sold, so publicly owned buildings were less well represented. Second, it was more comprehensive in places with active real-estate markets, so rural areas were less well represented. Third, it was more comprehensive for buildings more likely to be valuable assets, so there was some bias toward buildings located in shopping centers or other desirable locations. For future studies, it may be worthwhile to create a more comprehensive database of buildings by combining data sources such as local tax-assessor and public building inventories to use as a starting point for meter assignment. The feasibility of this task will likely depend on the geographic extent of the AMI data to be assigned.

The second requirement for a meter to be matched to a building was that both addresses matched. Prior to matching, addresses from both datasets were standardized to use common abbreviations (e.g., “St.,” “ST,” and “Street” were all standardized to “St”). We found that many utilities appended additional information useful to a person looking for the physical meter to the end of the address (e.g., “123 Main St Shed”). This additional information was removed prior to matching. In every dataset, some region-specific quirks were identified and the matching process was improved. One of the issues identified but not resolved was the problem of different datasets using different street names (e.g., “Broadway” in one dataset and “State Road 55” in the other). For future studies, it may be desirable to use some robust geocoding-based process to match building addresses to meter addresses. This will likely depend on the willingness of the utilities to share the meter addresses.



**Figure 217. Comparison of AMI EUI distributions using different filtering methods**

*Misclassification and outlier removal* is the process of attempting to find buildings that aren't actually of the type initially assigned based on their normalized energy consumption. The details of the process we developed and the methods for testing this process are documented in a forthcoming journal article. In summary, we selected a method that chose to reject buildings whose EUI fell outside the range of 0.33-times to 3-times the median for buildings of the same type. This method prioritized removing incorrectly classified buildings at the expense of removing buildings with unusually high or low energy consumption. While this method is good at removing outliers at the tails of the distribution, it cannot find systematic issues where the building area or the number of meters matched to a building is incorrect by a small amount.

Figure 217 compares the selected approach with an approach that removes buildings below the 5th percentile or above the 95th percentile found for that building type in the CBECS 2012 dataset. As shown, in some cases, such as Region 2b, both methods have a similar result, but in other cases such as Region 4b, even with a relatively high number of samples, the inclusion or exclusion of a few high-EUI points can shift the mean significantly. Overall, we observed that **the selected AMI filtering approach systematically shifts the mean of the AMI lower than what the full population would show**, and that **this effect is more pronounced when the number of AMI datapoints is low**.

In order to check the reasonableness of the AMI after the misclassification and outlier removal process was applied, we sought other points of comparison. The most geographically comprehensive dataset that includes information about both building type and energy consumption was the EIA CBECS survey. This dataset has the benefits of having been conducted to gather building energy information, so building type classification is well understood and believed to be reliable. Additionally, the energy consumption data associated with each building went through a thorough review process. As a point of comparison for the AMI, one downside of CBECS is that the latest data are from 2012, whereas the AMI data are primarily from 2019. Based on a comparison of the regional CBSA from 2014 and 2019, we know that some major end uses such as lighting have likely decreased between 2012 and 2018, so in general we would expect the AMI data to be below the CBECS data. Additionally, the CBECS data represents an entire census division, whereas the AMI data typically represents a concentrated subset of that census division, so intra-census-division differences could explain some difference between the datasets.

To evaluate the similarity between CBECS and the AMI, a Kolmogorov–Smirnov (K-S) test was used to compare the EUI distributions by building type for each region, where CBECS data were taken from the same census division as the AMI. Because of the laxity of the K-S test at low sample sizes, we did not perform the comparison where the number of samples from either dataset was less than 10. Figure 218 shows the results of the K-S test. The distance metric reported in the figure is the K-S statistic. This metric is reported directly to allow users to evaluate the distance for their own use cases.

<b>K-S Test Matrix</b> Target metric = distance AMI filter = LowEnd+3xMed CBECs weight = False	Small Office	Medium Office	Large Office	Strip Mall	Retail	Warehouse	Full Service Restaurant	Quick Service Restaurant	Small Hotel	Large Hotel	Outpatient
<b>region 1:</b> Fort Collins, CO	Distance = 0.13 AMI = 313 CBECs = 46	Distance = 0.2 AMI = 23 CBECs = 12	Low Sample AMI = 4 CBECs = 18	Distance = 0.4 AMI = 156 CBECs = 21	Distance = 0.26 AMI = 126 CBECs = 23	Distance = 0.43 AMI = 112 CBECs = 47	Distance = 0.42 AMI = 61 CBECs = 17	Low Sample AMI = 26 CBECs = 6	Low Sample AMI = 5 CBECs = 7	Low Sample AMI = 8 CBECs = 13	Distance = 0.52 AMI = 78 CBECs = 30
<b>region 2a:</b> Seattle, WA	Distance = 0.28 AMI = 480 CBECs = 95	Distance = 0.42 AMI = 64 CBECs = 60	Distance = 0.19 AMI = 105 CBECs = 43	Distance = 0.63 AMI = 561 CBECs = 52	Distance = 0.21 AMI = 304 CBECs = 57	Distance = 0.25 AMI = 410 CBECs = 163	Distance = 0.2 AMI = 107 CBECs = 40	Distance = 0.21 AMI = 26 CBECs = 18	Distance = 0.33 AMI = 19 CBECs = 13	Distance = 0.33 AMI = 25 CBECs = 26	Distance = 0.46 AMI = 105 CBECs = 46
<b>region 2b:</b> Portland, OR	Distance = 0.12 AMI = 250 CBECs = 95	Distance = 0.4 AMI = 10 CBECs = 60		Distance = 0.66 AMI = 889 CBECs = 52	Distance = 0.22 AMI = 926 CBECs = 57	Distance = 0.29 AMI = 1938 CBECs = 163	Distance = 0.2 AMI = 308 CBECs = 40	Distance = 0.27 AMI = 123 CBECs = 18	Distance = 0.33 AMI = 54 CBECs = 13	Distance = 0.32 AMI = 79 CBECs = 26	Distance = 0.29 AMI = 456 CBECs = 46
<b>region 3a:</b> Portland, ME	Distance = 0.31 AMI = 15 CBECs = 26	Distance = 0.21 AMI = 24 CBECs = 15	Low Sample AMI = 9 CBECs = 11	Distance = 0.39 AMI = 32 CBECs = 13	Low Sample AMI = 9 CBECs = 14	Distance = 0.36 AMI = 12 CBECs = 23	Low Sample AMI = 8 CBECs = 12				
<b>region 3b:</b> State of Vermont	Distance = 0.25 AMI = 261 CBECs = 26	Distance = 0.28 AMI = 58 CBECs = 15	Low Sample AMI = 3 CBECs = 11	Distance = 0.68 AMI = 151 CBECs = 13	Distance = 0.3 AMI = 56 CBECs = 14	Distance = 0.26 AMI = 158 CBECs = 23	Distance = 0.51 AMI = 32 CBECs = 12	Low Sample AMI = 16 CBECs = 3	Low Sample AMI = 2 CBECs = 1	Low Sample AMI = 107 CBECs = 3	Distance = 0.49 AMI = 35 CBECs = 12
<b>region 3c:</b> Cherryland, MI	Distance = 0.12 AMI = 17 CBECs = 72			Distance = 1.0 AMI = 11 CBECs = 32	Low Sample AMI = 3 CBECs = 52	Distance = 0.25 AMI = 63 CBECs = 78	Low Sample AMI = 3 CBECs = 32		Low Sample AMI = 3 CBECs = 2		Low Sample AMI = 1 CBECs = 46
<b>region 4b:</b> Chattanooga, TN	Distance = 0.12 AMI = 552 CBECs = 38	Distance = 0.55 AMI = 120 CBECs = 12	Low Sample AMI = 33 CBECs = 6	Distance = 0.56 AMI = 437 CBECs = 15	Distance = 0.26 AMI = 323 CBECs = 18	Distance = 0.23 AMI = 490 CBECs = 31	Distance = 0.42 AMI = 86 CBECs = 18	Distance = 0.41 AMI = 108 CBECs = 11	Low Sample AMI = 20 CBECs = 4	Low Sample AMI = 75 CBECs = 9	Distance = 0.46 AMI = 155 CBECs = 14
<b>region 4c:</b> Tallahassee, FL	Distance = 0.36 AMI = 918 CBECs = 126	Distance = 0.37 AMI = 214 CBECs = 45	Distance = 0.32 AMI = 24 CBECs = 84	Distance = 0.44 AMI = 173 CBECs = 89	Distance = 0.15 AMI = 322 CBECs = 59	Distance = 0.21 AMI = 346 CBECs = 148	Distance = 0.27 AMI = 123 CBECs = 42	Distance = 0.14 AMI = 95 CBECs = 28	Distance = 0.52 AMI = 25 CBECs = 14	Distance = 0.41 AMI = 33 CBECs = 38	Distance = 0.38 AMI = 155 CBECs = 44
<b>region 4d:</b> Horry County, SC	Distance = 0.28 AMI = 71 CBECs = 126	Low Sample AMI = 2 CBECs = 45		Distance = 0.5 AMI = 41 CBECs = 89	Distance = 0.25 AMI = 49 CBECs = 59	Distance = 0.24 AMI = 43 CBECs = 148	Low Sample AMI = 8 CBECs = 42	Low Sample AMI = 6 CBECs = 28	Low Sample AMI = 2 CBECs = 14		Low Sample AMI = 7 CBECs = 44

<b>Color Legend</b>	Strongest Agreement Between AMI & CBECs	Weakest Agreement Between AMI & CBECs	Not enough CBECs or AMI to Test Agreement (N < 10)	No AMI
---------------------	--	--	---	--------

Figure 218. Comparison of EUI distributions between regional AMI data and the corresponding census division from CBECs 2012

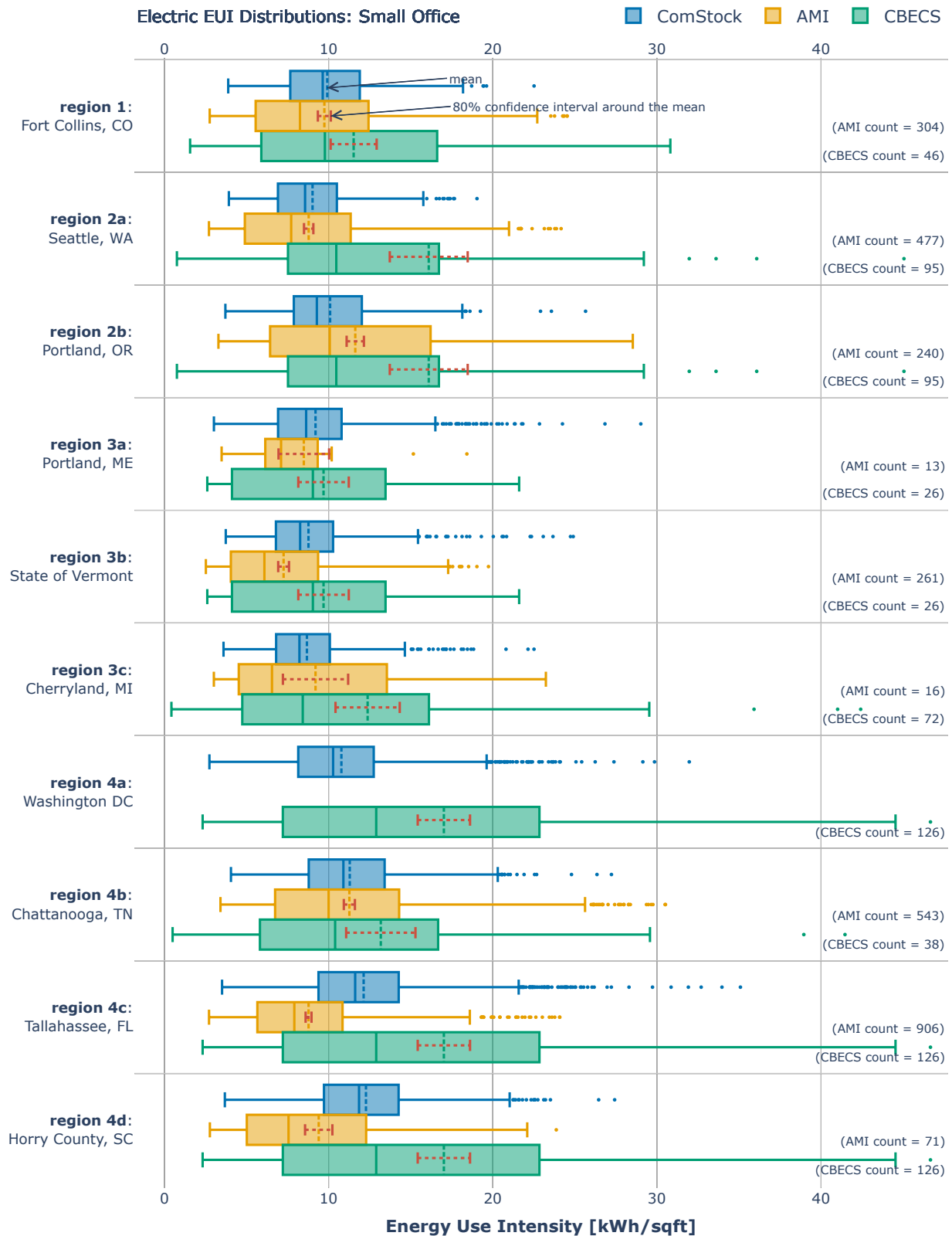


A review of Figure 218 illustrates the challenges associated with comparing ComStock to multiple datasets. For some regions or building types, the AMI and CBECS agree well, but in others they disagree. Rather than choosing a single “best” dataset for comparison, we decided to present a comparison of ComStock to both the AMI and CBECS whenever possible. Because the magnitude of the AMI sometimes disagrees with CBECS, the load profiles comparisons are presented in both raw form, reflecting the magnitude and shape of the AMI data, as well as normalized form, where both ComStock and the AMI data are normalized to an annual profile which sums to one. This removes a difference in magnitude from the comparison.

Additionally, after reviewing the AMI to CBECS comparison, it became clear that any evaluation must be done separately for each building type because some regions have low or missing AMI data for one or more building types, and the sample sizes between building types can vary widely even within a single region. **Note** that in the following figures, there is no AMI for some building types in some regions, and no EUI distributions are shown for Region 4a, Washington D.C., because the data were provided in aggregated bins from which no individual building EUIs could be extracted. Additionally, where the number of AMI samples is one or two, the data are hidden because the confidence intervals are so wide that they make the scale unreadable for all other regions.

#### *Small Office*

For small offices, the EUI distributions for ComStock, AMI, and CBECS for most regions overlap significantly, although the ComStock distributions are more tightly centered around the median. There is a noticeable difference in the means, where the AMI tends to be lower than both ComStock and CBECS. The normalized load profiles show that the overall load shape is a good match for the AMI data, and the non-normalized profiles show that in some regions, ComStock’s magnitude is slightly too low.



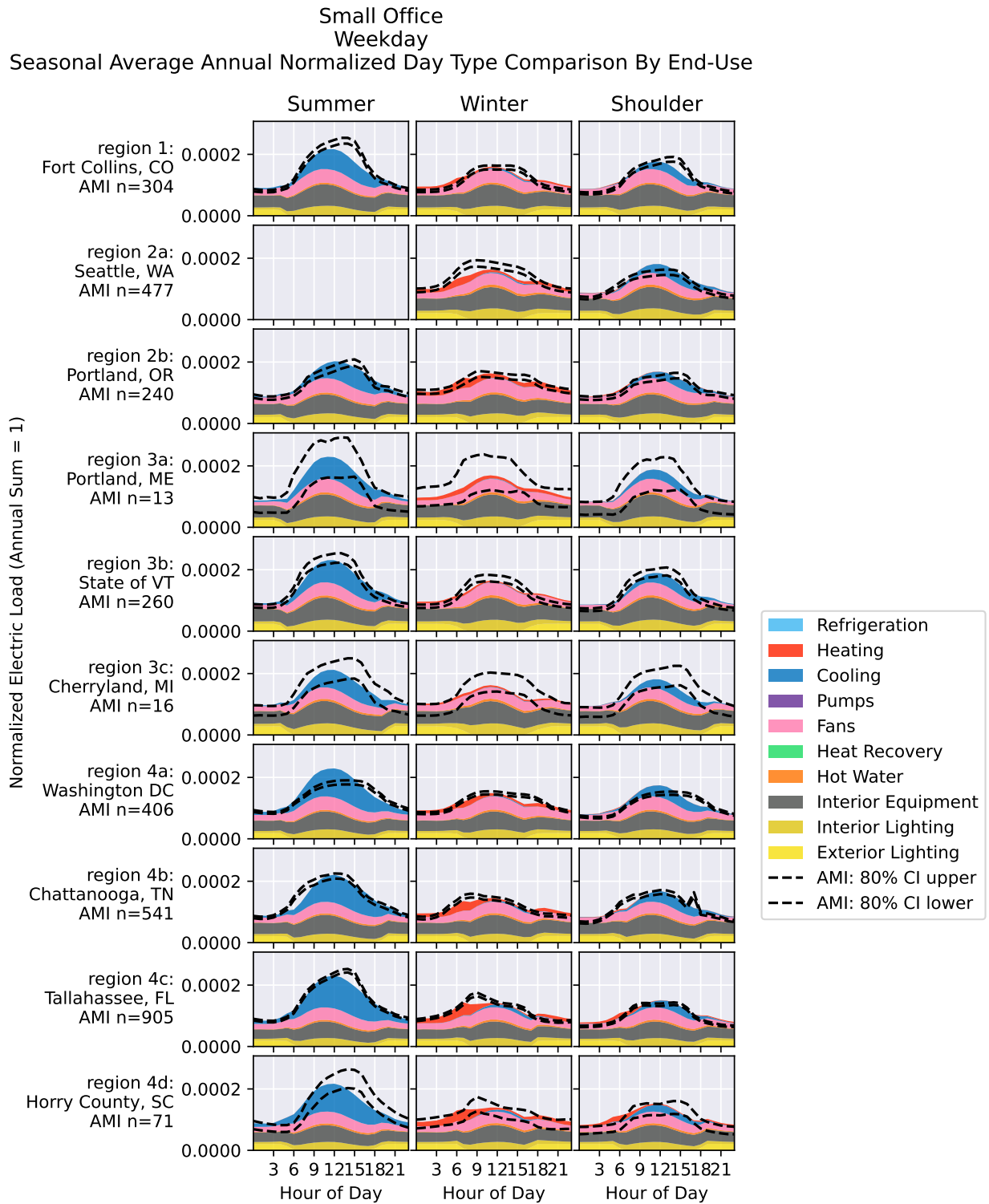


Figure 220. Small office weekday seasonal average annual normalized day type comparison by end use

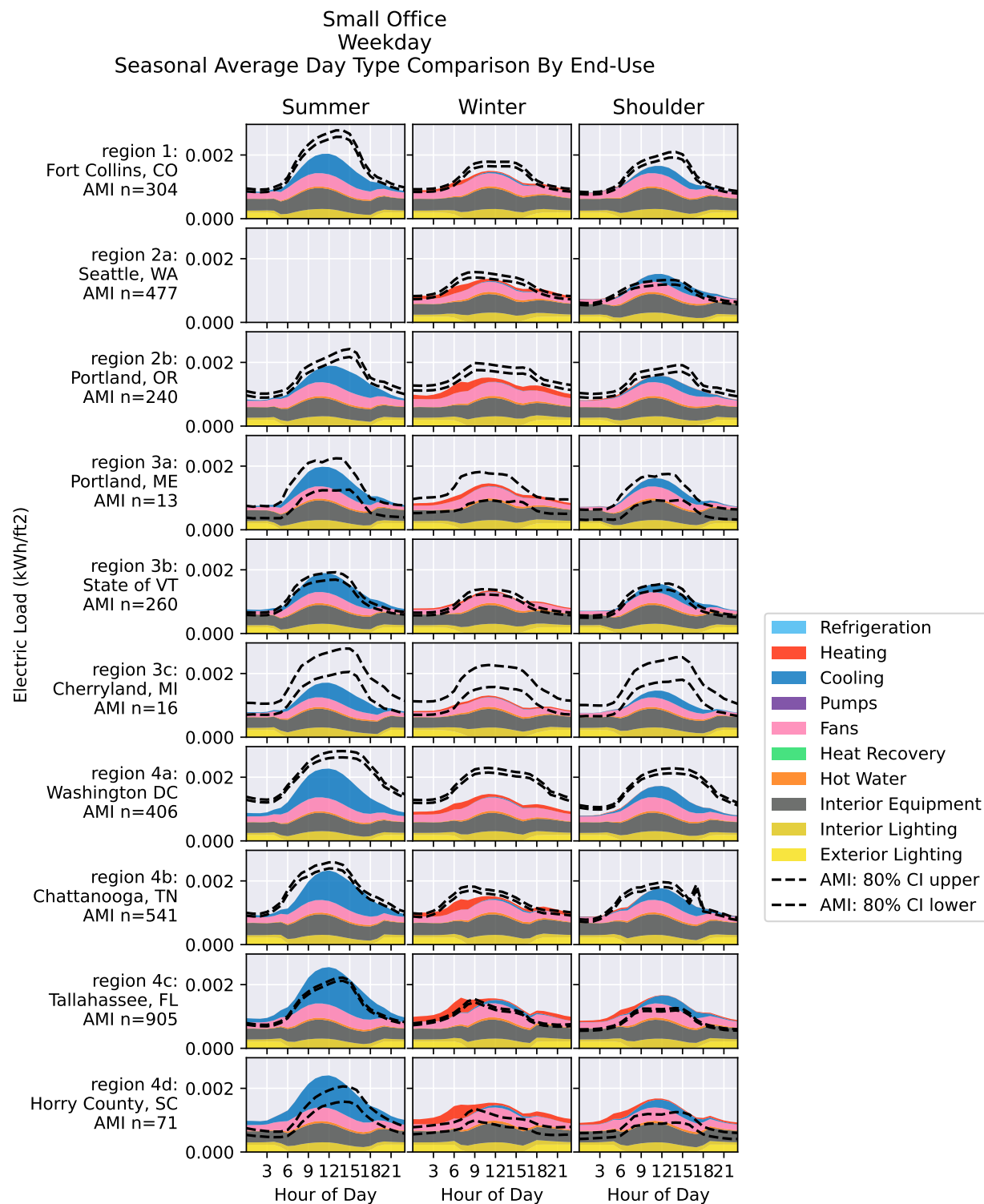


Figure 221. Small office weekday seasonal average day type comparison by end use

Small Office  
Weekend  
Seasonal Average Annual Normalized Day Type Comparison By End-Use

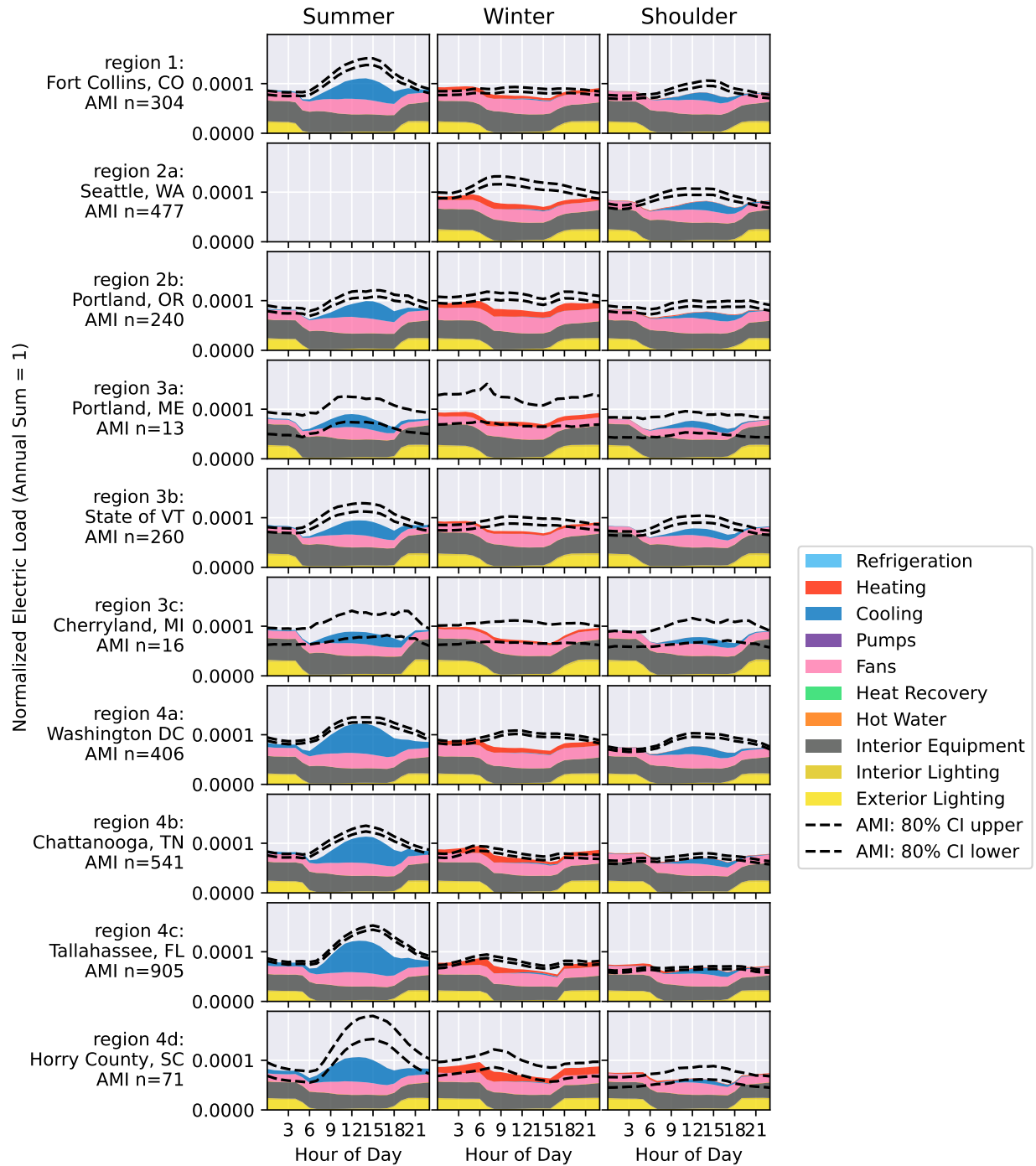
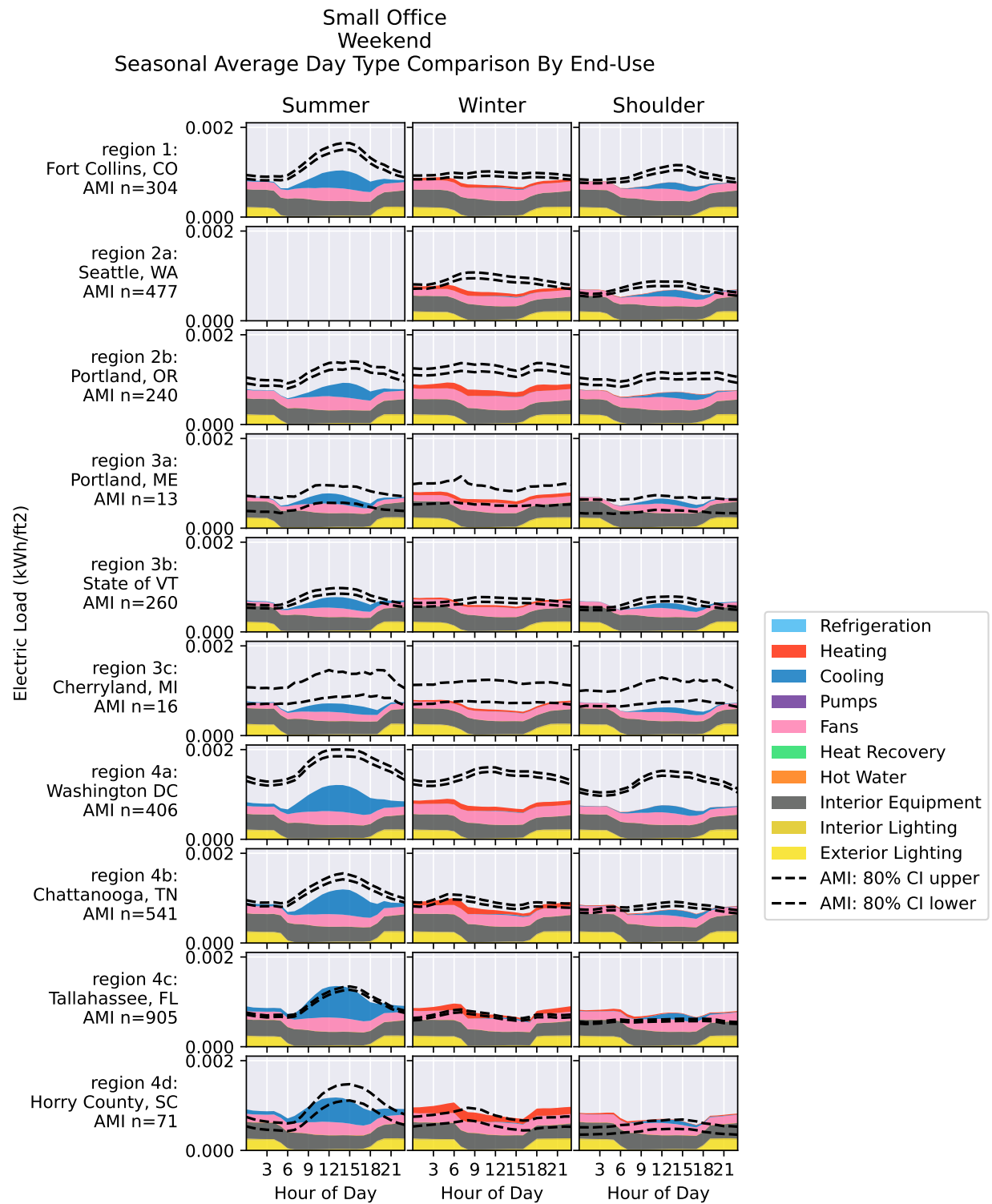


Figure 222. Small office weekend seasonal average annual normalized day type comparison by end use

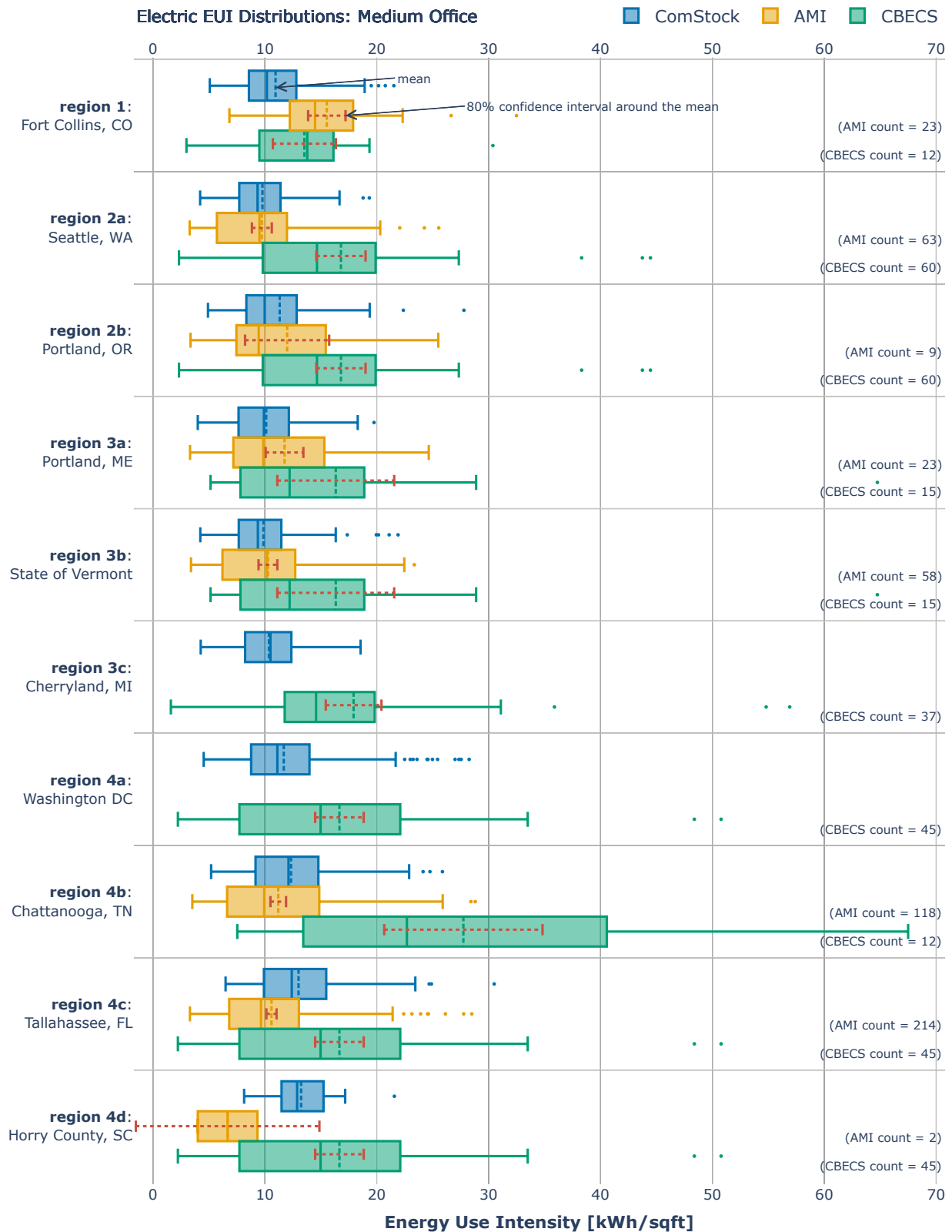


**Figure 223. Small office weekend seasonal average day type comparison by end use**

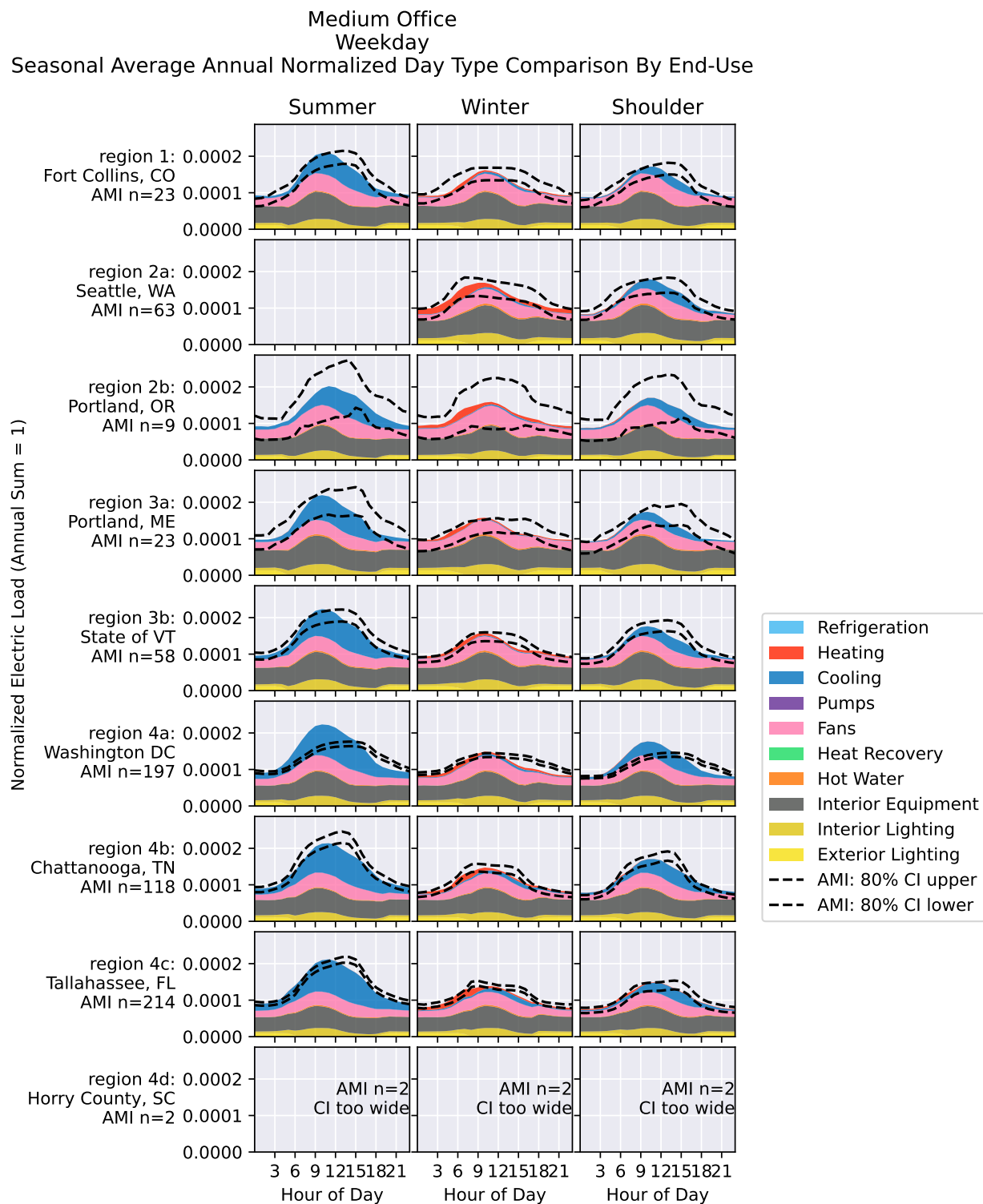
### *Medium Office*

For medium offices, the EUI distributions for ComStock, AMI, and CBECS for most regions overlap significantly, although the ComStock distributions are more tightly centered around the median and tend to be lower than CBECS and closer to the AMI data. The normalized load profiles show that the overall load shape is a good match for the AMI data, and the non-normalized profiles show that in some regions, ComStock's magnitude is slightly too low. ComStock appears to underestimate the magnitude of the load on weekends.

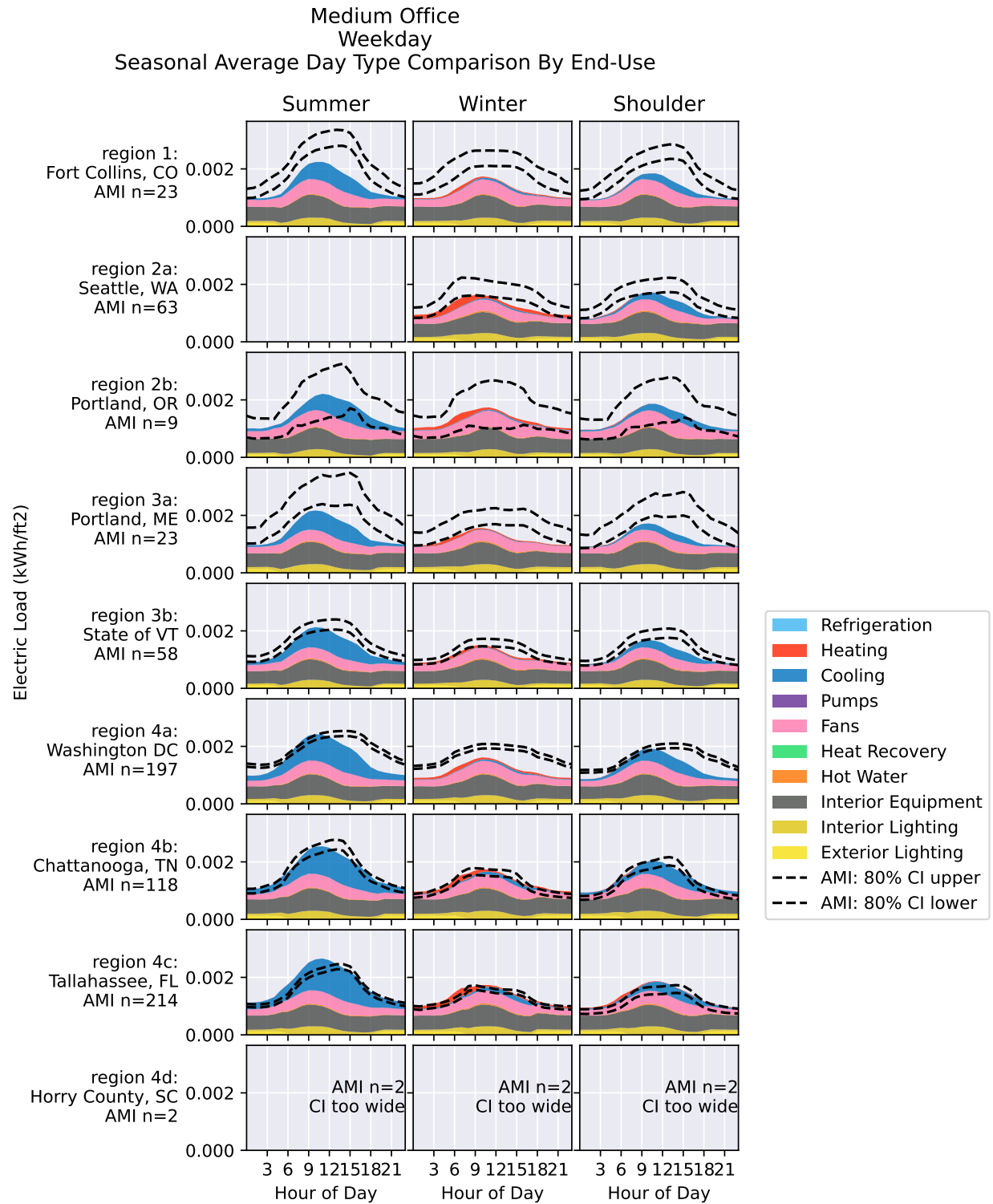




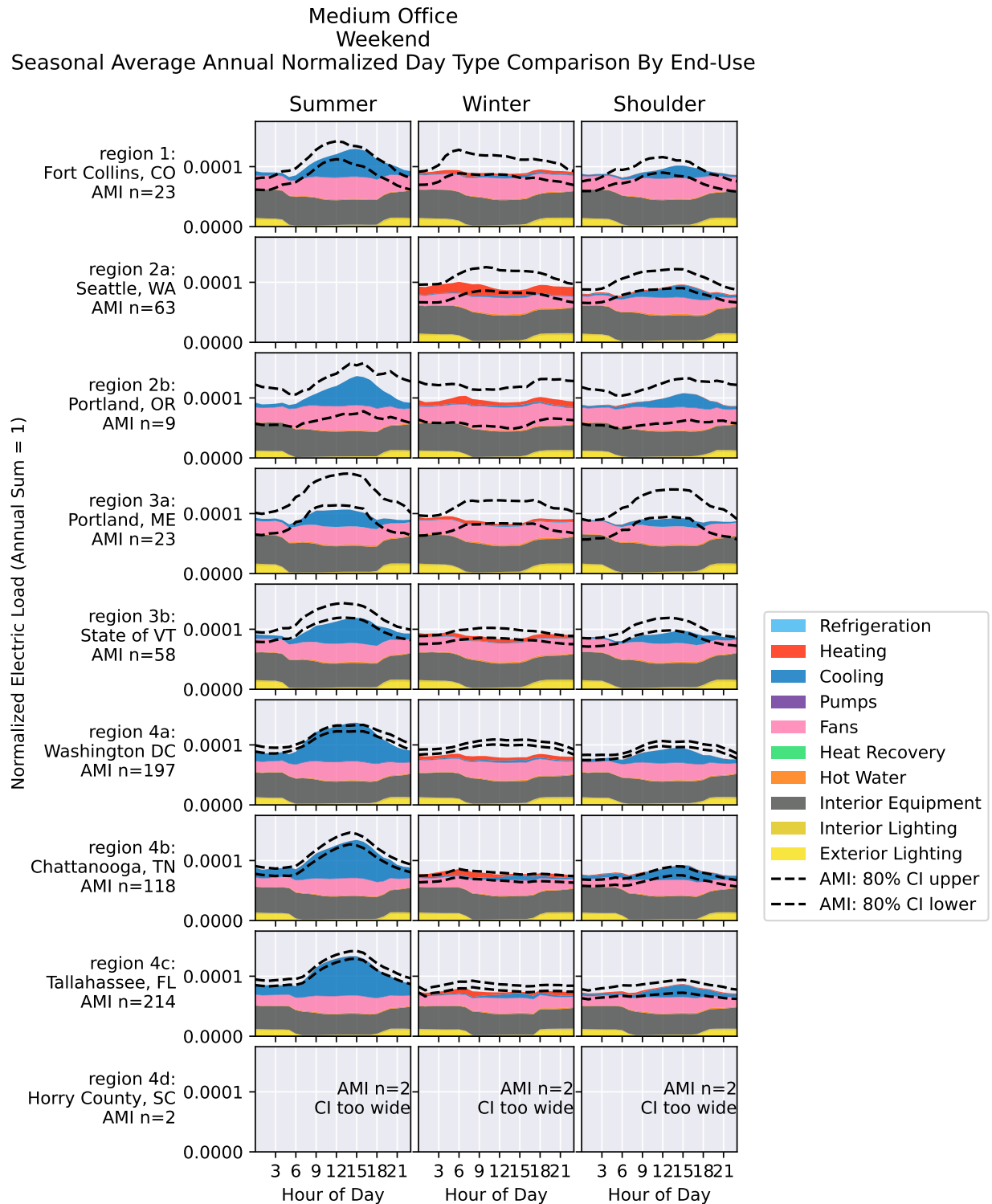
**Figure 224. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for medium offices**



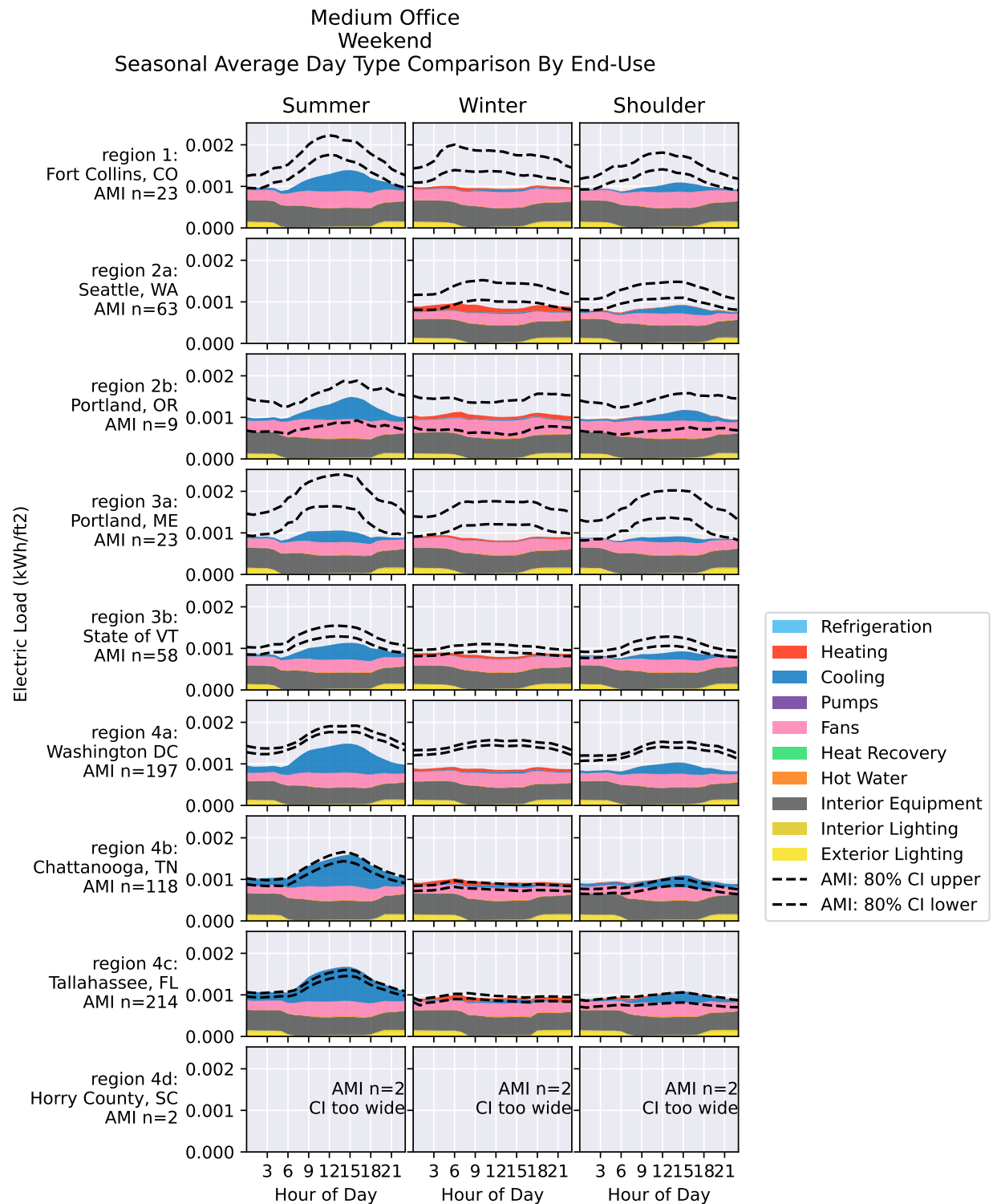
**Figure 225. Medium office weekday seasonal average annual normalized day type comparison by end use**



**Figure 226. Medium office weekday seasonal average day type comparison by end use**



**Figure 227. Medium office weekend seasonal average annual normalized day type comparison by end use**



**Figure 228. Medium office weekend seasonal average day type comparison by end use**

### *Large Office*

For large offices, the AMI and CBECS sample sizes tend to be lower, so the confidence intervals are wide. For most regions, the ComStock mean EUI falls within the confidence intervals of the other two datasets. The normalized load profiles show that the overall load shape is a good match for the AMI data, and the non-normalized profiles show that in some regions, ComStock's magnitude is slightly too low, particularly in winter, suggesting missing electric reheat. ComStock appears to underestimate the magnitude of the load on weekends in Washington D.C., suggesting less office usage reduction than other regions experience.

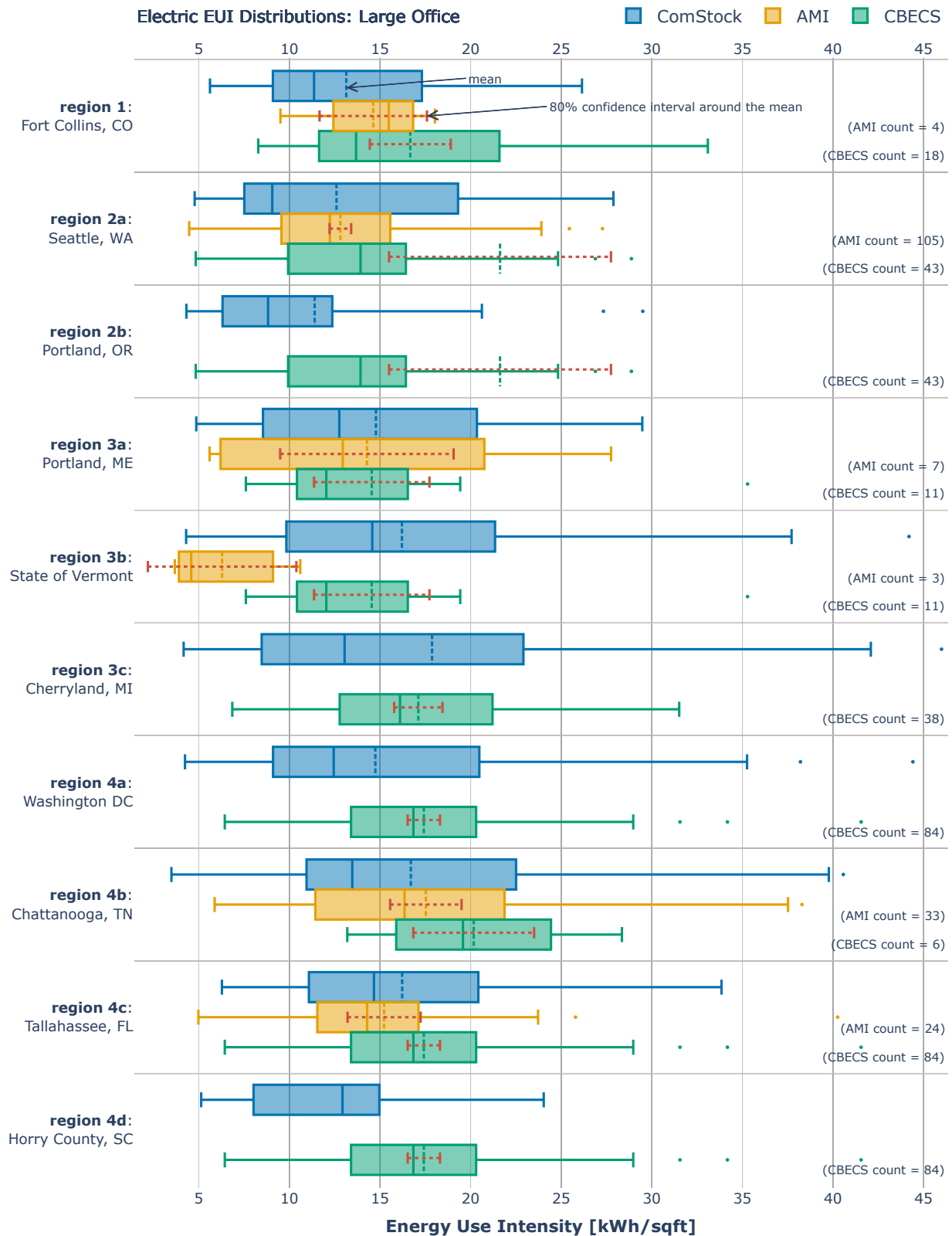


Figure 229. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for large offices



Large Office  
Weekday  
Seasonal Average Annual Normalized Day Type Comparison By End-Use

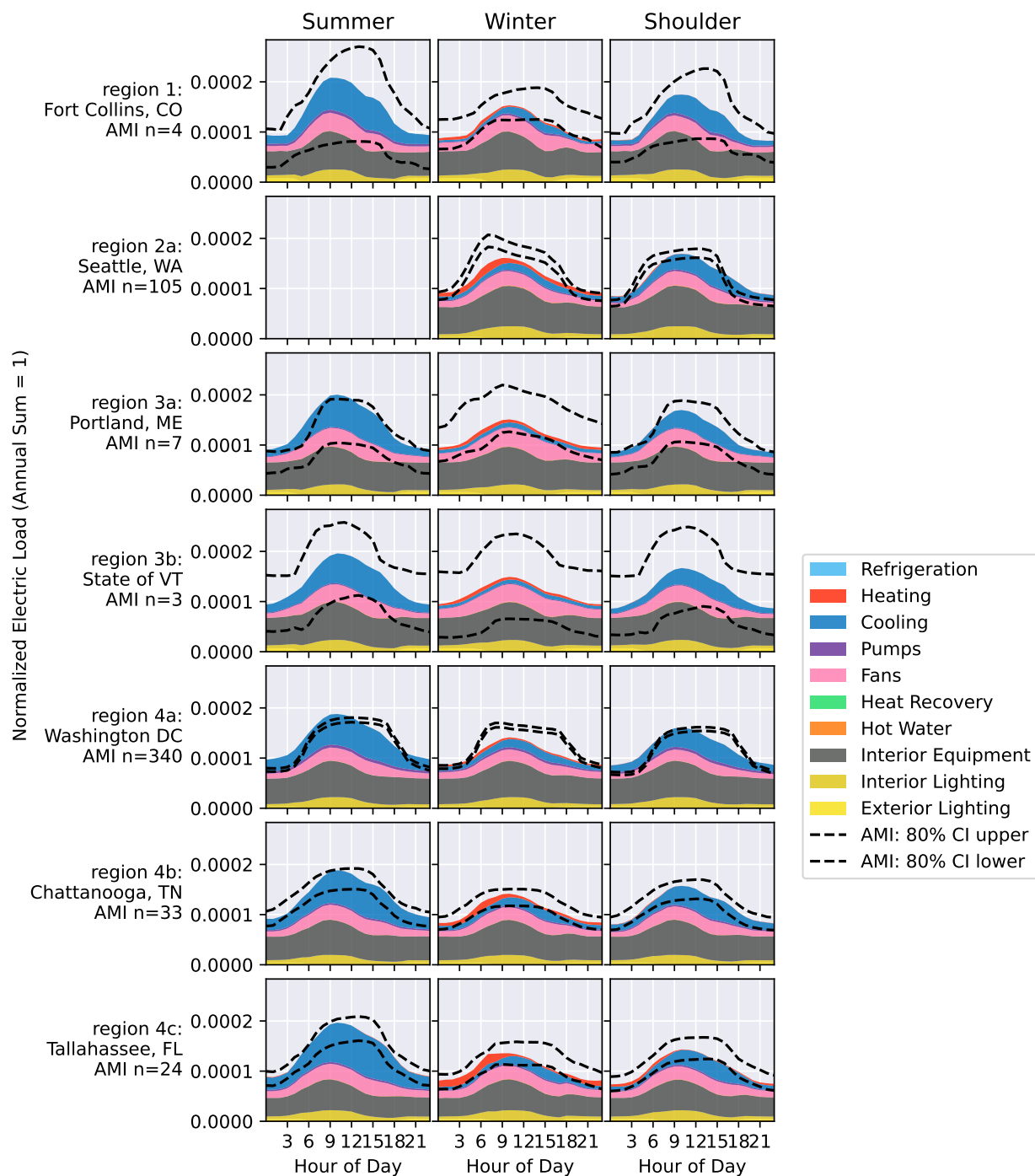


Figure 230. Large office weekday seasonal average annual normalized day type comparison by end use

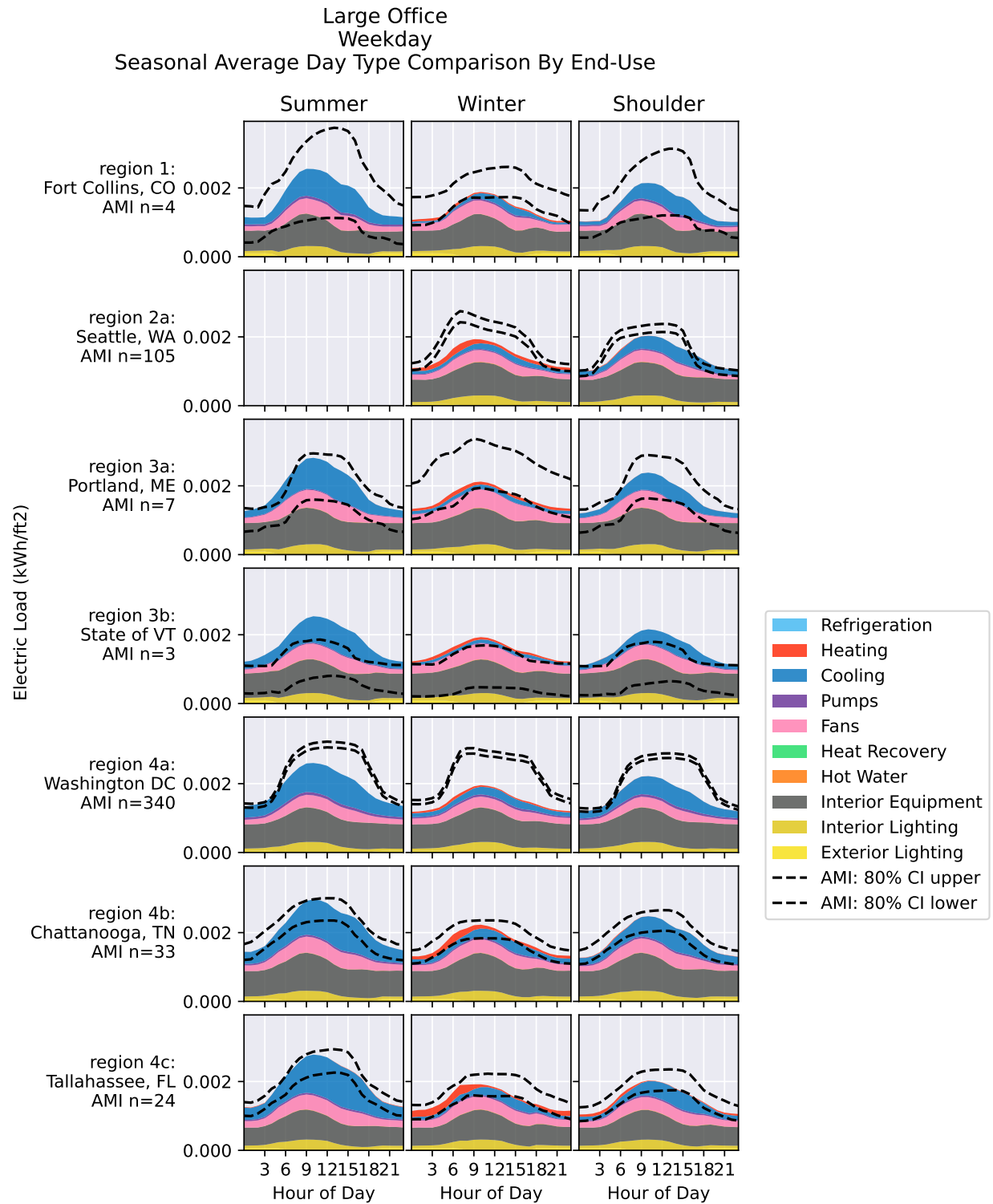


Figure 231. Large office weekday seasonal average day type comparison by end use

Large Office  
Weekend  
Seasonal Average Annual Normalized Day Type Comparison By End-Use

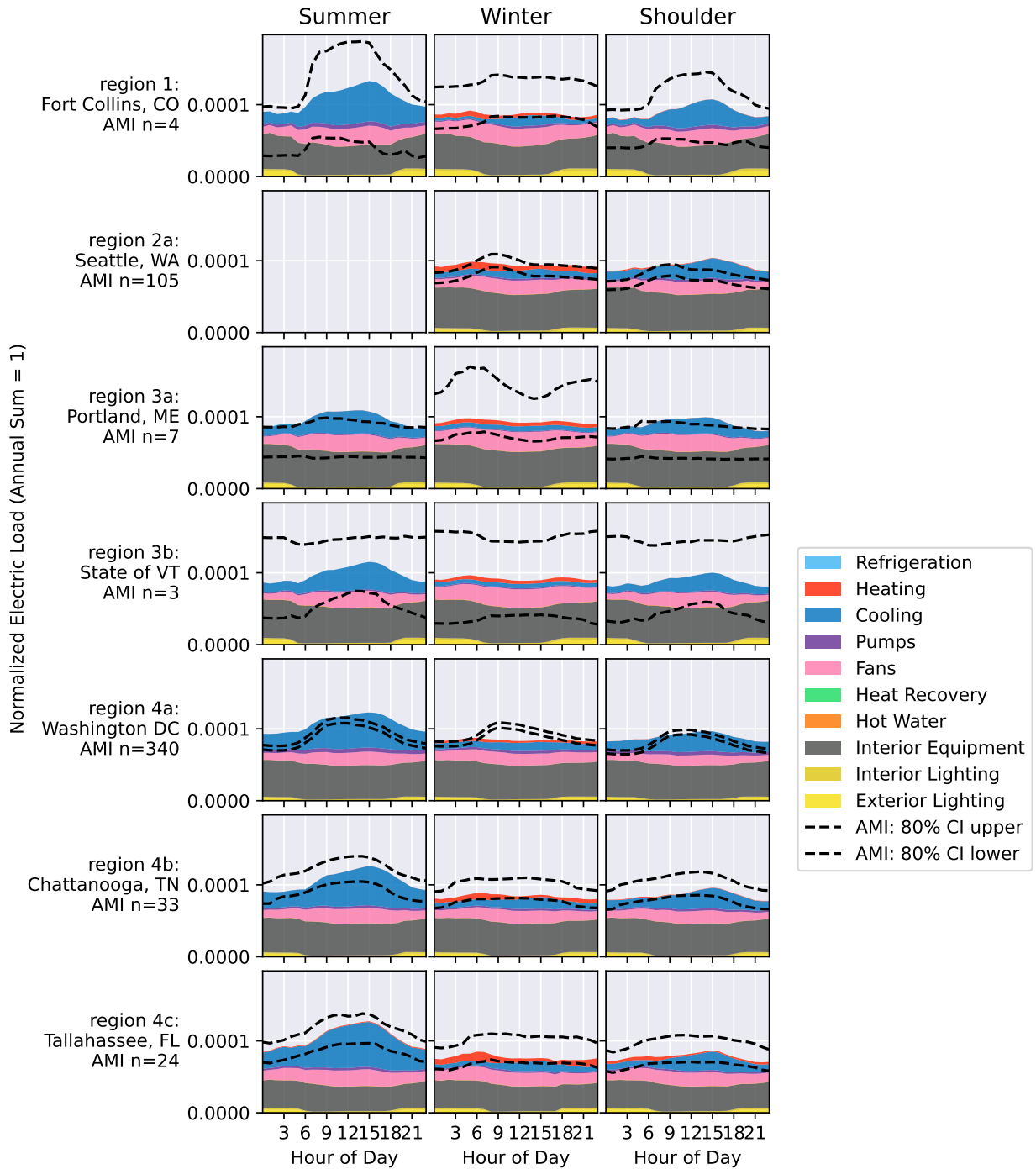
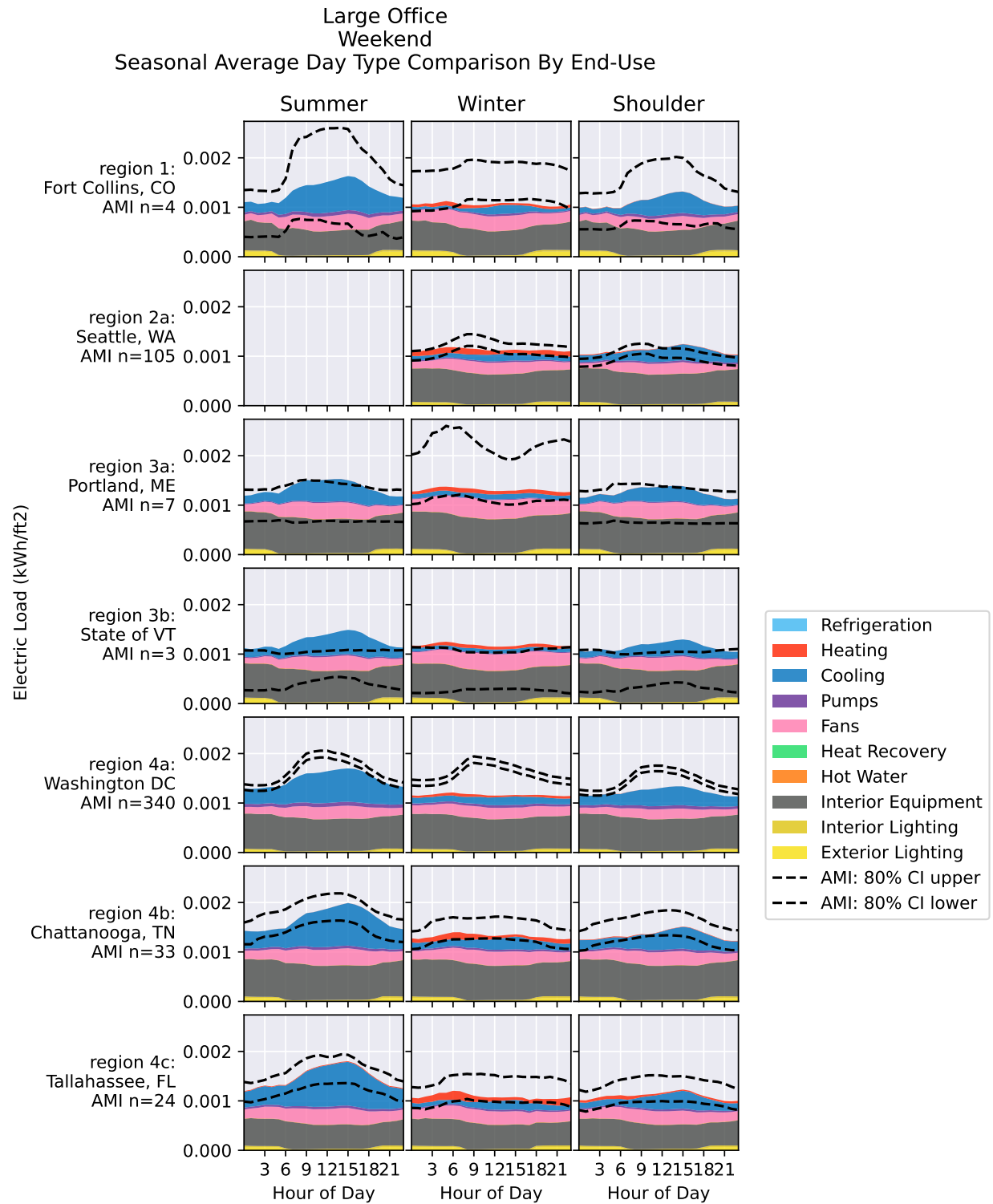


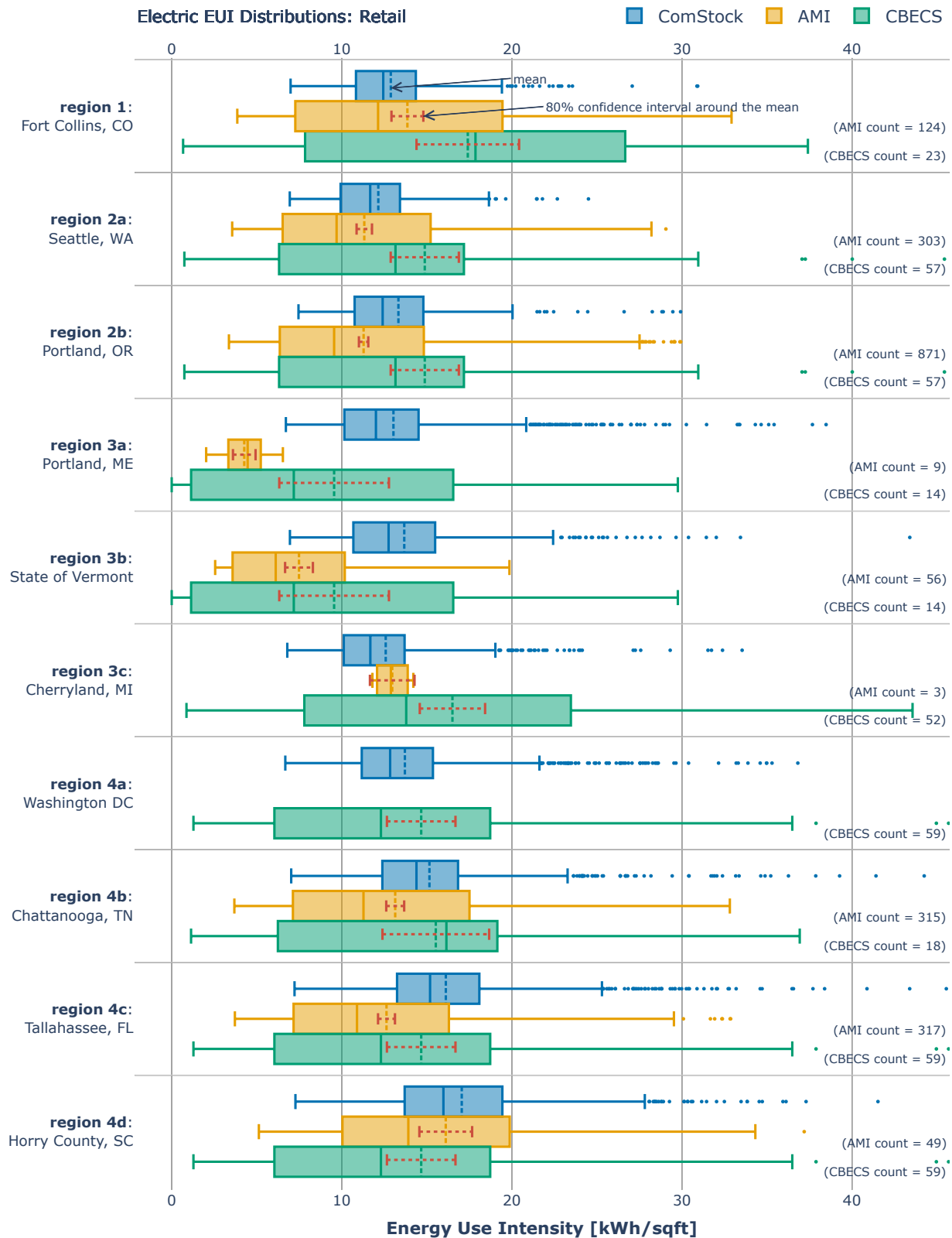
Figure 232. Large office weekend seasonal average annual normalized day type comparison by end use



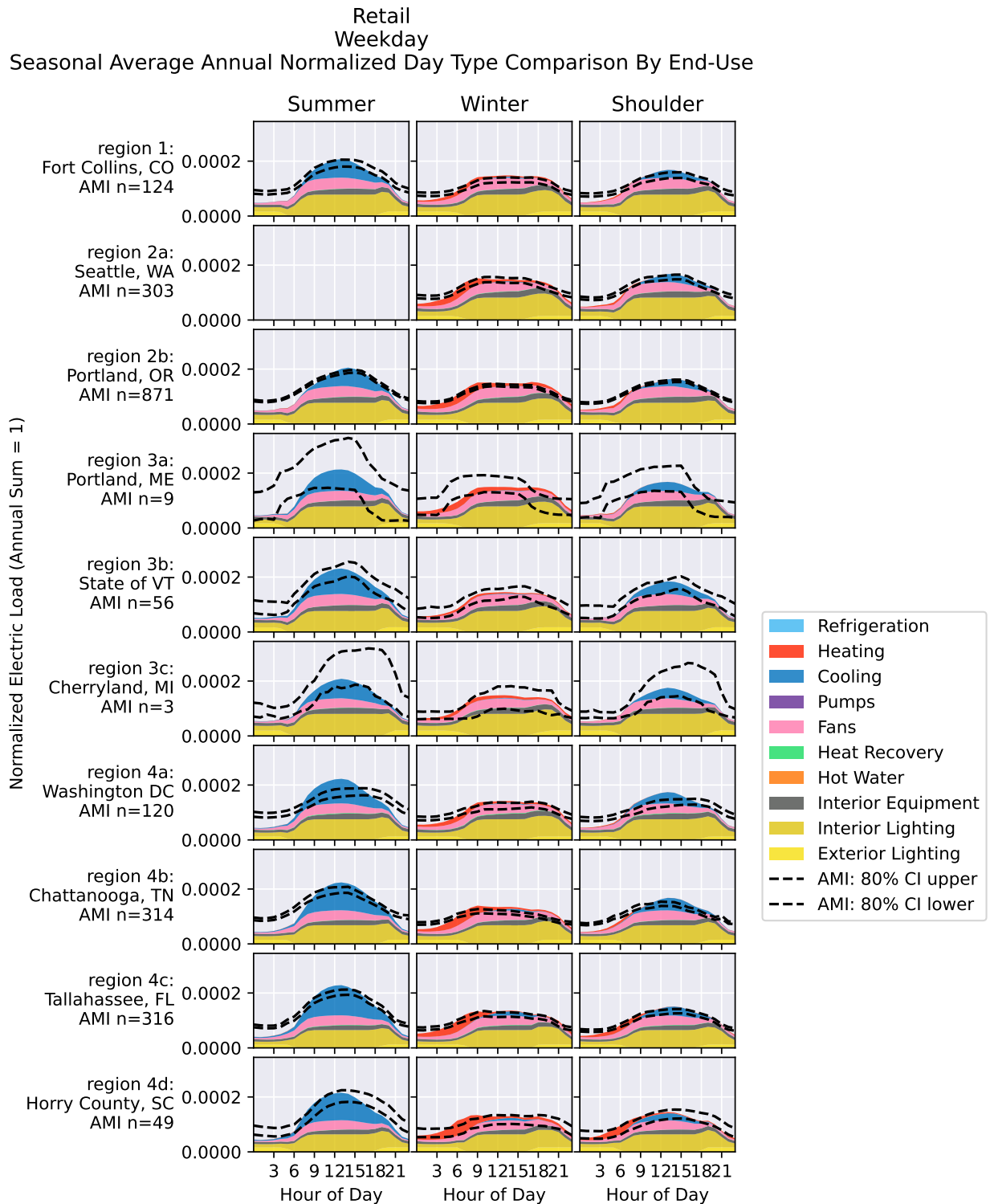
**Figure 233. Large office weekend seasonal average day type comparison by end use**

### *Retail*

For retail buildings, the ComStock EUI distributions are narrower than the AMI or CBECS for regions where sample sizes are higher, but the distributions do all tend to overlap. Both the normalized and non-normalized profiles show that ComStock has good agreement during daytime operation, but that it is missing load at night on both weekends and weekdays. Based on the magnitude of the missing load, this could be nighttime HVAC operation, although this was already adjusted based on available data during the calibration process.

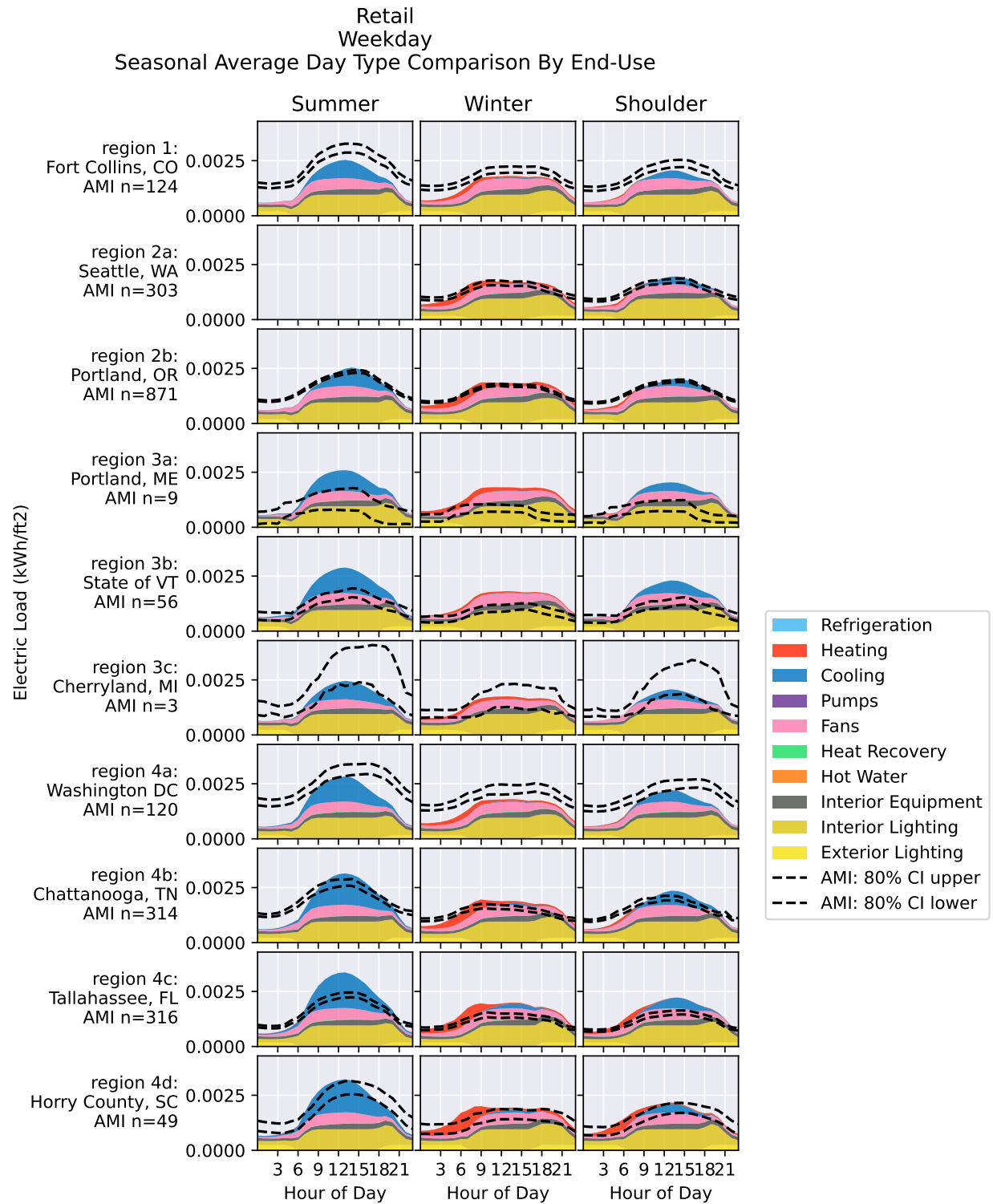


**Figure 234. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for retail buildings**



**Figure 235. Retail weekday seasonal average annual normalized day type comparison by end use**





**Figure 236. Retail weekday seasonal average day type comparison by end use**

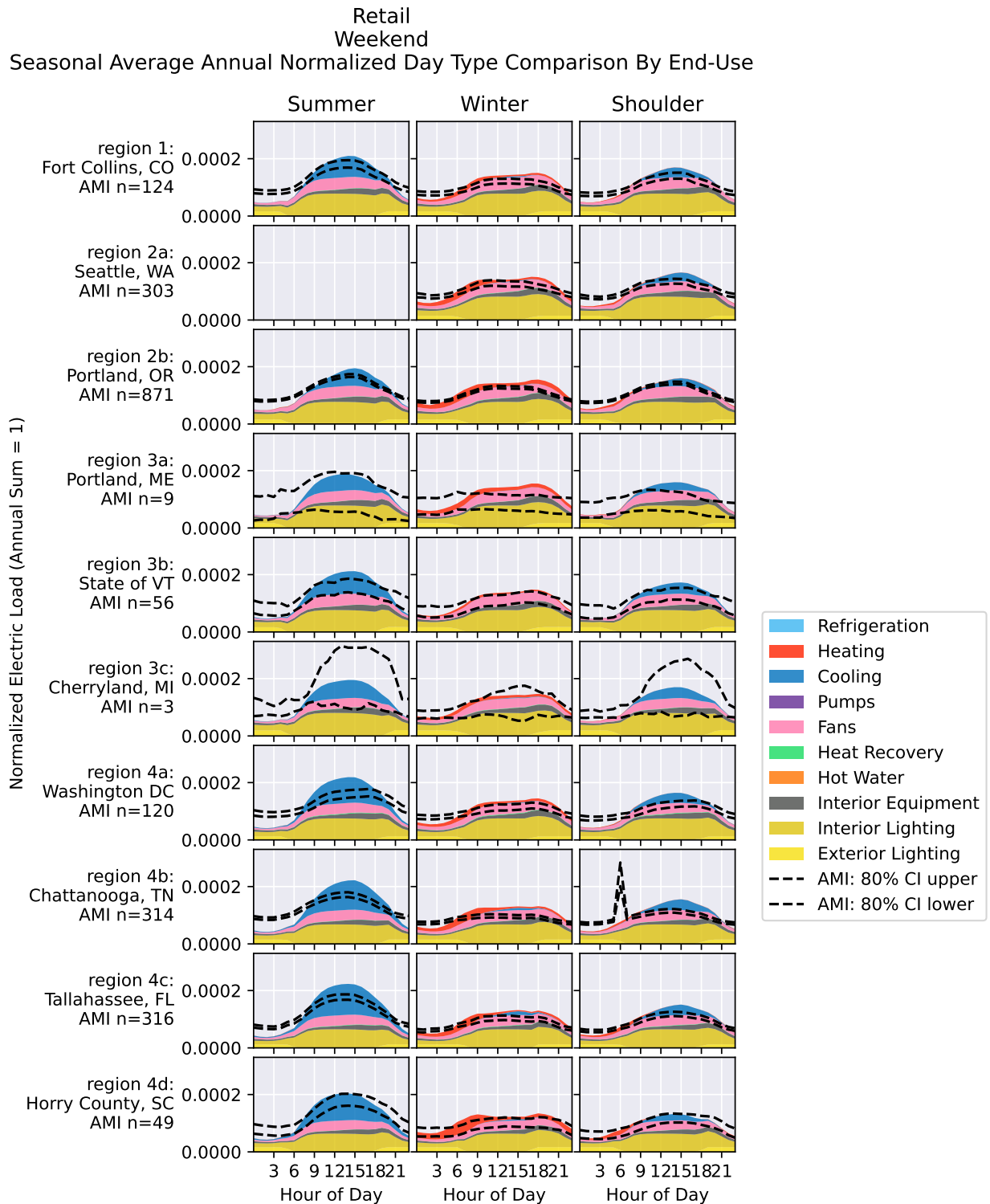
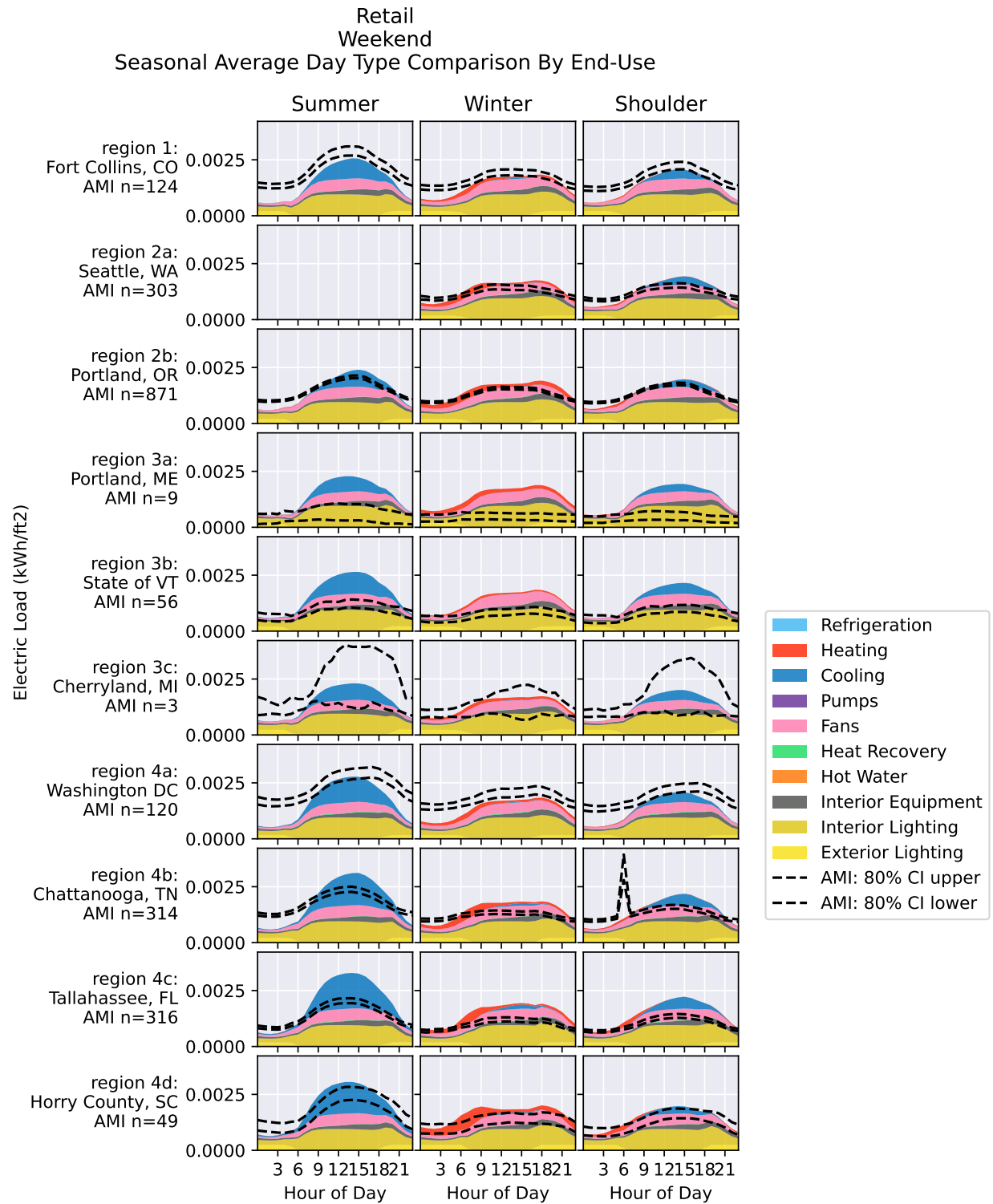


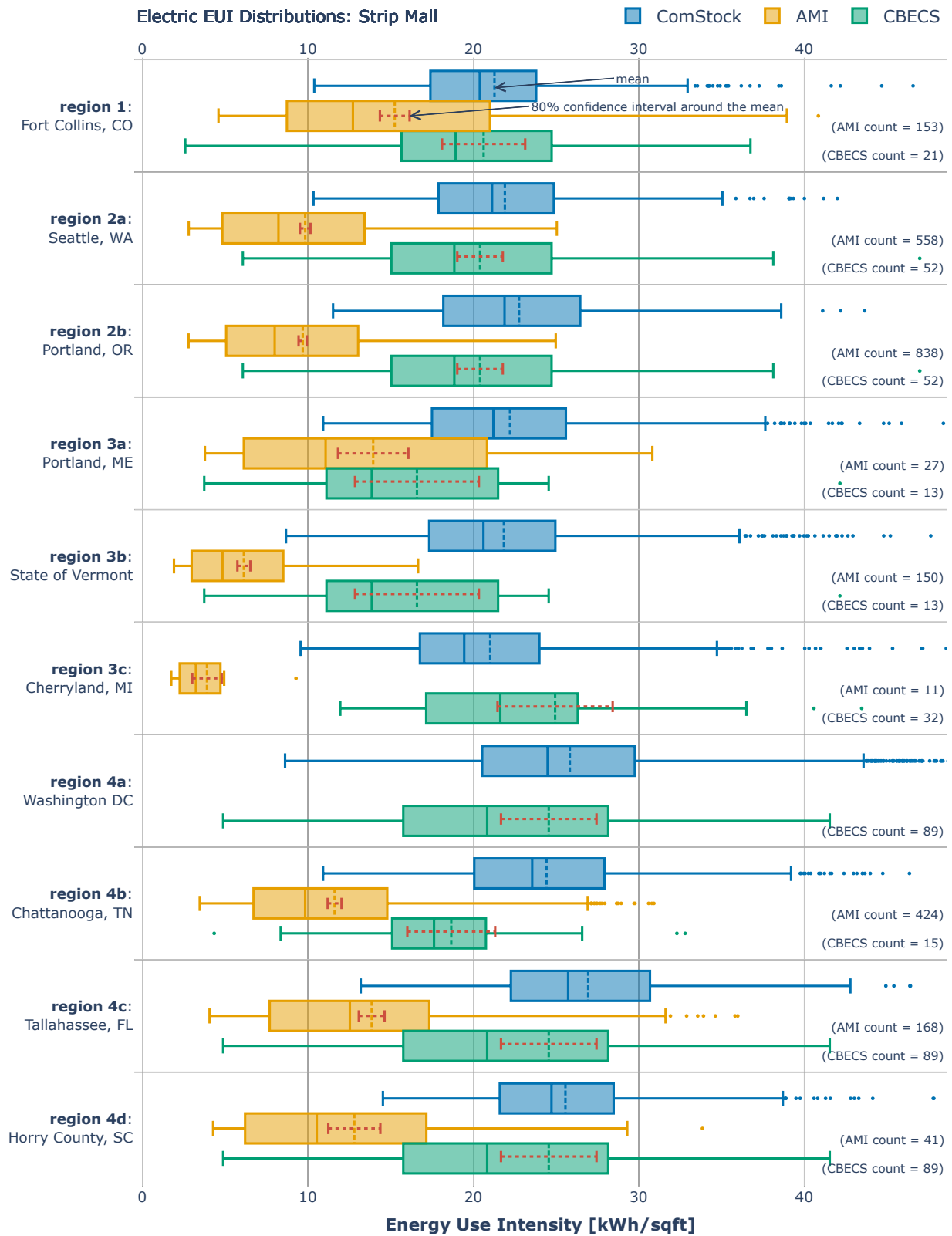
Figure 237. Retail weekend seasonal average annual normalized day type comparison by end use



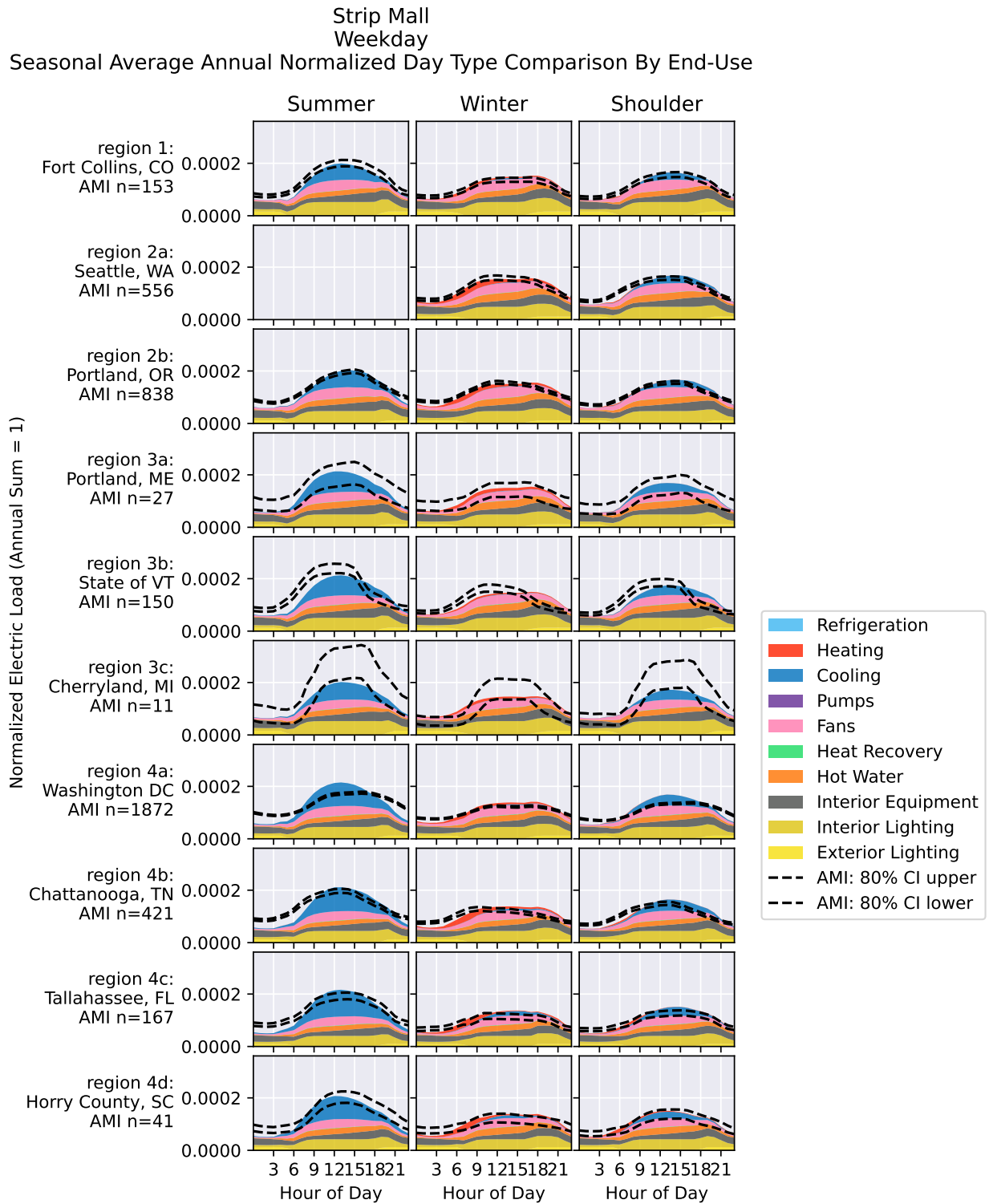
**Figure 238. Retail weekend seasonal average day type comparison by end use**

### *Strip Mall*

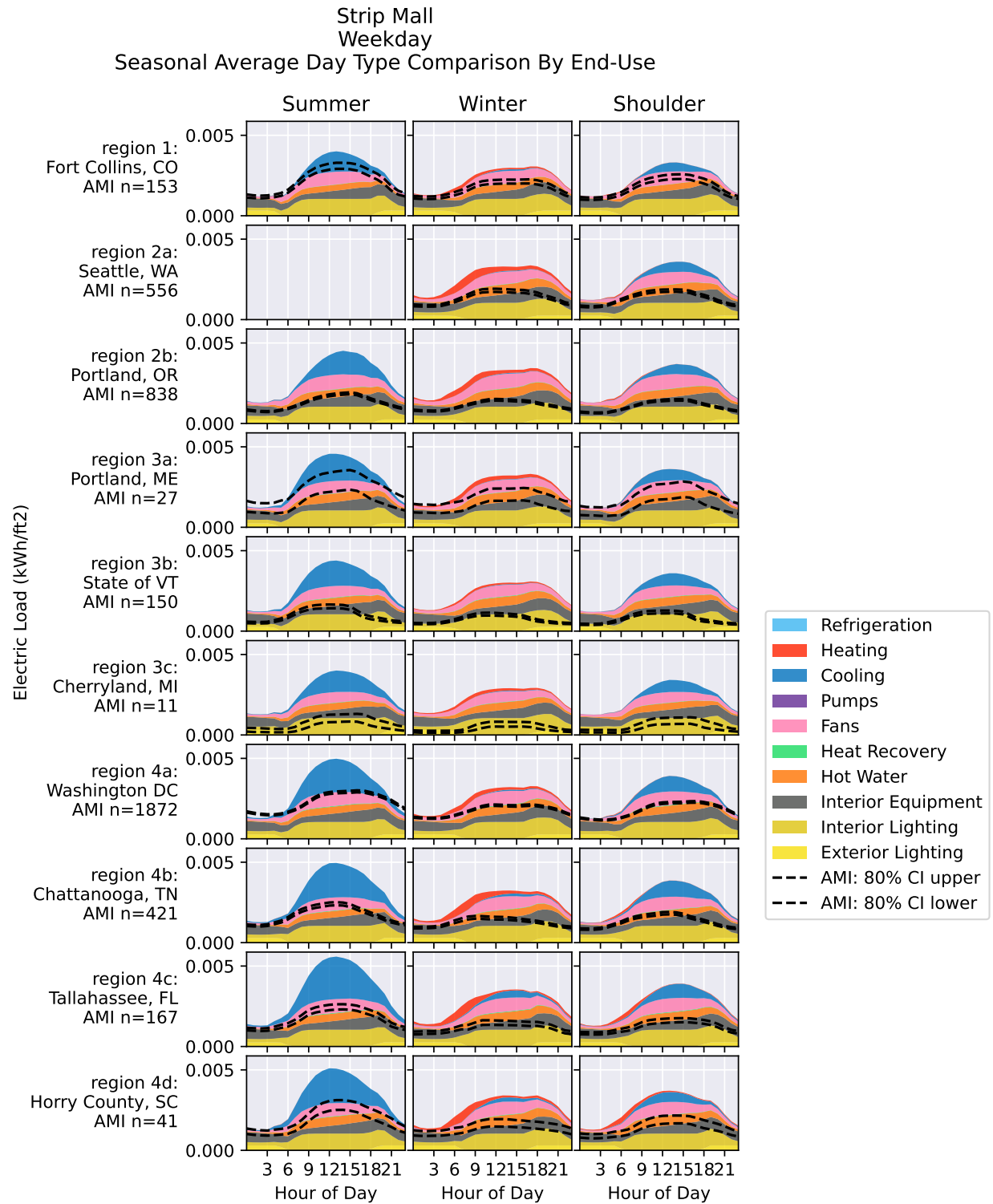
Strip malls are perhaps the most challenging building type to classify, as there is a range in the makeup of the stores in this building type, which leads to a wide range in valid EUIs. We found that the 3X median filtering approach tended to remove strip malls that contained mostly restaurants. This is a good example of a building type where the AMI and the CBECS data tell different stories. During the calibration process, we determined that because most strip malls do in fact contain significant amounts of restaurant space, it was important to reflect this in the models. The result is that the EUIs for ComStock the EUIs in CBECS but overestimate the AMI magnitude. As in retail buildings, the normalized load shape profiles match well during the day but underestimate load at night. For the non-normalized profiles, ComStock overestimates the AMI magnitude, which we believe is more an artifact of the outlier/misclassification method than a true reflection of the load. This is corroborated by the ComStock to CBECS annual comparisons in Figure 191, which do not show ComStock being too low by roughly half as the non-normalized AMI comparison would suggest.



**Figure 239. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for strip malls**

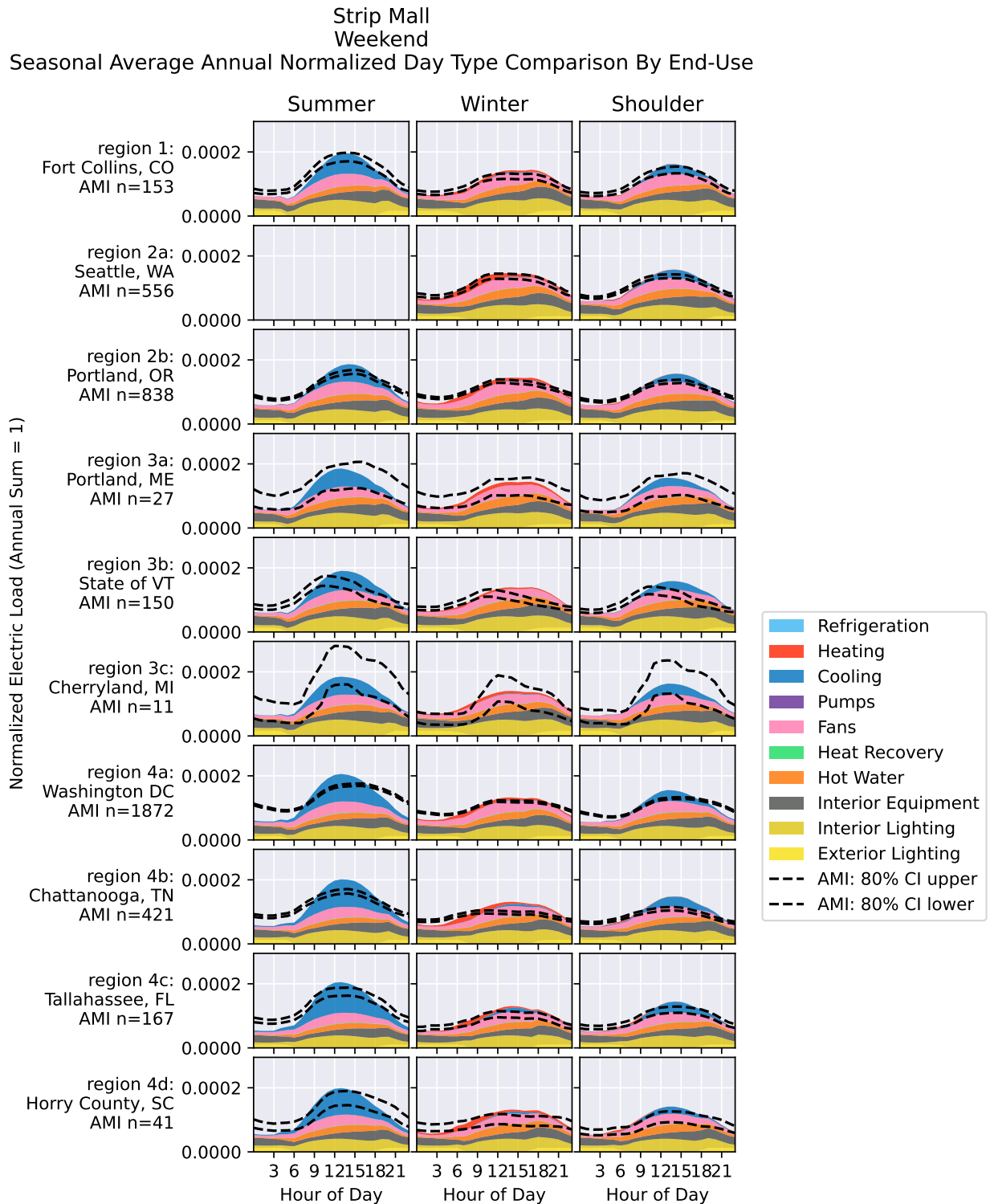


**Figure 240. Strip mall weekday seasonal average annual normalized day type comparison by end use**

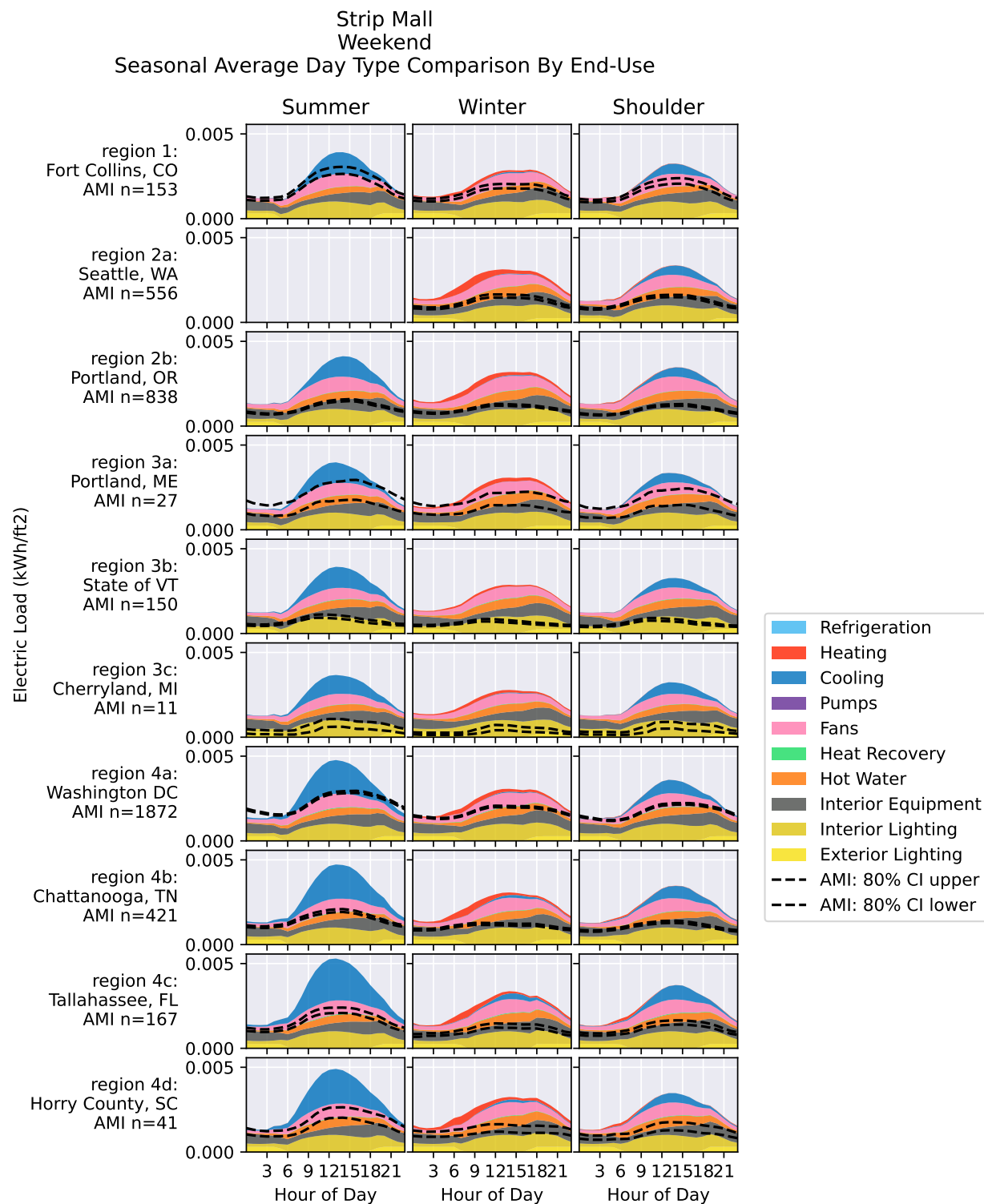


**Figure 241. Strip mall weekday seasonal average day type comparison by end use**





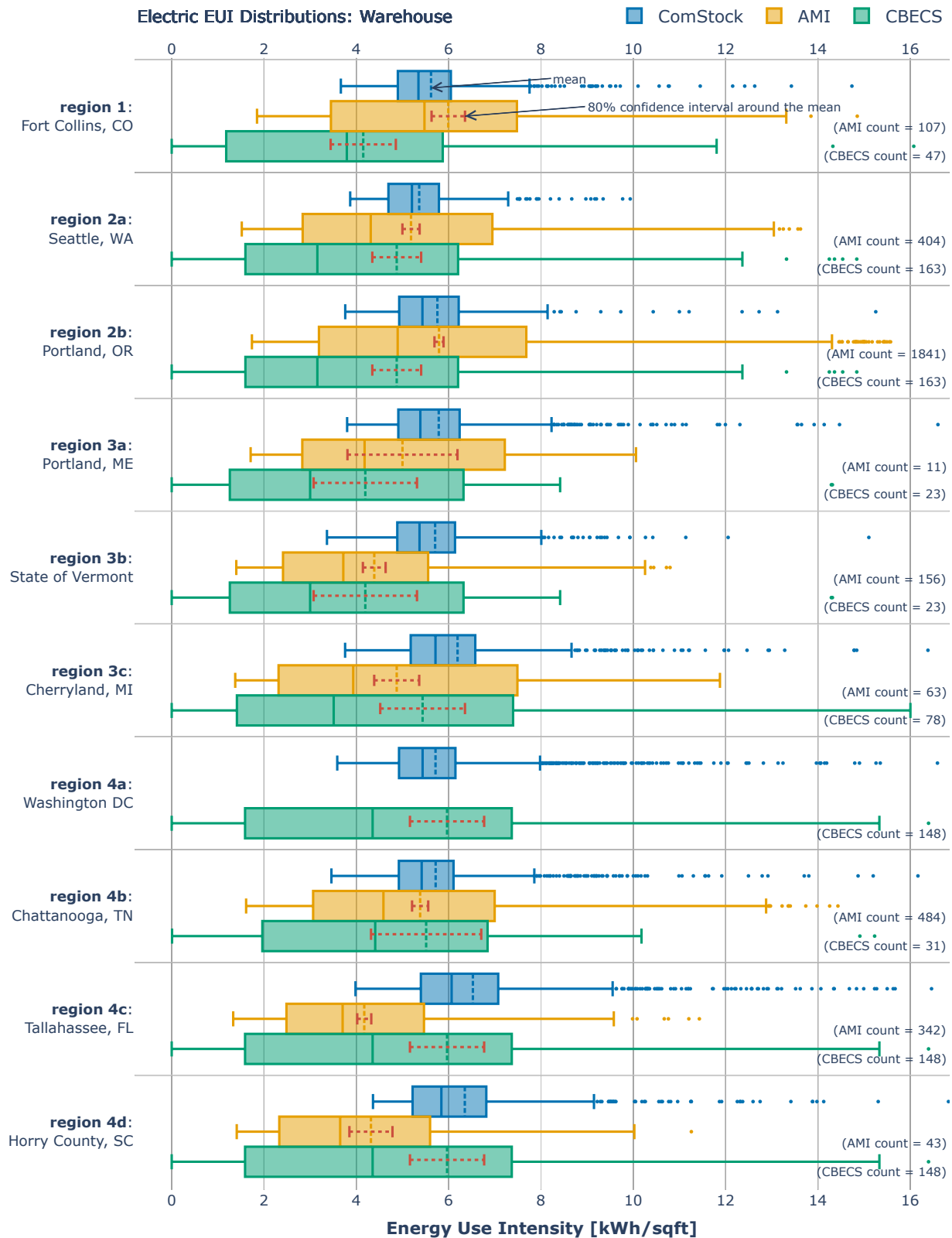
**Figure 242. Strip mall weekend seasonal average annual normalized day type comparison by end use**



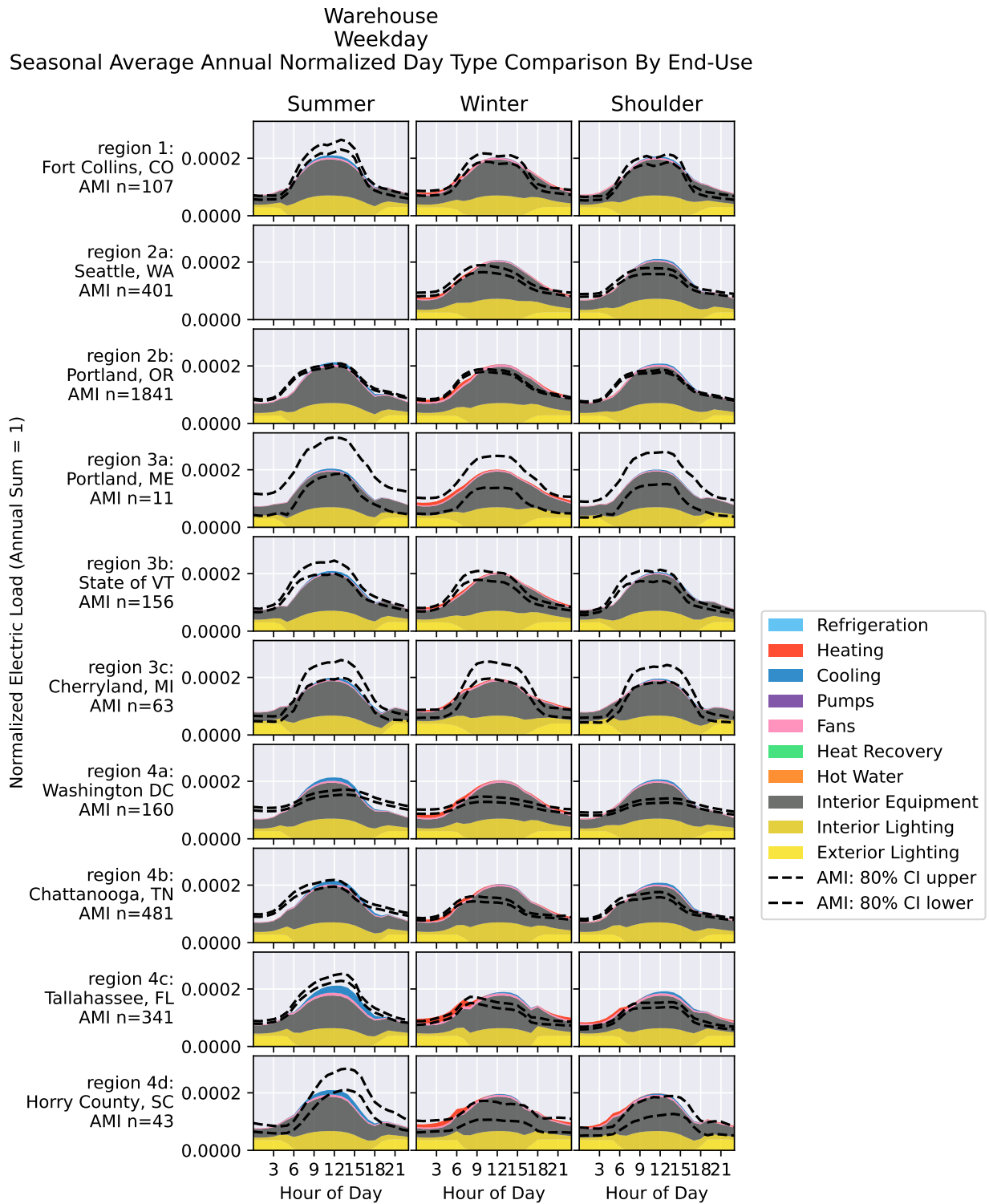
**Figure 243. Strip mall weekend seasonal average day type comparison by end use**

### *Warehouse*

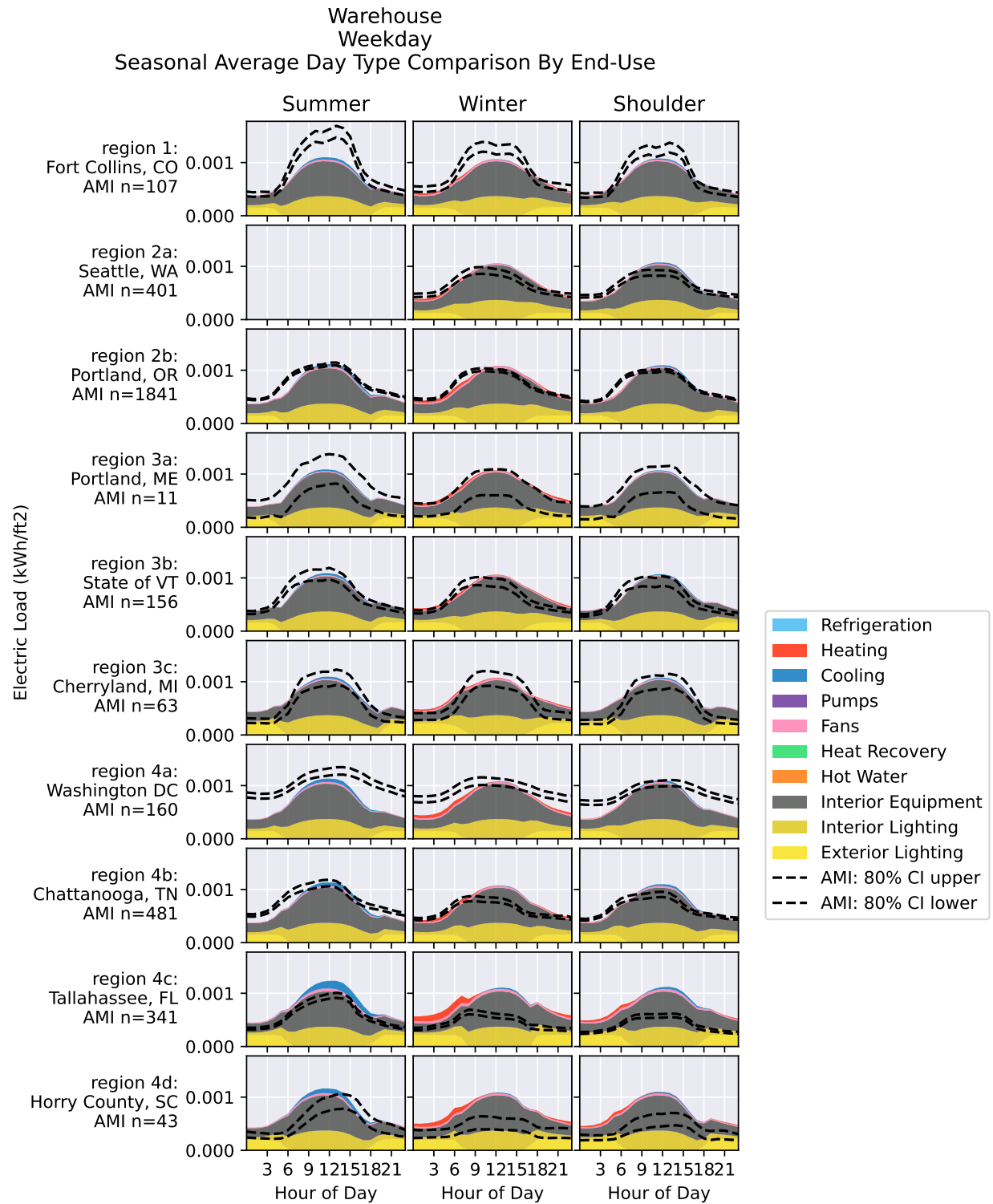
Warehouses are a building type that spans a wide range of energy usage patterns, from self-storage facilities that use very little energy to showrooms with some light manufacturing and office space included. ComStock represents a narrower portion of this spectrum, as evidenced by the tighter distributions. ComStock does, however, tend to align well with the confidence intervals for CBECS. ComStock shows good agreement with both the normalized and non-normalized weekday load profiles. However, it appears that ComStock underestimates the load on weekends, indicating that hours of operation for weekends are significantly underestimated.



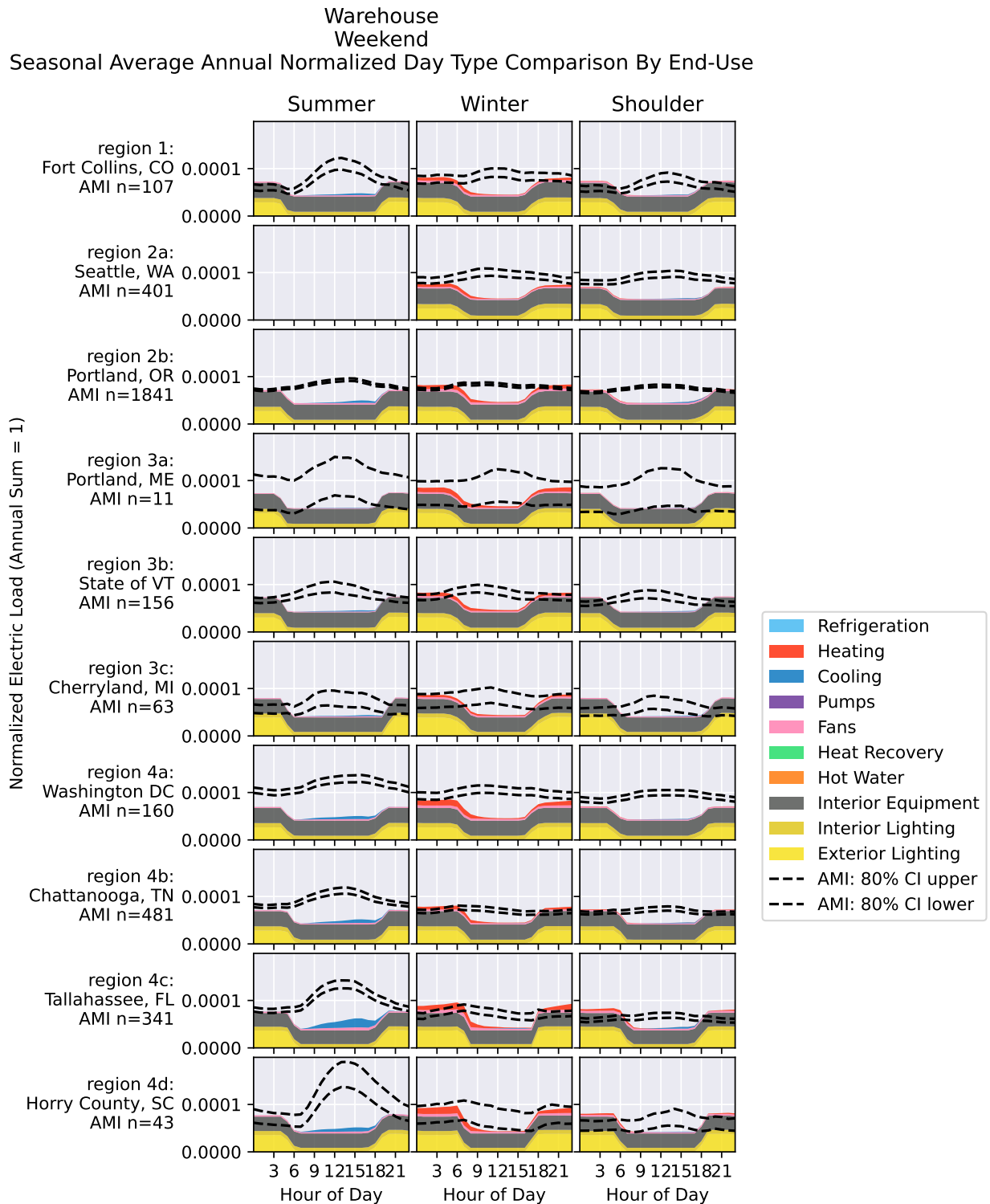
**Figure 244. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for warehouses**



**Figure 245. Warehouse weekday seasonal average annual normalized day type comparison by end use**

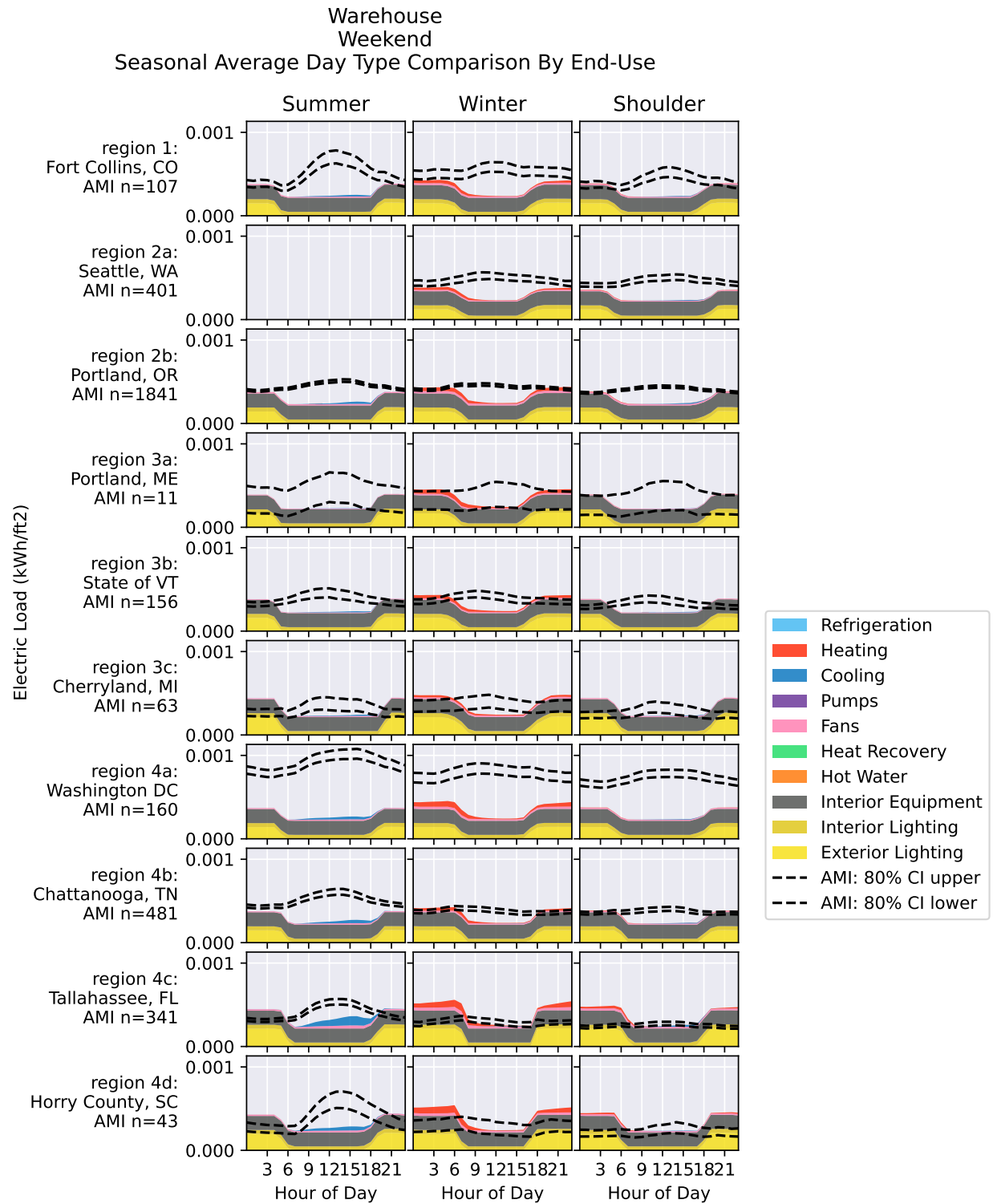


**Figure 246. Warehouse weekday seasonal average day type comparison by end use**



**Figure 247. Warehouse weekend seasonal average annual normalized day type comparison by end use**

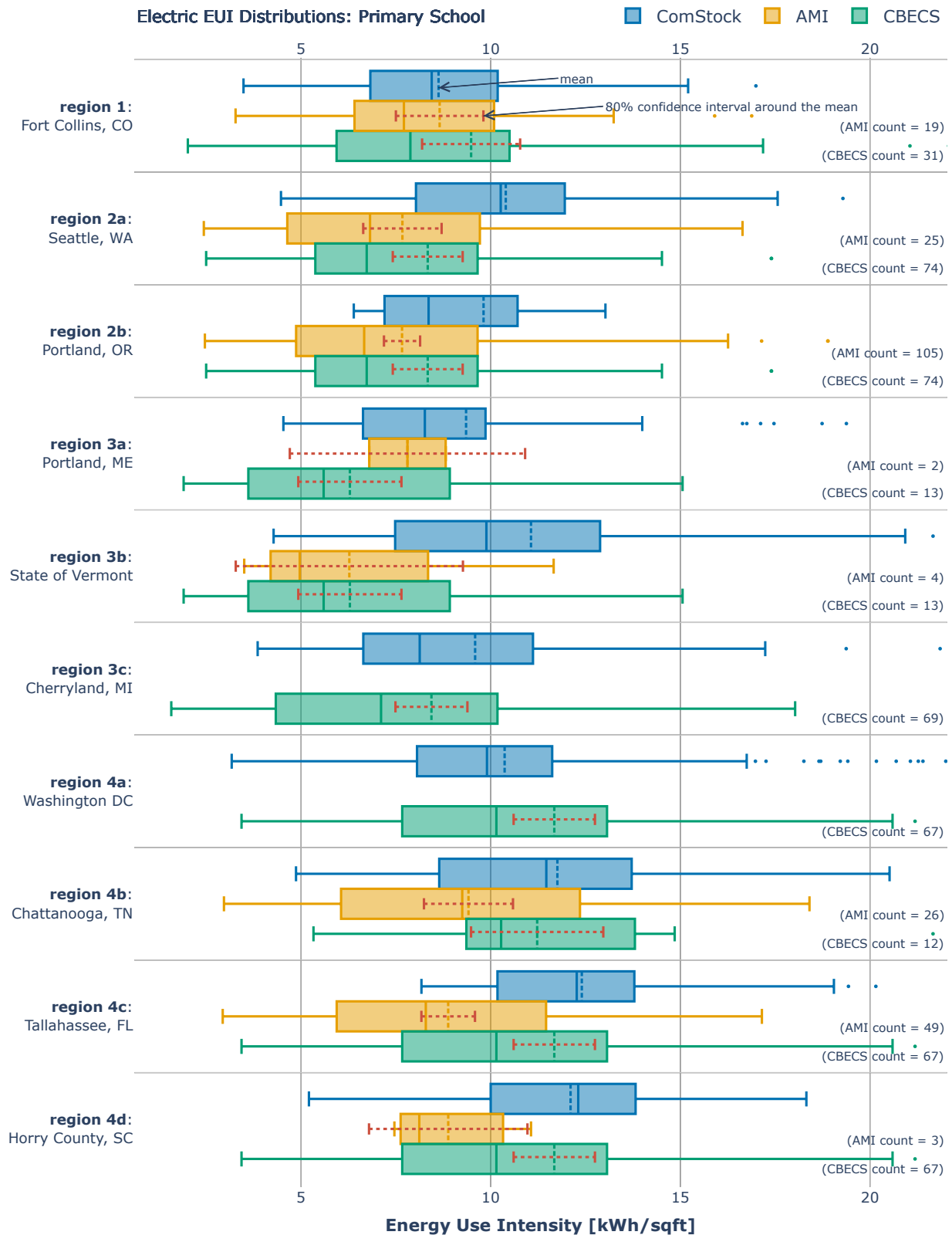




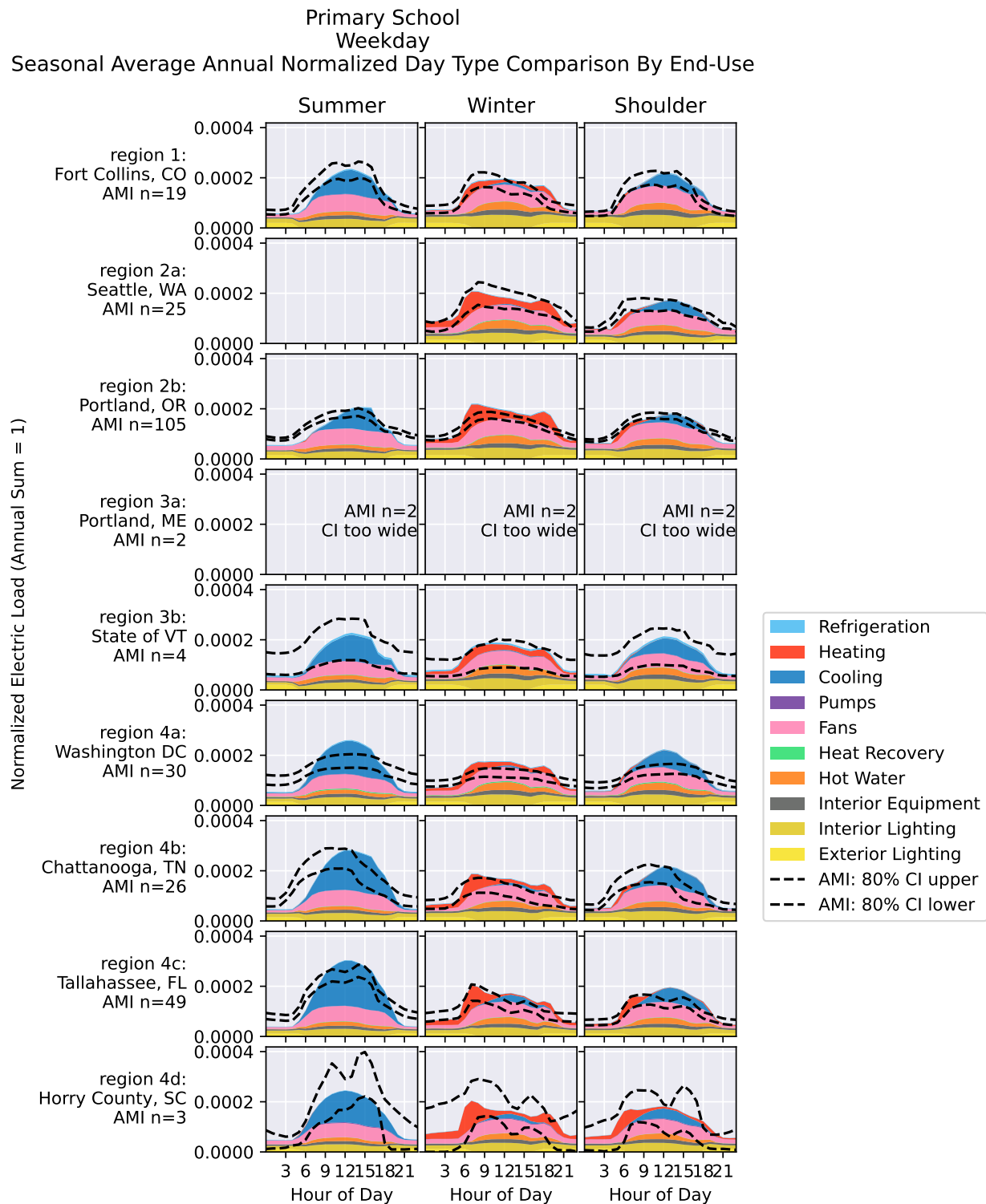
**Figure 248. Warehouse weekend seasonal average day type comparison by end use**

### *Primary School*

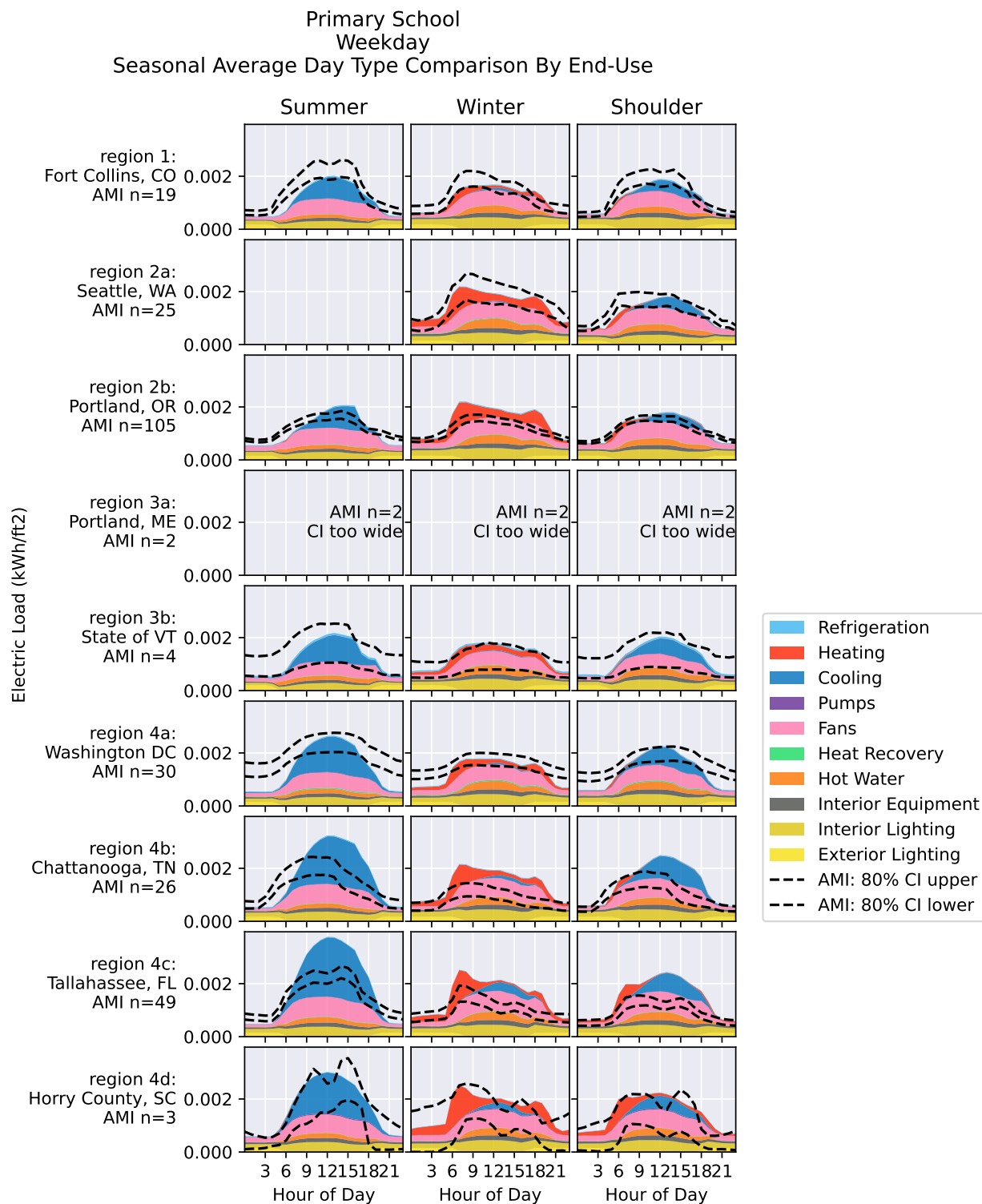
For primary schools, the EUIs for ComStock, AMI, and CBECS align well and the confidence intervals overlap. ComStock matches both the normalized and non-normalized weekday profiles well, with perhaps some underestimate in nighttime load. ComStock underestimates weekend operation based on both normalized and non-normalized load profile comparisons.



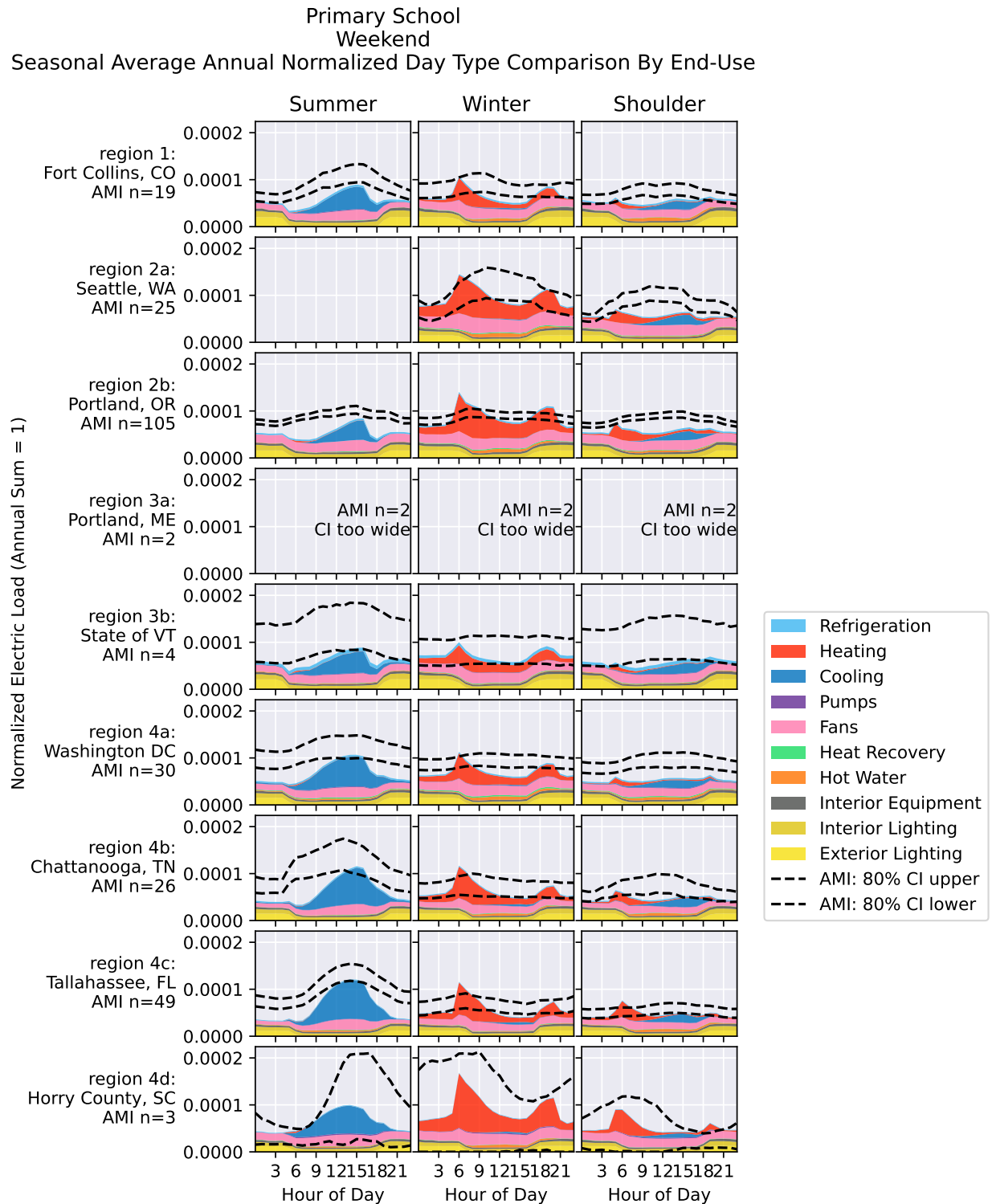
**Figure 249. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for primary schools**



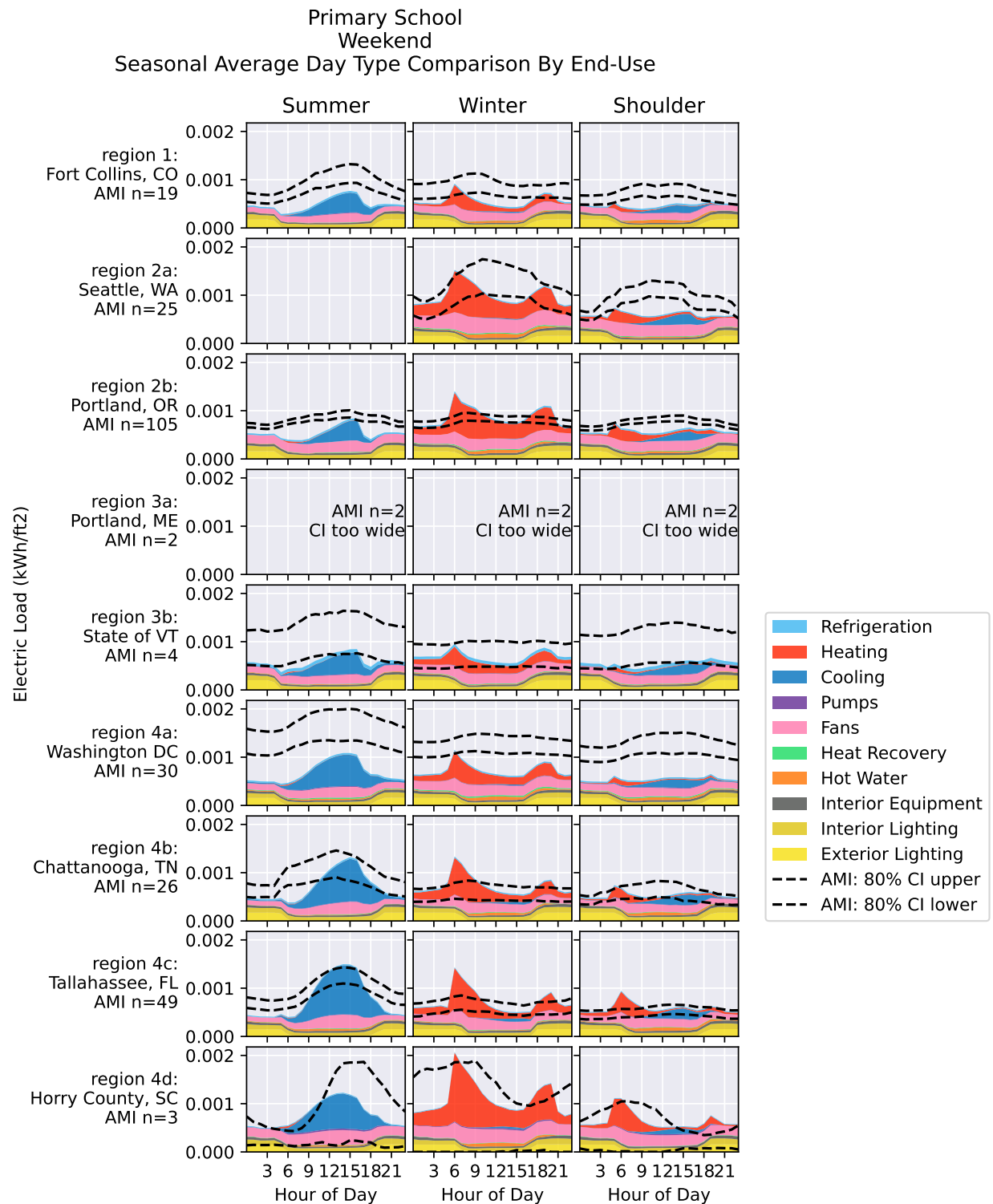
**Figure 250. Primary school weekday seasonal average annual normalized day type comparison by end use**



**Figure 251. Primary school weekday seasonal average day type comparison by end use**



**Figure 252. Primary school weekend seasonal average annual normalized day type comparison by end use**

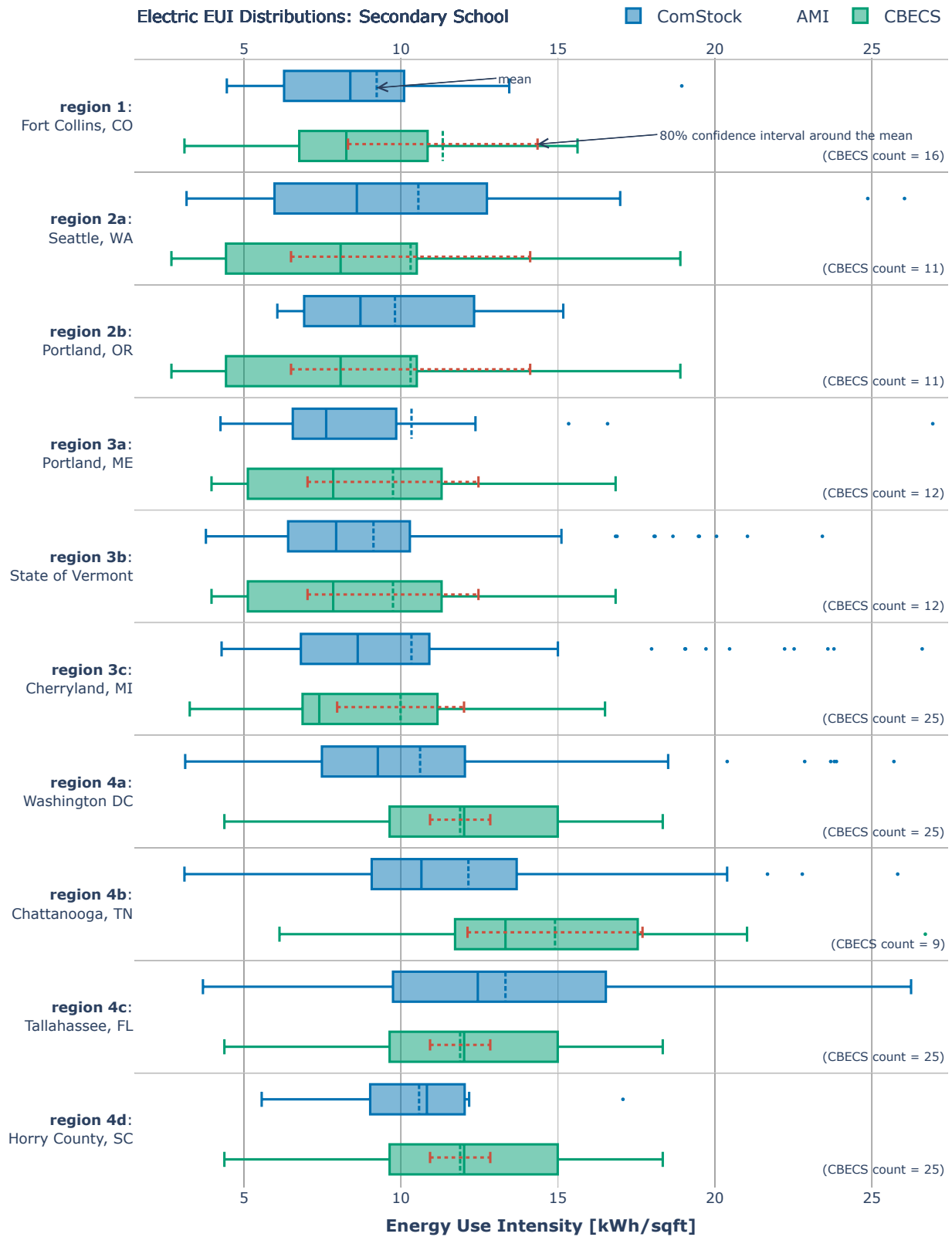


**Figure 253. Primary school weekend seasonal average day type comparison by end use**



### *Secondary School*

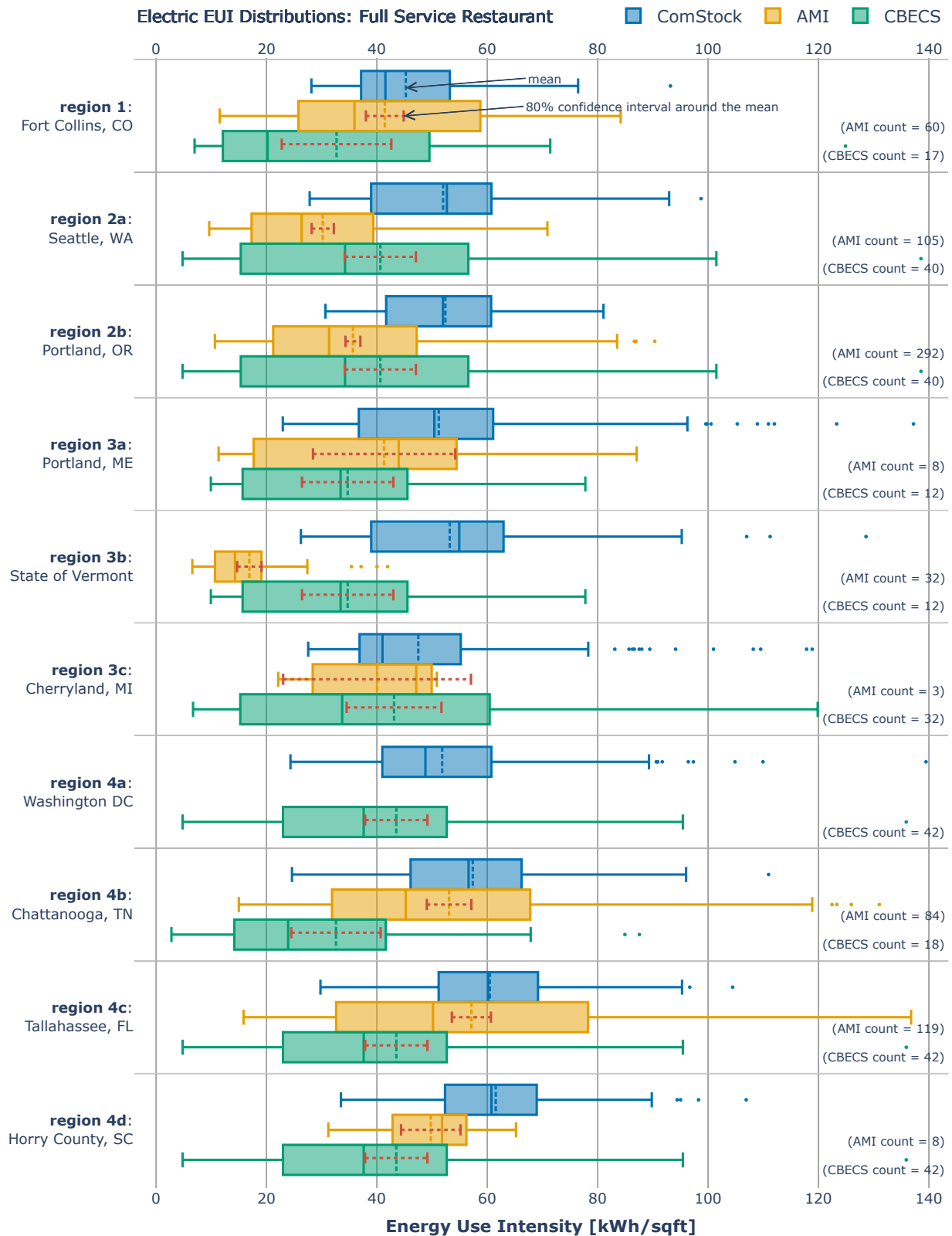
For secondary schools, there were no AMI data. Due to limitations in the building characteristics data, it was not possible to systematically differentiate between primary and secondary schools, but based on a review of sample buildings we concluded that most schools in the dataset were more like primary schools than secondary schools. The ComStock and CBECS distributions for this building type overlap broadly, with the ComStock mean generally falling within the CBECS confidence interval.



**Figure 254. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for secondary schools**

### *Full-Service Restaurant*

Full-service restaurants can cover a wide range of energy usage profiles, from breakfast-only to late-night bars. The relative comparison between the AMI and CBECS does not appear to have a strong pattern, but for regions with larger AMI samples the distributions overlap more. Broadly it appears that ComStock overestimates the mean or falls at the upper end of the EUI distributions. For weekends, where this type of building would be expected to be busiest, the normalized and non-normalized load profiles show good agreement for regions with larger samples. However, for these same regions, ComStock overestimates the weekday load profile magnitude. This could explain the overall high EUIs in ComStock.



**Figure 255. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for full-service restaurants**

Full Service Restaurant  
Weekday  
Seasonal Average Annual Normalized Day Type Comparison By End-Use

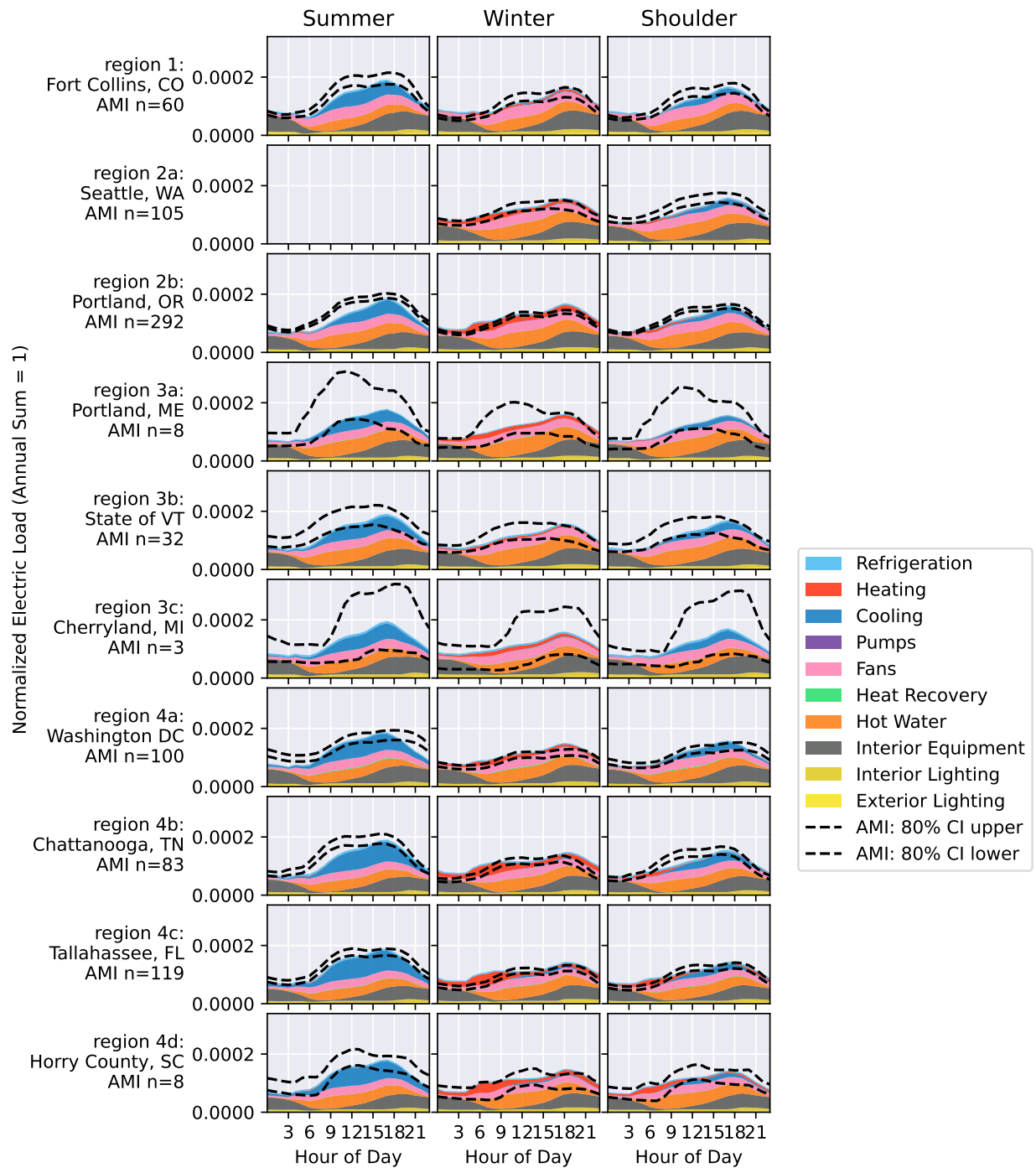
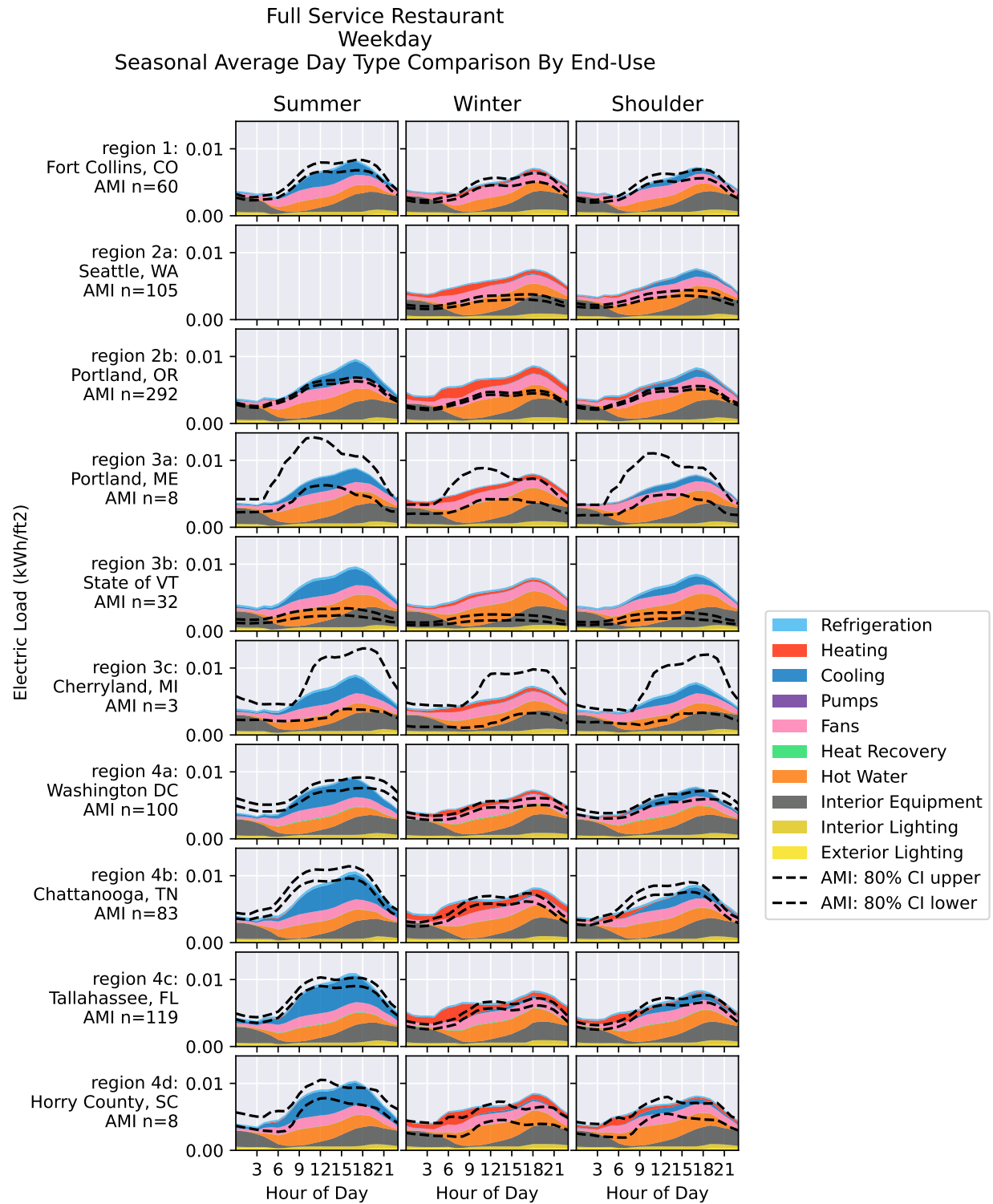


Figure 256. Full-service restaurant weekday seasonal average annual normalized day type comparison by end use



**Figure 257. Full-service restaurant weekday seasonal average day type comparison by end use**

Full Service Restaurant  
Weekend  
Seasonal Average Annual Normalized Day Type Comparison By End-Use

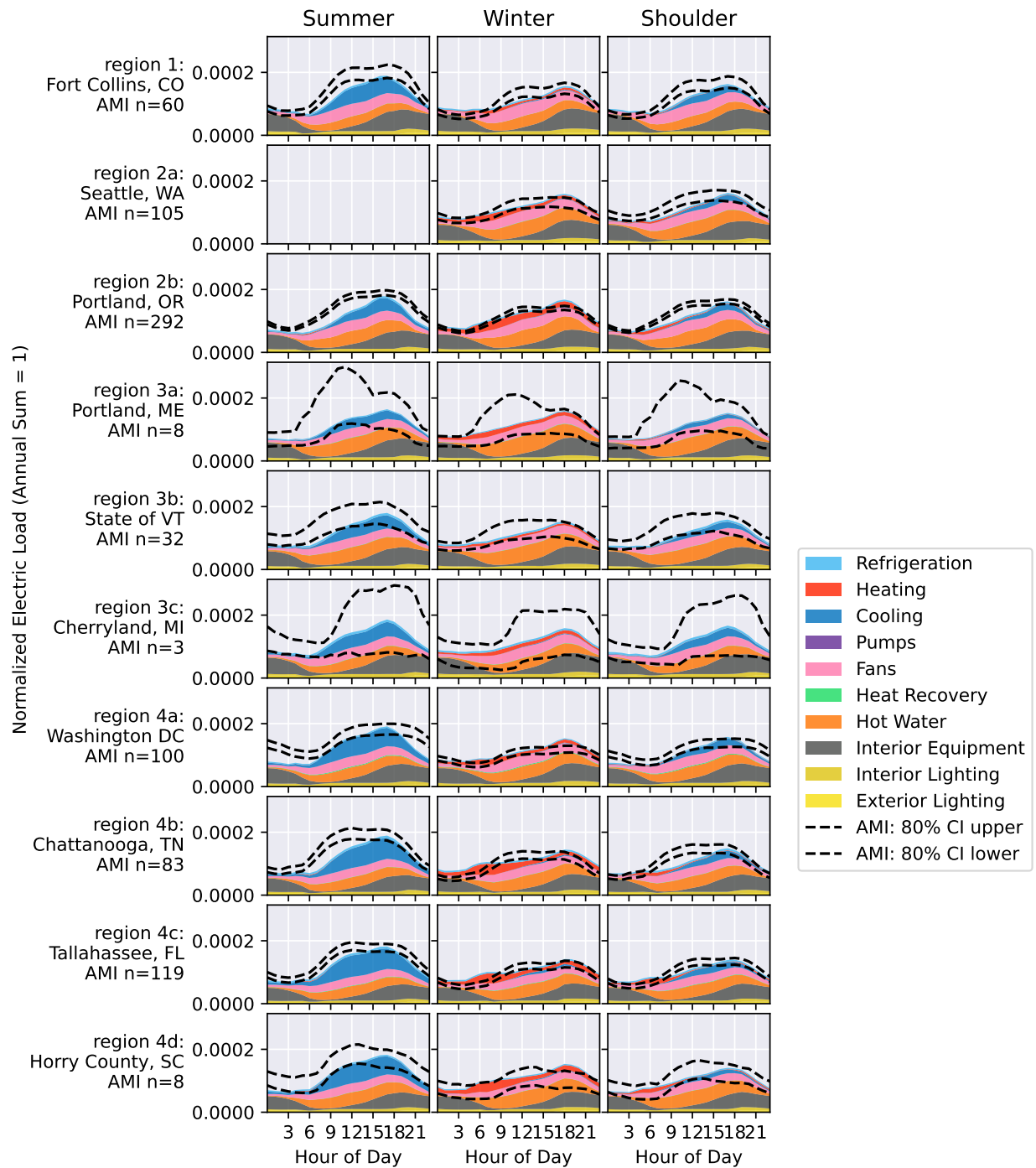
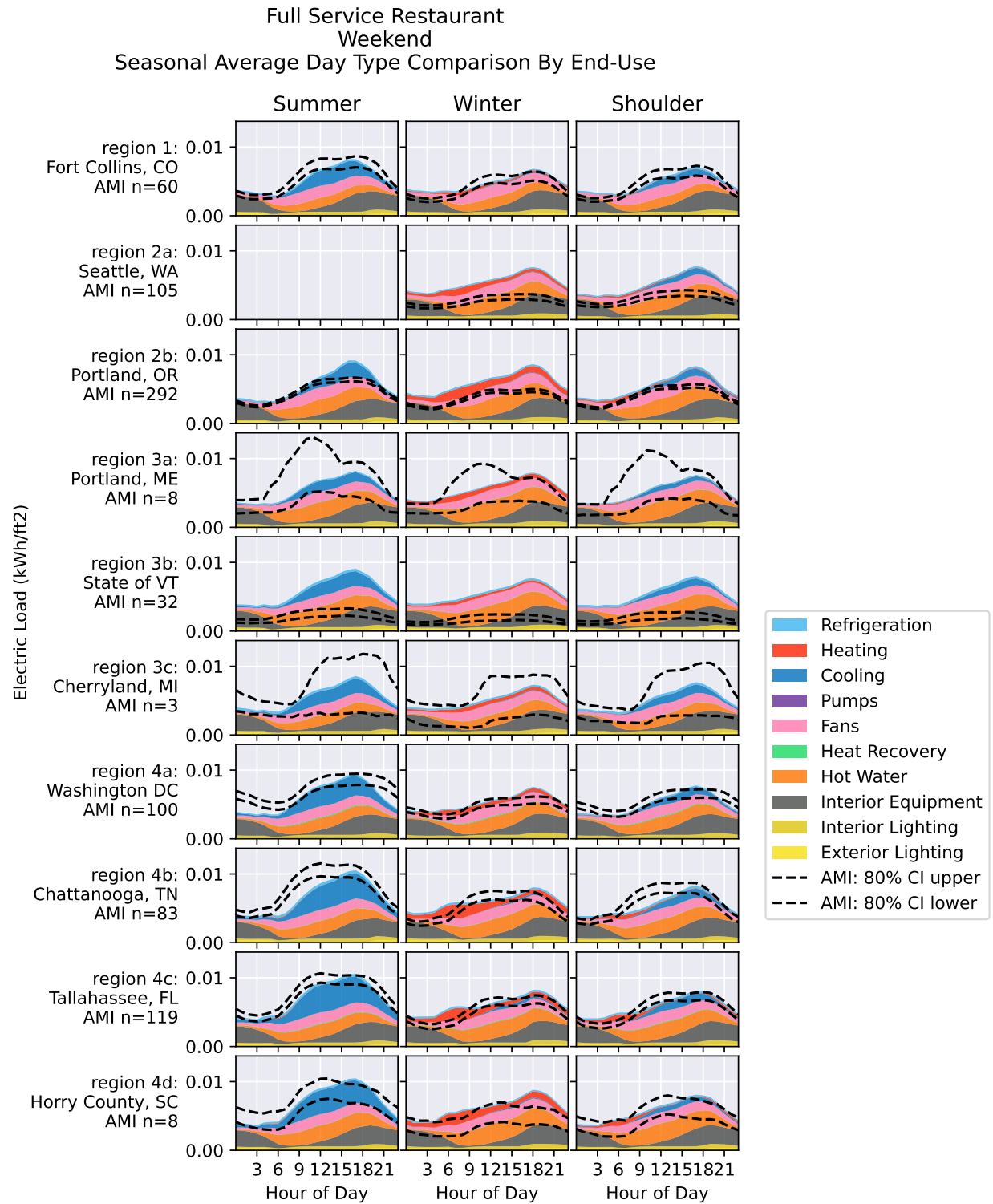


Figure 258. Full-service restaurant weekend seasonal average annual normalized day type comparison by end use

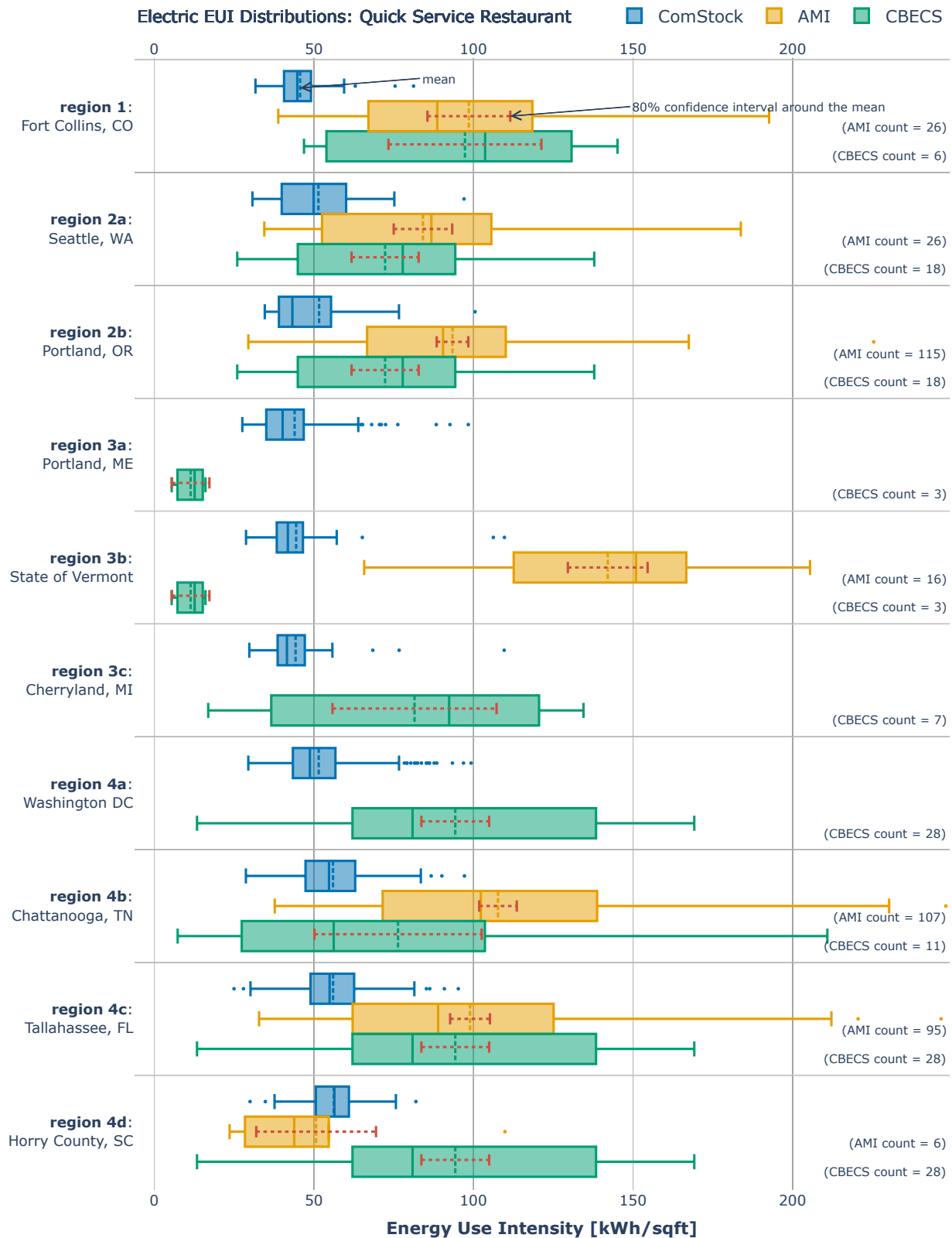




**Figure 259. Full-service restaurant weekend seasonal average day type comparison by end use**

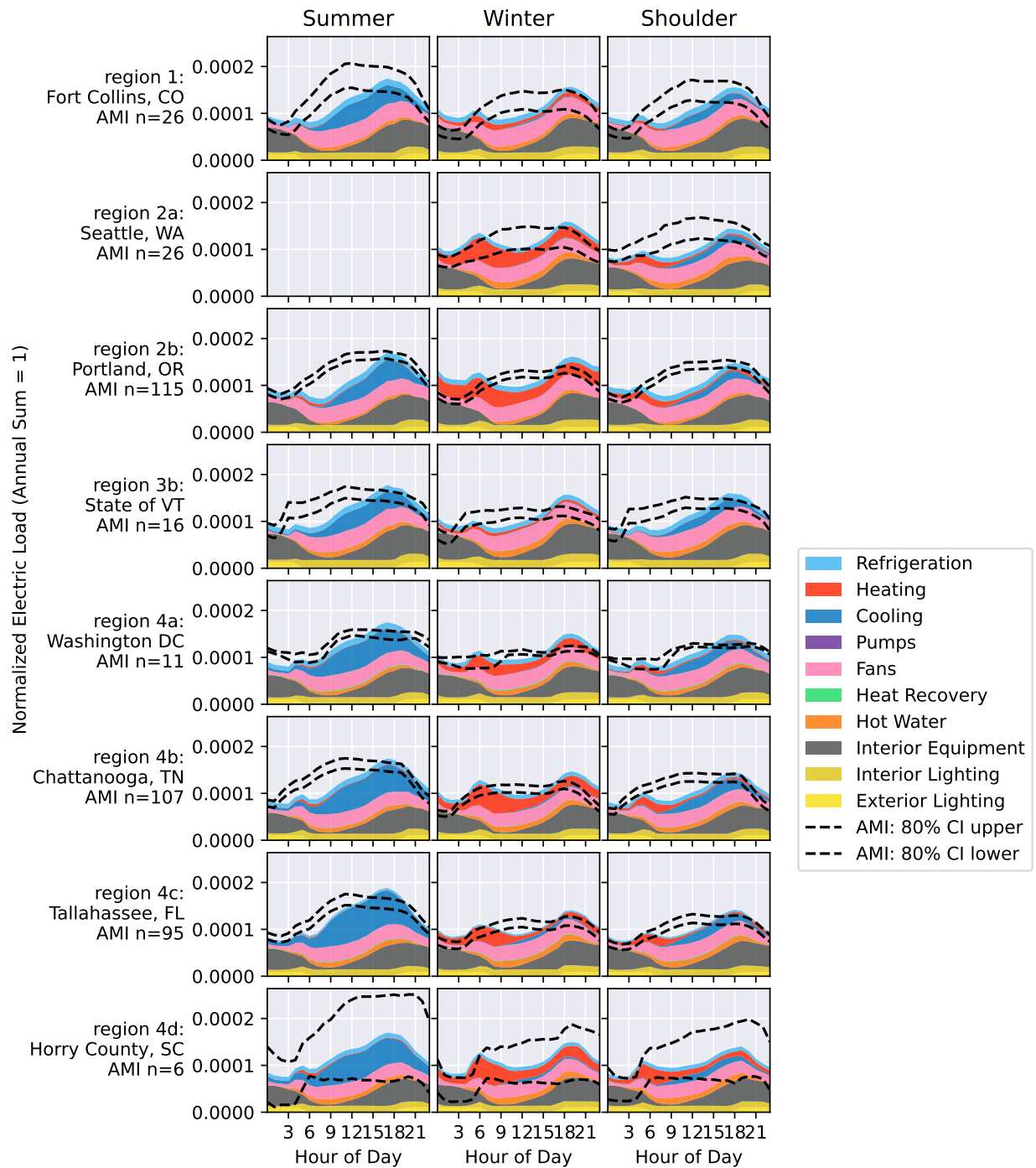
### *Quick-Service Restaurant*

For quick-service restaurants, ComStock has a narrower EUI distribution than both the AMI and CBECS across all regions. This distribution shows that ComStock is generally underestimating energy consumption, which is corroborated by the ComStock to CBECS annual comparisons in Figure 191. Looking at the normalized load profiles, it appears that ComStock is underestimating the morning load, suggesting that either the end-use data used to determine hours of operation or the internal equipment schedule is biased toward lunch/dinner restaurants. This underestimate appears on both weekdays and weekends and holds across regions.

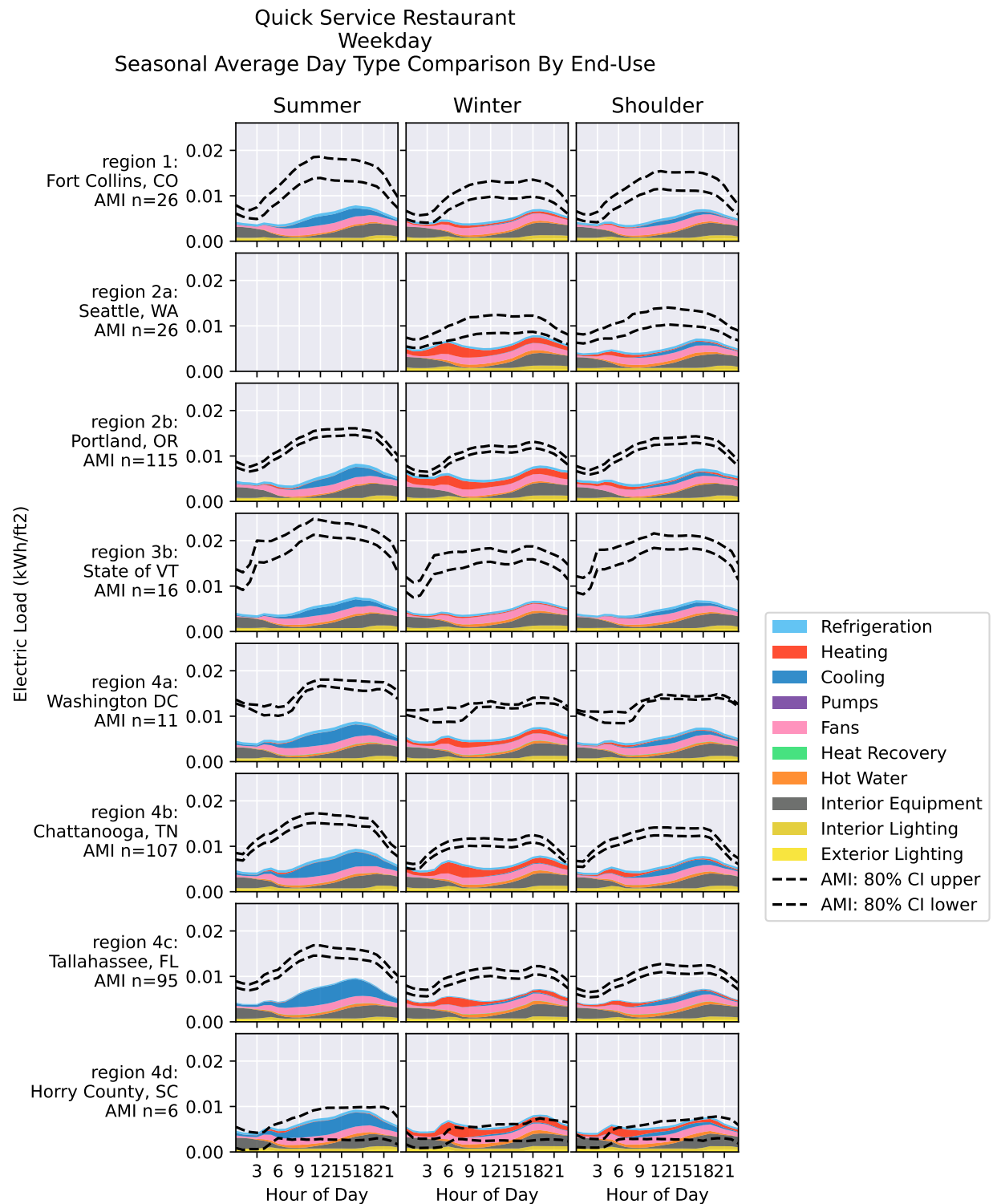


**Figure 260. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for quick-service restaurants**

Quick Service Restaurant  
Weekday  
Seasonal Average Annual Normalized Day Type Comparison By End-Use

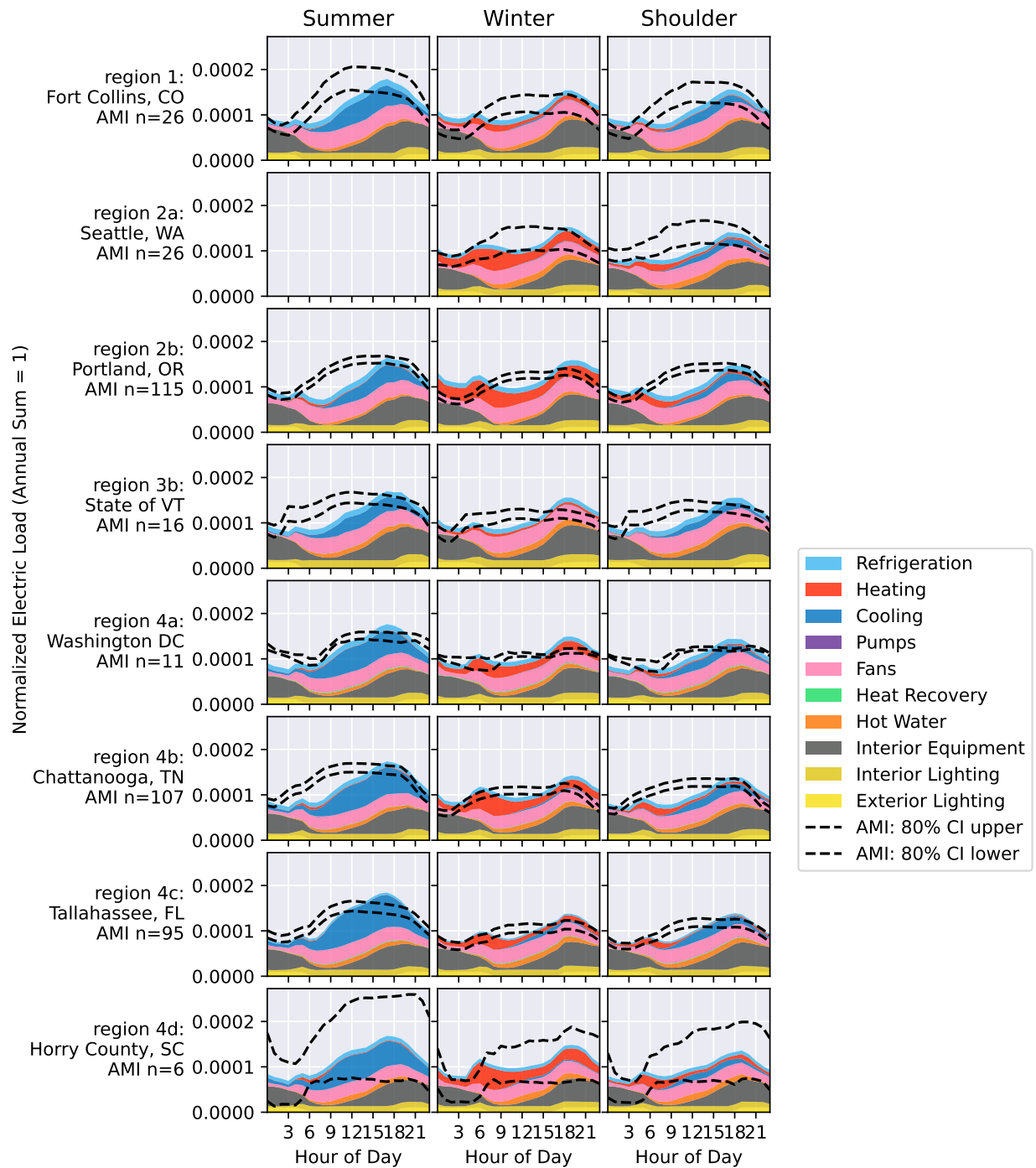


**Figure 261. Quick-service restaurant weekday seasonal average annual normalized day type comparison by end use**

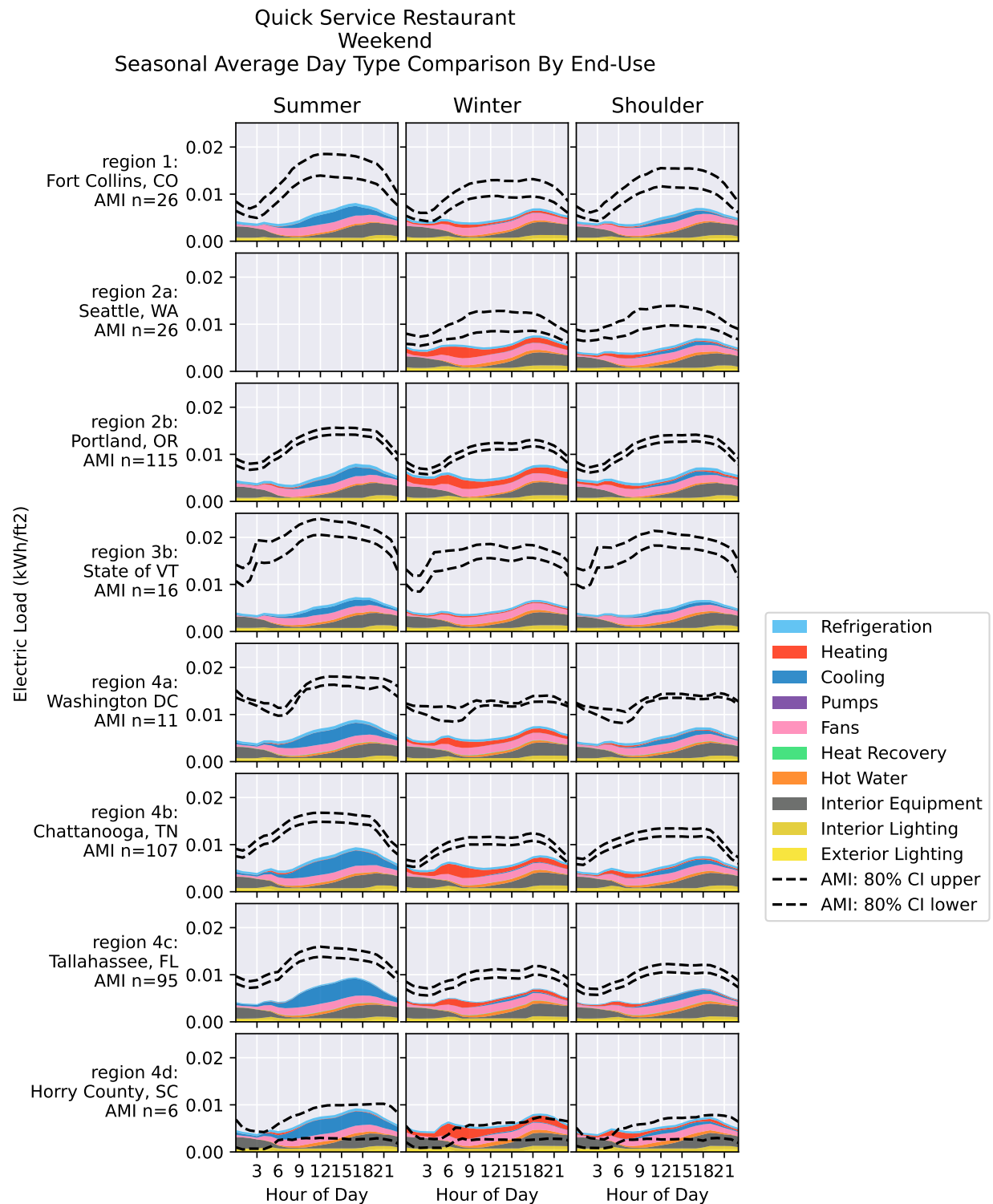


**Figure 262. Quick-service restaurant weekday seasonal average day type comparison by end use**

Quick Service Restaurant  
Weekend  
Seasonal Average Annual Normalized Day Type Comparison By End-Use



**Figure 263. Quick-service restaurant weekend seasonal average annual normalized day type comparison by end use**

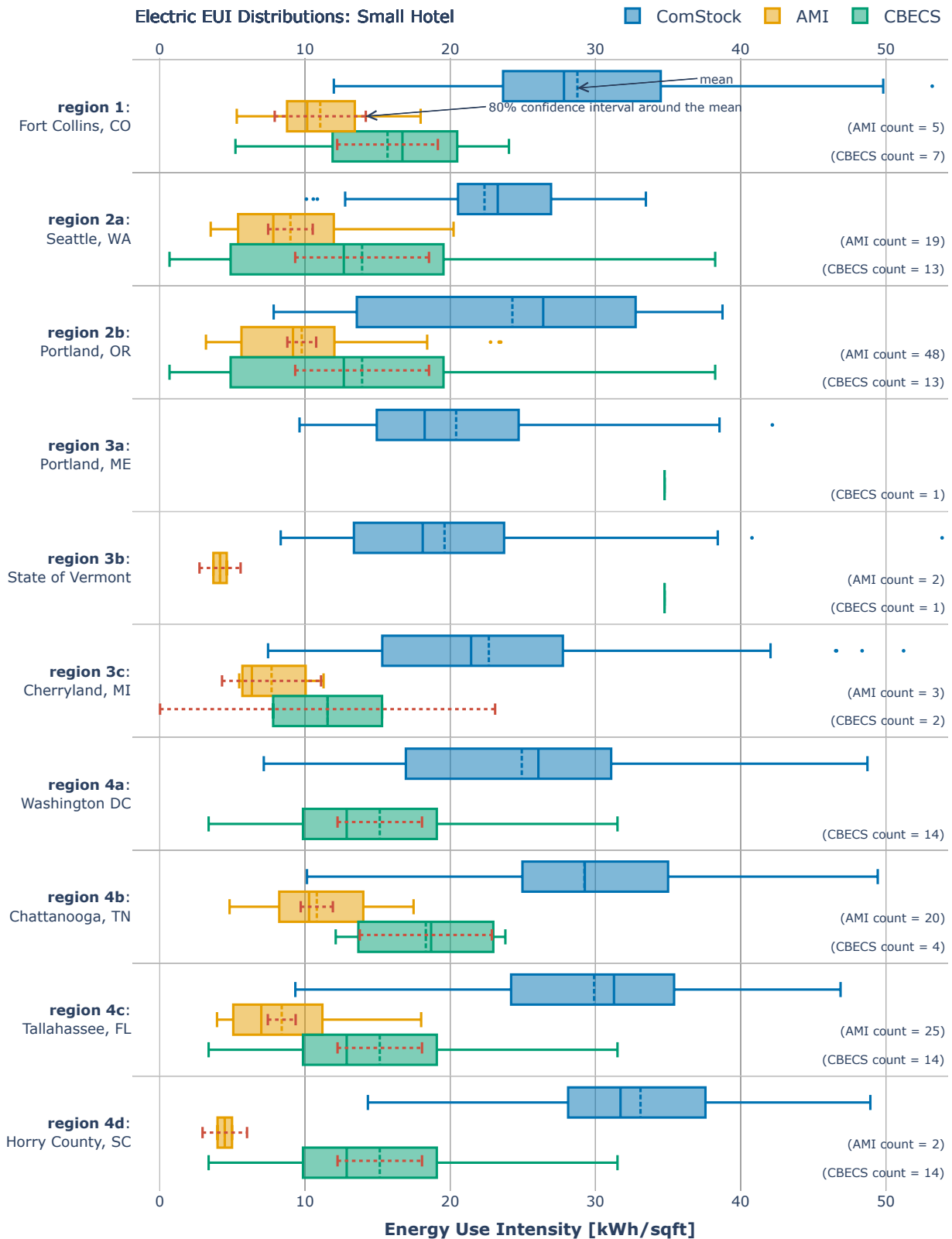


**Figure 264. Quick-service restaurant weekend seasonal average day type comparison by end use**

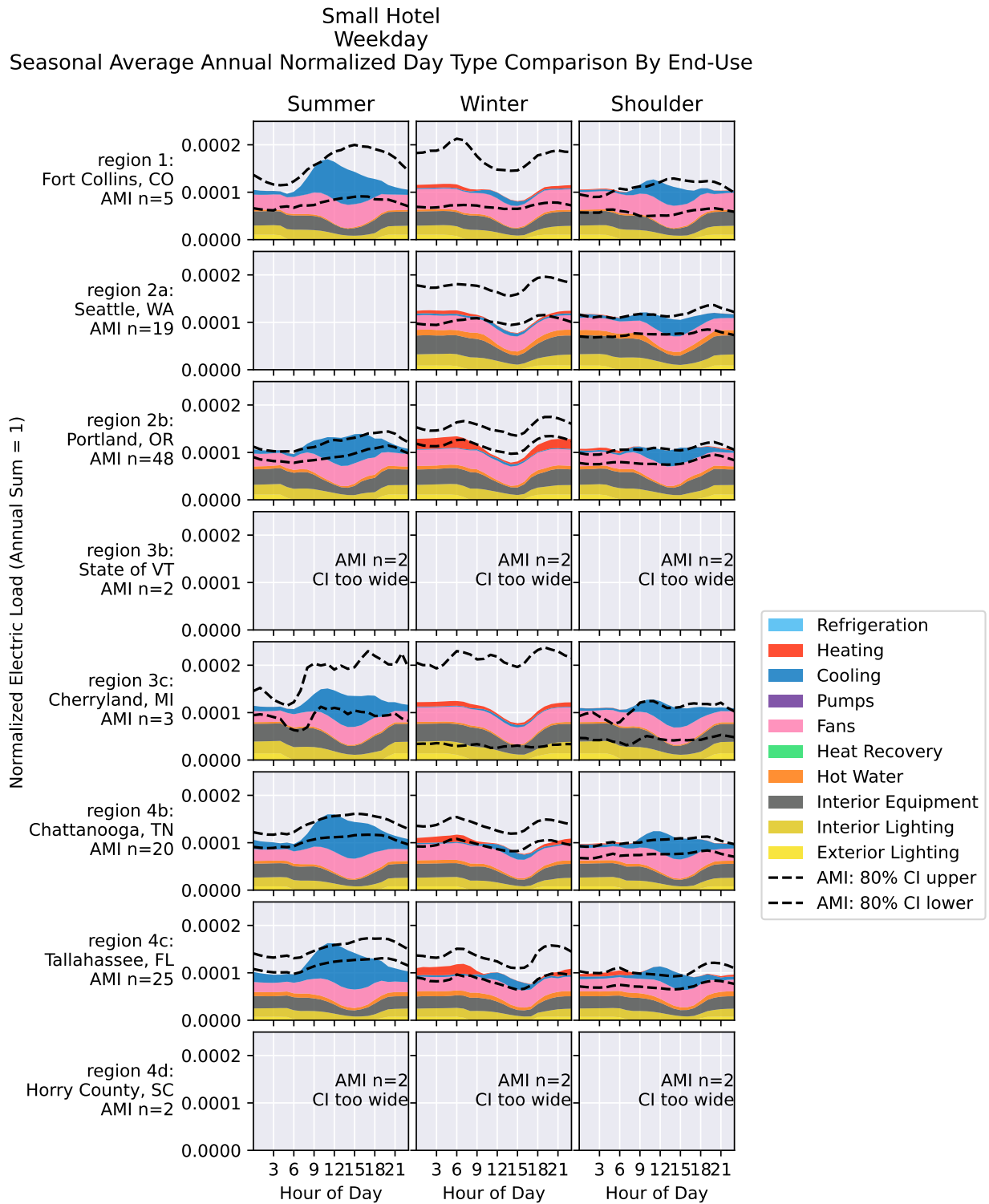


### *Small Hotel*

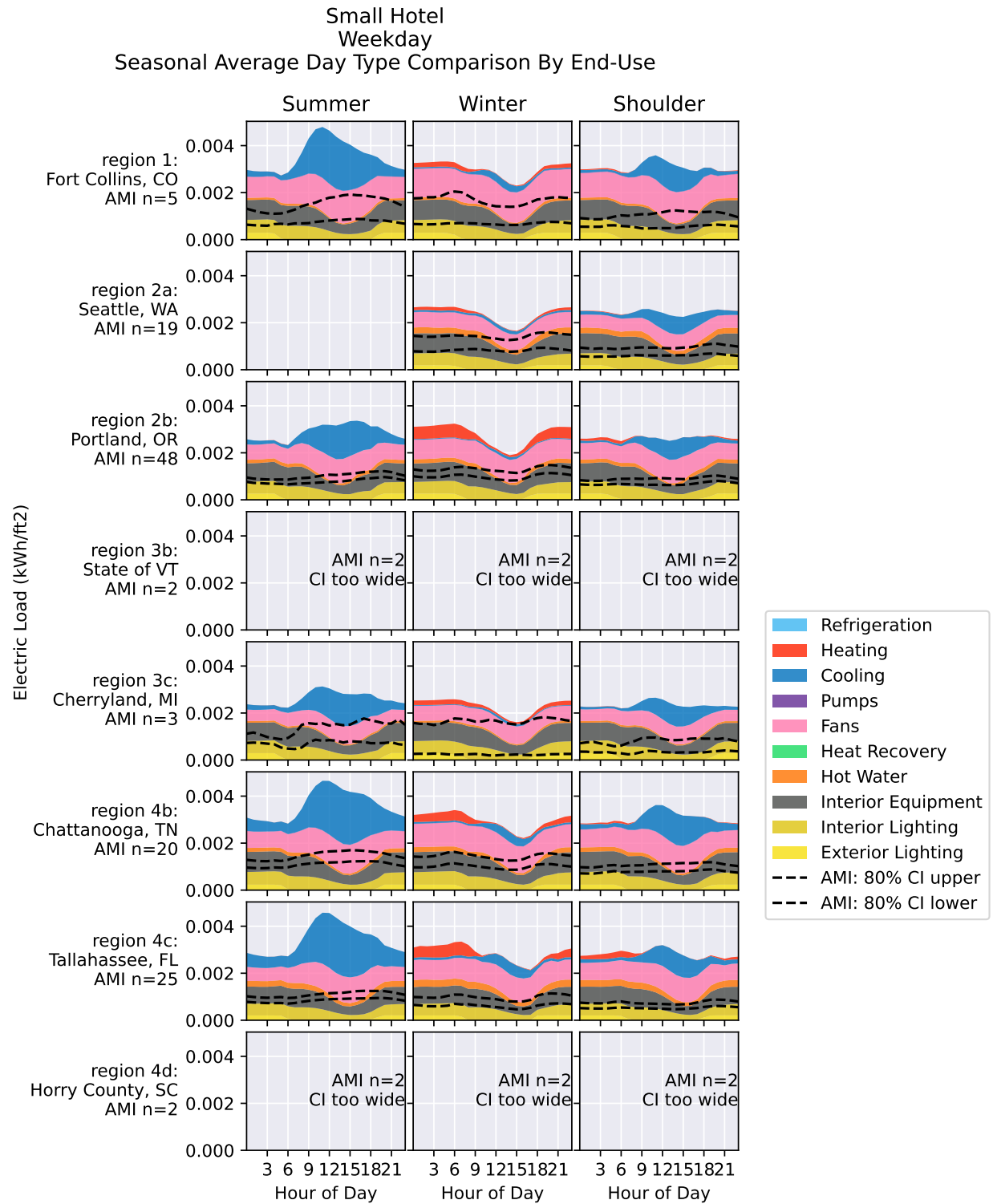
For small hotels, the EUI distributions show that ComStock generally overestimates the mean energy consumption. This is corroborated by the ComStock to CBECS annual comparisons in Figure 191. The normalized load profiles show that the AMI load shape is generally flatter than what ComStock models, indicating potential overestimation of a nighttime thermostat setback. Despite the wide confidence interval for this building type across regions, the overestimation appears likely to be a result of overestimation in several end uses.



**Figure 265. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for small hotels**



**Figure 266. Small hotel weekday seasonal average annual normalized day type comparison by end use**



**Figure 267. Small hotel weekday seasonal average day type comparison by end use**

Small Hotel  
Weekend  
Seasonal Average Annual Normalized Day Type Comparison By End-Use

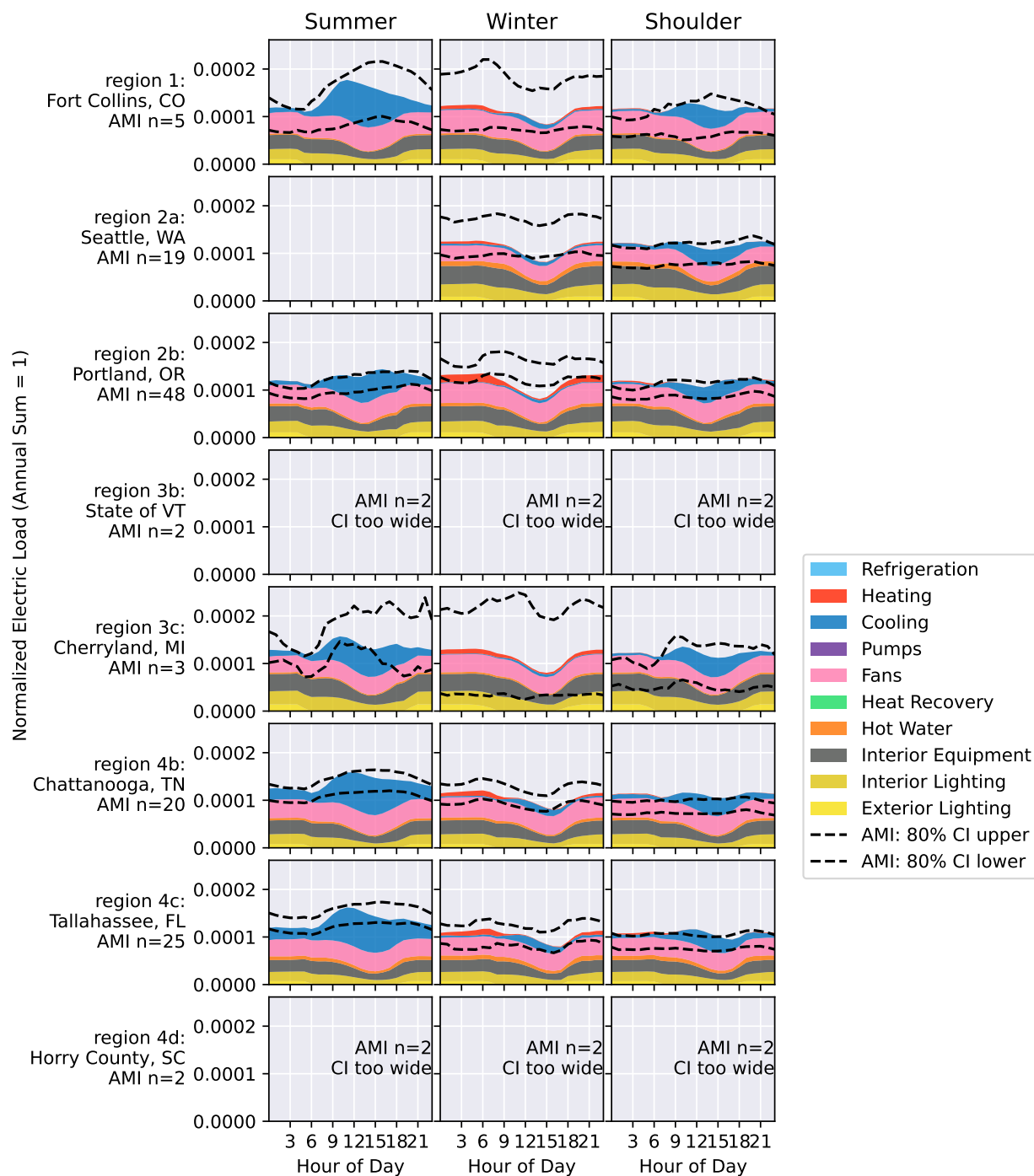
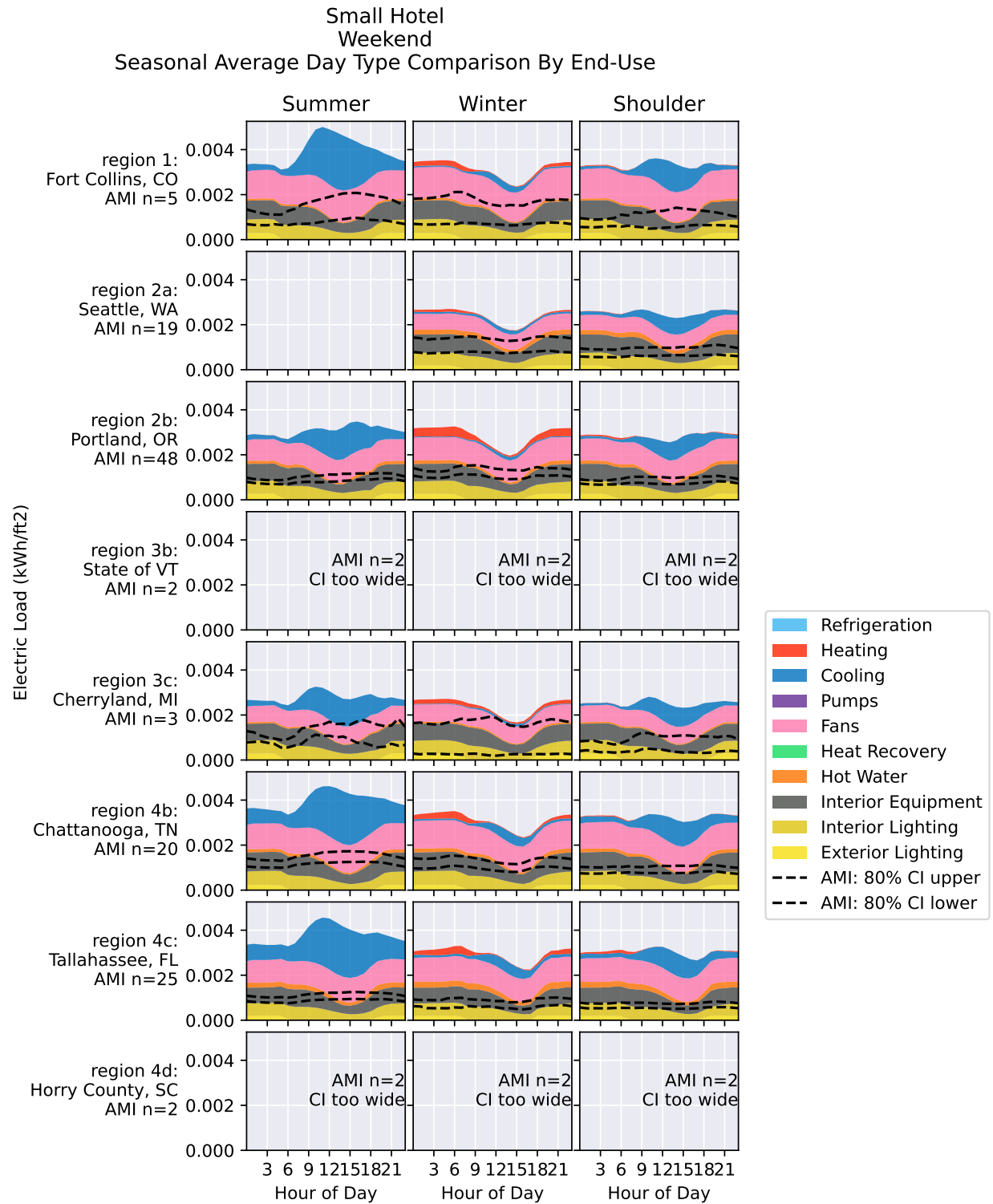


Figure 268. Small hotel weekend seasonal average annual normalized day type comparison by end use



**Figure 269. Small hotel weekend seasonal average day type comparison by end use**

### *Large Hotel*

For large hotels, the EUI distributions show that the AMI is generally lower than CBECS, but the distributions do overlap somewhat. In regions outside the Pacific Northwest, the ComStock mean falls inside the CBECS confidence interval. In the Pacific Northwest, ComStock appears to overestimate the mean EUI by a bit. At a national and annual scale, Figure 191 shows close agreement between ComStock and CBECS, indicating that ComStock is likely slightly overestimating the EUIs on the whole, but that regional differences do exist. The normalized load profiles show that ComStock is missing nighttime load, and that the load shape is generally too strongly morning peaking, indicating that perhaps ComStock is overestimating the prevalence of nighttime thermostat setbacks and potentially also overestimating the prevalence of electric water heating in this building type, which leads to a morning peak.



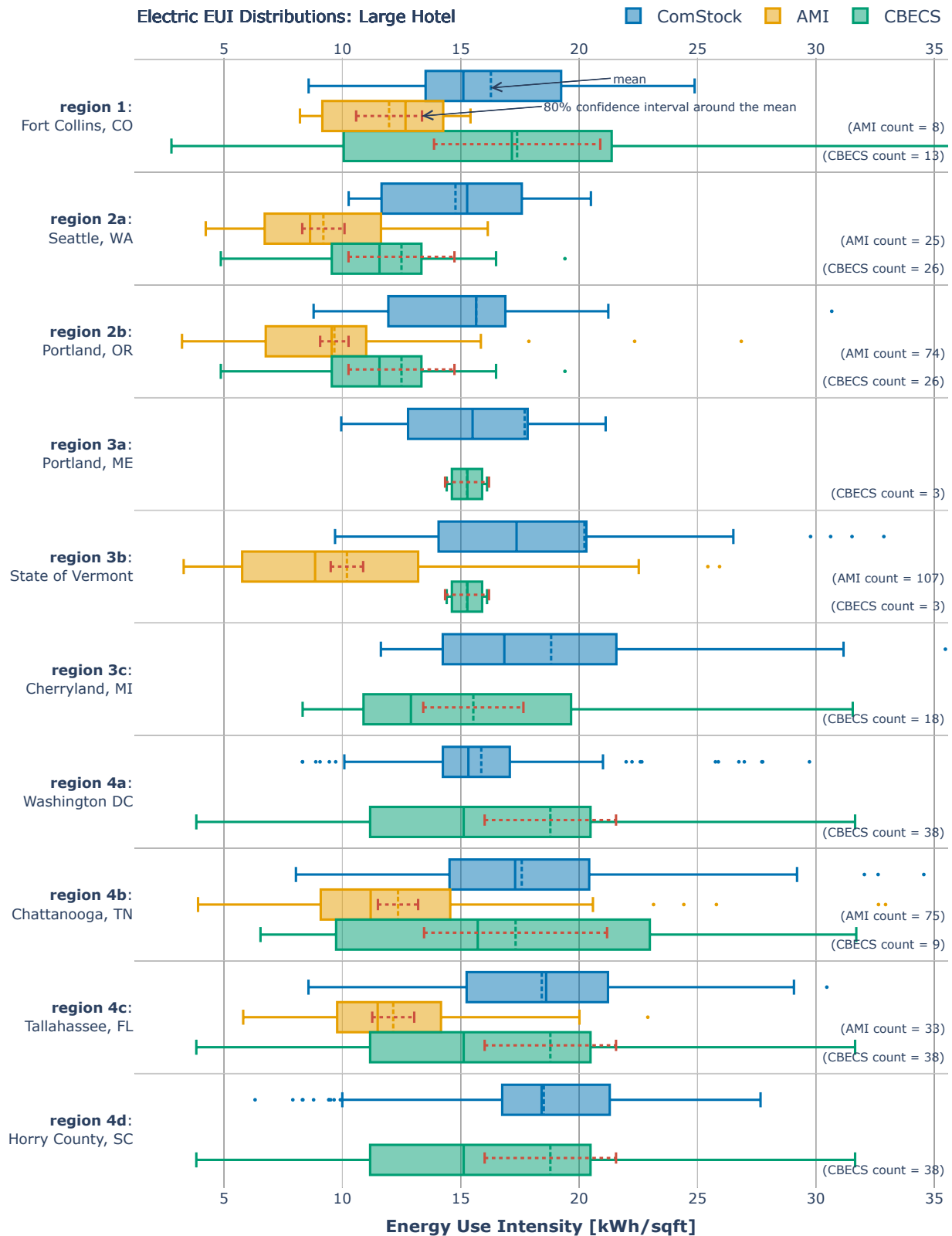


Figure 270. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for large hotels

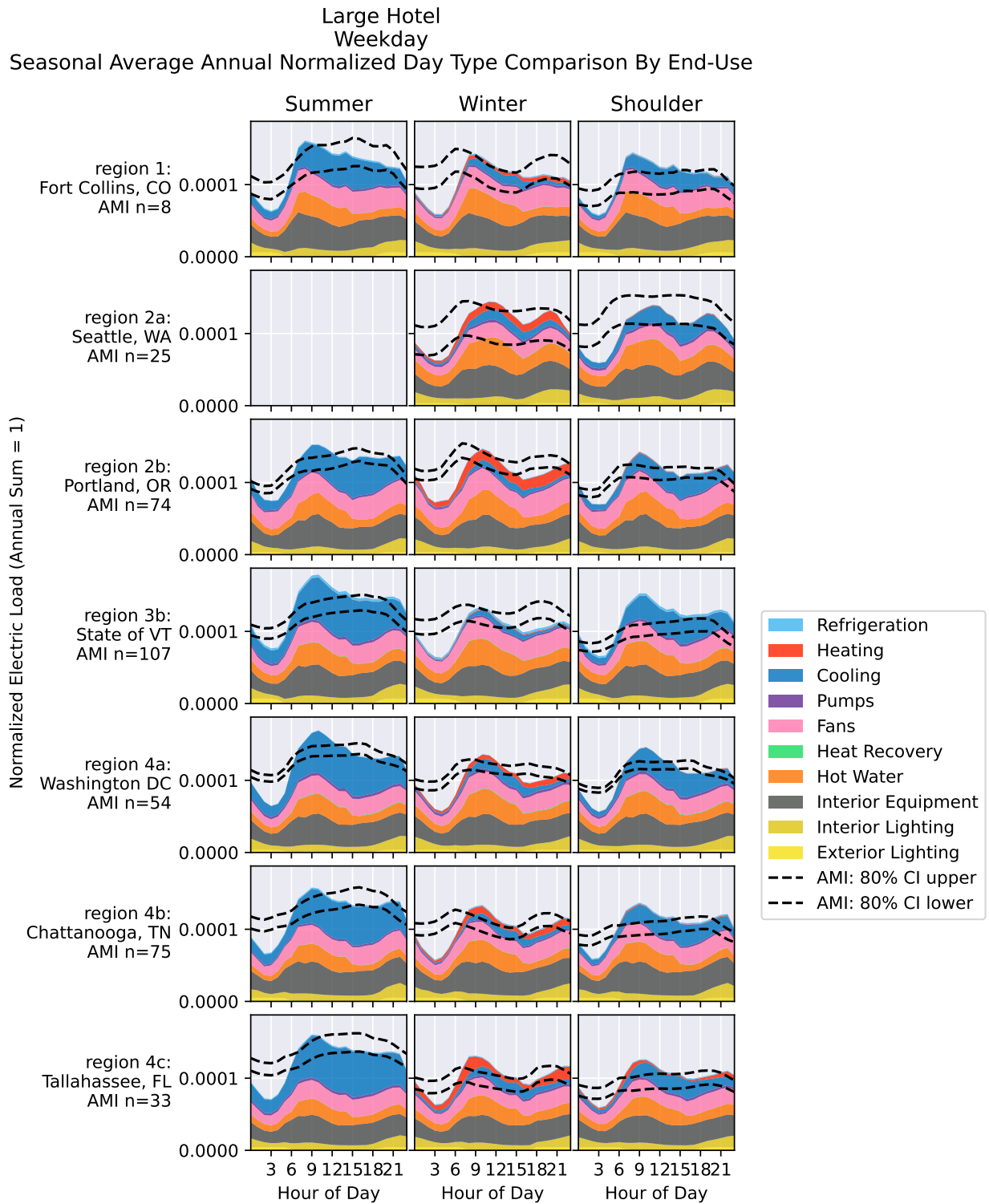
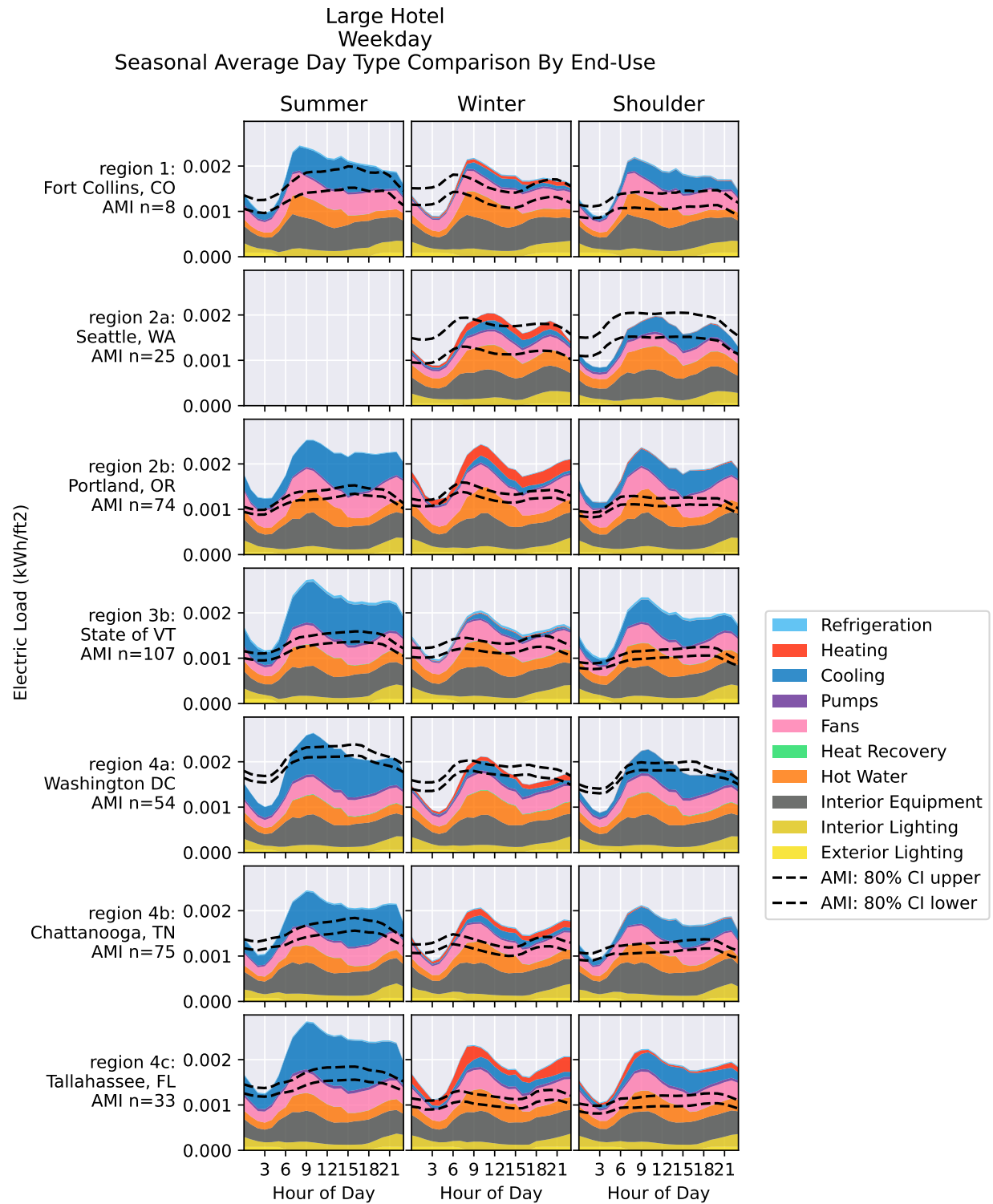


Figure 271. Large hotel weekday seasonal average annual normalized day type comparison by end use



**Figure 272. Large hotel weekday seasonal average day type comparison by end use**

Large Hotel  
Weekend  
Seasonal Average Annual Normalized Day Type Comparison By End-Use

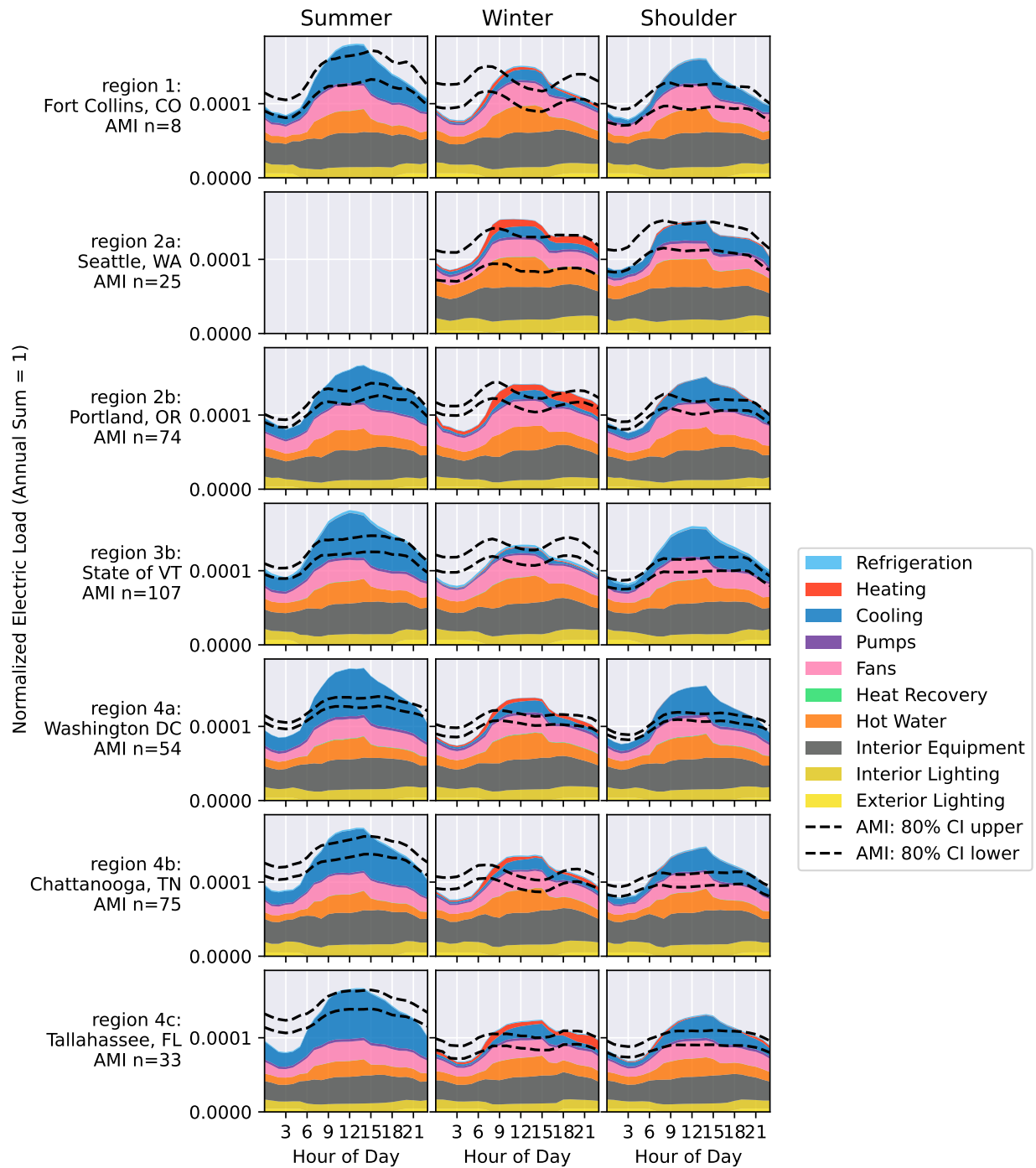
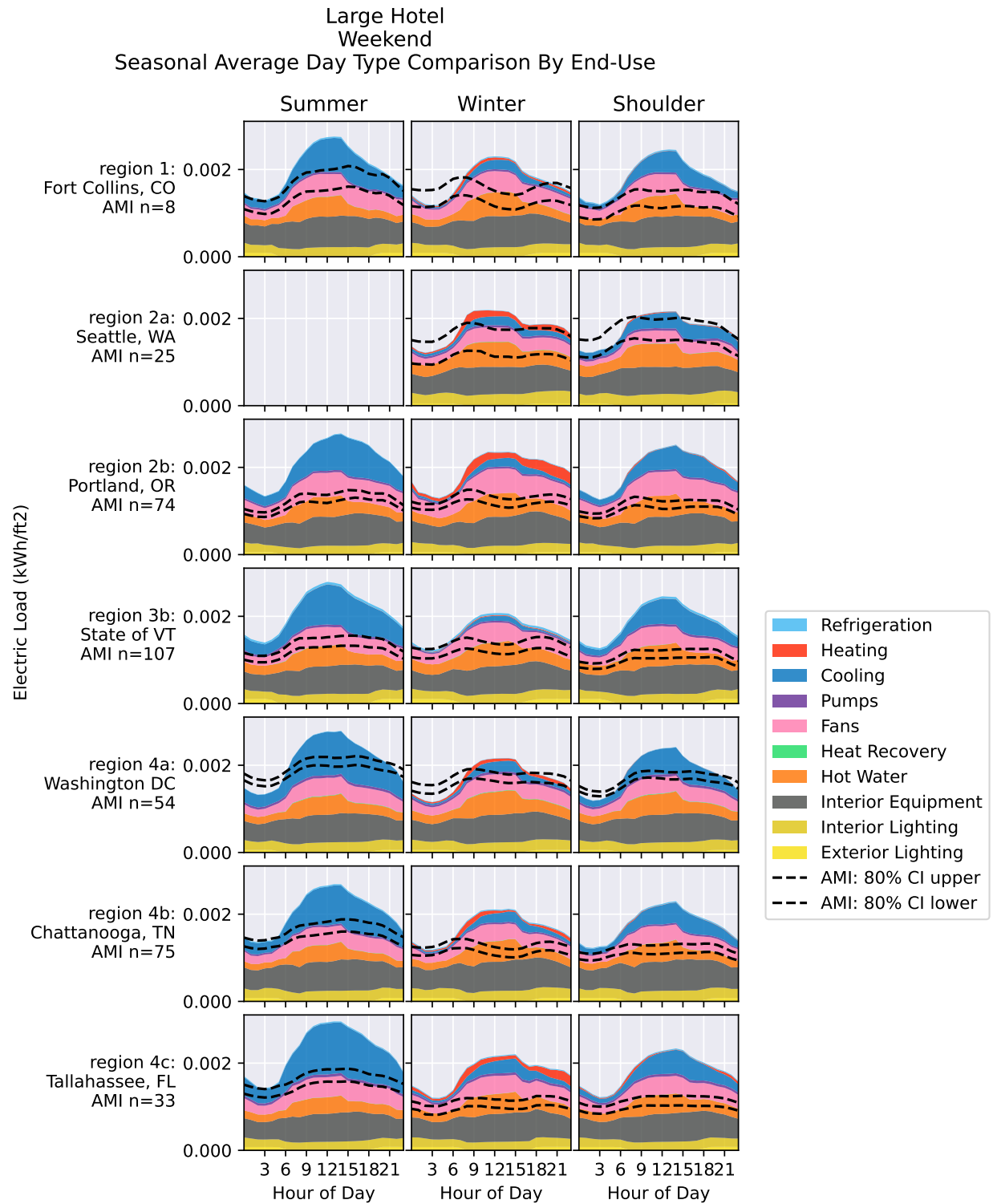


Figure 273. Large hotel weekend seasonal average annual normalized day type comparison by end use



**Figure 274. Large hotel weekend seasonal average day type comparison by end use**

### *Hospital*

For hospitals, the EUI distributions show that ComStock underestimates the EUI across all regions. This is corroborated at a national and annual scale by Figure 191. The AMI sample size was low across all regions, so the confidence intervals are very wide as a result. The limited normalized and non-normalized load profiles do indicate that the true load shape is flatter than what ComStock models.

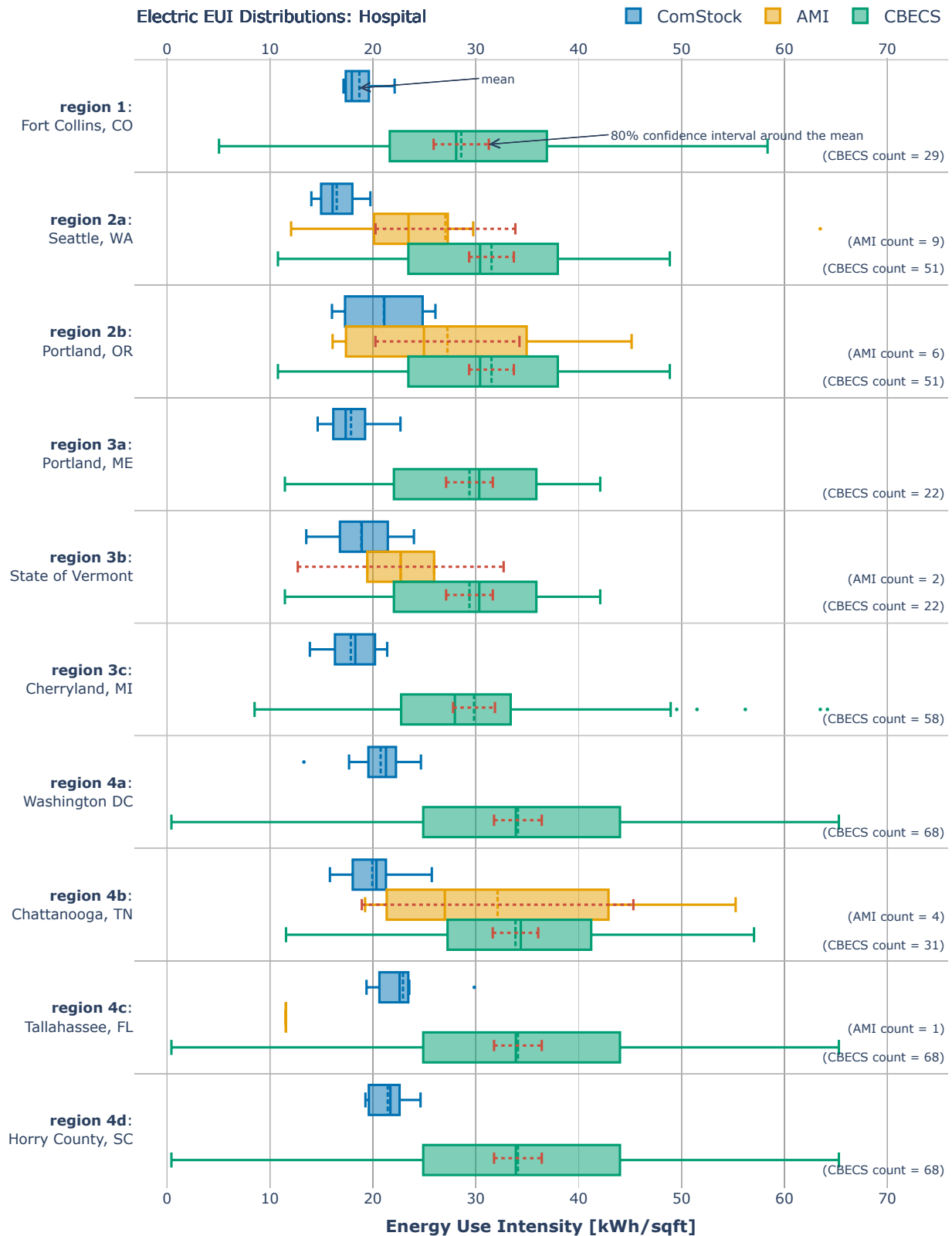
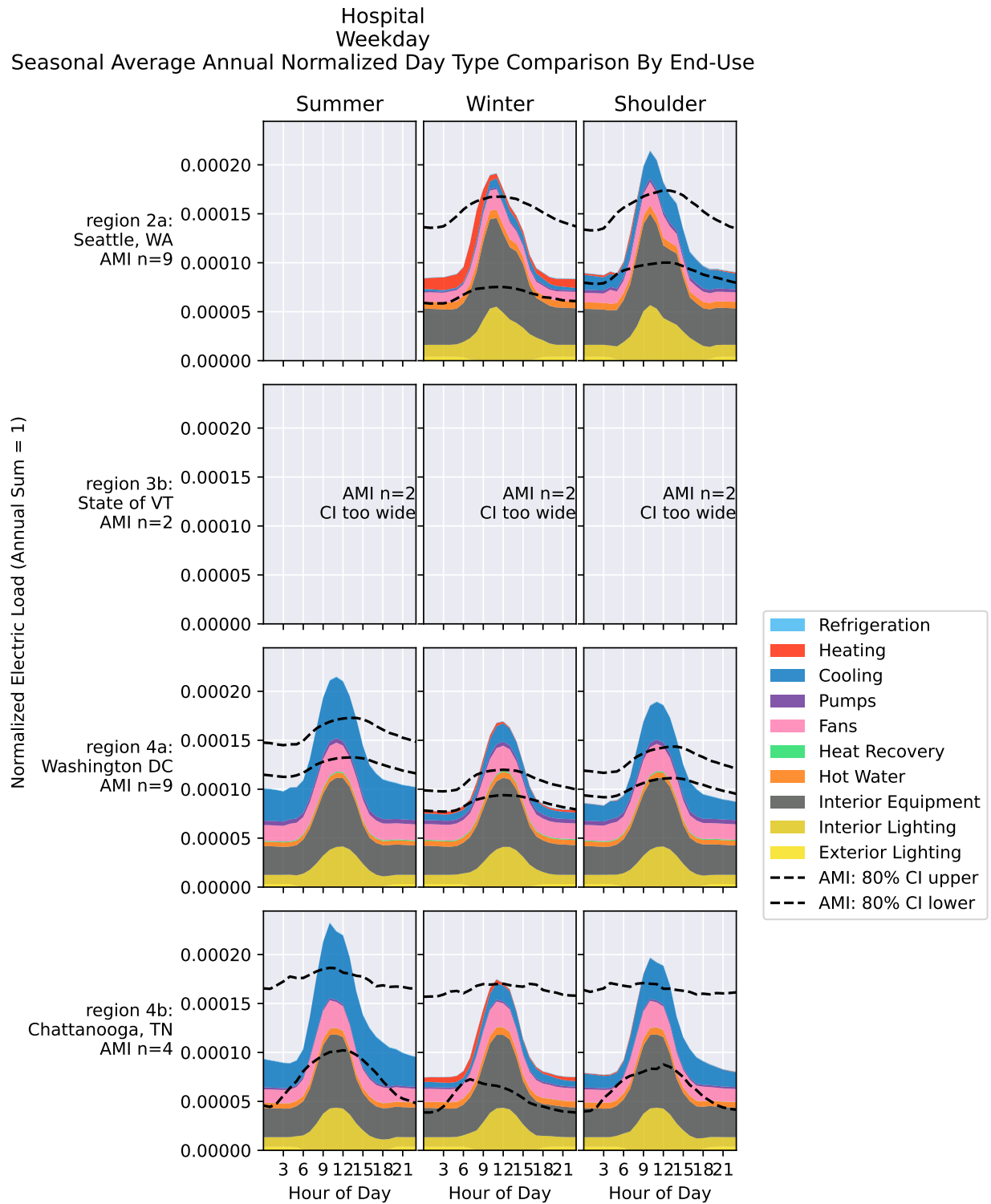
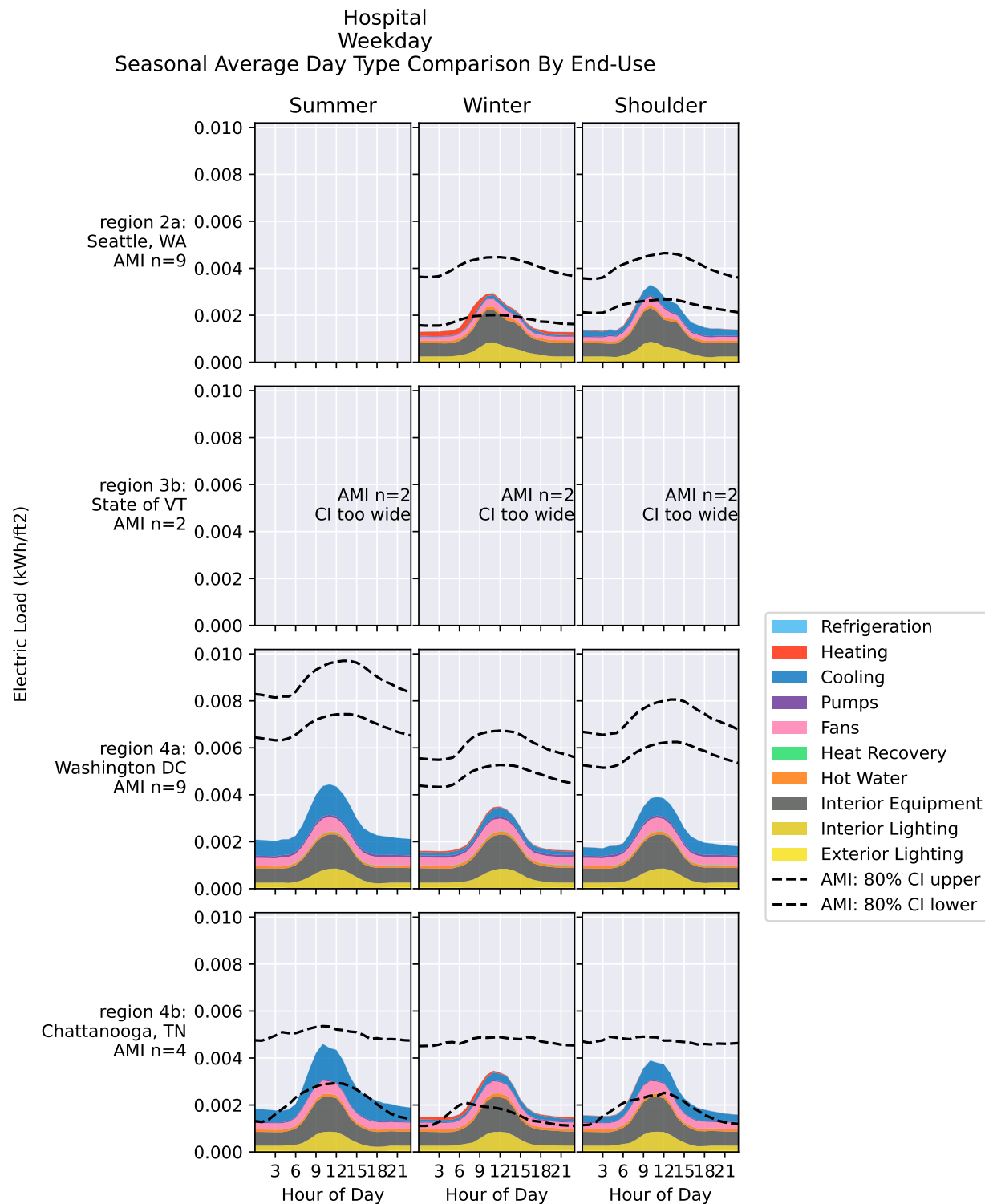


Figure 275. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for hospitals

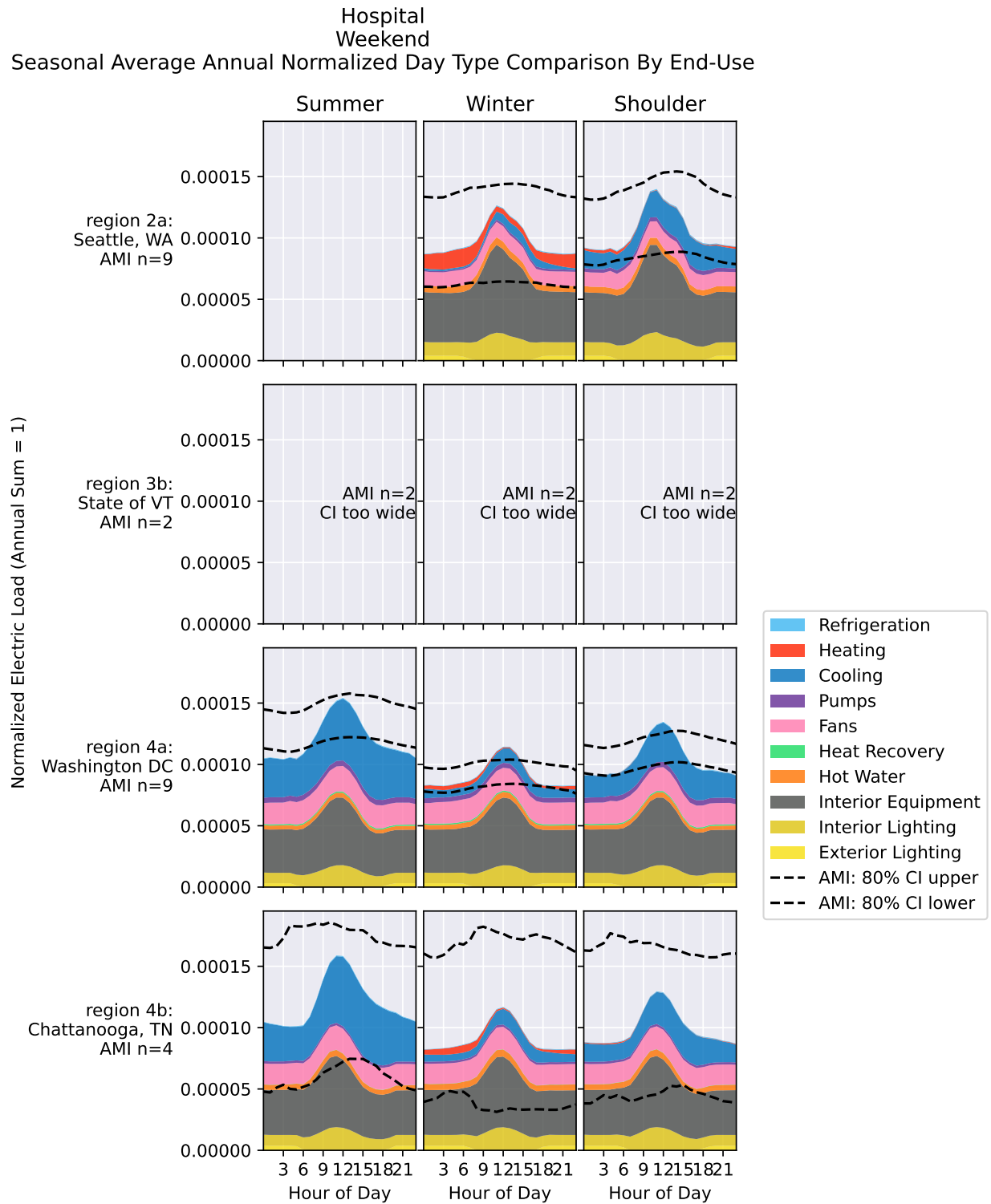




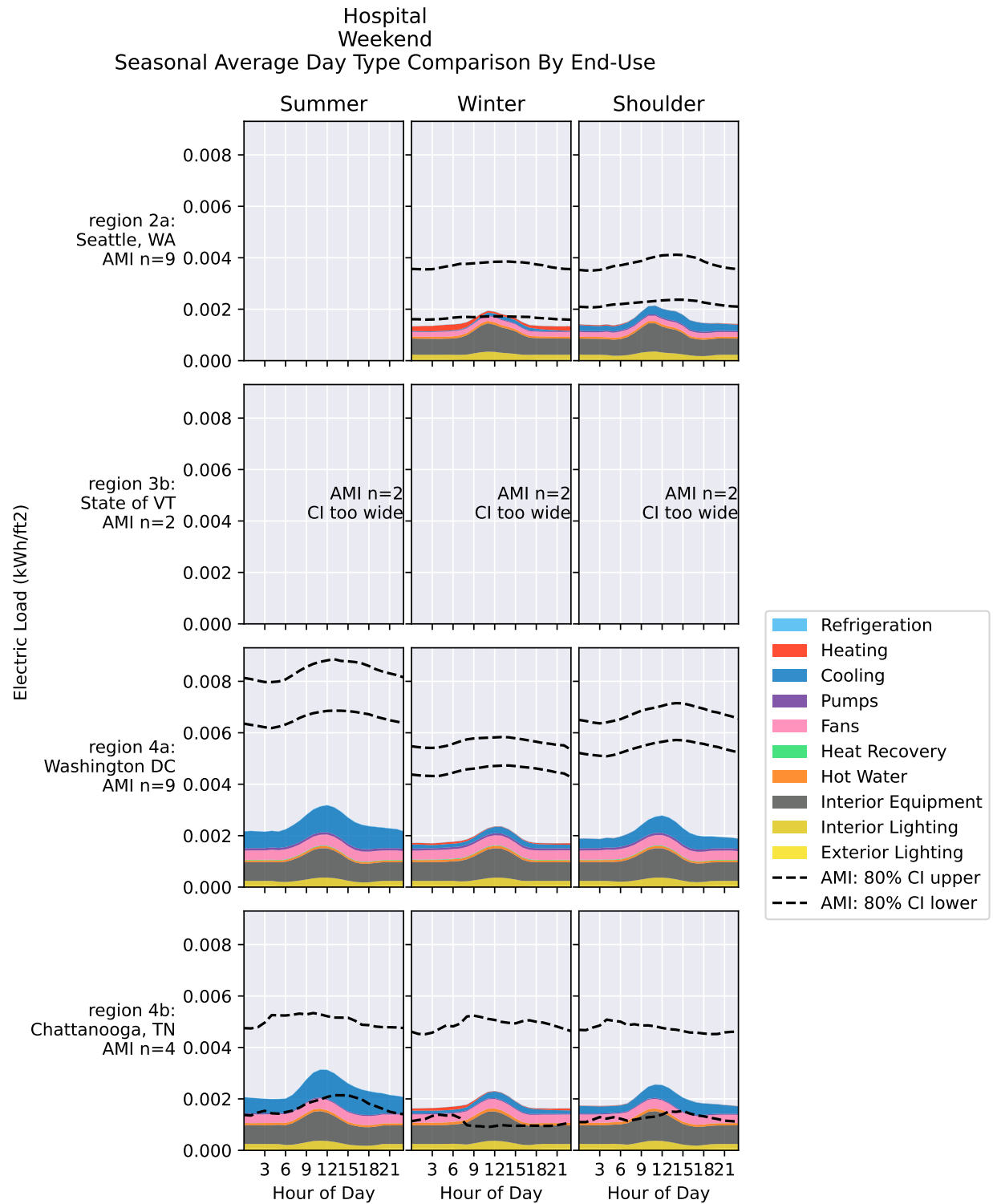
**Figure 276. Hospital weekday seasonal average annual normalized day type comparison by end use**



**Figure 277. Hospital weekday seasonal average day type comparison by end use**



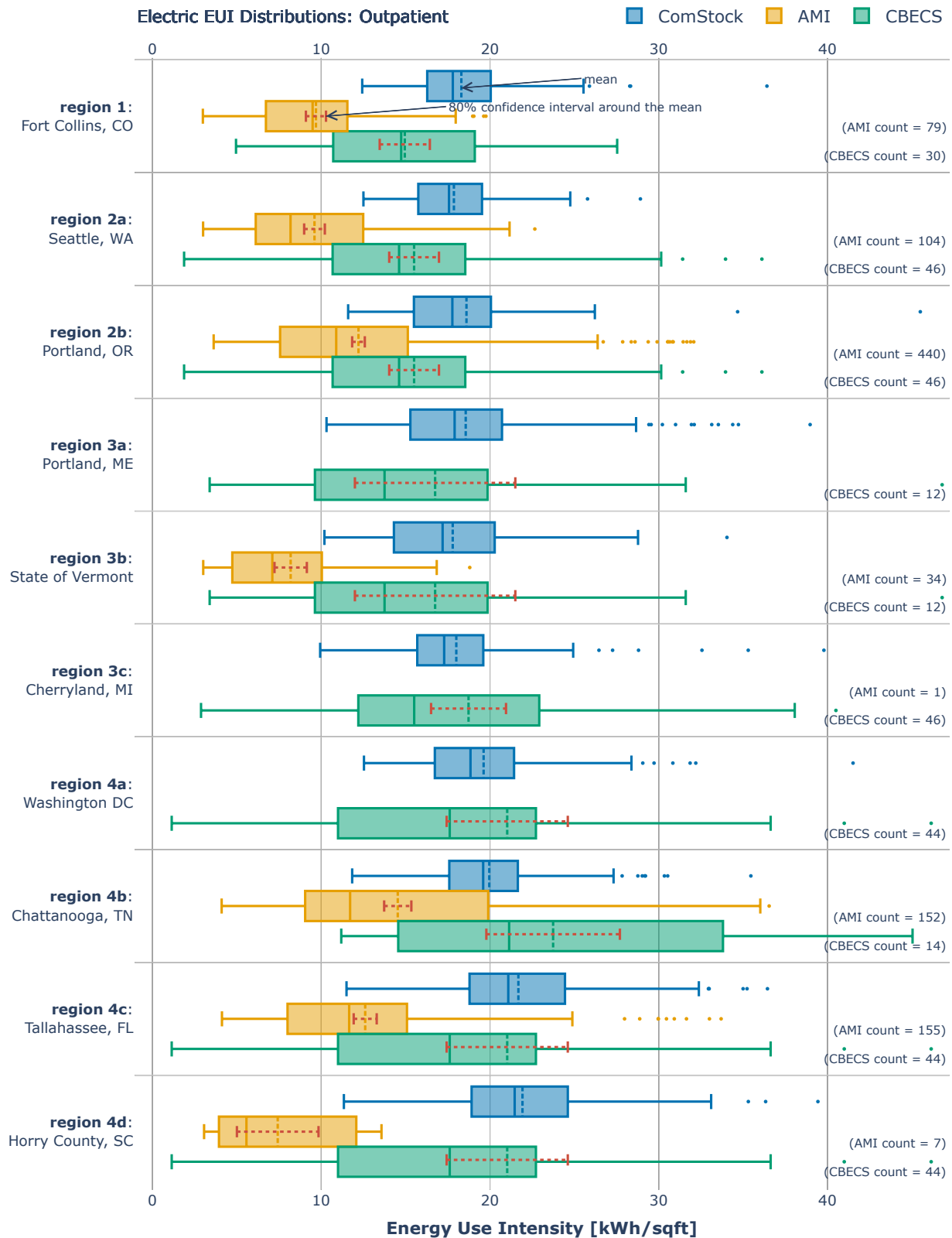
**Figure 278. Hospital weekend seasonal average annual normalized day type comparison by end use**



**Figure 279. Hospital weekend seasonal average day type comparison by end use**

### *Outpatient*

Outpatient buildings are another building type where the range of energy usage can vary widely between buildings. ComStock assumes a building where significant amounts of energy-consuming medical equipment is used, more typically of a large outpatient office building. However, the AMI includes many smaller medical practices that may include general practices that often do not perform major procedures. The EUI distributions show that ComStock is typically closer to the higher end of CBECS, but overestimates compared to the AMI. The normalized load shapes indicate that ComStock is missing nighttime load and is peakier than the AMI, perhaps indicating an overestimate of interior equipment that ramps up during the day and potentially an overestimate of nighttime HVAC reduction.



**Figure 280. Comparison of EUI distributions between regional AMI, CBECS 2012, and ComStock for outpatient buildings**

Outpatient  
Weekday  
Seasonal Average Annual Normalized Day Type Comparison By End-Use

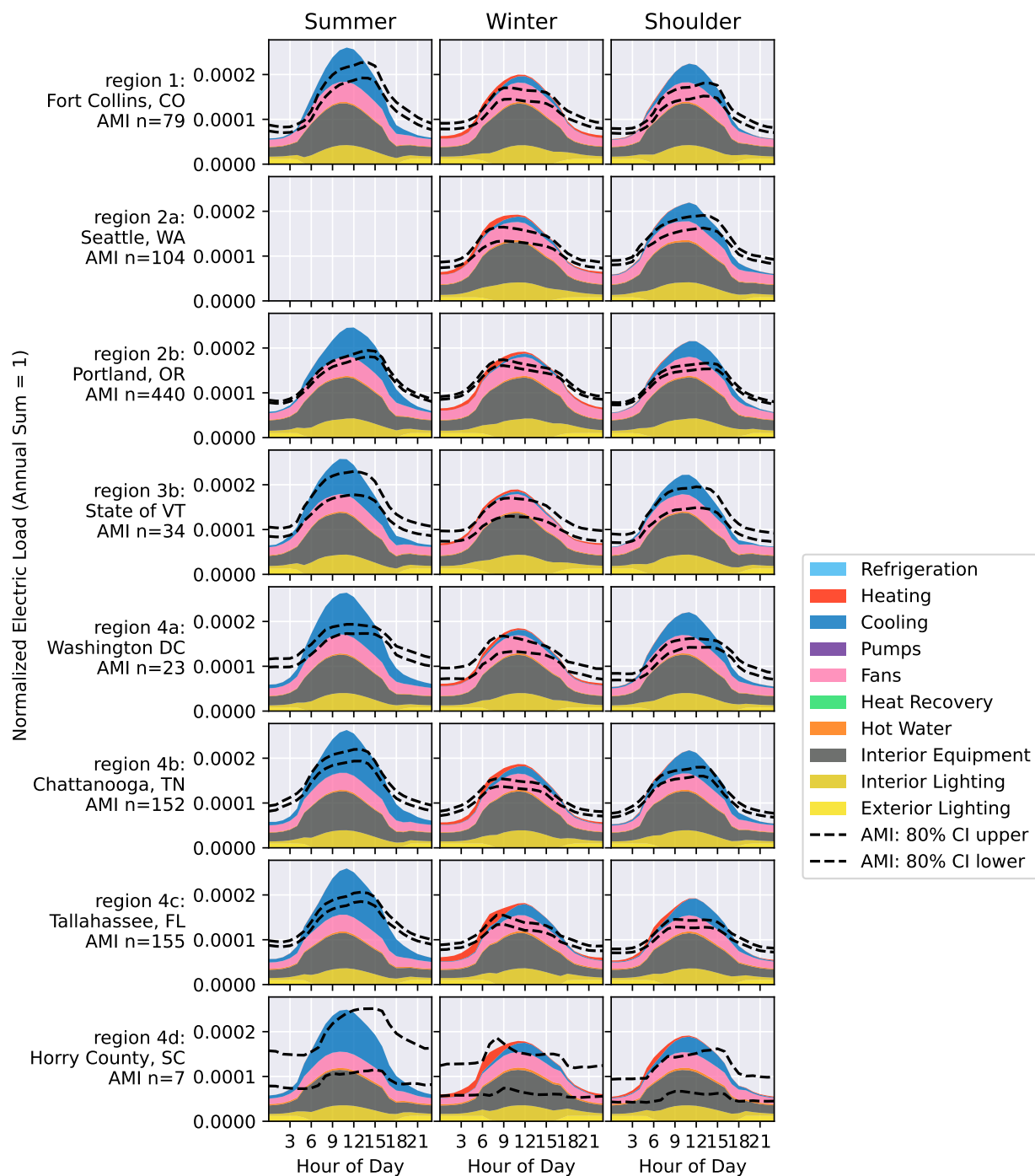


Figure 281. Outpatient weekday seasonal average annual normalized day type comparison by end use



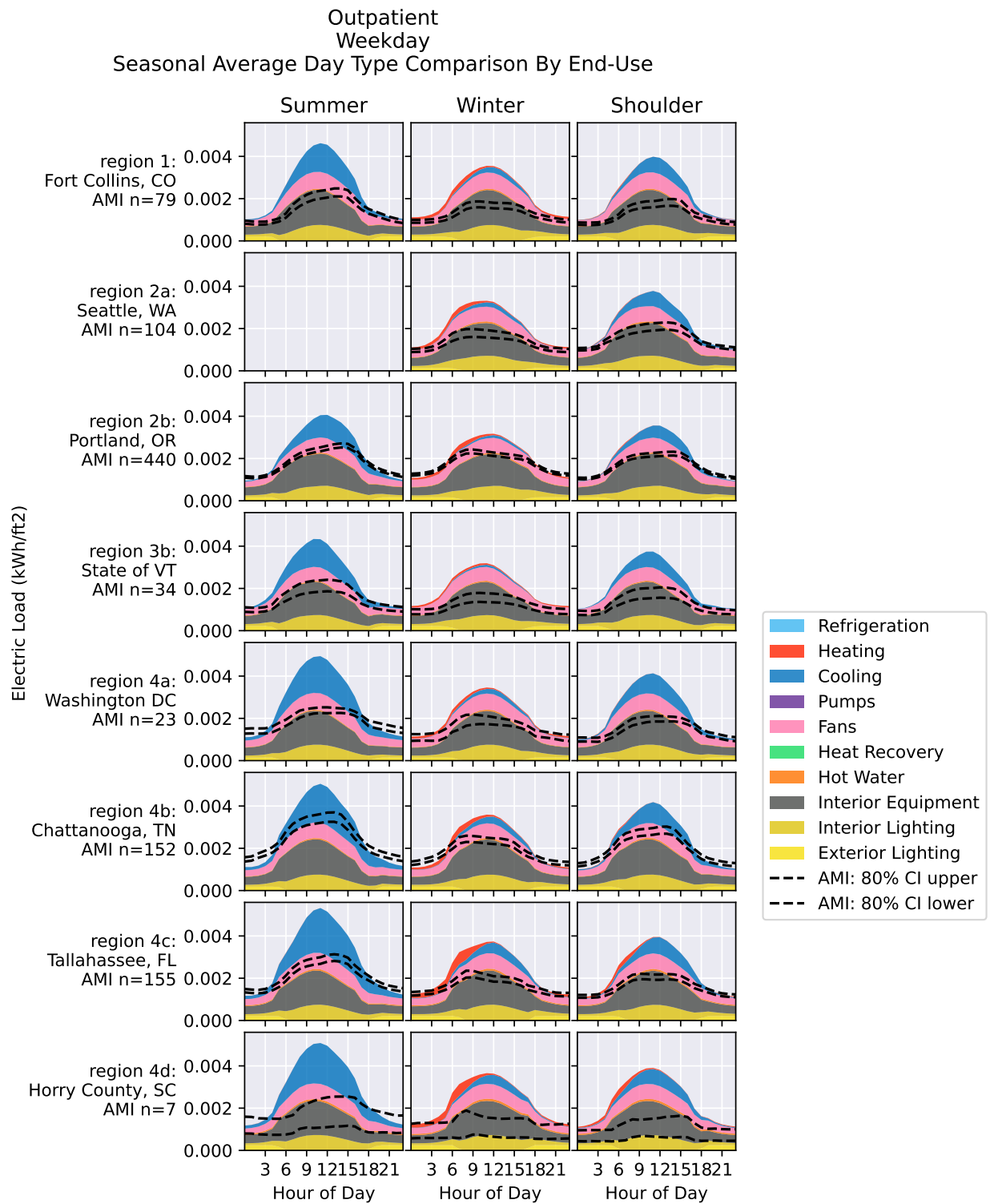
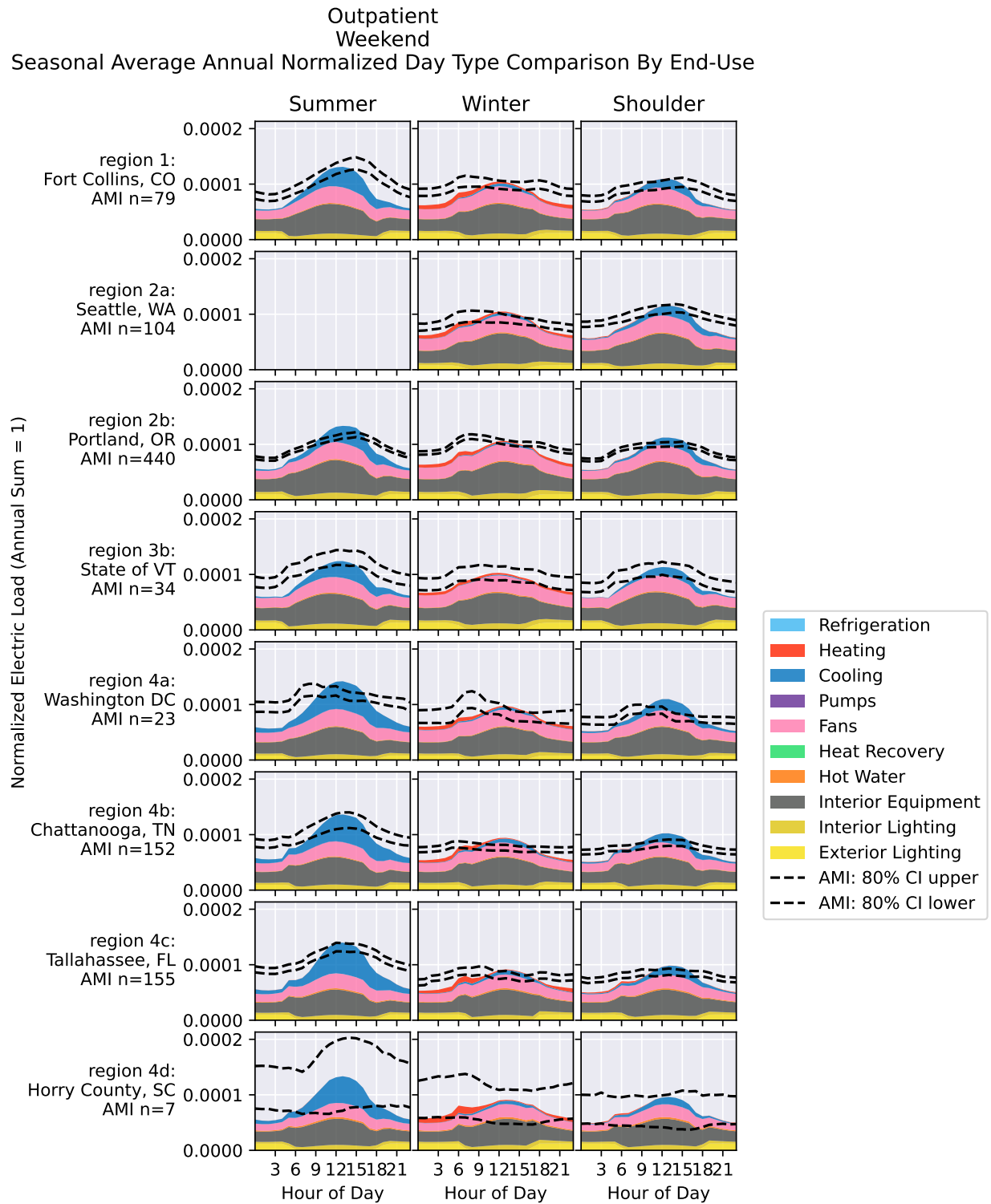


Figure 282. Outpatient weekday seasonal average day type comparison by end use



**Figure 283. Outpatient weekend seasonal average annual normalized day type comparison by end use**

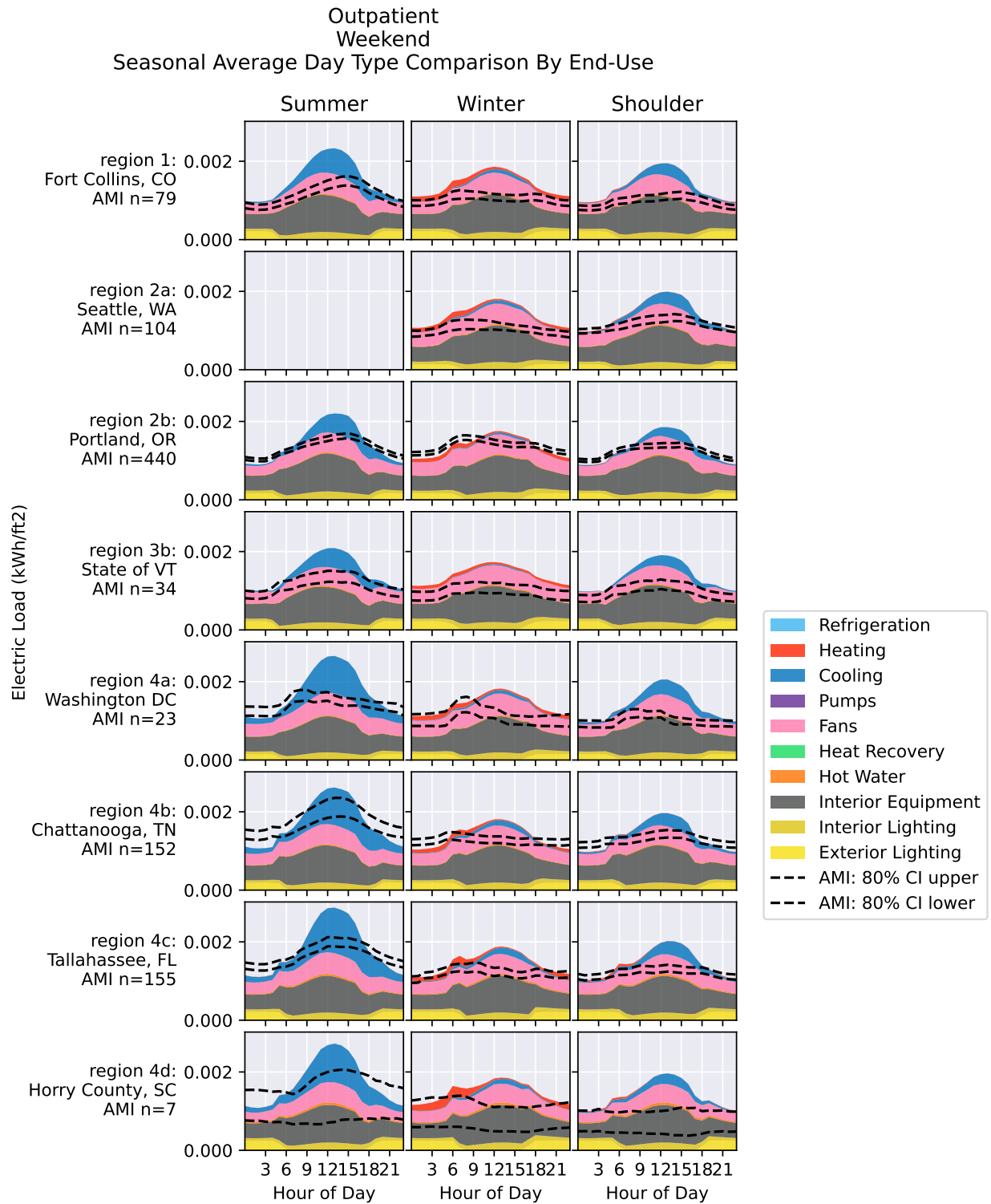


Figure 284. Outpatient weekend seasonal average day type comparison by end use

#### **4.2.3 Final Commercial QOIs**

Because there are QOIs for each building type, showing the progress of commercial QOIs over time is impractical. Because not all regions included AMI for all building types, the QOI comparison for commercial is done for each building type. While we believe the shape of the AMI data to be robust, for some building types and some regions, the magnitude of the AMI data is believed to be biased. While we believe a normalized comparison to be a more accurate representation of reality, we show both the normalized and non-normalized magnitude QOI comparison in the following figures for transparency, as we did with the AMI load shape comparison figures in the previous section.

## Small Office

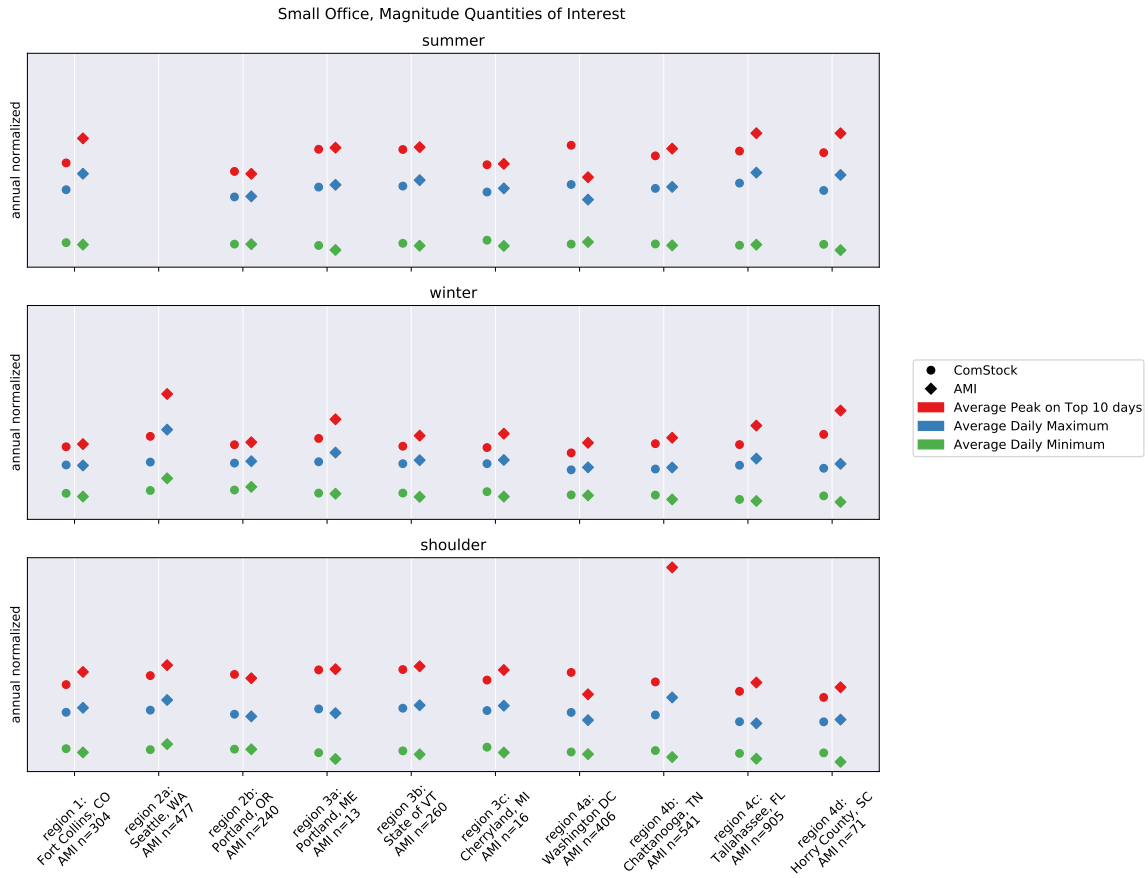


Figure 285. Normalized shape QOI comparison for small office.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	-6%	-15%	1%	10%	5%	8%	-11%	-17%	-3%
region2a	Seattle, WA	-12%	nan%	-30%	-12%	nan%	-21%	-8%	nan%	-30%
region2b	Portland, OR	3%	-1%	-2%	0%	0%	-6%	3%	2%	-3%
region3a	Portland, ME	6%	-2%	-11%	21%	13%	1%	-1%	-1%	-16%
region3b	State of VT	-4%	-6%	-5%	10%	6%	10%	-3%	-2%	-11%
region3c	Cherryland, MI	-6%	-4%	-5%	15%	15%	13%	-9%	-1%	-14%
region4a	Washington DC	11%	18%	-4%	6%	-5%	1%	23%	30%	-11%
region4b	Chattanooga, TN	-19%	-2%	-2%	21%	4%	12%	-52%	-5%	-6%
region4c	Tallahassee, FL	3%	-9%	-9%	18%	-2%	4%	-8%	-12%	-17%
region4d	Horry County, SC	-3%	-14%	-6%	34%	17%	18%	-10%	-13%	-19%

Figure 286. Normalized shape QOI error for small office.

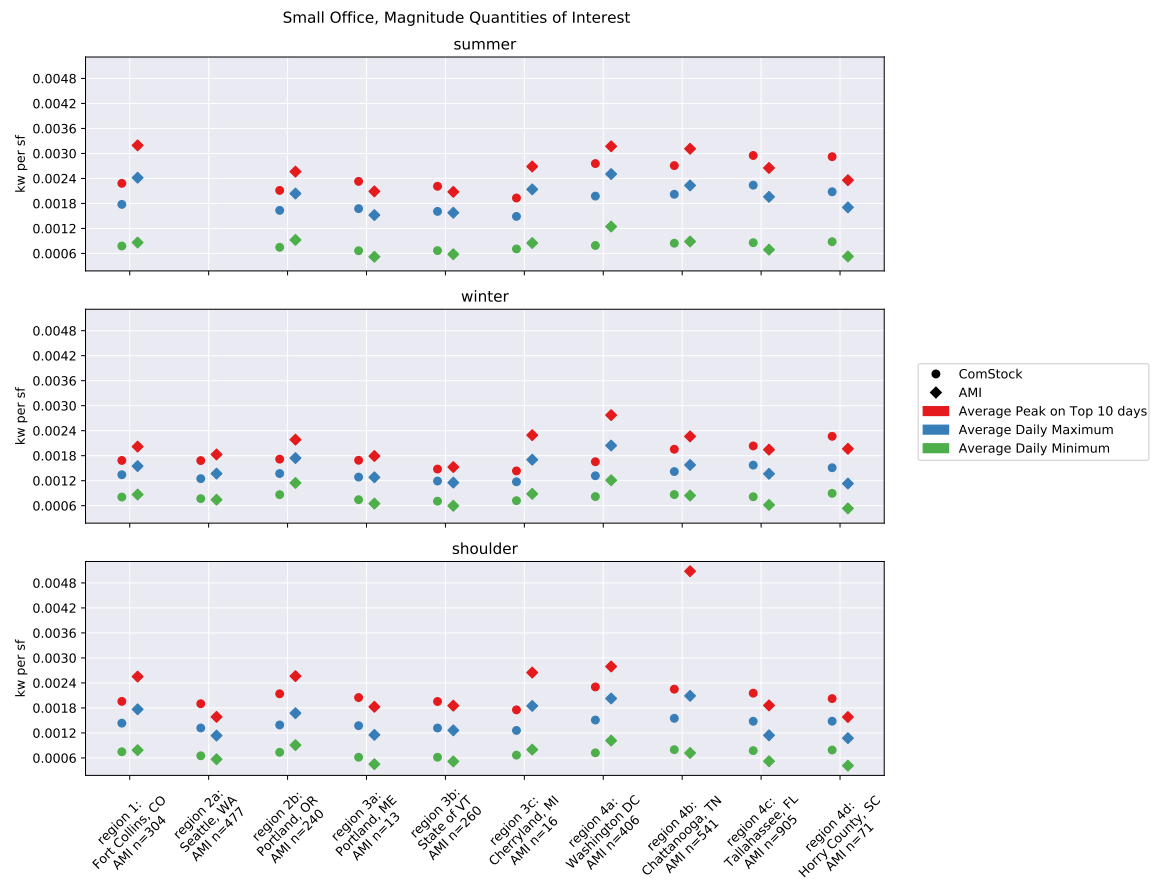


Figure 287. Magnitude QOI comparison for small office.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	-19%	-26%	-13%	-5%	-10%	-7%	-23%	-29%	-16%
region2a	Seattle, WA	16%	nan%	-9%	15%	nan%	4%	20%	nan%	-8%
region2b	Portland, OR	-17%	-20%	-21%	-19%	-19%	-24%	-17%	-18%	-21%
region3a	Portland, ME	19%	10%	0%	37%	28%	14%	12%	11%	-6%
region3b	State of VT	4%	2%	3%	19%	15%	19%	5%	6%	-3%
region3c	Cherryland, MI	-32%	-30%	-31%	-17%	-17%	-19%	-34%	-28%	-37%
region4a	Washington DC	-26%	-21%	-36%	-29%	-36%	-32%	-18%	-13%	-40%
region4b	Chattanooga, TN	-26%	-10%	-10%	11%	-5%	3%	-56%	-13%	-14%
region4c	Tallahassee, FL	29%	14%	15%	49%	24%	32%	16%	11%	5%
region4d	Horry County, SC	38%	22%	34%	91%	67%	68%	28%	24%	15%

Figure 288. Magnitude QOI error for small office.

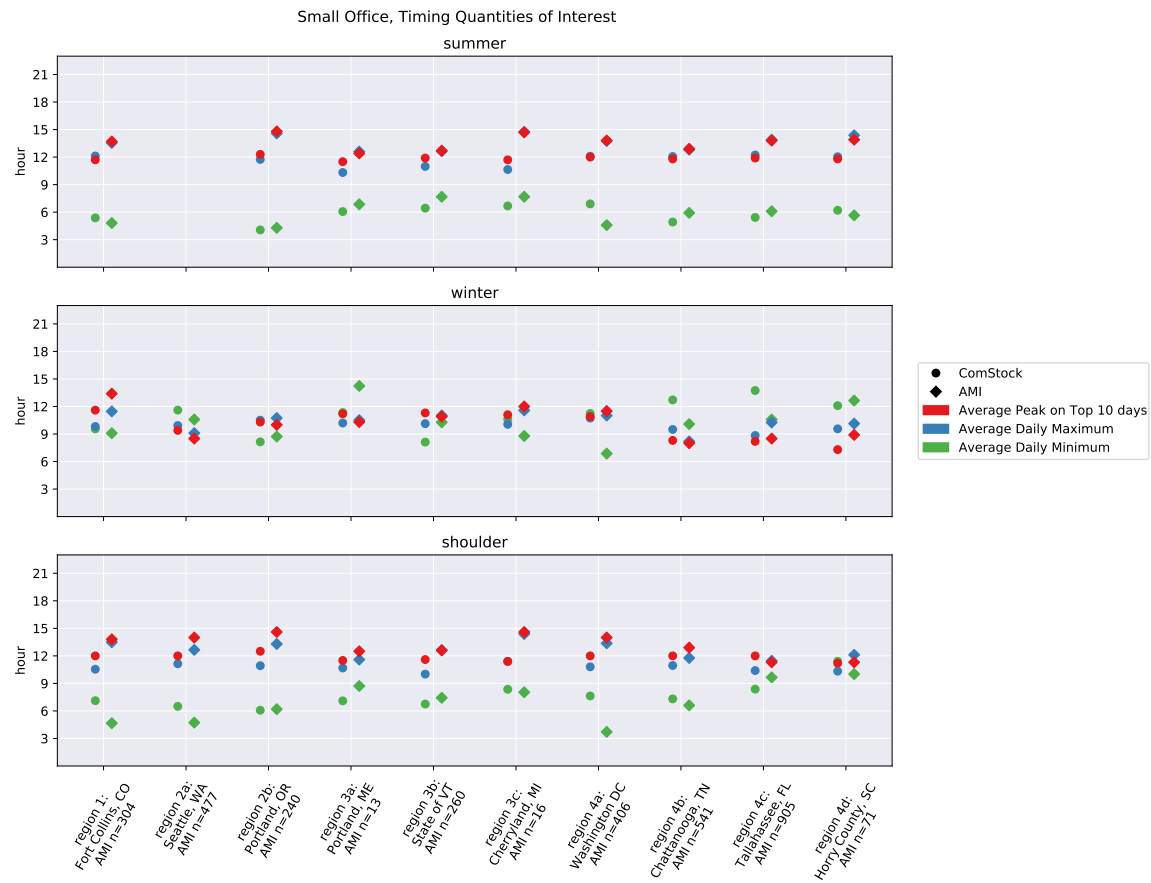


Figure 289. Timing QOI comparison for small office.

building_type		small_office								
qoi_category		timing								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	-2.9	-1.4	-1.6	2.5	0.6	0.5	-1.8	-2.0	-1.8
region2a	Seattle, WA	-1.5	nan	0.8	1.8	nan	1.0	-2.0	nan	0.9
region2b	Portland, OR	-2.4	-2.8	-0.2	-0.1	-0.2	-0.6	-2.1	-2.5	0.3
region3a	Portland, ME	-0.9	-2.3	-0.3	-1.6	-0.8	-2.9	-1.0	-0.9	0.9
region3b	State of VT	-2.6	-1.7	-0.9	-0.7	-1.2	-2.2	-1.0	-0.8	0.4
region3c	Cherryland, MI	-3.0	-4.1	-1.5	0.3	-1.0	1.8	-3.2	-3.0	-0.9
region4a	Washington DC	-2.6	-1.7	-0.3	3.9	2.3	4.4	-2.0	-1.8	-0.6
region4b	Chattanooga, TN	-0.8	-0.7	1.3	0.7	-1.0	2.6	-0.9	-1.1	0.3
region4c	Tallahassee, FL	-1.1	-1.7	-1.4	-1.3	-0.7	3.2	0.7	-1.9	-0.3
region4d	Horry County, SC	-1.8	-2.3	-0.6	1.4	0.5	-0.6	-0.1	-2.1	-1.6

Figure 290. Timing QOI error (hours) for small office.



## Medium Office

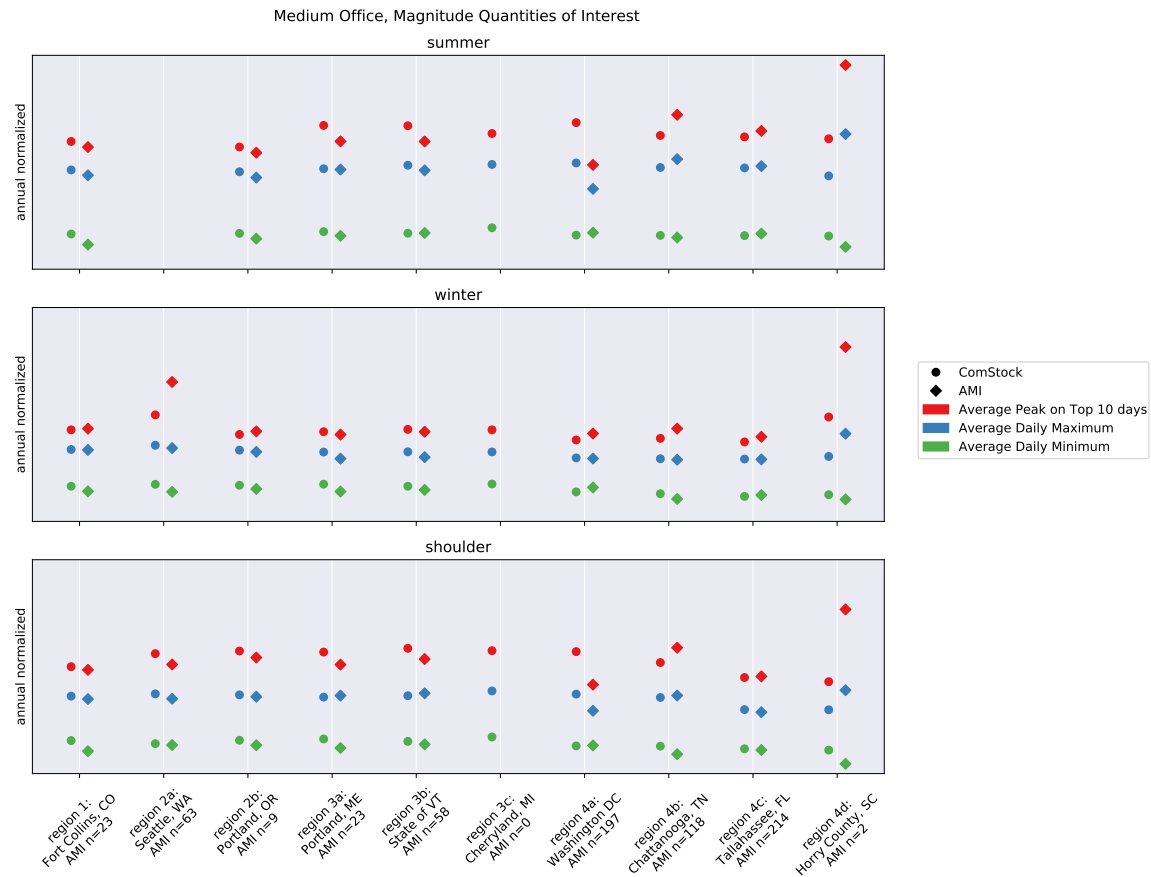


Figure 291. Normalized shape QOI comparison for medium office.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	3%	5%	0%	24%	23%	10%	3%	4%	-1%
region2a	Seattle, WA	5%	nan%	3%	3%	nan%	15%	8%	nan%	-21%
region2b	Portland, OR	2%	5%	2%	10%	10%	7%	5%	4%	-3%
region3a	Portland, ME	-2%	1%	8%	19%	8%	15%	10%	11%	3%
region3b	State of VT	-2%	4%	6%	6%	-1%	7%	8%	11%	2%
region4a	Washington DC	20%	26%	1%	-1%	-4%	-8%	30%	34%	-6%
region4b	Chattanooga, TN	-2%	-6%	1%	20%	4%	12%	-10%	-12%	-9%
region4c	Tallahassee, FL	3%	-1%	0%	3%	-4%	-3%	-1%	-4%	-5%
region4d	Horry County, SC	-19%	-27%	-21%	45%	25%	11%	-39%	-33%	-36%

Figure 292. Normalized shape QOI error for medium office.

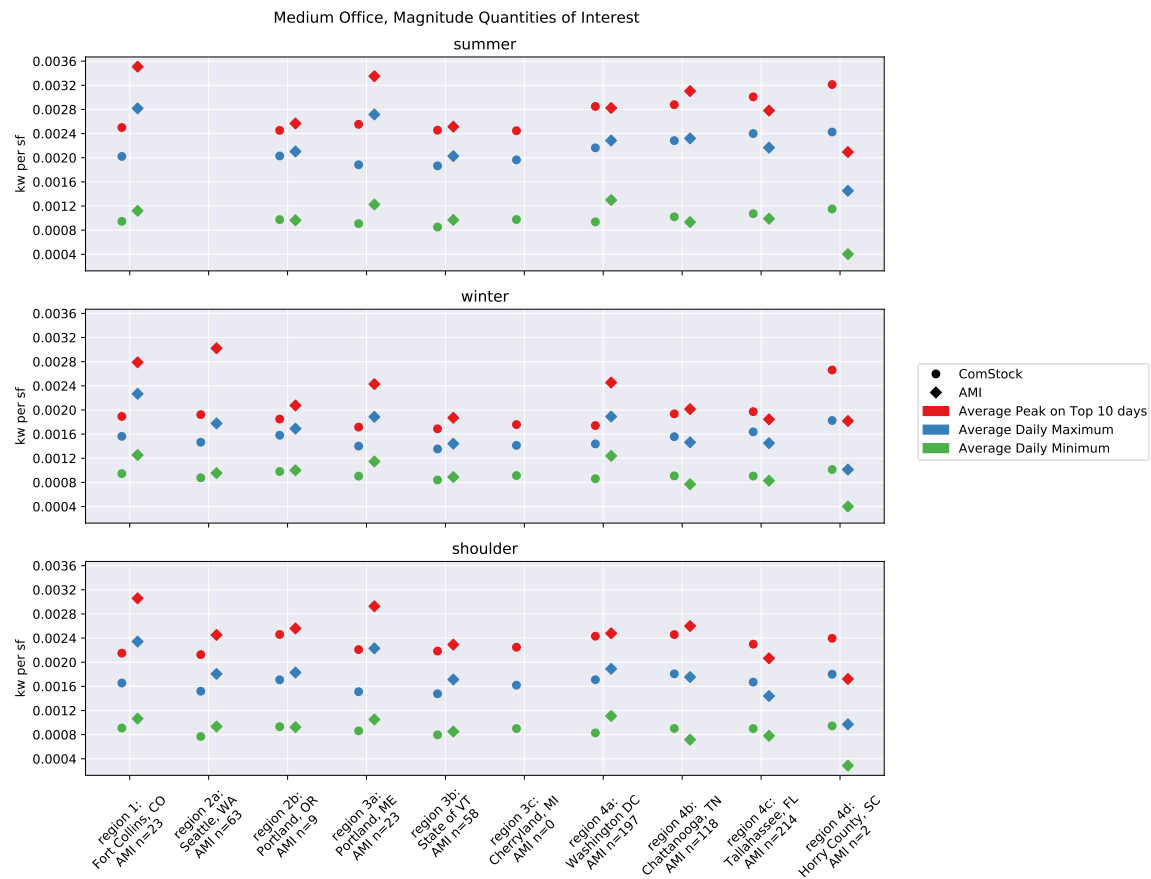


Figure 293. Magnitude QOI comparison for medium office.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	-29%	-28%	-31%	-15%	-16%	-25%	-30%	-29%	-32%
region2a	Seattle, WA	-16%	nan%	-18%	-18%	nan%	-8%	-13%	nan%	-36%
region2b	Portland, OR	-7%	-4%	-6%	1%	1%	-2%	-4%	-4%	-11%
region3a	Portland, ME	-32%	-31%	-26%	-18%	-26%	-21%	-25%	-24%	-29%
region3b	State of VT	-14%	-8%	-6%	-6%	-12%	-6%	-5%	-2%	-10%
region4a	Washington DC	-10%	-5%	-24%	-25%	-28%	-31%	-2%	1%	-29%
region4b	Chattanooga, TN	3%	-2%	6%	26%	9%	18%	-5%	-7%	-4%
region4c	Tallahassee, FL	16%	11%	13%	15%	8%	9%	11%	8%	7%
region4d	Horry County, SC	85%	67%	80%	230%	185%	153%	39%	53%	46%

Figure 294. Magnitude QOI error for medium office.

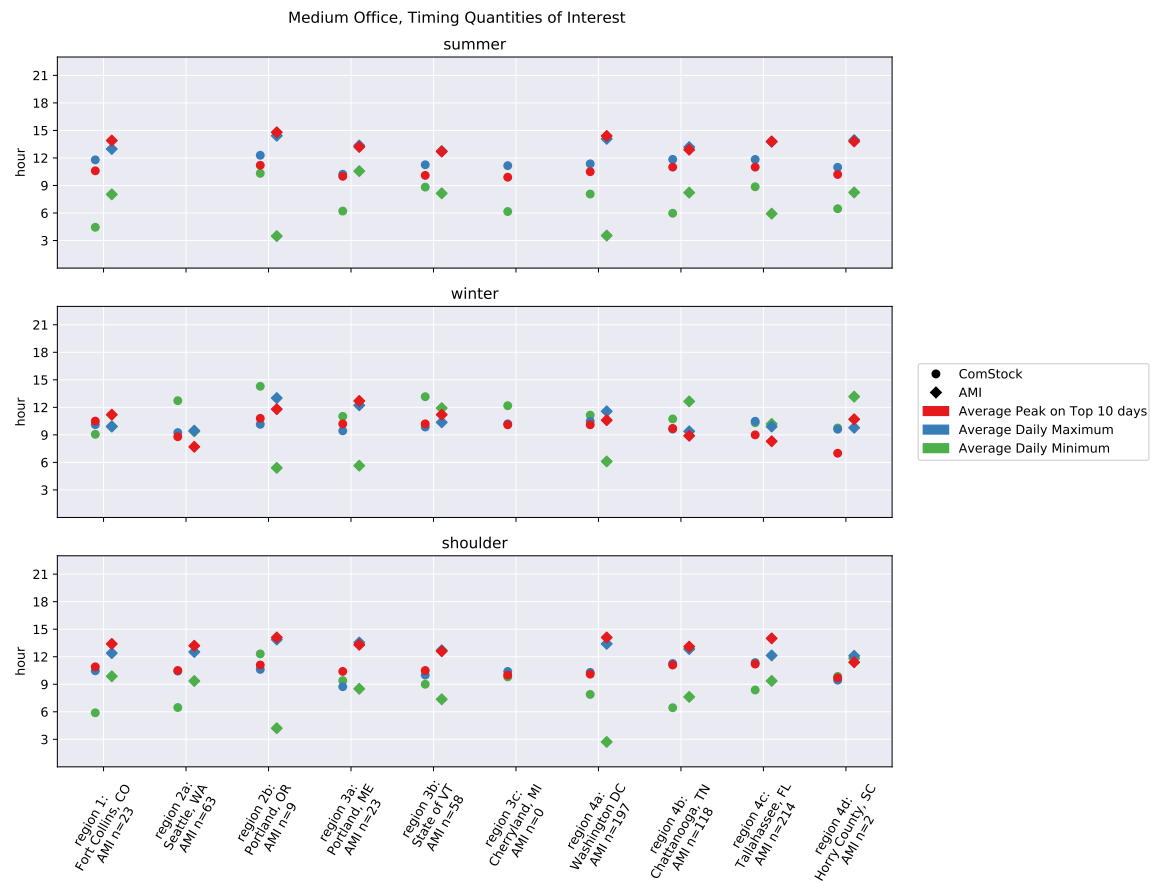


Figure 295. Timing QOI comparison for medium office.

building_type		medium_office								
qoi_category		timing								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	-1.9	-1.2	0.2	-4.0	-3.6	-0.8	-2.5	-3.3	-0.7
region2a	Seattle, WA	-2.1	nan	-0.2	-2.9	nan	3.3	-2.7	nan	1.1
region2b	Portland, OR	-3.3	-2.1	-2.9	8.1	6.8	8.9	-3.0	-3.6	-1.0
region3a	Portland, ME	-4.8	-3.1	-2.8	0.9	-4.4	5.4	-2.9	-3.2	-2.5
region3b	State of VT	-2.7	-1.5	-0.5	1.6	0.7	1.2	-2.1	-2.6	-1.0
region4a	Washington DC	-3.1	-2.7	-1.1	5.2	4.5	5.0	-4.0	-3.9	-0.5
region4b	Chattanooga, TN	-1.6	-1.3	0.2	-1.2	-2.2	-1.9	-2.0	-1.9	0.8
region4c	Tallahassee, FL	-0.8	-1.9	0.6	-1.0	2.9	0.1	-2.8	-2.8	0.7
region4d	Horry County, SC	-2.7	-3.0	-0.2	-2.0	-1.8	-3.4	-1.7	-3.6	-3.7

Figure 296. Timing QOI error (hours) for medium office.

Large Office

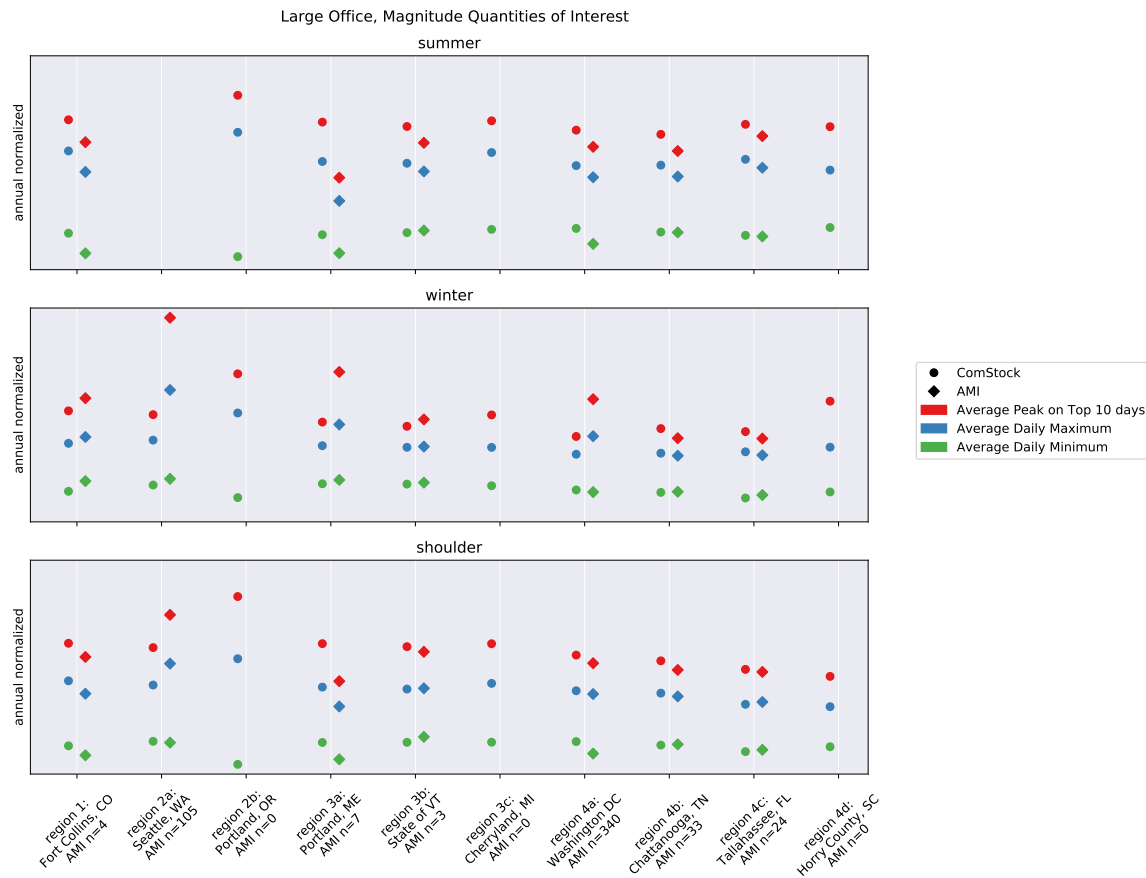


Figure 297. Normalized shape QOI comparison for large office.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	11%	16%	-5%	18%	39%	-14%	9%	14%	-8%
region2a	Seattle, WA	-15%	nan%	-30%	2%	nan%	-8%	-17%	nan%	-41%
region3a	Portland, ME	19%	38%	-16%	34%	36%	-5%	29%	44%	-27%
region3b	State of VT	-1%	6%	-1%	-8%	-3%	-2%	3%	10%	-5%
region4a	Washington DC	3%	9%	-15%	22%	26%	3%	6%	11%	-24%
region4b	Chattanooga, TN	3%	9%	2%	-1%	0%	-1%	7%	11%	8%
region4c	Tallahassee, FL	-2%	6%	3%	-3%	2%	-5%	2%	7%	6%

Figure 298. Normalized shape QOI error for large office.

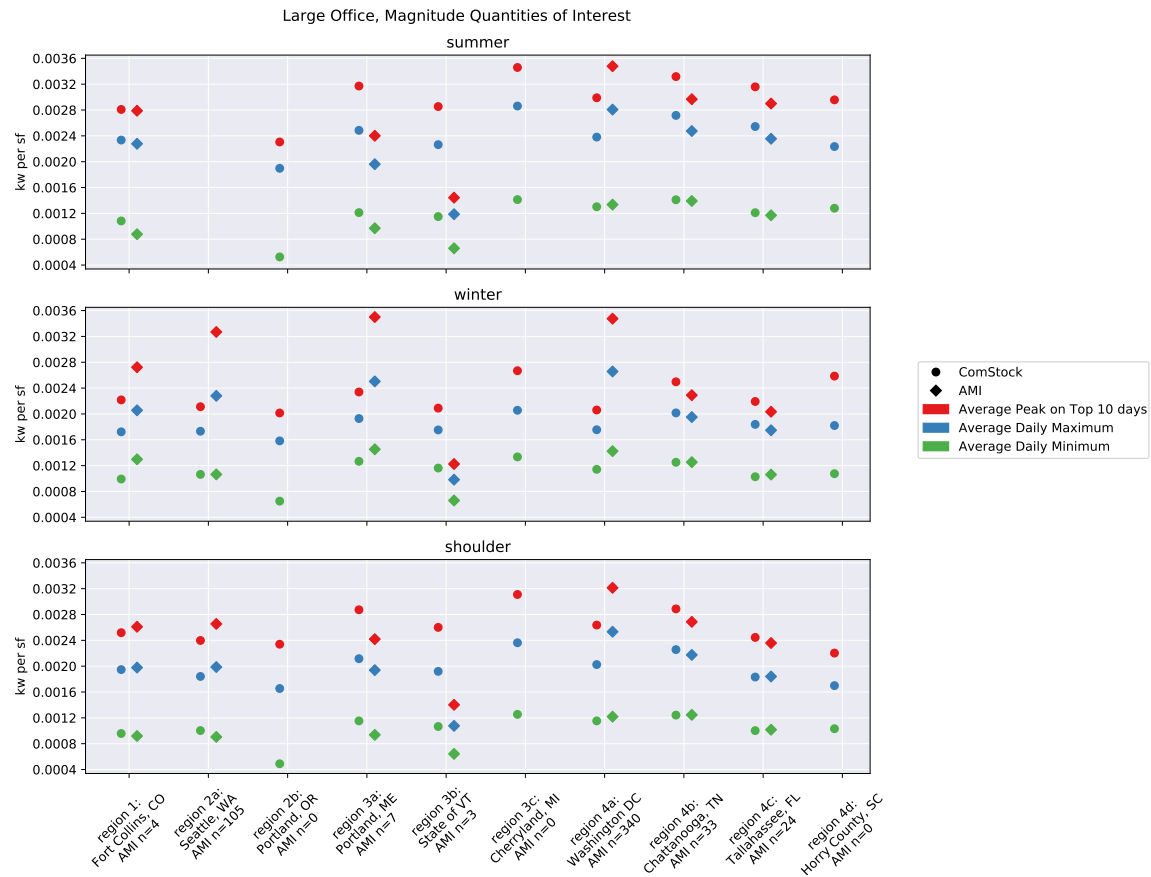


Figure 299. Magnitude QOI comparison for large office.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	-2%	3%	-16%	4%	23%	-24%	-4%	1%	-19%
region2a	Seattle, WA	-7%	nan%	-24%	11%	nan%	-0%	-10%	nan%	-35%
region3a	Portland, ME	9%	27%	-23%	23%	25%	-13%	19%	32%	-33%
region3b	State of VT	78%	91%	78%	66%	74%	76%	86%	98%	71%
region4a	Washington DC	-20%	-15%	-34%	-5%	-2%	-20%	-18%	-14%	-41%
region4b	Chattanooga, TN	4%	10%	3%	-0%	1%	-0%	7%	12%	9%
region4c	Tallahassee, FL	-0%	8%	5%	-1%	3%	-3%	4%	9%	8%

Figure 300. Magnitude QOI error for large office.

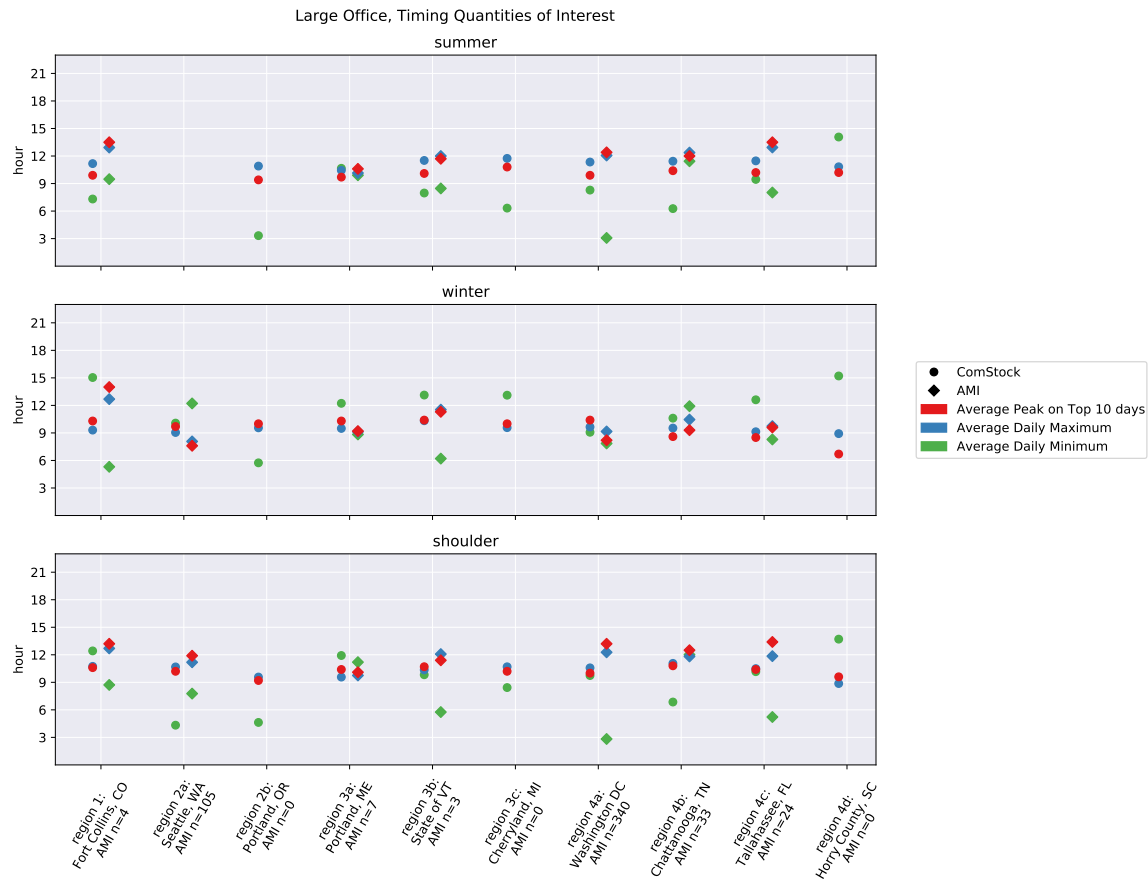


Figure 301. Timing QOI comparison for large office.

building_type		large_office								
qoi_category		timing								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	-2.0	-1.8	-3.4	3.7	-2.2	9.7	-2.6	-3.6	-3.7
region2a	Seattle, WA	-0.5	nan	1.0	-3.4	nan	-2.1	-1.7	nan	2.1
region3a	Portland, ME	-0.2	0.3	0.4	0.7	0.8	3.4	0.3	-0.9	1.1
region3b	State of VT	-1.7	-0.5	-1.2	4.1	-0.5	6.9	-0.7	-1.6	-0.9
region4a	Washington DC	-1.7	-0.7	0.5	6.9	5.2	1.2	-3.2	-2.5	2.2
region4b	Chattanooga, TN	-0.8	-0.9	-0.9	-5.2	-5.2	-1.3	-1.7	-1.6	-0.7
region4c	Tallahassee, FL	-1.4	-1.5	-0.6	4.9	1.4	4.3	-3.0	-3.3	-1.1

Figure 302. Timing QOI error (hours) for large office.

## Retail

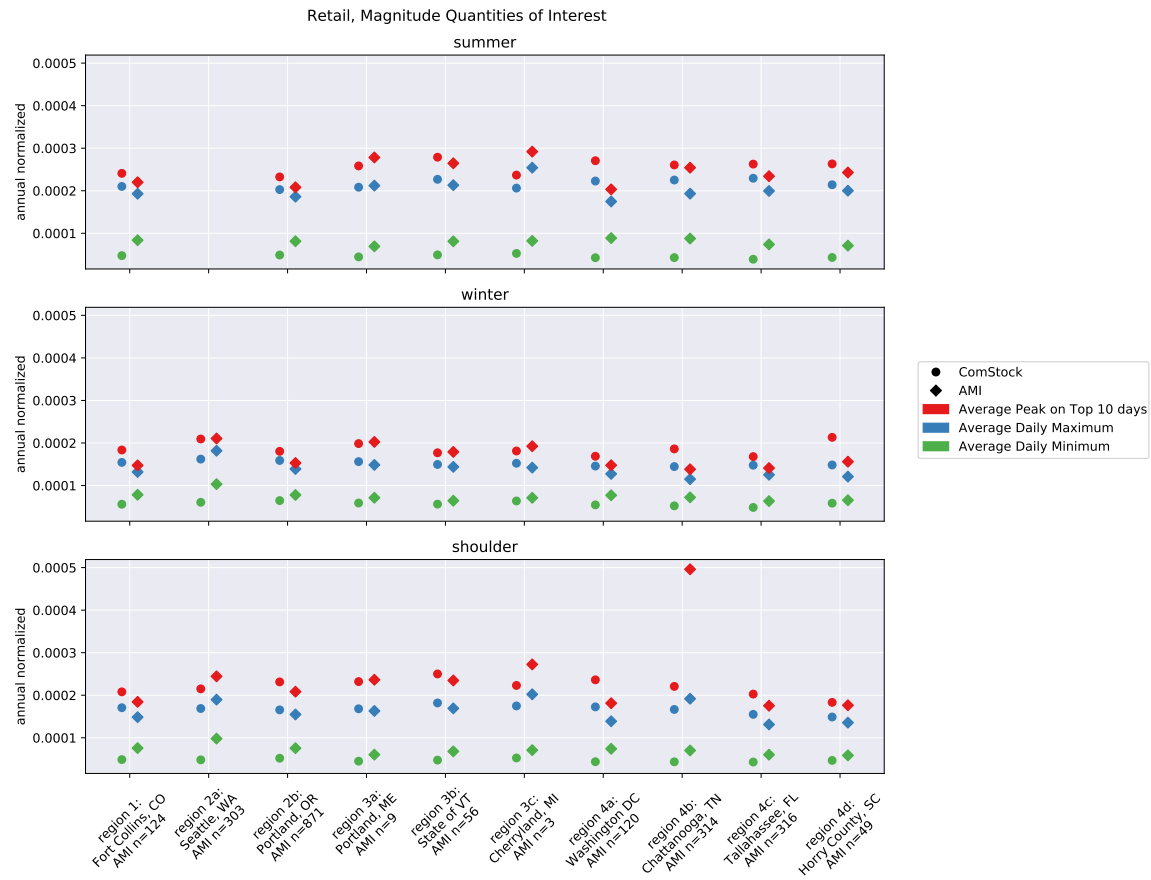


Figure 303. Normalized shape QOI comparison for retail.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	15%	9%	17%	-36%	-43%	-28%	13%	10%	25%
region2a	Seattle, WA	-11%	nan%	-11%	-51%	nan%	-41%	-12%	nan%	-1%
region2b	Portland, OR	7%	9%	14%	-31%	-40%	-17%	11%	12%	18%
region3a	Portland, ME	3%	-2%	5%	-25%	-36%	-17%	-2%	-7%	-2%
region3b	State of VT	8%	6%	4%	-30%	-39%	-13%	7%	5%	-1%
region3c	Cherryland, MI	-14%	-19%	7%	-26%	-36%	-10%	-18%	-19%	-6%
region4a	Washington DC	24%	28%	14%	-41%	-52%	-29%	30%	33%	14%
region4b	Chattanooga, TN	-13%	17%	26%	-38%	-51%	-28%	-55%	3%	35%
region4c	Tallahassee, FL	18%	15%	18%	-29%	-47%	-23%	16%	12%	19%
region4d	Horry County, SC	10%	7%	23%	-20%	-39%	-11%	4%	8%	37%

Figure 304. Normalized shape QOI error for retail.



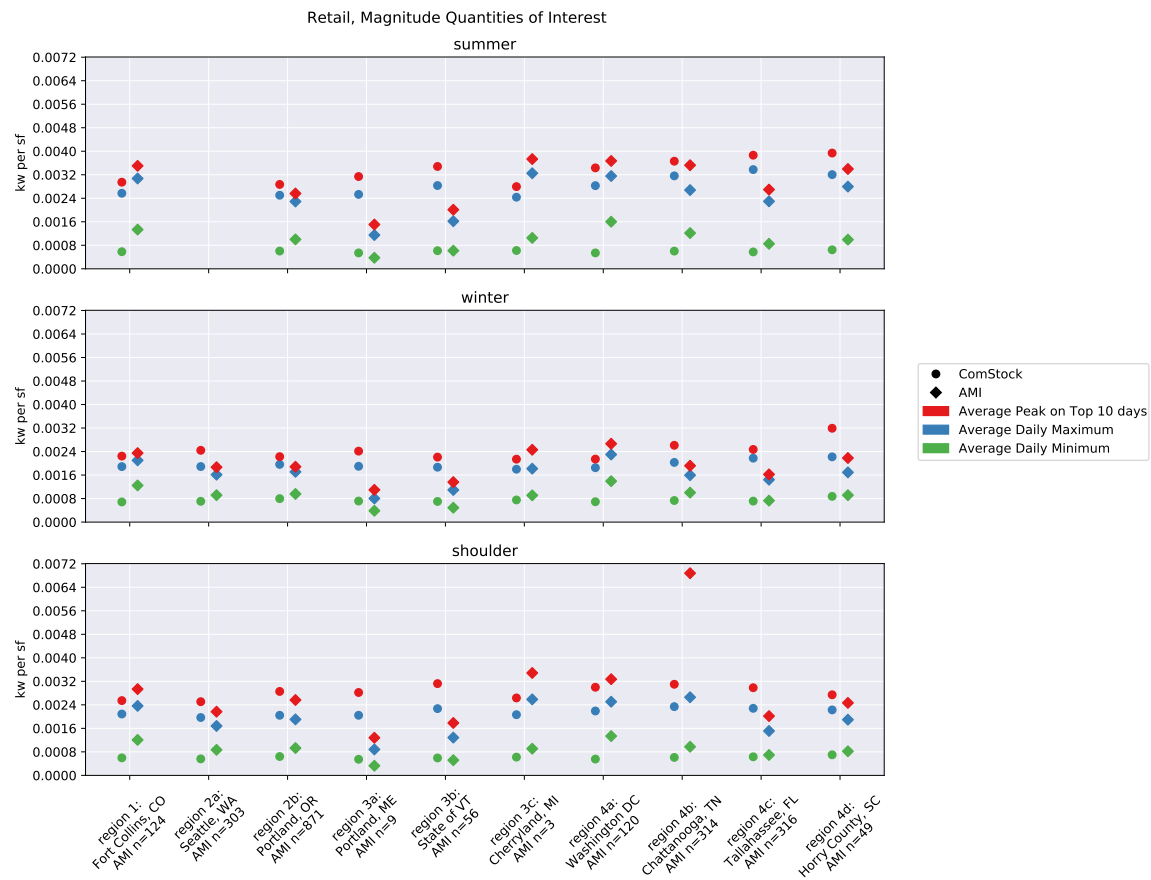


Figure 305. Magnitude QOI comparison for retail.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	-12%	-16%	-10%	-51%	-56%	-45%	-13%	-16%	-4%
region2a	Seattle, WA	17%	nan%	17%	-35%	nan%	-23%	16%	nan%	31%
region2b	Portland, OR	7%	9%	15%	-31%	-40%	-17%	11%	12%	18%
region3a	Portland, ME	132%	121%	136%	67%	44%	86%	120%	108%	120%
region3b	State of VT	77%	75%	71%	15%	-0%	44%	75%	73%	63%
region3c	Cherryland, MI	-20%	-25%	-1%	-32%	-41%	-17%	-24%	-25%	-13%
region4a	Washington DC	-13%	-10%	-20%	-59%	-66%	-50%	-8%	-6%	-20%
region4b	Chattanooga, TN	-12%	18%	27%	-37%	-51%	-27%	-55%	4%	36%
region4c	Tallahassee, FL	51%	47%	51%	-9%	-32%	-2%	48%	43%	52%
region4d	Horry County, SC	18%	15%	31%	-15%	-35%	-4%	11%	16%	46%

Figure 306. Magnitude QOI error for retail.

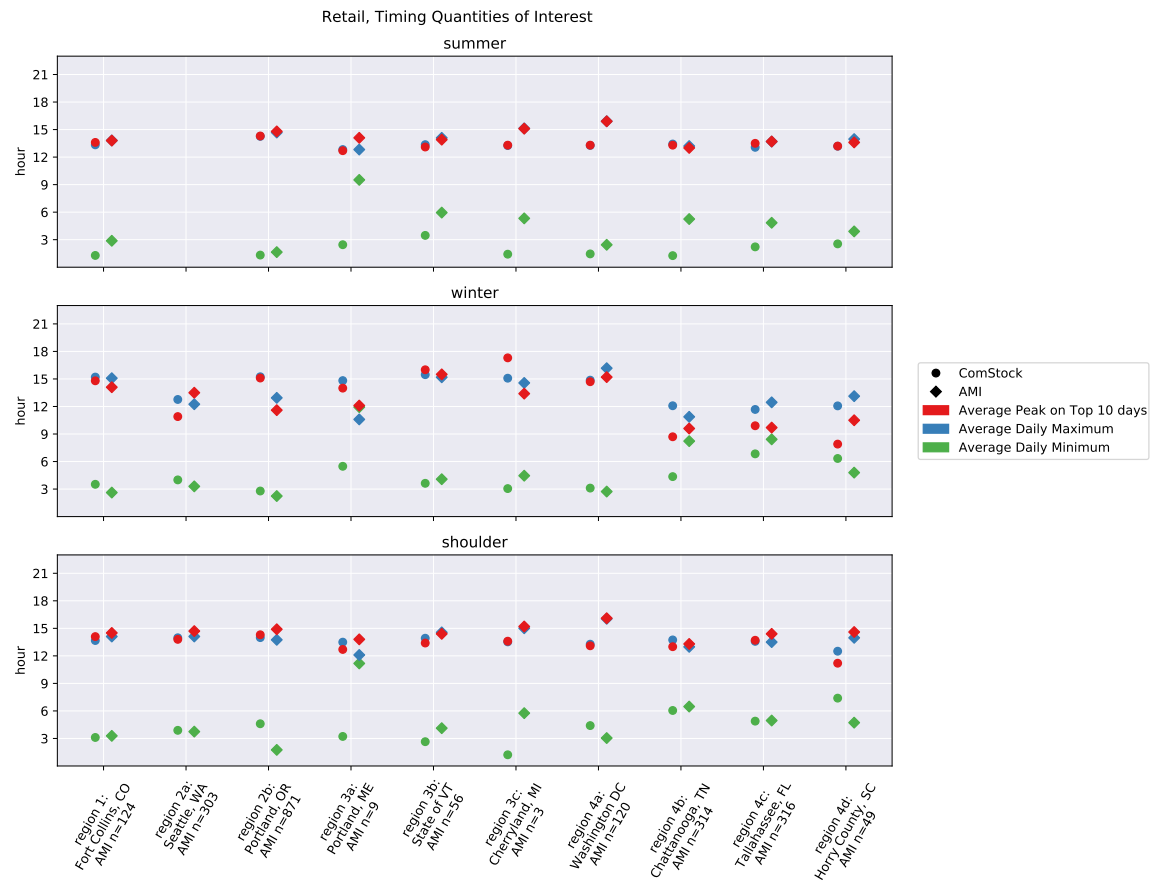


Figure 307. Timing QOI comparison for retail.

building_type		retail								
qoi_category		timing								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	-0.4	-0.5	0.1	-0.2	-1.6	0.9	-0.4	-0.2	0.7
region2a	Seattle, WA	-0.1	nan	0.5	0.1	nan	0.7	-0.9	nan	-2.6
region2b	Portland, OR	0.2	-0.4	2.3	2.8	-0.3	0.5	-0.6	-0.5	3.5
region3a	Portland, ME	1.4	0.0	4.2	-8.0	-7.1	-6.4	-1.1	-1.4	1.9
region3b	State of VT	-0.6	-0.7	0.3	-1.5	-2.5	-0.4	-1.0	-0.8	0.5
region3c	Cherryland, MI	-1.5	-1.9	0.5	-4.5	-3.9	-1.4	-1.6	-1.8	3.9
region4a	Washington DC	-2.8	-2.6	-1.3	1.4	-1.0	0.4	-3.0	-2.6	-0.5
region4b	Chattanooga, TN	0.8	0.2	1.2	-0.4	-4.0	-3.9	-0.3	0.3	-0.9
region4c	Tallahassee, FL	0.1	-0.6	-0.8	-0.1	-2.6	-1.6	-0.7	-0.2	0.2
region4d	Horry County, SC	-1.5	-0.8	-1.1	2.7	-1.4	1.5	-3.4	-0.4	-2.6

Figure 308. Timing QOI error (hours) for retail.

## Strip Mall

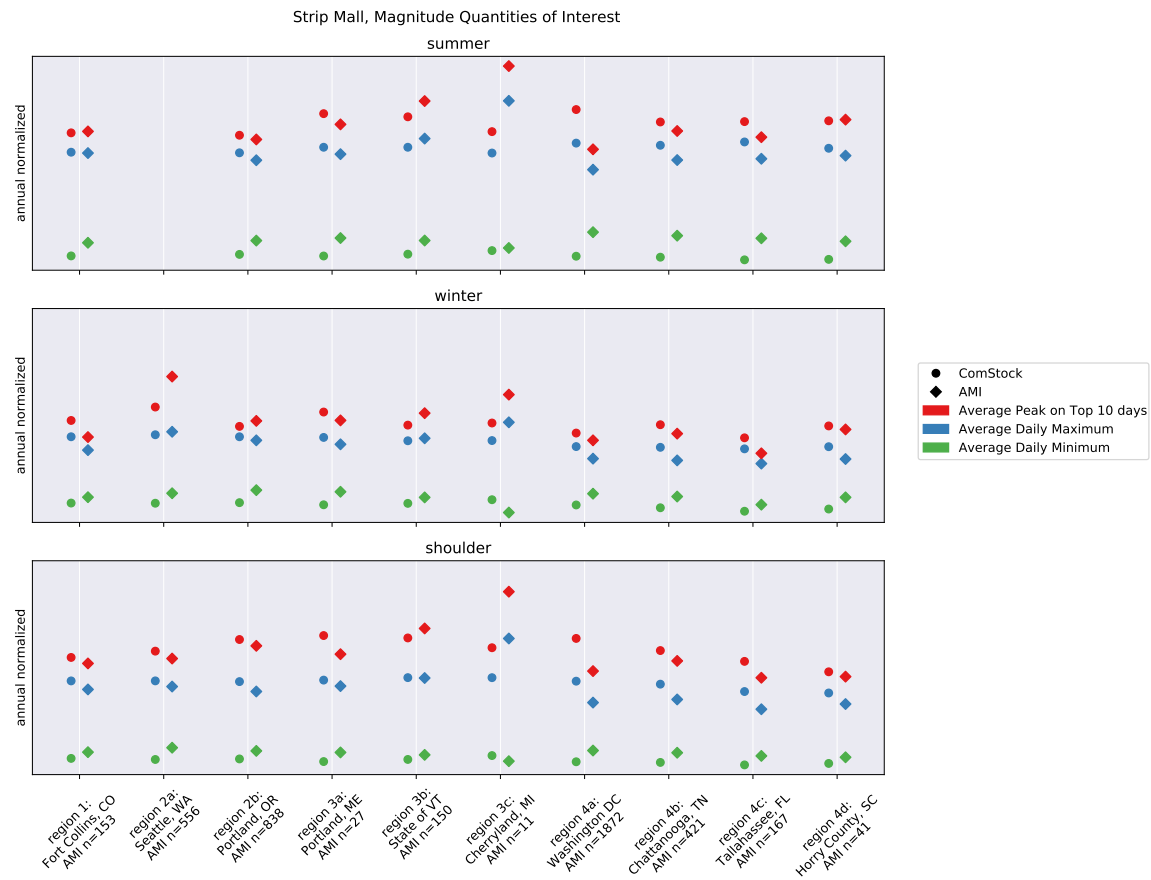


Figure 309. Normalized shape QOI comparison for strip mall.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	8%	1%	14%	-13%	-24%	-11%	4%	-1%	15%
region2a	Seattle, WA	5%	nan%	-3%	-22%	nan%	-18%	5%	nan%	-18%
region2b	Portland, OR	9%	5%	3%	-16%	-25%	-21%	4%	3%	-4%
region3a	Portland, ME	5%	5%	7%	-19%	-31%	-23%	13%	6%	7%
region3b	State of VT	0%	-5%	-2%	-10%	-24%	-12%	-6%	-8%	-9%
region3c	Cherryland, MI	-24%	-27%	-14%	14%	-5%	35%	-27%	-28%	-18%
region4a	Washington DC	22%	21%	13%	-22%	-37%	-20%	25%	27%	7%
region4b	Chattanooga, TN	15%	11%	15%	-20%	-35%	-21%	7%	5%	8%
region4c	Tallahassee, FL	19%	12%	17%	-20%	-37%	-15%	13%	10%	16%
region4d	Horry County, SC	11%	5%	14%	-14%	-33%	-23%	4%	-1%	3%

Figure 310. Normalized shape QOI error for strip mall.

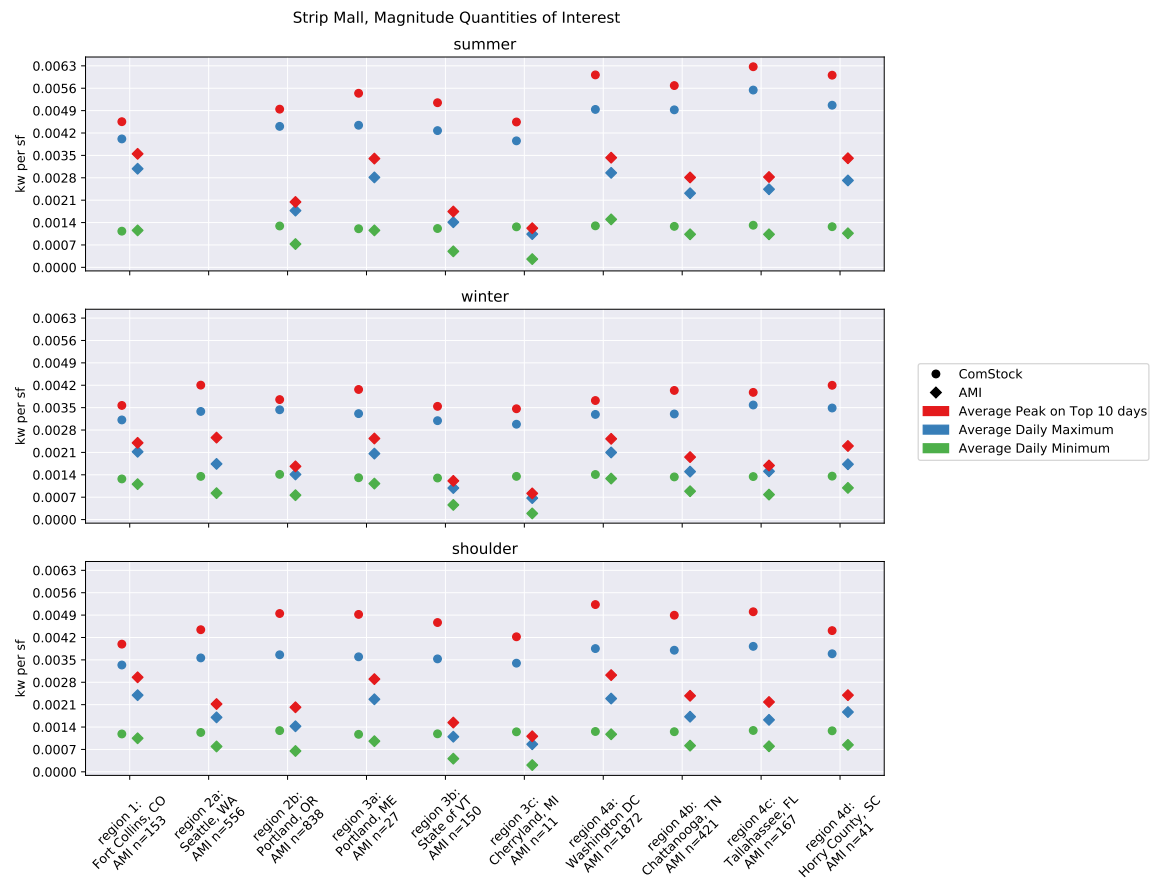


Figure 311. Magnitude QOI comparison for strip mall.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	39%	30%	47%	13%	-2%	15%	35%	28%	49%
region2a	Seattle, WA	109%	nan%	94%	55%	nan%	63%	110%	nan%	64%
region2b	Portland, OR	157%	148%	143%	98%	78%	85%	145%	142%	126%
region3a	Portland, ME	59%	58%	61%	22%	4%	16%	70%	60%	60%
region3b	State of VT	222%	203%	214%	189%	143%	184%	203%	195%	192%
region3c	Cherryland, MI	294%	281%	344%	492%	391%	603%	280%	271%	323%
region4a	Washington DC	68%	67%	57%	7%	-13%	10%	73%	76%	47%
region4b	Chattanooga, TN	121%	113%	120%	53%	24%	51%	106%	102%	107%
region4c	Tallahassee, FL	141%	127%	137%	62%	27%	72%	129%	122%	135%
region4d	Horry County, SC	98%	87%	102%	52%	19%	37%	84%	76%	83%

Figure 312. Magnitude QOI error for strip mall.

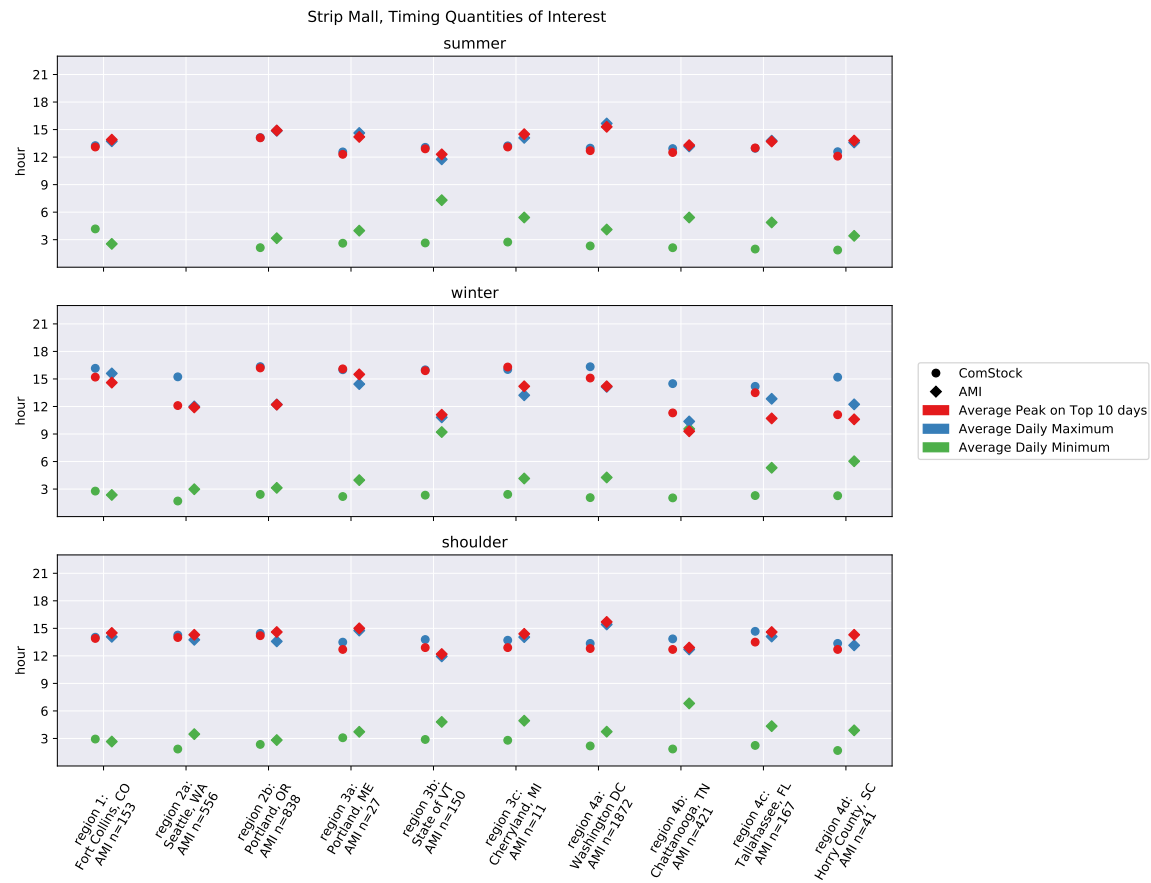


Figure 313. Timing QOI comparison for strip mall.

building_type		strip_mall								
qoi_category		timing								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	-0.0	-0.5	0.6	0.3	1.6	0.4	-0.6	-0.8	0.6
region2a	Seattle, WA	0.5	nan	3.2	-1.6	nan	-1.3	-0.3	nan	0.2
region2b	Portland, OR	0.9	-0.7	4.1	-0.5	-1.0	-0.7	-0.4	-0.8	4.0
region3a	Portland, ME	-1.3	-2.1	1.6	-0.6	-1.4	-1.8	-2.3	-1.9	0.6
region3b	State of VT	1.8	1.3	5.2	-1.9	-4.7	-6.9	0.7	0.6	4.8
region3c	Cherryland, MI	-0.3	-0.9	2.8	-2.1	-2.7	-1.7	-1.5	-1.4	2.1
region4a	Washington DC	-2.1	-2.7	2.2	-1.6	-1.8	-2.2	-2.9	-2.6	0.9
region4b	Chattanooga, TN	1.1	-0.2	4.1	-5.0	-3.3	-7.5	-0.2	-0.8	2.0
region4c	Tallahassee, FL	0.6	-0.8	1.4	-2.1	-2.9	-3.0	-1.1	-0.7	2.8
region4d	Horry County, SC	0.2	-1.0	3.0	-2.2	-1.5	-3.8	-1.6	-1.7	0.5

Figure 314. Timing QOI error (hours) for strip mall.

## Warehouse

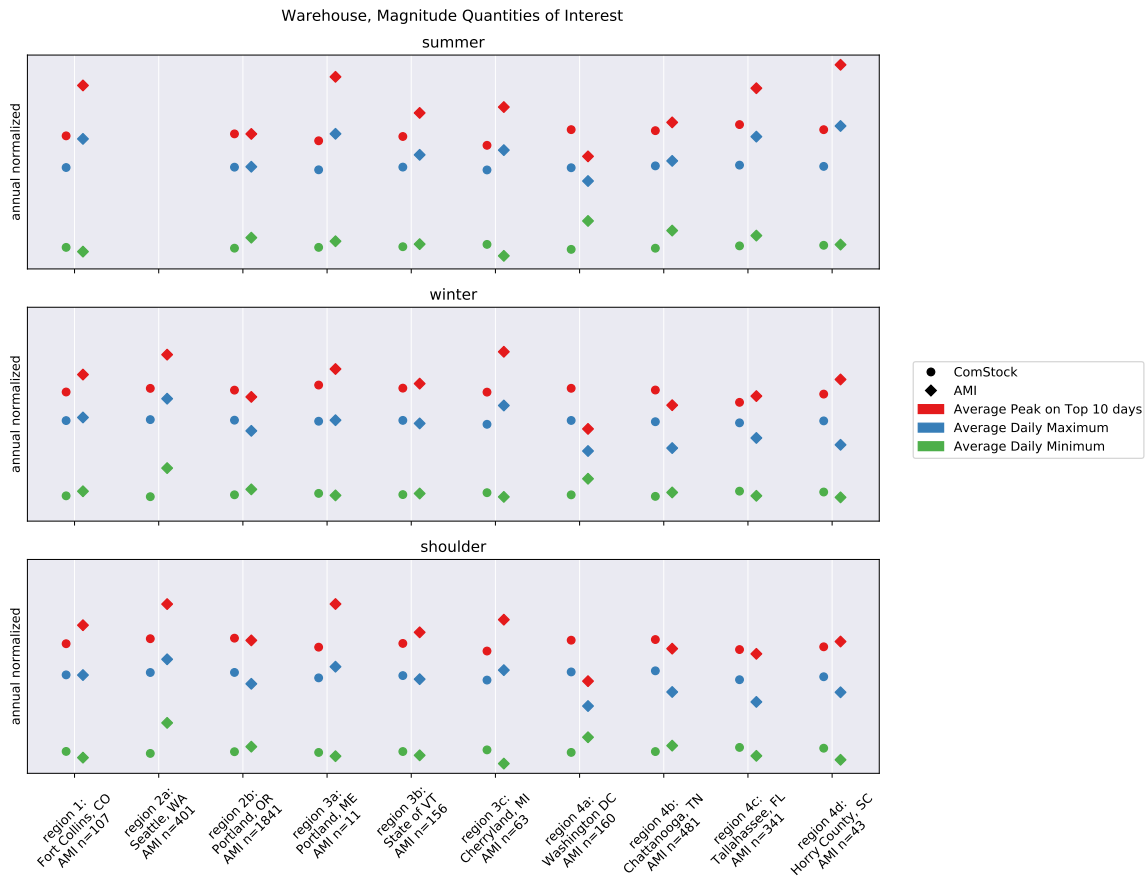


Figure 315. Normalized shape QOI comparison for warehouse.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	0%	-18%	-2%	15%	10%	-8%	-11%	-24%	-10%
region2a	Seattle, WA	-9%	nan%	-14%	-40%	nan%	-36%	-18%	nan%	-17%
region2b	Portland, OR	10%	-0%	9%	-9%	-18%	-9%	1%	0%	4%
region3a	Portland, ME	-8%	-22%	-1%	9%	-11%	4%	-22%	-29%	-9%
region3b	State of VT	3%	-9%	2%	9%	-5%	-2%	-7%	-13%	-3%
region3c	Cherryland, MI	-8%	-14%	-13%	38%	29%	8%	-17%	-20%	-21%
region4a	Washington DC	36%	12%	32%	-24%	-38%	-23%	35%	19%	34%
region4b	Chattanooga, TN	20%	-4%	27%	-11%	-27%	-7%	6%	-5%	11%
region4c	Tallahassee, FL	23%	-18%	14%	19%	-17%	9%	3%	-18%	-4%
region4d	Horry County, SC	14%	-24%	23%	29%	-1%	11%	-3%	-28%	-9%

Figure 316. Normalized shape QOI error for warehouse.

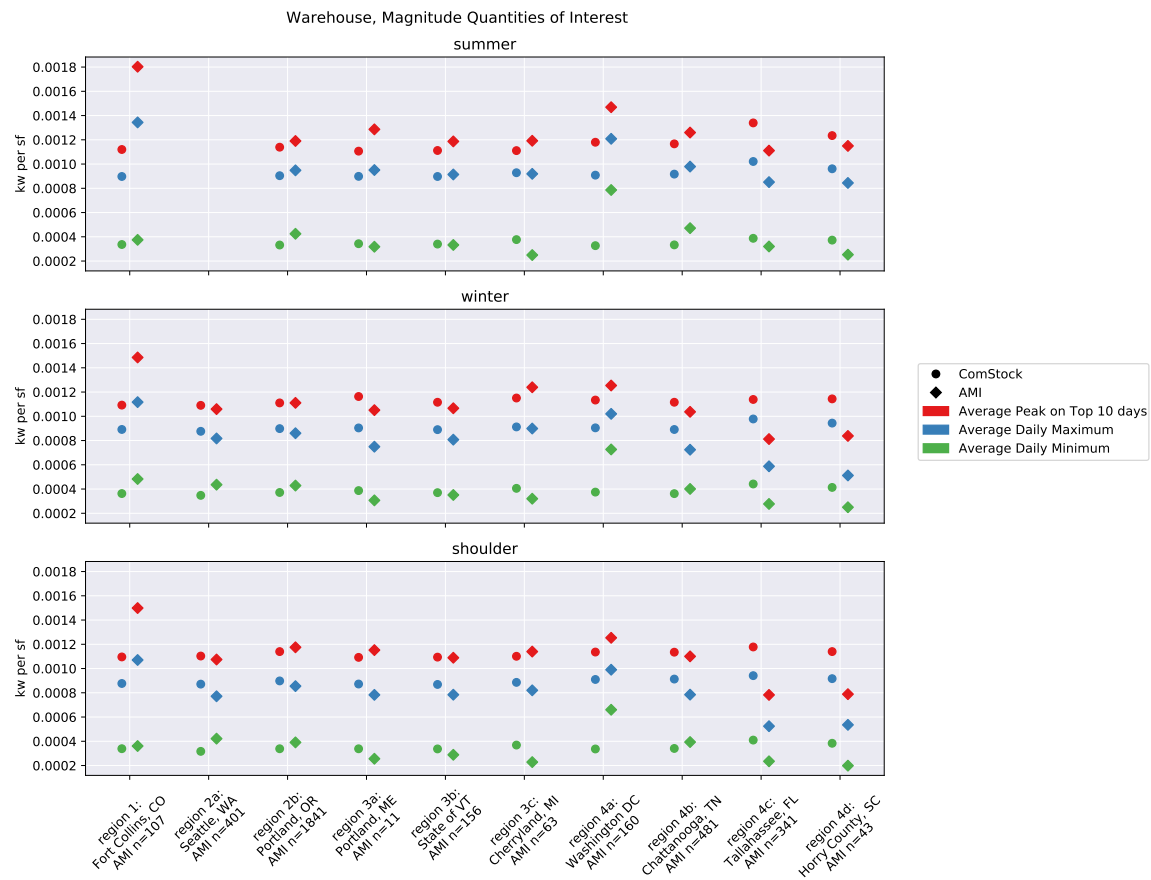


Figure 317. Magnitude QOI comparison for warehouse.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	-18%	-33%	-20%	-6%	-10%	-25%	-27%	-38%	-26%
region2a	Seattle, WA	13%	nan%	7%	-25%	nan%	-20%	3%	nan%	3%
region2b	Portland, OR	5%	-5%	4%	-13%	-22%	-13%	-3%	-4%	-0%
region3a	Portland, ME	11%	-5%	21%	32%	8%	26%	-5%	-14%	11%
region3b	State of VT	11%	-2%	10%	17%	2%	5%	0%	-6%	5%
region3c	Cherryland, MI	8%	1%	1%	62%	51%	27%	-3%	-7%	-7%
region4a	Washington DC	-8%	-25%	-11%	-49%	-58%	-48%	-9%	-20%	-10%
region4b	Chattanooga, TN	16%	-6%	23%	-13%	-29%	-10%	3%	-7%	8%
region4c	Tallahassee, FL	79%	20%	66%	74%	21%	59%	51%	21%	40%
region4d	Horry County, SC	71%	14%	84%	93%	47%	65%	45%	7%	36%

Figure 318. Magnitude QOI error for warehouse.



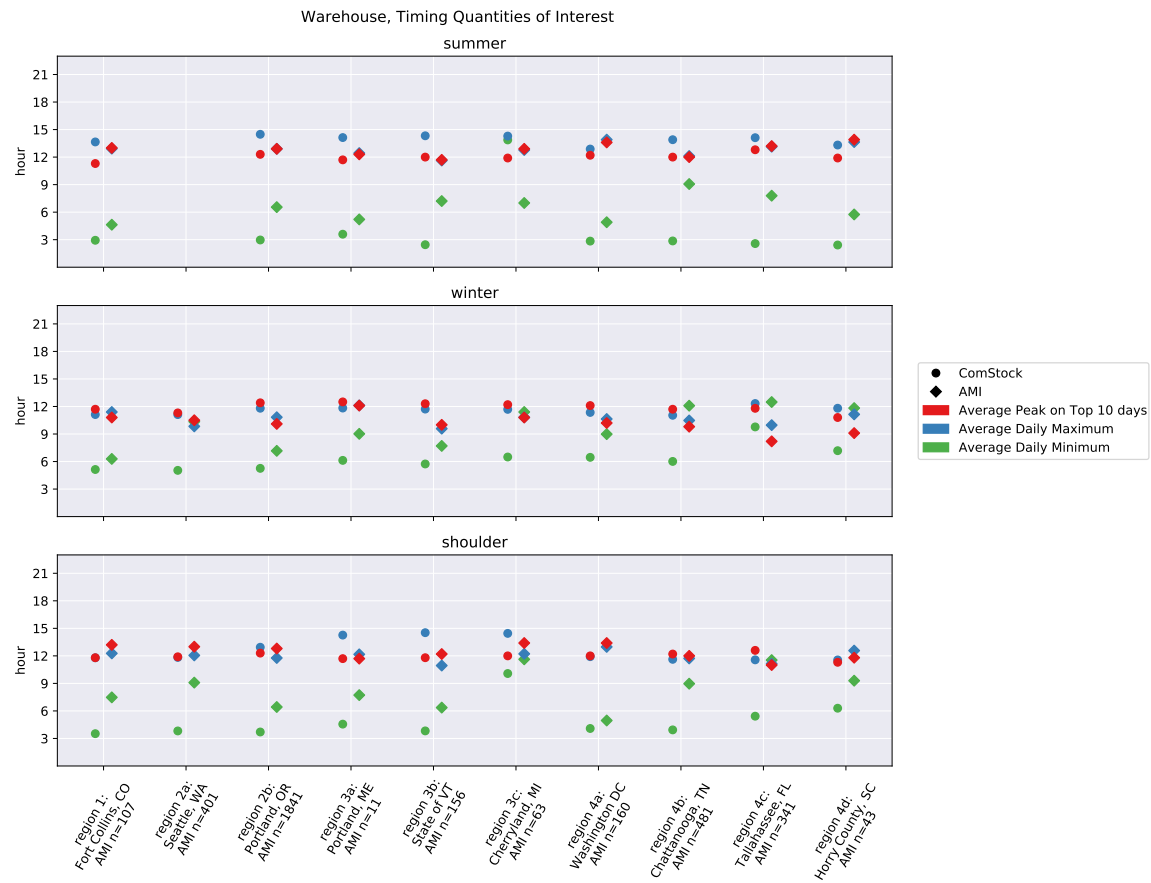


Figure 319. Timing QOI comparison for warehouse.

building_type		warehouse								
qoi_category		timing								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	-0.5	0.7	-0.3	-4.0	-1.7	-1.2	-1.4	-1.7	0.9
region2a	Seattle, WA	-0.2	nan	1.3	-5.3	nan	-5.3	-1.1	nan	0.8
region2b	Portland, OR	1.2	1.6	1.0	-2.7	-3.6	-1.9	-0.5	-0.6	2.3
region3a	Portland, ME	2.1	1.7	-0.3	-3.2	-1.6	-2.9	0.0	-0.6	0.4
region3b	State of VT	3.6	2.7	2.1	-2.5	-4.8	-2.0	-0.4	0.3	2.3
region3c	Cherryland, MI	2.2	1.5	0.9	-1.5	6.9	-4.9	-1.4	-1.0	1.4
region4a	Washington DC	-1.1	-1.0	0.7	-0.9	-2.1	-2.5	-1.4	-1.4	1.9
region4b	Chattanooga, TN	-0.1	1.8	0.6	-5.0	-6.2	-6.1	0.2	0.0	1.9
region4c	Tallahassee, FL	0.4	1.0	2.4	-6.1	-5.2	-2.7	1.6	-0.4	3.6
region4d	Horry County, SC	-1.0	-0.3	0.7	-3.0	-3.3	-4.6	-0.5	-2.0	1.7

Figure 320. Timing QOI error (hours) for warehouse.

## Primary School

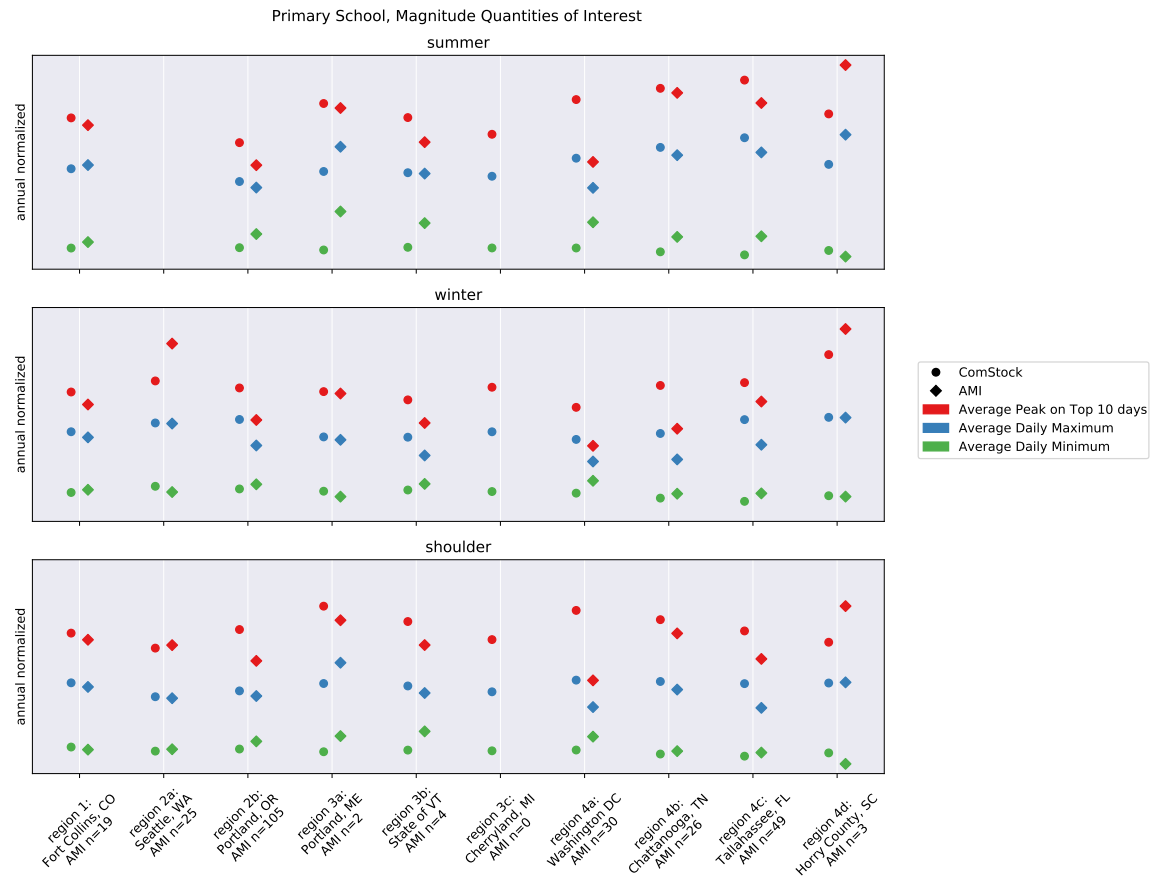


Figure 321. Normalized shape QOI comparison for primary school.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	5%	-3%	6%	9%	-19%	-8%	5%	5%	10%
region2a	Seattle, WA	2%	nan%	1%	-7%	nan%	17%	-2%	nan%	-20%
region2b	Portland, OR	6%	7%	32%	-21%	-34%	-11%	27%	21%	30%
region3a	Portland, ME	-18%	-19%	3%	-37%	-62%	19%	9%	3%	1%
region3b	State of VT	8%	1%	26%	-40%	-48%	-15%	18%	19%	22%
region4a	Washington DC	38%	35%	34%	-32%	-50%	-27%	72%	56%	48%
region4b	Chattanooga, TN	9%	7%	39%	-11%	-41%	-14%	9%	3%	45%
region4c	Tallahassee, FL	35%	12%	31%	-14%	-50%	-25%	23%	13%	15%
region4d	Horry County, SC	-1%	-21%	0%	77%	36%	3%	-21%	-23%	-13%

Figure 322. Normalized shape QOI error for primary school.

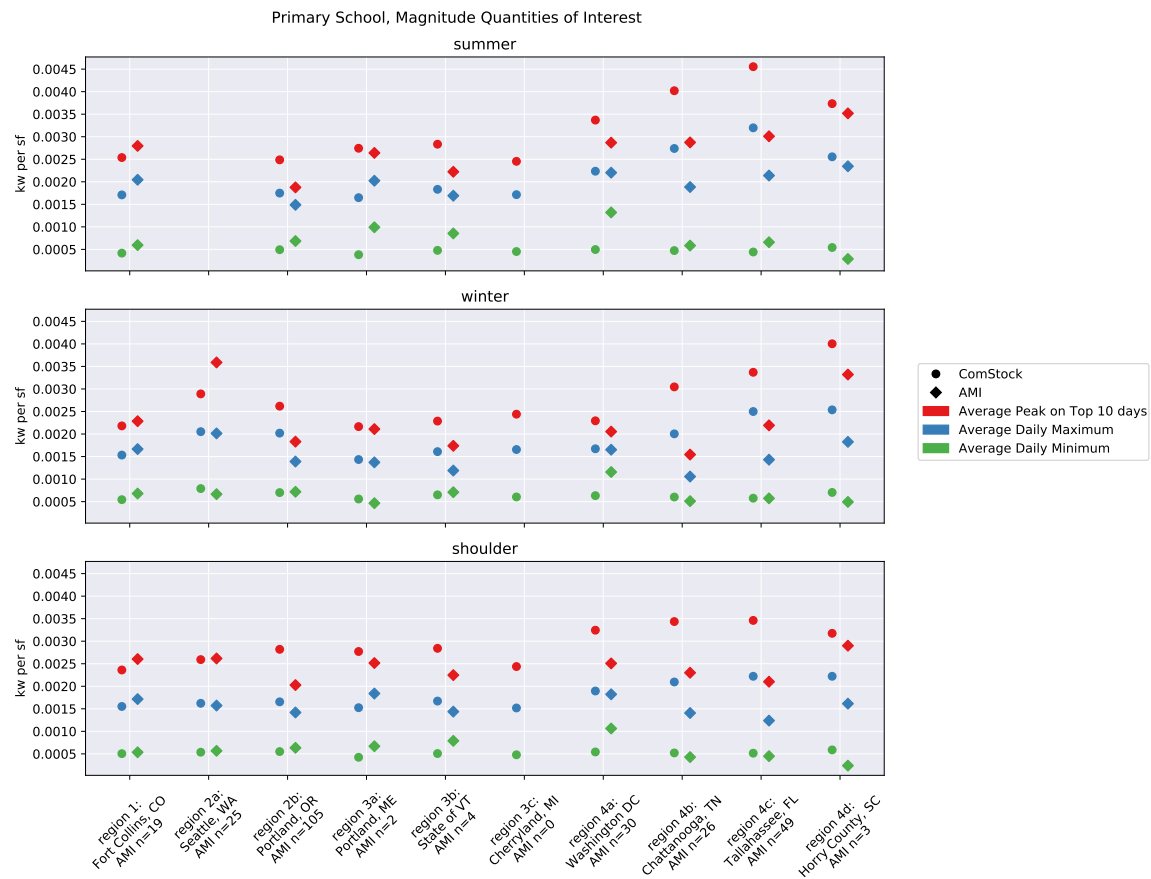


Figure 323. Magnitude QOI comparison for primary school.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	-10%	-16%	-8%	-6%	-30%	-20%	-9%	-9%	-5%
region2a	Seattle, WA	3%	nan%	2%	-5%	nan%	19%	-1%	nan%	-19%
region2b	Portland, OR	17%	18%	45%	-13%	-28%	-2%	39%	33%	43%
region3a	Portland, ME	-17%	-19%	5%	-37%	-61%	20%	10%	4%	3%
region3b	State of VT	16%	8%	35%	-36%	-44%	-8%	26%	28%	32%
region4a	Washington DC	4%	1%	1%	-49%	-62%	-45%	29%	17%	12%
region4b	Chattanooga, TN	49%	45%	90%	22%	-19%	18%	49%	40%	97%
region4c	Tallahassee, FL	79%	50%	75%	15%	-33%	1%	65%	51%	54%
region4d	Horry County, SC	38%	9%	39%	146%	88%	43%	9%	6%	21%

Figure 324. Magnitude QOI error for primary school.

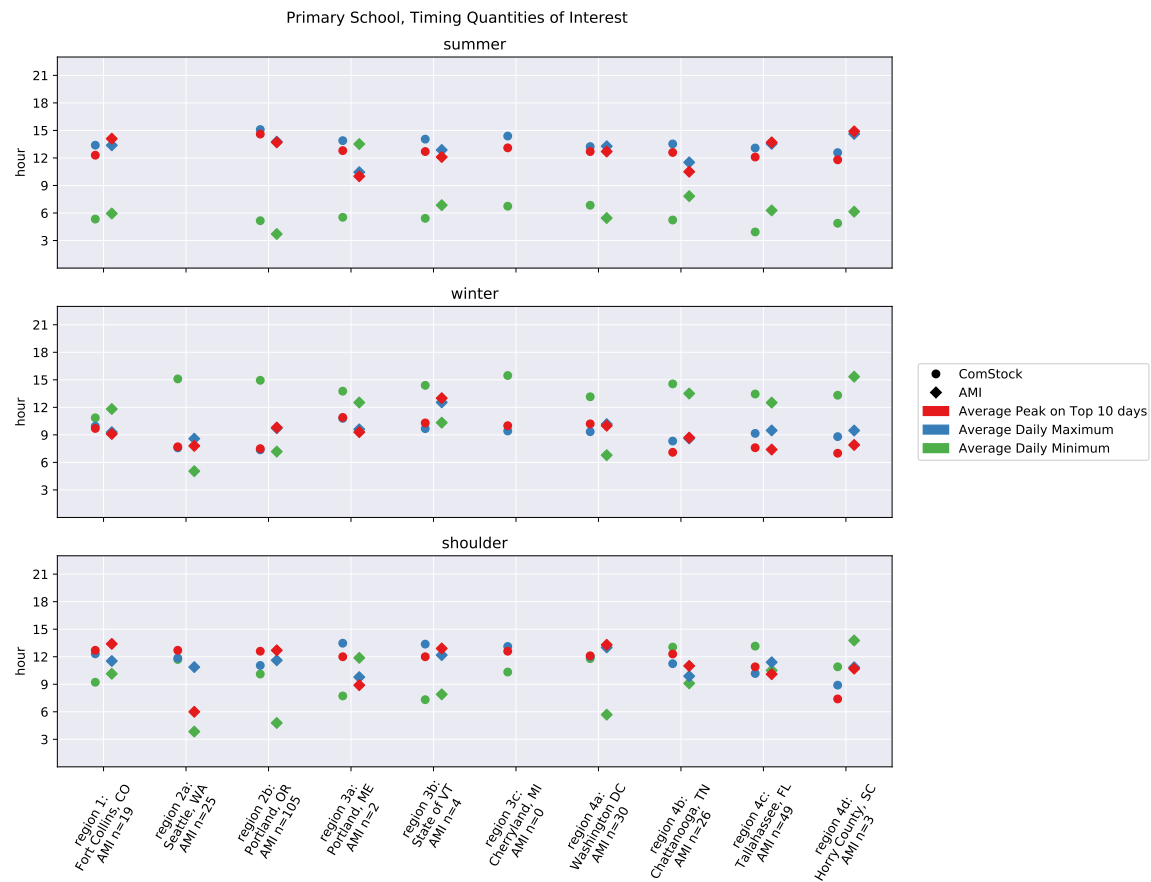


Figure 325. Timing QOI comparison for primary school.

building_type		primary_school									
qoi_category		timing									
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum			
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter	
region	region_name										
region1	Fort Collins, CO	0.8	0.0	0.7	-0.9	-0.6	-1.0	-0.7	-1.8	0.6	
region2a	Seattle, WA	1.0	nan	-1.0	7.9	nan	10.1	6.7	nan	-0.1	
region2b	Portland, OR	-0.6	1.3	-2.4	5.3	1.5	7.8	-0.1	0.9	-2.3	
region3a	Portland, ME	3.7	3.4	1.2	-4.2	-8.0	1.2	3.1	2.8	1.6	
region3b	State of VT	1.2	1.2	-2.9	-0.6	-1.4	4.1	-0.9	0.6	-2.7	
region4a	Washington DC	-1.0	-0.0	-0.8	6.1	1.4	6.4	-1.2	0.0	0.2	
region4b	Chattanooga, TN	1.4	2.0	-0.3	4.0	-2.6	1.1	1.3	2.1	-1.6	
region4c	Tallahassee, FL	-1.2	-0.5	-0.3	2.6	-2.3	0.9	0.8	-1.6	0.2	
region4d	Horry County, SC	-2.0	-2.0	-0.7	-2.9	-1.3	-2.0	-3.3	-3.1	-0.9	

Figure 326. Timing QOI error (hours) for primary school.

*Secondary School*

As previously described, all school AMI were classified as Primary School.

### Full-Service Restaurant

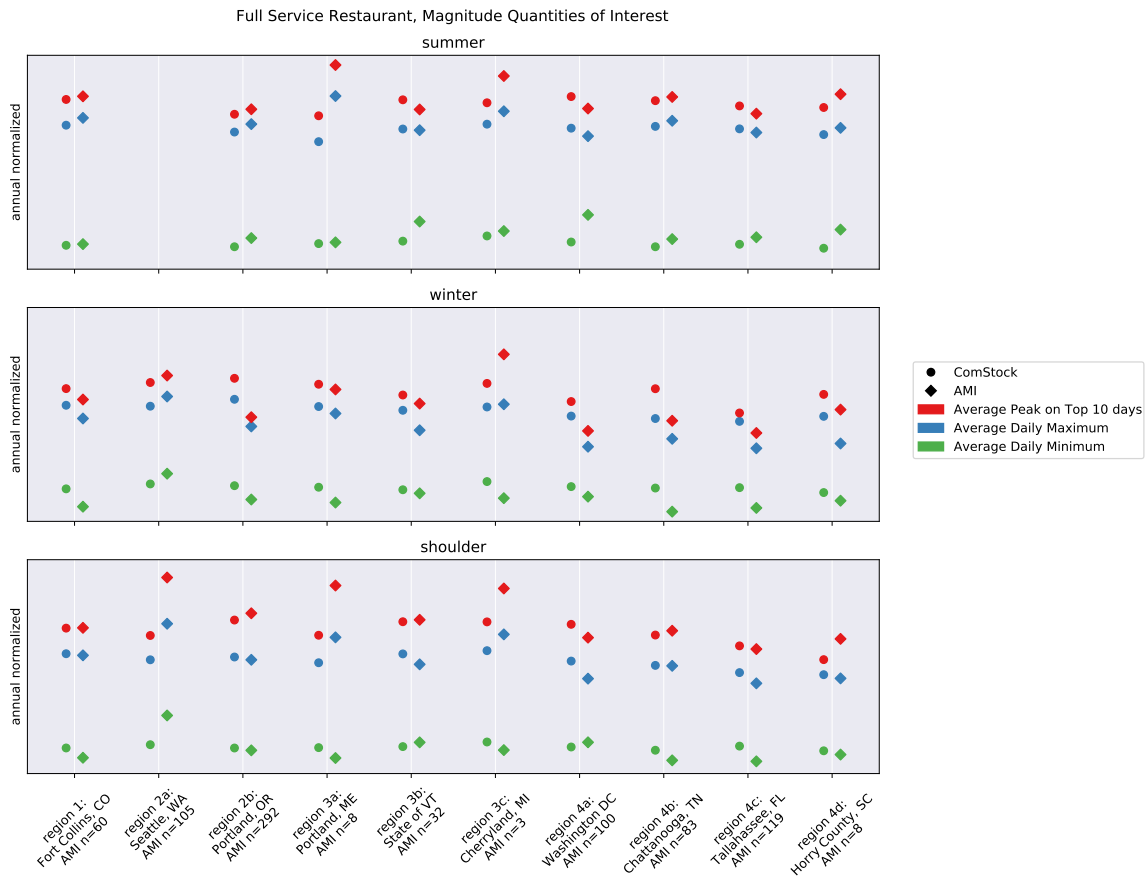


Figure 327. Normalized shape QOI comparison for full service restaurant.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	1%	-4%	9%	18%	-2%	34%	-0%	-1%	7%
region2a	Seattle, WA	-19%	nan%	-6%	-30%	nan%	-12%	-25%	nan%	-4%
region2b	Portland, OR	2%	-4%	20%	4%	-13%	23%	-3%	-3%	27%
region3a	Portland, ME	-15%	-22%	5%	19%	-2%	27%	-22%	-21%	3%
region3b	State of VT	7%	1%	15%	-6%	-23%	5%	-1%	5%	5%
region3c	Cherryland, MI	-9%	-7%	-2%	13%	-6%	27%	-15%	-12%	-14%
region4a	Washington DC	13%	5%	27%	-7%	-29%	16%	8%	6%	23%
region4b	Chattanooga, TN	0%	-3%	17%	20%	-11%	50%	-2%	-2%	23%
region4c	Tallahassee, FL	8%	2%	24%	30%	-10%	39%	2%	4%	16%
region4d	Horry County, SC	3%	-4%	23%	7%	-24%	14%	-12%	-6%	10%

Figure 328. Normalized shape QOI error for full service restaurant.

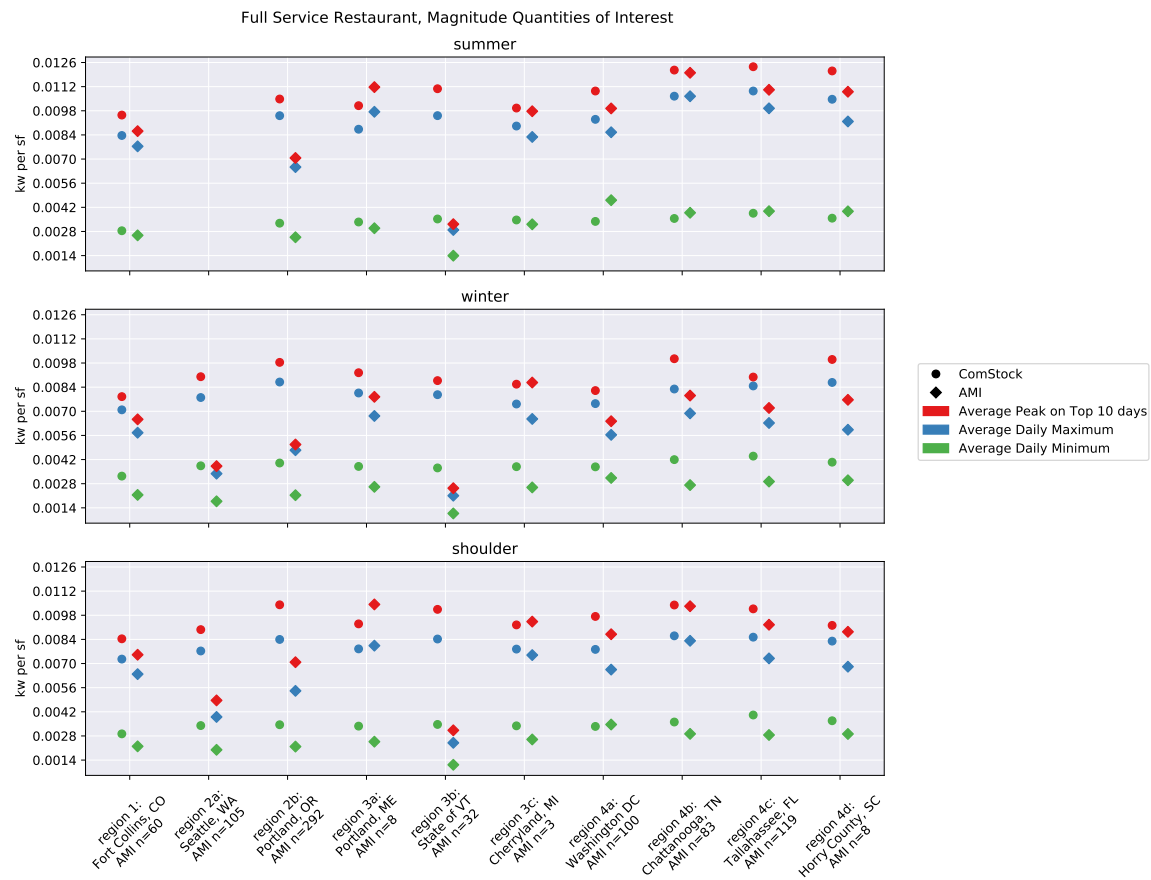


Figure 329. Magnitude QOI comparison for full service restaurant.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	14%	8%	23%	33%	10%	51%	12%	11%	20%
region2a	Seattle, WA	98%	nan%	131%	70%	nan%	116%	84%	nan%	136%
region2b	Portland, OR	55%	46%	83%	58%	33%	87%	47%	48%	94%
region3a	Portland, ME	-2%	-10%	20%	37%	12%	45%	-11%	-10%	18%
region3b	State of VT	251%	230%	278%	208%	153%	245%	224%	244%	246%
region3c	Cherryland, MI	5%	8%	13%	30%	8%	47%	-2%	2%	-1%
region4a	Washington DC	18%	9%	32%	-3%	-27%	21%	12%	10%	28%
region4b	Chattanooga, TN	3%	0%	21%	24%	-9%	54%	1%	1%	27%
region4c	Tallahassee, FL	17%	10%	34%	41%	-3%	50%	10%	12%	25%
region4d	Horry County, SC	22%	14%	46%	26%	-10%	35%	4%	11%	31%

Figure 330. Magnitude QOI error for full service restaurant.



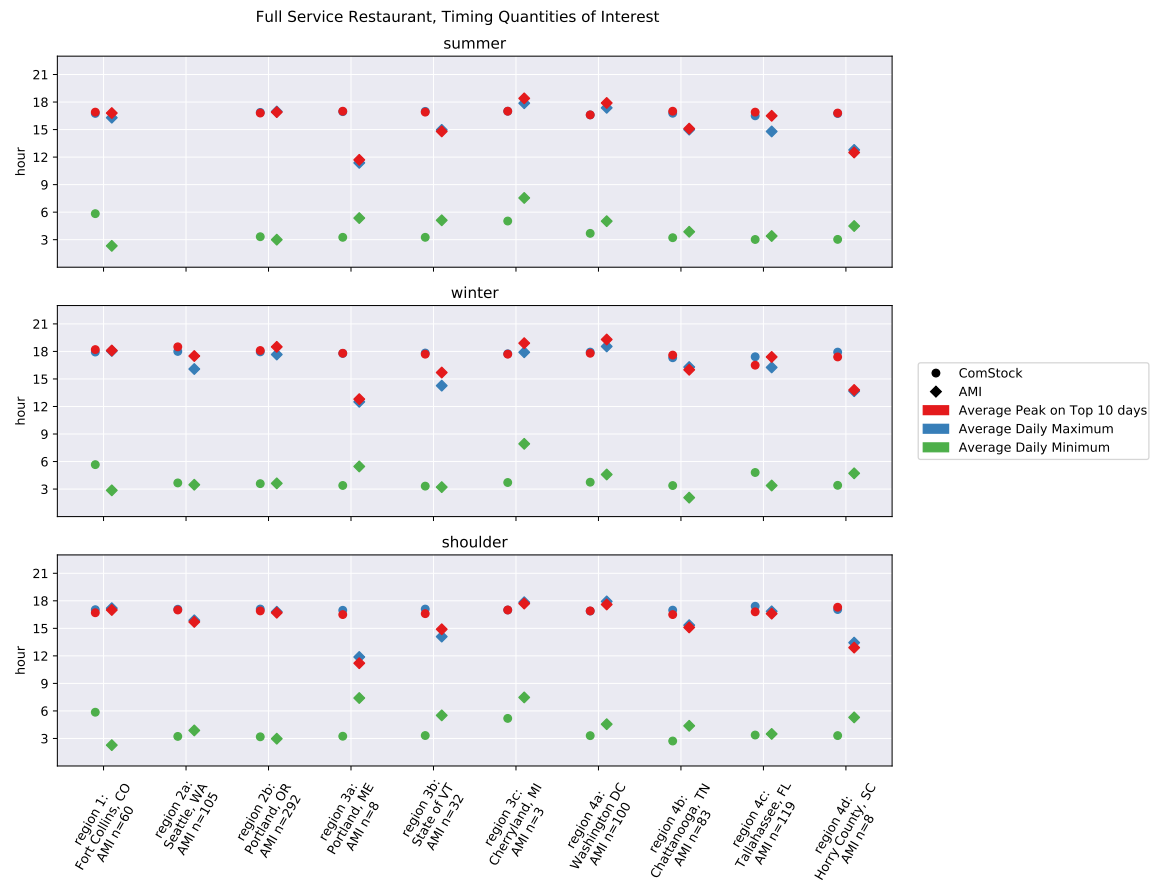


Figure 331. Timing QOI comparison for full service restaurant.

building_type		full_service_restaurant								
qoi_category		timing								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	-0.2	0.5	-0.1	3.6	3.5	2.8	-0.3	0.1	0.1
region2a	Seattle, WA	1.2	nan	1.9	-0.7	nan	0.2	1.3	nan	1.0
region2b	Portland, OR	0.3	-0.1	0.3	0.2	0.3	-0.0	0.2	-0.1	-0.4
region3a	Portland, ME	5.1	5.6	5.3	-4.2	-2.1	-2.1	5.3	5.3	5.0
region3b	State of VT	3.0	2.0	3.6	-2.2	-1.9	0.1	1.7	2.1	2.0
region3c	Cherryland, MI	-0.8	-0.9	-0.2	-2.3	-2.5	-4.2	-0.7	-1.4	-1.2
region4a	Washington DC	-1.1	-0.8	-0.6	-1.2	-1.3	-0.8	-0.7	-1.3	-1.5
region4b	Chattanooga, TN	1.6	1.8	1.0	-1.7	-0.6	1.3	1.4	1.9	1.6
region4c	Tallahassee, FL	0.6	1.7	1.2	-0.1	-0.4	1.4	0.2	0.4	-0.9
region4d	Horry County, SC	3.6	4.0	4.3	-2.0	-1.5	-1.3	4.4	4.3	3.6

Figure 332. Timing QOI error (hours) for full service restaurant.

## Quick-Service Restaurant

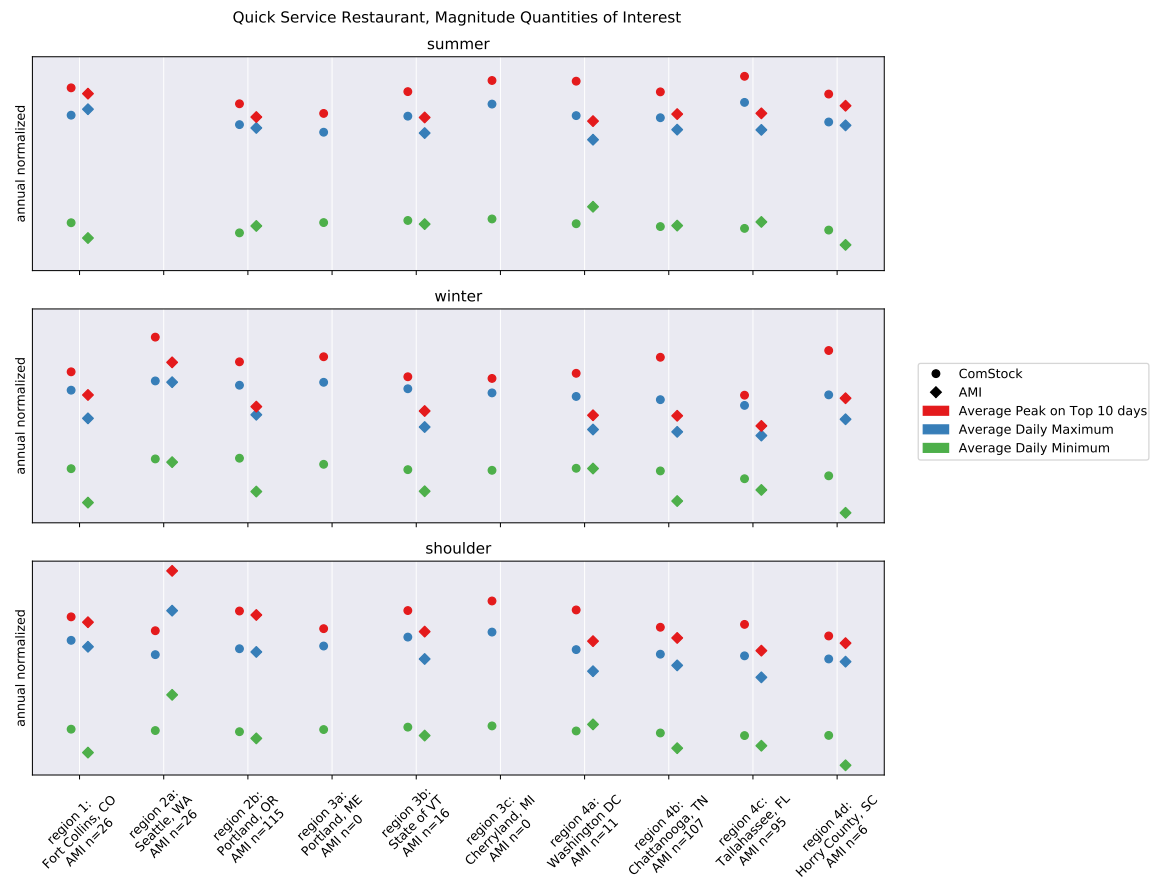


Figure 333. Normalized shape QOI comparison for quick service restaurant.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	4%	-3%	19%	37%	21%	55%	3%	3%	14%
region2a	Seattle, WA	-21%	nan%	1%	-29%	nan%	3%	-24%	nan%	13%
region2b	Portland, OR	2%	2%	20%	9%	-8%	46%	2%	7%	28%
region3b	State of VT	14%	9%	28%	10%	4%	30%	11%	13%	22%
region4a	Washington DC	15%	14%	25%	-7%	-16%	0%	18%	21%	28%
region4b	Chattanooga, TN	7%	7%	24%	22%	-1%	48%	6%	11%	40%
region4c	Tallahassee, FL	16%	15%	24%	14%	-7%	15%	16%	19%	22%
region4d	Horry County, SC	2%	2%	17%	59%	22%	72%	4%	6%	29%

Figure 334. Normalized shape QOI error for quick service restaurant.

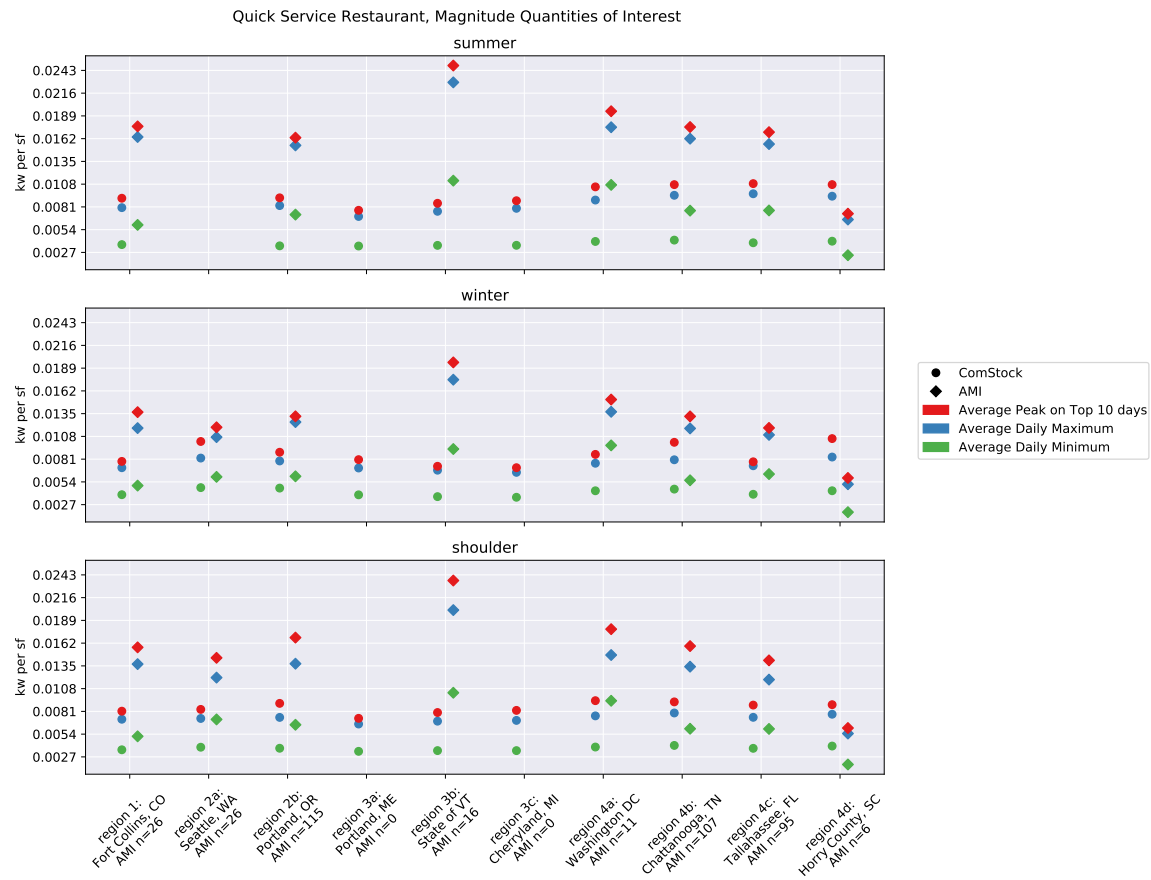


Figure 335. Magnitude QOI comparison for quick service restaurant.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	-48%	-51%	-40%	-31%	-39%	-22%	-48%	-48%	-43%
region2a	Seattle, WA	-40%	nan%	-23%	-46%	nan%	-21%	-42%	nan%	-14%
region2b	Portland, OR	-46%	-46%	-37%	-43%	-52%	-23%	-46%	-44%	-32%
region3b	State of VT	-66%	-67%	-61%	-67%	-68%	-61%	-66%	-66%	-63%
region4a	Washington DC	-49%	-49%	-45%	-59%	-63%	-55%	-48%	-46%	-43%
region4b	Chattanooga, TN	-41%	-41%	-32%	-33%	-46%	-19%	-42%	-39%	-23%
region4c	Tallahassee, FL	-38%	-38%	-33%	-38%	-50%	-38%	-37%	-36%	-34%
region4d	Horry County, SC	42%	42%	63%	121%	70%	140%	45%	47%	80%

Figure 336. Magnitude QOI error for quick service restaurant.

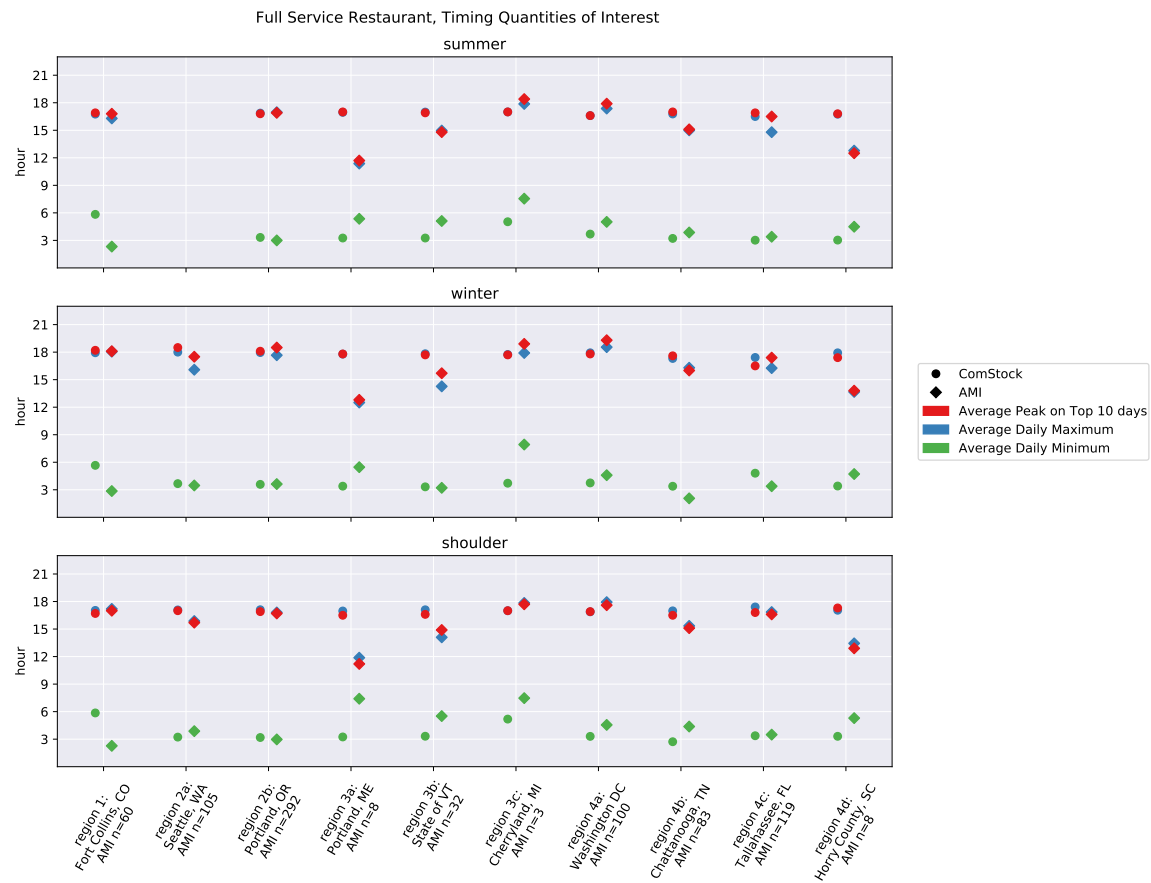


Figure 337. Timing QOI comparison for full service restaurant.

building_type		quick_service_restaurant								
qoi_category		timing								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	4.4	4.8	2.5	5.6	3.5	5.6	4.2	4.8	5.0
region2a	Seattle, WA	4.0	nan	4.0	0.3	nan	1.0	4.4	nan	4.5
region2b	Portland, OR	1.6	1.5	1.2	6.7	5.7	7.4	1.2	1.9	1.9
region3b	State of VT	4.8	5.5	2.7	4.8	2.1	6.0	5.3	5.6	1.7
region4a	Washington DC	1.0	2.7	0.3	-0.1	-2.4	-1.4	1.7	1.2	0.3
region4b	Chattanooga, TN	2.9	5.0	-0.7	-0.1	0.3	3.0	5.1	5.8	-5.5
region4c	Tallahassee, FL	3.2	4.9	0.6	1.9	0.2	3.2	4.5	5.5	-1.2
region4d	Horry County, SC	-1.8	-1.4	-1.8	-0.1	-1.0	0.7	-4.7	-2.7	-9.0

Figure 338. Timing QOI error (hours) for quick service restaurant.

## Small Hotel

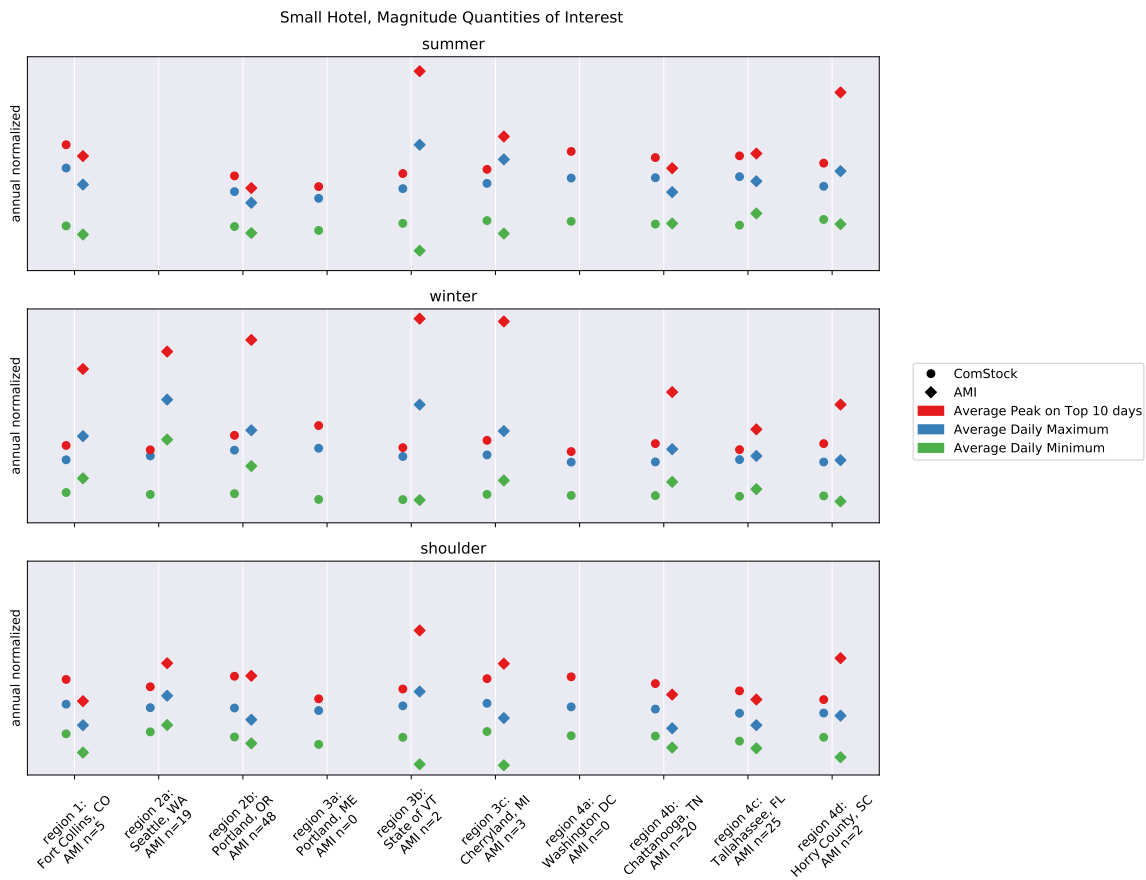


Figure 339. Normalized shape QOI comparison for small hotel.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	26%	14%	-20%	36%	13%	-19%	21%	8%	-42%
region2a	Seattle, WA	-11%	nan%	-37%	-9%	nan%	-48%	-17%	nan%	-49%
region2b	Portland, OR	14%	11%	-16%	10%	10%	-32%	-0%	11%	-45%
region3b	State of VT	-13%	-28%	-35%	66%	55%	1%	-34%	-45%	-55%
region3c	Cherryland, MI	17%	-17%	-20%	85%	19%	-19%	-11%	-20%	-51%
region4b	Chattanooga, TN	25%	13%	-12%	20%	-1%	-19%	10%	8%	-32%
region4c	Tallahassee, FL	15%	4%	-4%	13%	-13%	-11%	8%	-2%	-16%
region4d	Horry County, SC	3%	-12%	-2%	42%	6%	11%	-28%	-34%	-26%

Figure 340. Normalized shape QOI error for small hotel.

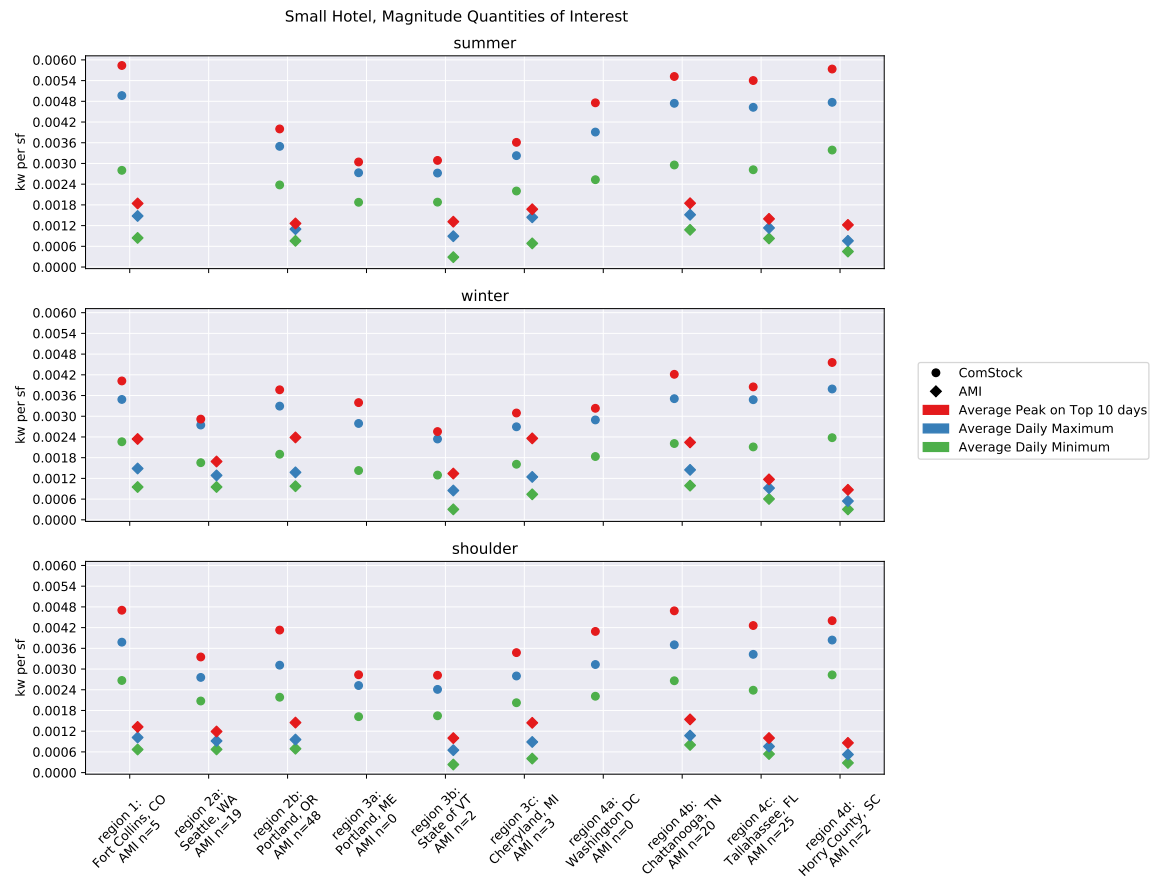


Figure 341. Magnitude QOI comparison for small hotel.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	271%	236%	135%	299%	232%	138%	255%	217%	72%
region2a	Seattle, WA	200%	nan%	113%	209%	nan%	74%	181%	nan%	73%
region2b	Portland, OR	225%	218%	140%	216%	214%	95%	185%	217%	58%
region3b	State of VT	271%	205%	176%	604%	556%	328%	182%	135%	91%
region3c	Cherryland, MI	215%	124%	117%	398%	222%	118%	141%	116%	31%
region4b	Chattanooga, TN	246%	213%	143%	232%	174%	123%	204%	199%	88%
region4c	Tallahassee, FL	352%	308%	279%	343%	240%	249%	326%	287%	229%
region4d	Horry County, SC	633%	529%	598%	908%	655%	688%	411%	370%	424%

Figure 342. Magnitude QOI error for small hotel.

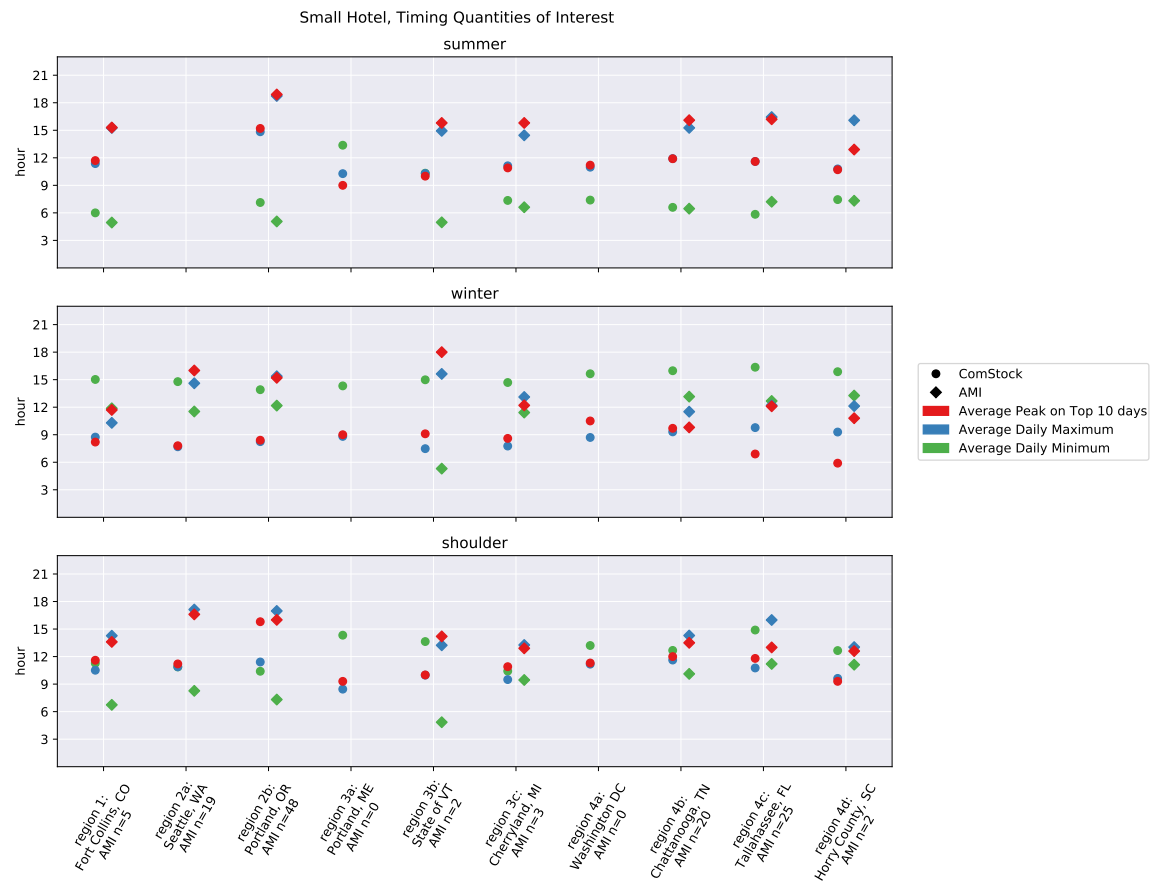


Figure 343. Timing QOI comparison for small hotel.

building_type		small_hotel								
qoi_category		timing								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	-3.8	-3.9	-1.6	4.5	1.1	3.1	-2.0	-3.6	-3.5
region2a	Seattle, WA	-6.2	nan	-6.9	2.6	nan	3.3	-5.4	nan	-8.2
region2b	Portland, OR	-5.6	-3.9	-7.1	3.1	2.1	1.7	-0.2	-3.7	-6.8
region3b	State of VT	-3.3	-4.6	-8.1	8.8	5.4	9.7	-4.2	-5.8	-8.9
region3c	Cherryland, MI	-3.8	-3.3	-5.3	1.0	0.7	3.3	-2.0	-4.9	-3.6
region4b	Chattanooga, TN	-2.7	-3.3	-2.2	2.6	0.1	2.8	-1.5	-4.2	-0.1
region4c	Tallahassee, FL	-5.2	-4.8	-2.4	3.7	-1.4	3.7	-1.2	-4.6	-5.2
region4d	Horry County, SC	-3.4	-5.3	-2.8	1.5	0.1	2.6	-3.3	-2.2	-4.9

Figure 344. Timing QOI error (hours) for small hotel.



Large Hotel

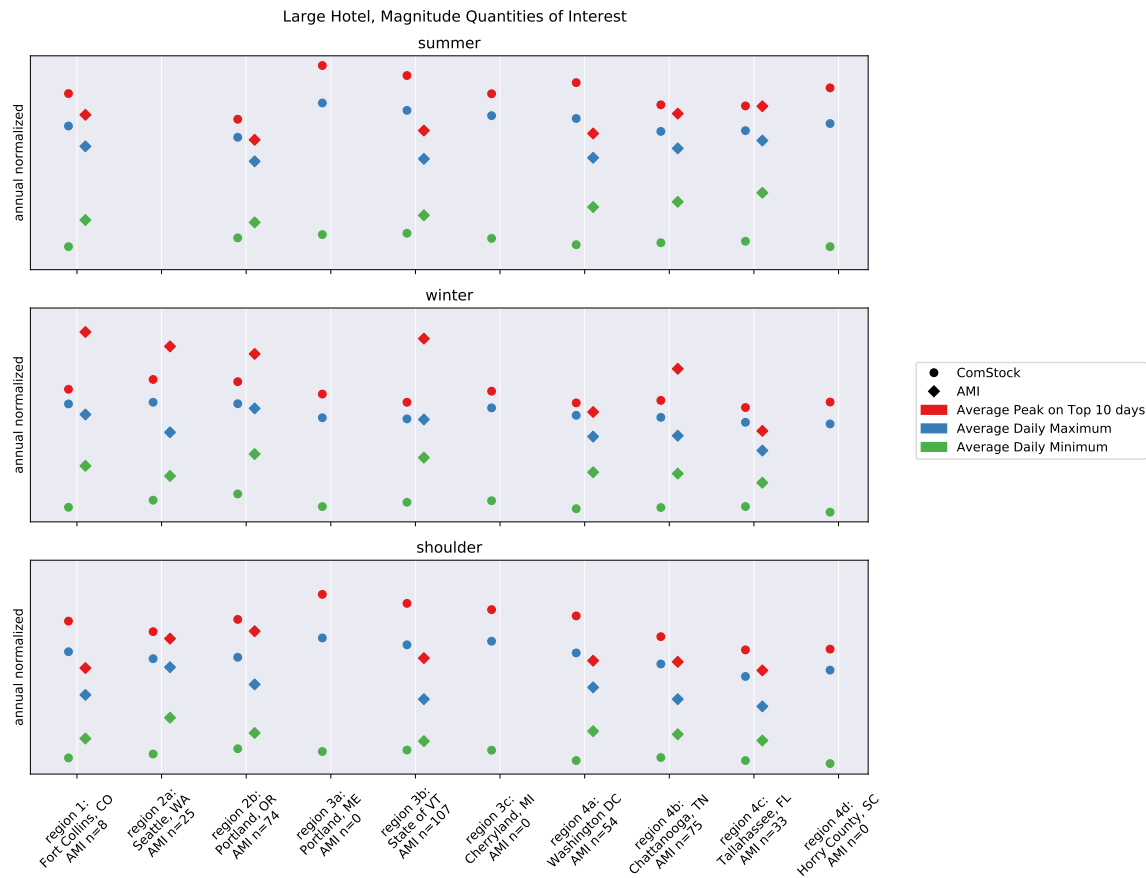


Figure 345. Normalized shape QOI comparison for large hotel.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	30%	11%	6%	-20%	-24%	-35%	28%	10%	-23%
region2a	Seattle, WA	5%	nan%	20%	-30%	nan%	-22%	4%	nan%	-14%
region2b	Portland, OR	18%	14%	3%	-15%	-14%	-31%	6%	11%	-12%
region3b	State of VT	39%	28%	0%	-9%	-15%	-35%	31%	27%	-26%
region4a	Washington DC	23%	23%	14%	-28%	-30%	-32%	25%	26%	5%
region4b	Chattanooga, TN	26%	9%	12%	-23%	-31%	-31%	14%	4%	-15%
region4c	Tallahassee, FL	23%	5%	21%	-21%	-35%	-23%	12%	0%	15%

Figure 346. Normalized shape QOI error for large hotel.

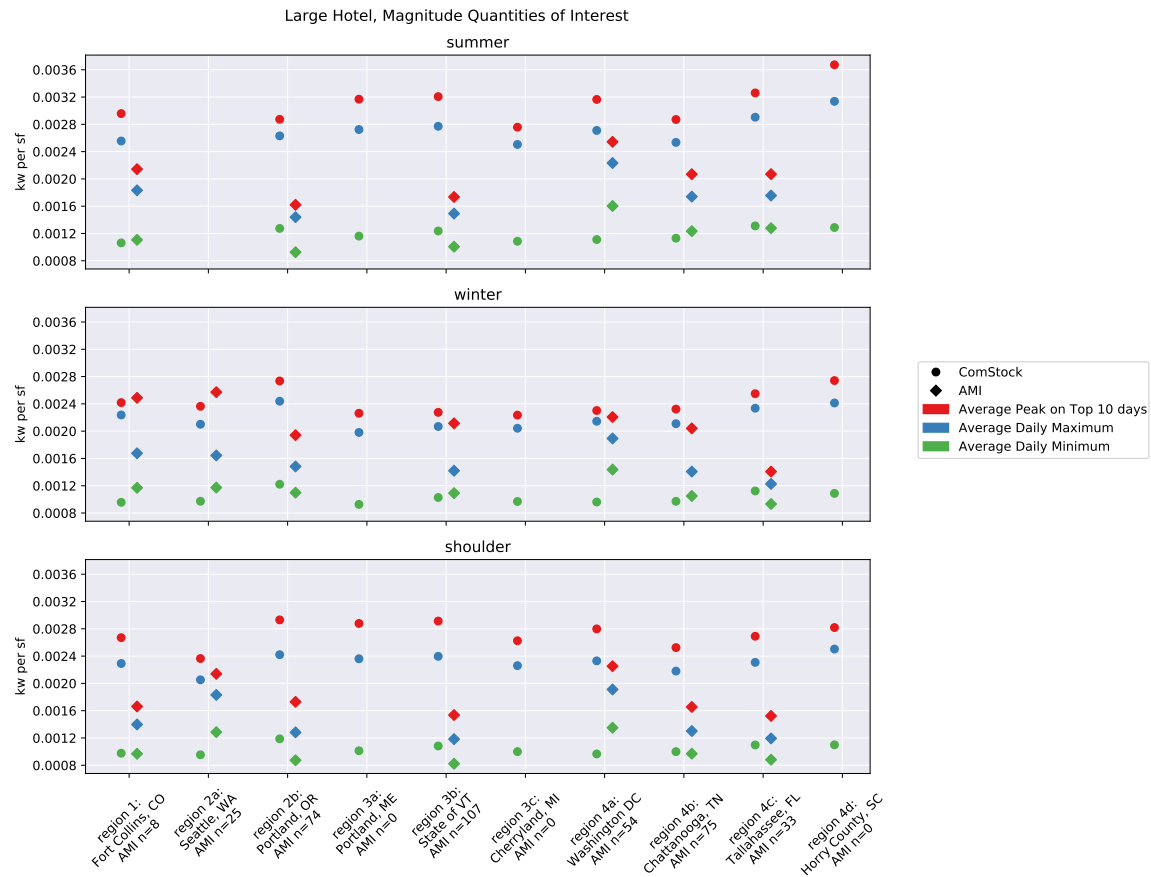


Figure 347. Magnitude QOI comparison for large hotel.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	64%	39%	34%	1%	-4%	-18%	61%	38%	-3%
region2a	Seattle, WA	12%	nan%	28%	-26%	nan%	-17%	11%	nan%	-8%
region2b	Portland, OR	89%	83%	65%	36%	38%	11%	69%	77%	41%
region3b	State of VT	103%	86%	46%	32%	23%	-6%	90%	85%	8%
region4a	Washington DC	22%	21%	13%	-28%	-31%	-33%	24%	24%	4%
region4b	Chattanooga, TN	68%	46%	50%	3%	-8%	-7%	53%	39%	14%
region4c	Tallahassee, FL	93%	65%	90%	25%	3%	20%	77%	58%	81%

Figure 348. Magnitude QOI error for large hotel.

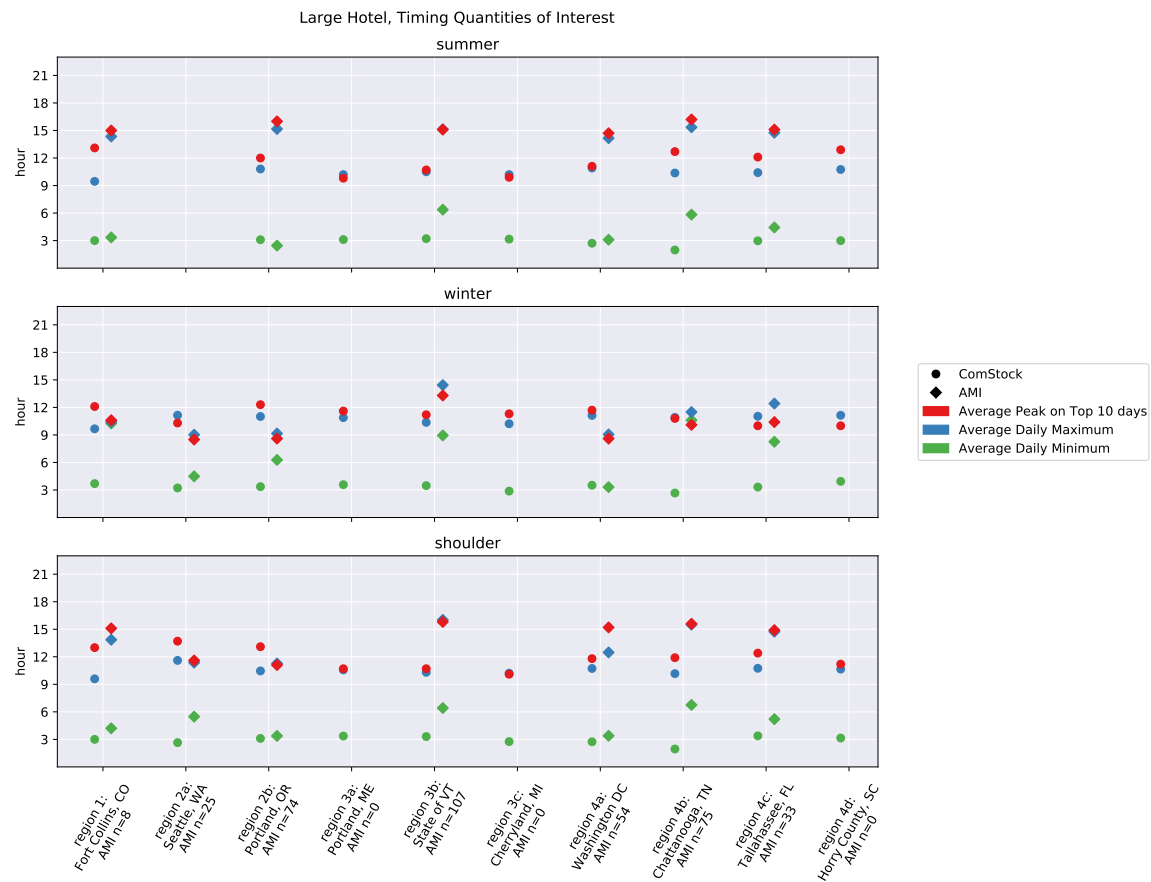


Figure 349. Timing QOI comparison for large hotel.

building_type		large_hotel								
qoi_category		timing								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	-4.3	-4.9	-0.8	-1.2	-0.3	-6.6	-2.1	-1.9	1.5
region2a	Seattle, WA	0.2	nan	2.1	-2.8	nan	-1.3	2.1	nan	1.8
region2b	Portland, OR	-0.8	-4.4	1.9	-0.3	0.6	-2.9	2.0	-4.0	3.7
region3b	State of VT	-5.7	-4.6	-4.1	-3.1	-3.2	-5.5	-5.1	-4.4	-2.1
region4a	Washington DC	-1.7	-3.3	2.1	-0.7	-0.4	0.2	-3.4	-3.6	3.1
region4b	Chattanooga, TN	-5.3	-5.0	-0.6	-4.8	-3.9	-7.9	-3.7	-3.5	0.7
region4c	Tallahassee, FL	-4.0	-4.3	-1.4	-1.8	-1.5	-4.9	-2.5	-3.0	-0.4

Figure 350. Timing QOI error (hours) for large hotel.

Hospital

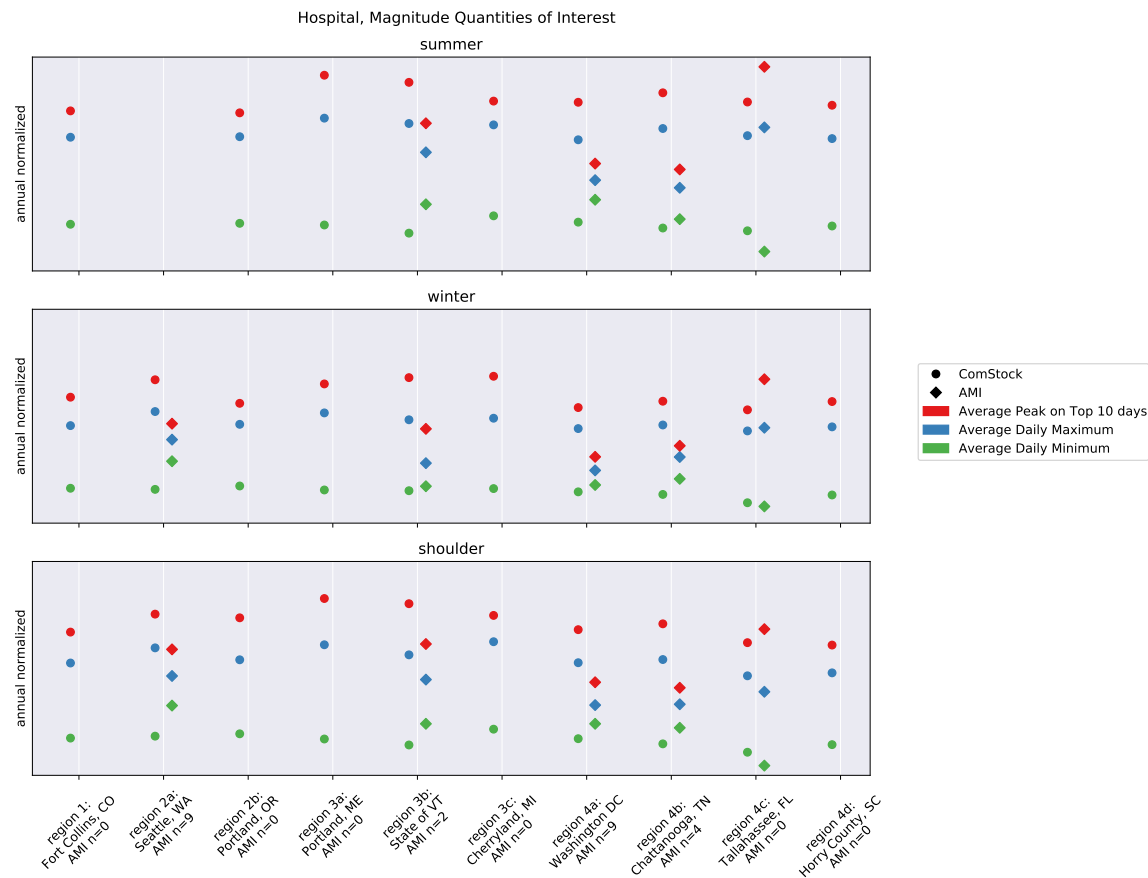


Figure 351. Normalized shape QOI comparison for hospital.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region2a	Seattle, WA	21%	nan%	24%	-30%	nan%	-30%	22%	nan%	33%
region3b	State of VT	19%	19%	47%	-25%	-29%	-6%	25%	23%	40%
region4a	Washington DC	41%	33%	49%	-18%	-22%	-10%	42%	44%	50%
region4b	Chattanooga, TN	43%	51%	32%	-20%	-11%	-20%	53%	57%	41%
region4c	Tallahassee, FL	14%	-5%	-2%	32%	40%	7%	-8%	-15%	-17%

Figure 352. Normalized shape QOI error for hospital.

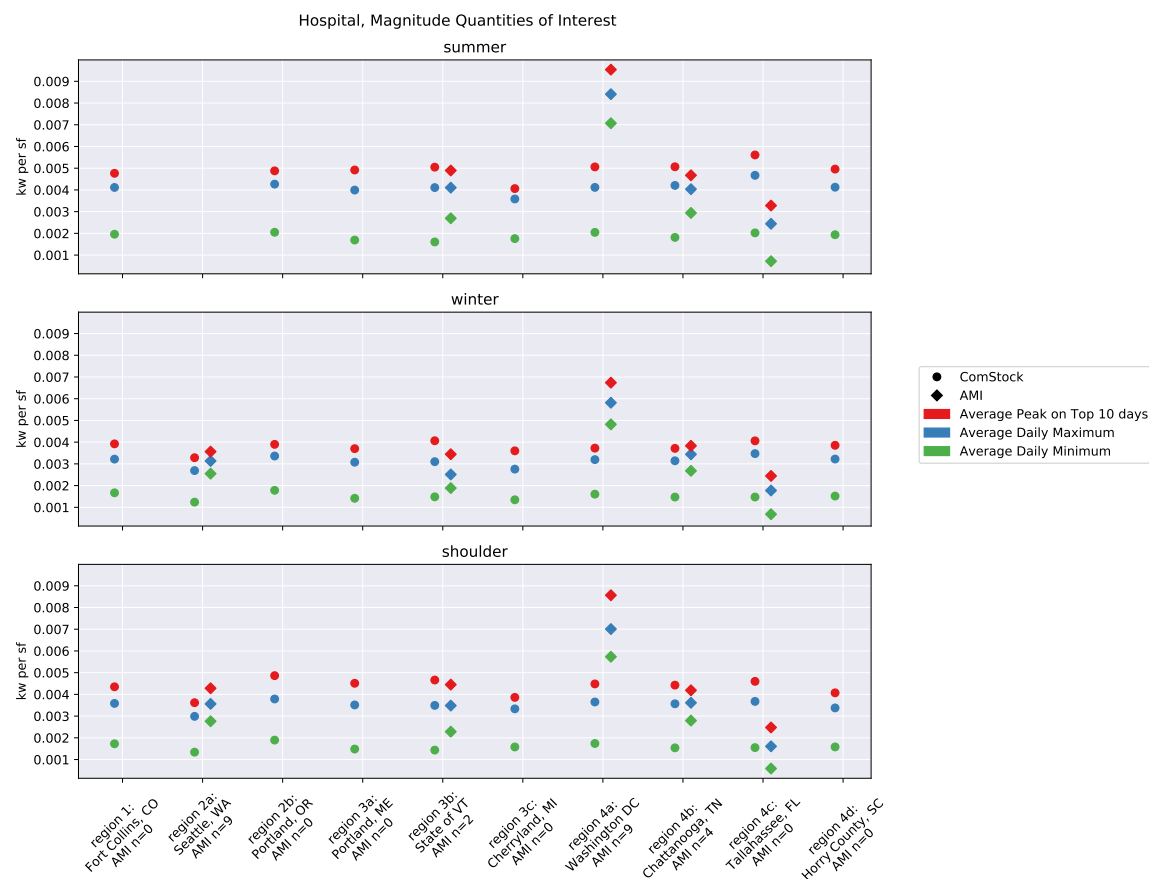


Figure 353. Magnitude QOI comparison for hospital.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region2a	Seattle, WA	-16%	nan%	-14%	-52%	nan%	-51%	-15%	nan%	-8%
region3b	State of VT	0%	0%	23%	-37%	-40%	-21%	5%	3%	18%
region4a	Washington DC	-48%	-51%	-45%	-70%	-71%	-67%	-48%	-47%	-45%
region4b	Chattanooga, TN	-1%	4%	-9%	-45%	-38%	-45%	6%	8%	-3%
region4c	Tallahassee, FL	129%	91%	96%	165%	181%	116%	86%	71%	66%

Figure 354. Magnitude QOI error for hospital.

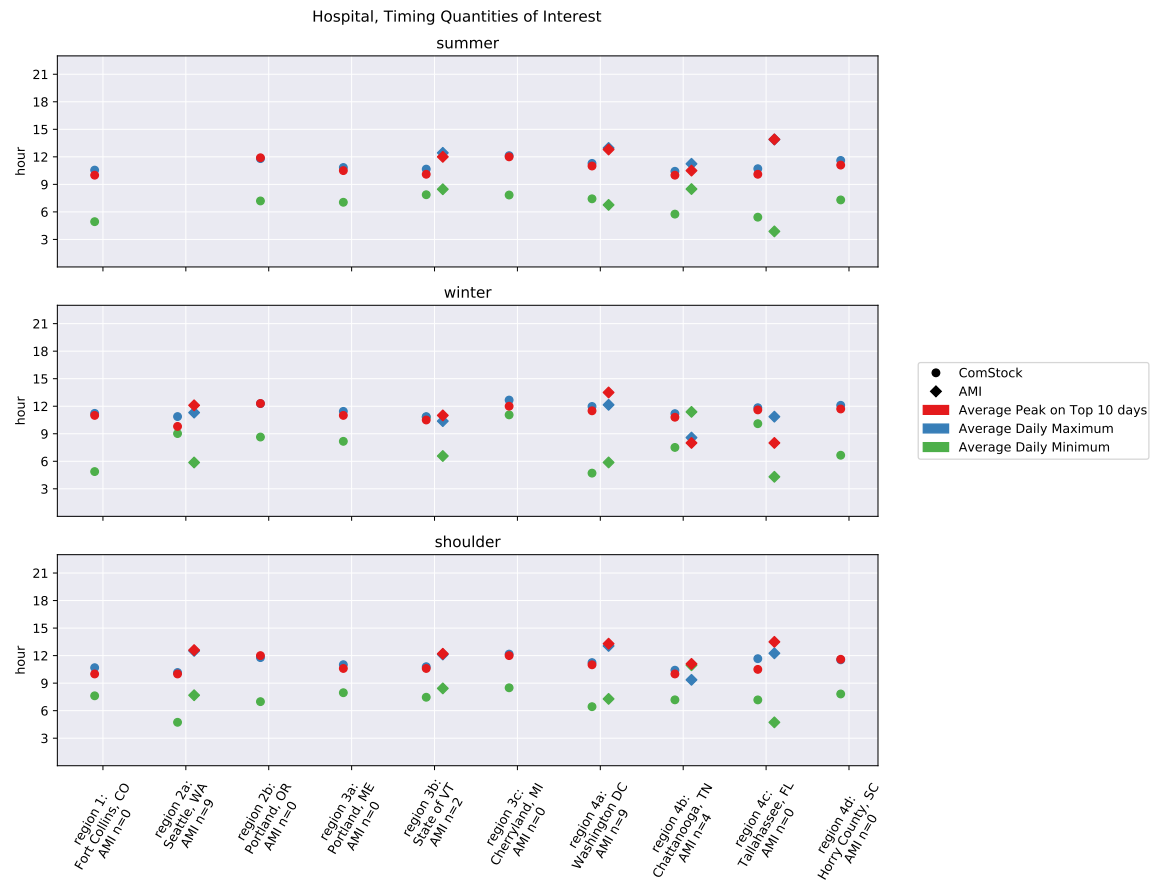


Figure 355. Timing QOI comparison for hospital.

building_type		hospital								
qoi_category		timing								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region2a	Seattle, WA	-2.4	nan	-0.4	-3.0	nan	3.2	-2.6	nan	-2.3
region3b	State of VT	-1.3	-1.8	0.5	-1.0	-0.6	4.2	-1.6	-1.9	-0.5
region4a	Washington DC	-1.8	-1.7	-0.2	-0.8	0.7	-1.2	-2.3	-1.8	-2.0
region4b	Chattanooga, TN	1.0	-0.8	2.6	-3.8	-2.7	-3.9	-1.1	-0.5	2.8
region4c	Tallahassee, FL	-0.6	-3.2	1.0	2.4	1.5	5.8	-3.0	-3.8	3.6

Figure 356. Timing QOI error (hours) for hospital.

## Outpatient

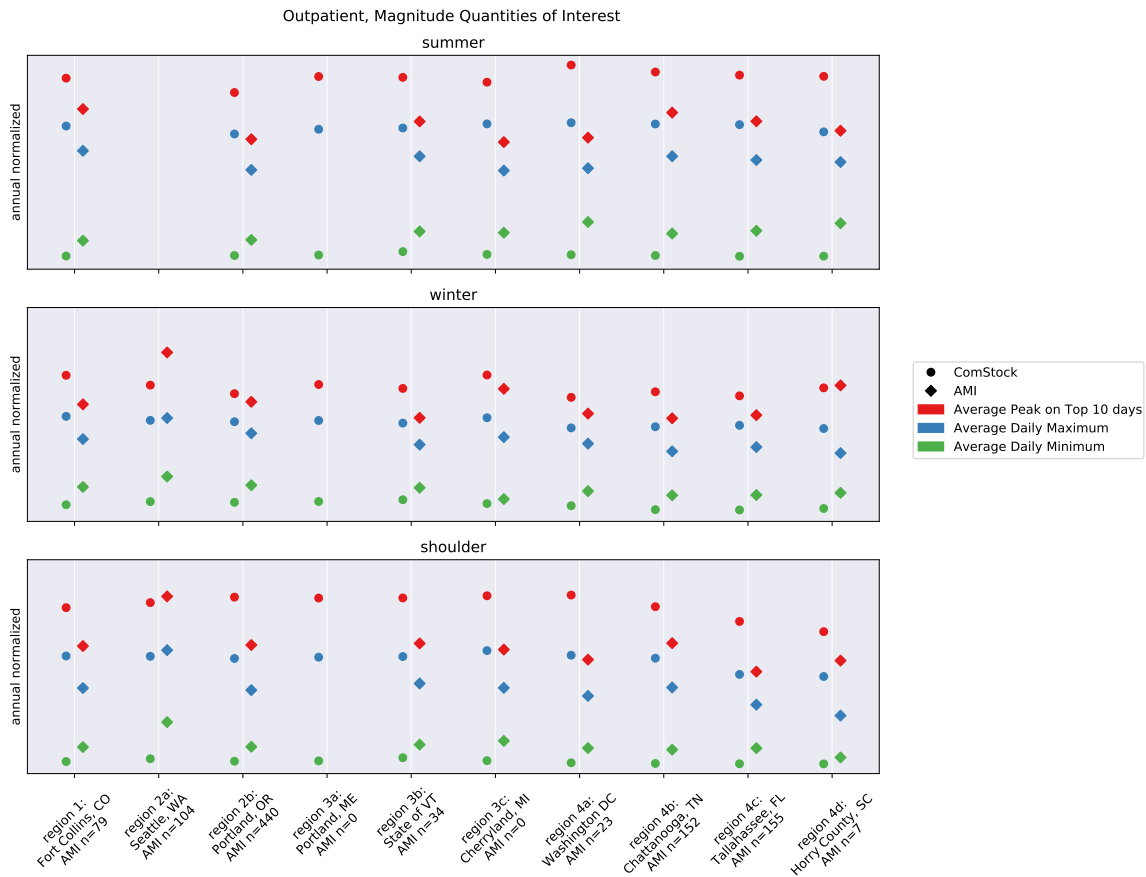


Figure 357. Normalized shape QOI comparison for outpatient.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	28%	17%	20%	-26%	-26%	-28%	24%	16%	20%
region2a	Seattle, WA	-4%	nan%	-2%	-45%	nan%	-34%	-3%	nan%	-16%
region2b	Portland, OR	28%	28%	10%	-26%	-26%	-26%	30%	29%	5%
region3b	State of VT	22%	20%	20%	-22%	-30%	-19%	28%	25%	22%
region3c	Cherryland, MI	32%	36%	17%	-32%	-33%	-9%	35%	38%	9%
region4a	Washington DC	38%	35%	15%	-27%	-43%	-24%	45%	45%	12%
region4b	Chattanooga, TN	25%	23%	25%	-26%	-34%	-26%	23%	22%	20%
region4c	Tallahassee, FL	31%	25%	21%	-28%	-37%	-27%	38%	26%	14%
region4d	Horry County, SC	45%	22%	25%	-14%	-43%	-27%	20%	32%	-1%

Figure 358. Normalized shape QOI error for outpatient.



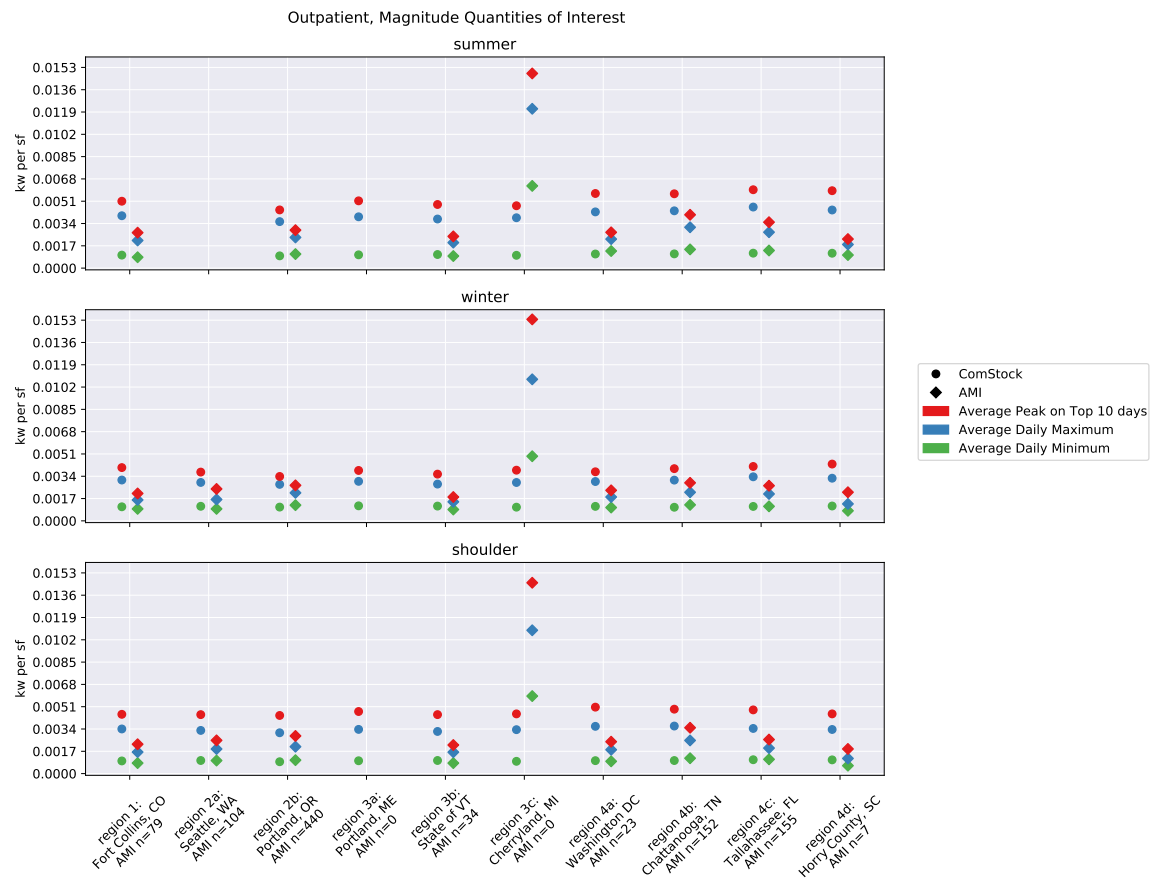


Figure 359. Magnitude QOI comparison for outpatient.

qoi_category		magnitude								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	107%	89%	95%	21%	19%	17%	102%	89%	94%
region2a	Seattle, WA	75%	nan%	80%	0%	nan%	21%	77%	nan%	53%
region2b	Portland, OR	52%	52%	30%	-12%	-13%	-12%	54%	53%	25%
region3b	State of VT	97%	93%	93%	25%	13%	30%	106%	101%	96%
region3c	Cherryland, MI	-69%	-68%	-73%	-84%	-84%	-79%	-69%	-68%	-75%
region4a	Washington DC	98%	94%	65%	6%	-17%	9%	108%	109%	61%
region4b	Chattanooga, TN	43%	40%	43%	-15%	-24%	-15%	40%	39%	37%
region4c	Tallahassee, FL	77%	70%	64%	-3%	-15%	-1%	87%	71%	55%
region4d	Horry County, SC	191%	146%	152%	73%	14%	47%	142%	166%	98%

Figure 360. Magnitude QOI error for outpatient.

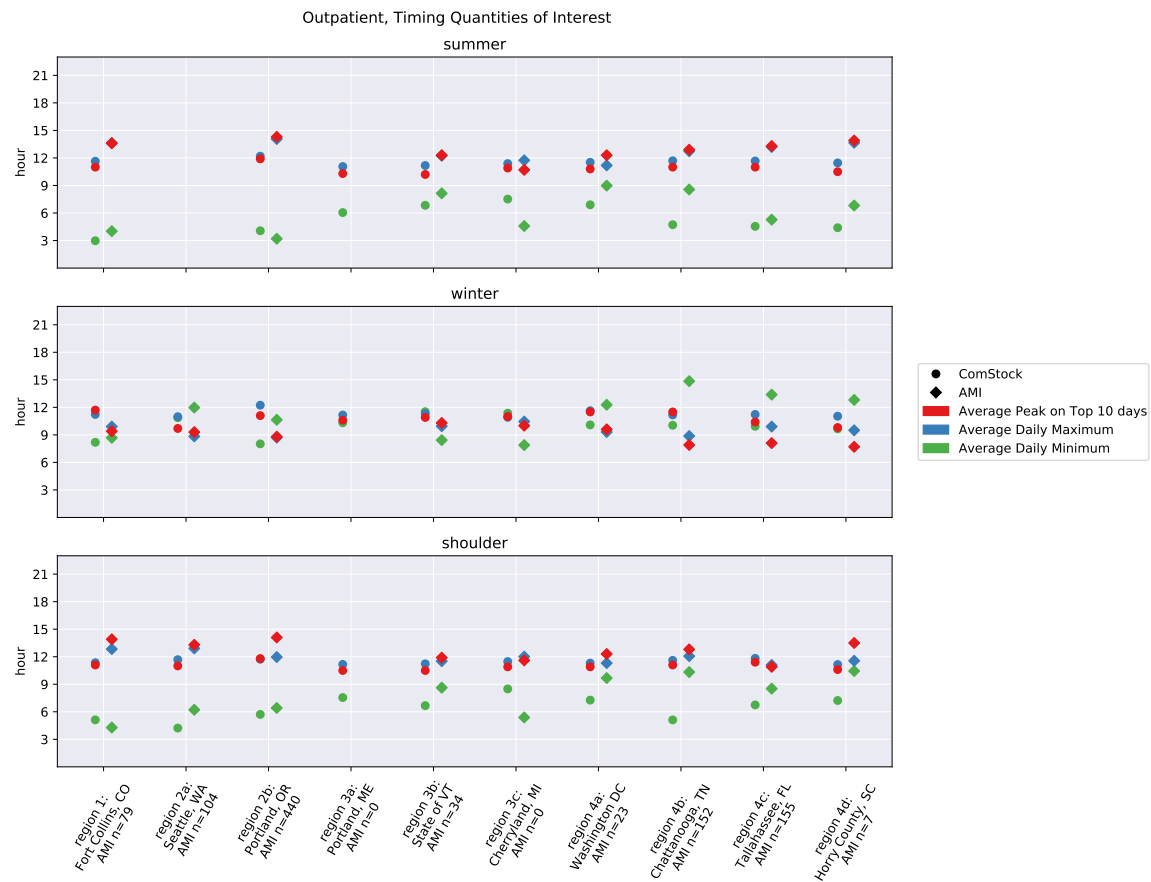


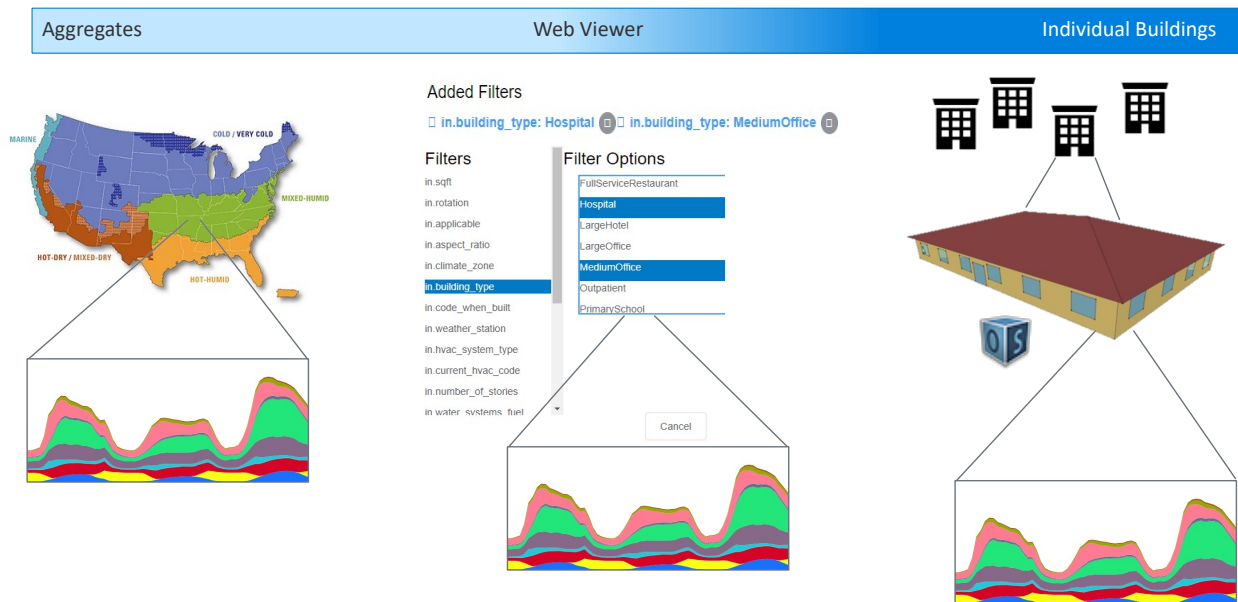
Figure 361. Timing QOI comparison for outpatient.

building_type		outpatient								
qoi_category		timing								
qoi_type		average_daily_maximum			average_daily_minimum			top10_peak_maximum		
season		shoulder	summer	winter	shoulder	summer	winter	shoulder	summer	winter
region	region_name									
region1	Fort Collins, CO	-1.5	-2.0	1.3	0.8	-1.0	-0.5	-2.8	-2.6	2.3
region2a	Seattle, WA	-1.2	nan	2.1	-2.0	nan	-1.1	-2.3	nan	0.4
region2b	Portland, OR	-0.2	-1.9	3.5	-0.7	0.9	-2.6	-2.3	-2.4	2.3
region3b	State of VT	-0.3	-1.1	1.4	-2.0	-1.3	3.1	-1.4	-2.1	0.6
region3c	Cherryland, MI	-0.5	-0.4	0.5	3.1	2.9	3.5	-0.7	0.2	1.0
region4a	Washington DC	0.0	0.3	2.3	-2.4	-2.1	-2.2	-1.4	-1.5	1.9
region4b	Chattanooga, TN	-0.5	-1.0	2.3	-5.2	-3.8	-4.8	-1.7	-1.9	3.6
region4c	Tallahassee, FL	0.7	-1.5	1.3	-1.8	-0.7	-3.5	0.5	-2.3	2.3
region4d	Horry County, SC	-0.4	-2.2	1.5	-3.2	-2.4	-3.2	-2.9	-3.4	2.1

Figure 362. Timing QOI error (hours) for outpatient.

### 4.3 Description of Published Dataset

At the most fundamental level, the dataset created by this project is simply the output of approximately 900,000 (550,000 ResStock plus 350,000 ComStock) building energy models. The output of each building energy model is one year of energy consumption at a 15-minute interval, separated by end-use category. However, this fundamental level is not appropriate for many of the use cases identified earlier in the project (Mims Frick et al. 2019). For this reason, we decided to take the dataset and transform it to make it accessible in three different ways, with the goal of meeting the needs of many users and use cases, as shown in Figure 363.



**Figure 363. One dataset, published in multiple ways. The project results have been published in three different ways to meet the needs of a variety of users and use cases.**

#### Aggregate data

- **Online:** see README.md file at OpenEI.org

For many of the use cases we identified, users wanted EULPs that reflect the sum or average of all buildings of a given type in a given geographic area. We refer to these as “aggregate” load profiles. The choice of building types reflects the building types represented natively by ResStock and ComStock. The preferred geographic breakdown depended on the use case. To support many different use cases, aggregates for the following geographic resolutions were created:

- 16 ASHRAE/IECC climate zones
- 5 DOE Building America climate zones
- 8 Electric System Independent System Operator (ISO) regions
- 2,400 U.S. Census Public Use Microdata Areas (PUMA)
- 3,000 U.S. counties

Each aggregate load profile (combination of building type and geographic area) is represented by a single comma separated value (CSV) file, which can be opened by common data analysis tools such as Microsoft Excel. There is a column for each end use and fuel type, and a row for each 15-minute time period. In addition to the timeseries data, each file also contains a list of identifiers for all of the building energy models included in the aggregate. These identifiers can be cross-referenced against a separate tab separated value (TSV) file which contains the characteristics (age, type, height, etc.) of each building energy model. In this manner it is possible to understand the building stock in the aggregate.

**One important note about county-level aggregates.** County-level aggregates are provided primarily to facilitate further aggregation to larger geographic areas such as utility service territories. ResStock and ComStock attempt to model buildings in proportion to the distribution found in reality. Heavily populated urban counties contain many buildings, so ResStock and ComStock include many models in these areas and the aggregate load profiles for these areas are robust. However, rural counties in sparsely populated areas do not contain many buildings. Thus ResStock and ComStock do not include many building energy models in these areas, and the aggregate load profiles for these areas are less robust. We **highly recommend** rolling up aggregates to include **a minimum of 1,000 models** to increase the robustness of the load profiles (see the *Uncertainty Due to Model Sample Size* section for further discussion).

#### *Web data viewers*

- **Online:** Residential building data viewer: [resstock.nrel.gov](https://resstock.nrel.gov)
- **Online:** Commercial building data viewer: [comstock.nrel.gov](https://comstock.nrel.gov)

For other use cases, users wanted to be able to quickly filter, slice, combine, visualize, and download the results in custom ways. Because of the size of the datasets, this is only possible using big data technologies and skill sets that put this out of reach for many users. To address this challenge, we developed online data viewers to handle the big data for users behind the scenes. These data viewers are available at [comstock.nrel.gov](https://comstock.nrel.gov) and [resstock.nrel.gov](https://resstock.nrel.gov). Figure 364 shows a screenshot of the data viewer.



Figure 364. An example screenshot from the comstock.nrel.gov data viewer

### Individual Building Results

- **Online:** see README.md file at OpenEI.org

As previously described, the raw dataset is a group of several hundred thousand files, each containing the outputs of an individual building energy model. Combined, these files add up to 20 terabytes. While processing these results using conventional desktop is impractical, there are several cloud service providers whose services make the required computing power and querying technology available to those with the technical skill set. Additionally, some users may have in-house access to advanced computing resources, or want to download just a small subset of individual building load profiles for their own custom use cases. To facilitate these use cases, the raw individual building results, along with the corresponding building characteristics, have been published to a publicly available website. They may be downloaded directly from this website, or may be queried in-place using big data technologies.

### Building Energy Models

- **Online:** see README.md file at OpenEI.org

The large number of building energy models used by ResStock and ComStock to represent the building stock makes it impractical for most organizations to run the full set themselves. However, some users expressed the desire to have access to the models for other use cases, such as modeling a smaller geographic area using a subset of the models or serving as a starting point for other modeling efforts. For these reasons, the individual building energy models have been published online. These models are available in the OpenStudio format, which can be converted to the EnergyPlus for simulation. The file of building characteristics allows users to easily identify buildings with a target set of properties to download and use. See the *Sample Size Recommendations for Users* section for recommendations on the number of models to use for representing subsets of the building stock.

## 4.4 Uncertainty Quantification Results

### 4.4.1 Uncertainty Propagation Results—Commercial

The detailed UQ results for ComStock are shown in the following tables. Table 25 and Table 26 summarize the uncertainty of energy-related QOIs, and the unit is percentage change of energy. For example, the uncertainty range of annual energy consumption of South Atlantic census division is  $[100\% - 2.9\%, 100\% + 2.9\%]$  of the ComStock estimated value. Table 27 and Table 28 summarize the uncertainty of timing QOIs, and the unit is change of time in hours. For example, the uncertainty range of summer daily peak time in the South Atlantic census division is  $[-0.66, +0.66]$  (equivalent to  $\pm 40$  minutes) to be added to its original value.

The uncertainty varies by region, building type, and QOI. We can see that the regions of West North Central, New England, East South Central, and Pacific have relatively large QOI uncertainty; similarly, the building types Secondary School, Primary School, Outpatient and Hospital, Large Office, and Large Hotel also have relatively large QOI uncertainty. In terms of comparison among seasons, summer is the season with the least uncertainty and winter is the season with the largest uncertainty. Additionally, the uncertainty of annual energy is smaller than the uncertainty by season.

**Table 25. Uncertainty (%) of ComStock Energy QOIs by Census Division**

Census Division	Annual	Summer Base	Summer Daily Peak	Summer Top 10 Peak	Winter Base	Winter Daily Peak	Winter Top 10 Peak	Shoulder Base	Shoulder Daily Peak	Shoulder Top 10 Peak
Atlantic	2.89	3.1	3.35	4.04	3.75	4.16	4.8	3.67	3.6	4.28
Mid-Atlantic	4.04	4.22	5.76	6.25	4.88	6.22	6.55	4.87	6.05	6.08
East North Central	6.09	4.64	8.47	8.19	4.75	8.61	8.43	5.62	9.2	8.97
Pacific	5.43	6.5	6.99	7.32	5.65	6.61	6.95	5.29	7.16	7.29
West South Central	3.88	3.32	4.25	4.42	3.54	4.19	4.83	3.27	4.17	4.39
West North Central	6.28	4.31	9.69	9.86	4.76	10.52	11.06	4.83	10.44	11.01
Mountain	4.28	3.18	6.27	6.61	3.53	6.66	7.56	3.21	6.36	6.75
New England	5.71	5.21	8.63	8.58	5.73	8.24	7.89	4.96	8.07	7.74
East South Central	3.52	3.88	4.17	4.61	4.26	4.73	5.54	4.25	4.4	4.78

**Table 26. Uncertainty (%) of ComStock Energy QOIs by Building Type**

Building Type	Annual	Summer Base	Summer Daily Peak	Summer Top 10 Peak	Winter Base	Winter Daily Peak	Winter Top 10 Peak	Shoulder Base	Shoulder Daily Peak	Shoulder Top 10 Peak
Full Service Restaurant	2.58	3.72	4.36	4.92	4.88	5.52	6.27	4.02	4.24	4.73
Strip Mall	4.84	4.71	6.59	6.64	5.23	7.25	7.34	5.6	7.56	7.82
Warehouse	3.84	5.03	4.83	5.39	4.51	4.67	5.14	4.96	5.02	5.13
Retail	3.52	4.37	5.78	6.36	5.04	6.4	7.5	3.92	5.9	6.46
Small Office	5.3	2.97	6.78	6.65	3.23	6.88	6.67	3.09	6.76	6.47
Primary School	8.03	5.53	9.32	9.85	6.12	8.69	9.59	6.07	9.66	10.2
Medium Office	5.36	2.9	7.15	7.7	3.44	7.49	8.31	3.62	7.1	8.17
Small Hotel	4.51	5.25	7.6	8.87	4.35	7.22	8.4	5.4	7.53	8.41
Secondary School	7.86	5.2	8.55	8.48	5.93	7.86	9.18	5.75	8.3	9.1
Outpatient	6.32	4.2	6.37	6.78	3.64	6.35	7.21	3.84	5.68	6.19
Quick Service Restaurant	5.03	4.31	5.18	5.33	5.02	5.68	5.79	5.45	5.44	5.93
Large Office	3.2	3.34	4.03	4.58	3.45	3.8	4.45	2.24	3.82	4.96
Large Hotel	3.66	3.37	4.43	4.62	3.39	5.14	6.34	4.02	4.82	5.44
Hospital	4.58	3.37	4.45	5.16	3.85	5.73	6.6	3.12	4.64	5.21

**Table 27. Uncertainty (hours) of ComStock Timing QOIs by Census Division**

Census Division	Summer Daily Peak Time	Summer Top10 Peak Time	Winter Daily Peak Time	Winter Top10 Peak Time	Shoulder Daily Peak Time	Shoulder Top10 Peak Time
South Atlantic	0.66	1.02	0.64	1.03	0.58	1.08
Mid-Atlantic	0.58	1.09	0.72	1.12	0.63	1.09
East North Central	0.66	1.19	0.7	1.14	0.66	1.14
Pacific	0.62	1.08	0.57	1.05	0.63	1.06
West South Central	0.53	1.01	0.58	0.94	0.58	1
West North Central	0.69	1.19	0.76	1.26	0.69	1.28
Mountain	0.63	1.13	0.73	1.21	0.64	1.2
New England	0.68	1.16	0.73	1.13	0.69	1.18
East South Central	0.69	1.06	0.66	1.08	0.68	1.13



**Table 28. Uncertainty (hours) of ComStock Timing QOIs by Building Type**

Building Type	Summer Daily Peak Time	Summer Top10 Peak Time	Winter Daily Peak Time	Winter Top10 Peak Time	Shoulder Daily Peak Time	Shoulder Top10 Peak Time
Full Service Restaurant	0.71	1.3	0.68	1.33	0.68	1.13
Strip Mall	0.58	1.02	0.63	1.07	0.64	1.15
Warehouse	0.57	1.12	0.68	1.03	0.6	1.07
Retail	0.59	0.97	0.66	1.04	0.61	1.08
Small Office	0.67	1.01	0.58	1	0.58	0.98
Primary School	0.78	1.42	0.8	1.3	0.79	1.31
Medium Office	0.79	1.31	0.8	1.29	0.75	1.35
Small Hotel	0.79	1.26	0.87	1.47	0.78	1.37
Secondary School	0.83	1.45	0.84	1.35	0.8	1.32
Outpatient	0.7	1.29	0.69	1.32	0.72	1.38
Quick Service Restaurant	0.79	1.48	0.81	1.23	0.74	1.27
Large Office	0.83	1.42	0.75	1.29	0.79	1.4
Large Hotel	0.82	1.35	0.74	1.36	0.81	1.37
Hospital	0.7	1.24	0.73	1.37	0.69	1.29

#### 4.4.2 Uncertainty Propagation Results—Residential

The detailed UQ results for ResStock runs are presented in the following tables. Table 29 summarizes uncertainty ranges of energy magnitude by census division and Table 30 by building type. The unit is percentage change of energy. For example, the uncertainty range of annual energy consumption of Single Family Detached Homes is  $[100\% \pm 4.2\%]$  of the ResStock estimated value. Table 31 and Table 32 summarize uncertainty of the timing QOIs, and the unit is change of time in hours. For example, the uncertainty range of summer daily peak time in the Single Family Detached Homes is  $[-0.9, +0.9]$  (equivalent to  $\pm 54$  minutes).

**Table 29. Uncertainty (%) of ResStock Energy QOIs by Census Division**

Census Division	Annual	Summer Base	Summer Daily Peak	Summer Top 10 Peak	Winter Base	Winter Daily Peak	Winter Top 10 Peak	Shoulder Base	Shoulder Daily Peak	Shoulder Top 10 Peak
South Atlantic	2.6	1.6	4.5	4.2	1.6	4.7	4.2	1.4	4.3	4.7
Mid-Atlantic	4.8	2.0	7.9	8.0	2.3	7.9	7.8	2.0	8.2	8.5
East North Central	5.6	2.4	5.7	5.4	2.8	5.6	5.1	2.6	5.9	6.1
Pacific	2.4	1.3	5.7	4.7	1.4	6.3	3.5	1.5	5.5	4.5
West South Central	3.4	1.5	3.1	3.5	1.6	3.0	3.2	1.6	3.3	3.7
West North Central	5.5	2.8	6.9	6.4	3.4	8.1	7.5	2.9	6.8	6.6
Mountain	4.2	1.8	4.8	5.9	1.5	5.1	6.1	1.6	4.9	5.9
New England	4.8	2.0	7.1	9.0	1.8	6.9	8.9	1.6	6.8	8.6
East South Central	3.2	1.2	3.4	3.8	1.1	3.4	4.1	1.2	3.7	4.3

**Table 30. Uncertainty (%) of ResStock Energy QOIs by Building Type**

Building Type	Annual	Summer Base	Summer Daily Peak	Summer Top 10 Peak	Winter Base	Winter Daily Peak	Winter Top 10 Peak	Shoulder Base	Shoulder Daily Peak	Shoulder Top 10 Peak
Single-Family Detached	4.2	1.7	4.8	4.8	1.9	4.8	4.2	1.7	4.7	5.0
Single-Family Attached	4.0	1.9	7.2	7.3	1.9	7.2	7.1	2.1	7.1	7.4
Mobile Home	4.1	2.6	4.9	4.6	2.4	4.8	4.4	2.4	4.2	4.1
Multifamily, 2–4 Units	3.4	2.4	7.1	8.3	2.7	7.8	9.0	2.3	7.8	9.0
Multifamily, 5+ Units	3.0	2.0	6.1	5.7	1.9	6.7	6.4	2.0	6.7	6.5

**Table 31. Uncertainty (hours) of ResStock Timing QOIs by Census Division**

Census Division	Summer Daily Peak Time	Summer Top10 Peak Time	Winter Daily Peak Time	Winter Top10 Peak Time	Shoulder Daily Peak Time	Shoulder Top10 Peak Time
South Atlantic	0.7	1.1	0.8	1.3	0.8	1.4
Mid-Atlantic	0.9	1.6	0.9	1.6	0.9	1.4
East North Central	0.7	1.3	0.6	1.1	0.7	1.2
Pacific	0.7	1.2	0.7	1.3	0.7	1.1
West South Central	0.8	1.4	0.7	1.2	0.8	1.5
West North Central	0.8	1.4	0.9	1.3	0.7	1.4
Mountain	0.7	1.4	0.8	1.4	0.8	1.5
New England	0.7	1.2	0.7	1.3	0.7	1.2
East South Central	0.7	1.3	0.8	1.3	0.8	1.4

**Table 32. Uncertainty (hours) of ResStock Timing QOIs by Building Type**

Building Type	Summer Daily Peak Time	Summer Top 10 Peak Time	Winter Daily Peak Time	Winter Top 10 Peak Time	Shoulder Daily Peak Time	Shoulder Top 10 Peak Time
Single-Family Detached	0.8	1.4	0.8	1.4	0.8	1.4
Single-Family Attached	0.6	1.0	0.7	1.1	0.6	1.0
Mobile Home	0.6	1.1	0.7	1.1	0.7	1.1
Multifamily, 2–4 Units	0.7	1.3	0.7	1.2	0.7	1.1
Multifamily, 5+ Units	0.7	1.1	0.8	1.2	0.8	1.2

## 5 Discussion and Conclusions

### 5.1 Interpreting and Using Results

A forthcoming companion report, *End-Use Load Profiles for the U.S. Building Stock: Applications and Opportunities*, provides—for a general audience—a detailed look at how the EULP dataset can be used for a number of example applications. Thus, this report does not include discussion of all the opportunities for applying the data and considerations for those applications. In this section, we summarize the ways that users can understand the accuracy and uncertainty of the results, to inform the level of confidence they should have when applying the data.

#### 5.1.1 How Accurate Are the Results?

In Section 4, we presented extensive comparisons between the ResStock and ComStock end-use load profile outputs and all of the empirical data sources we had access to. The comparisons focused on average 24-hour profiles and QOIs, such as peak magnitude and timing, that are most relevant for the use cases identified in Mims Frick et al. (2019).

In general, the accuracy of results is highly dependent on the sector, building type, and validation dataset. Accuracy also depends on whether users are applying just the load shape, or both the shape and magnitude in their application. We recommend that users of data review the relevant results in Section 4 to understand the known accuracy and uncertainty of the data they wish to apply.

#### Uncertainty in Empirical Data

One important point to remember when reviewing these comparisons is that there can be a large degree of uncertainty in the empirical data that outputs are compared to. As discussed in detail in *Biases in Commercial AMI Data*, a number of factors contributed to relatively large uncertainty in the magnitude of the commercial AMI EUI profiles by building type. We have greater confidence in the profile *shapes* and thus focus on comparing normalized shapes of AMI profiles to model outputs (e.g., Figure 220), and rely on EIA and CBECS data for EUI magnitude comparisons. For the commercial AMI comparisons, we represent AMI data uncertainty by plotting the lower and upper bounds of the 80% confidence interval of the mean at each timestep. Building types with large numbers of samples (like small offices) have narrower confidence intervals, and types with fewer samples (like hospitals) have wider confidence intervals. There is less uncertainty in the annual CBECS and monthly EIA data, although the CBECS data are from 2012 and the EIA data have large uncertainty in how commercial and industrial customers are differentiated by utilities when submitting data to EIA. The format of the commercial end-use datasets (e.g., BAS data) and because they were used to derive model inputs such as schedules, meant that the end-use data were not well-suited for use in evaluating final model accuracy.

Averages and sums of residential AMI data typically have low uncertainty because (1) there is usually a large enough sample (or entire population) of customer meters, (2) disaggregation into building use types is not necessary, and (3) meters typically correspond to dwelling units. However, if it is not a complete sample, then residential AMI data may have some sampling bias. We found that there was geographic sampling bias in the ComEd AMI data because their AMI meters were still being rolled out in the years for which we had data (2016 and 2017). Thus, we switched to using ComEd LRD, which was known to be a statistical representation of customers, for most of our comparisons. Residential AMI data can also have bias due to erroneous datapoints and the cleaning process used to handle those erroneous data. For the residential LRD comparisons in Section 4.1.3, we typically do not know the sample size, and assume  $\pm 10\%$  for the purpose of the comparison figures, based on conversations with our TAG on typically LRD sample sizes and standard deviations. Like LRD, residential end-use datasets have much smaller sample sizes than AMI data. When comparing residential end-use datasets to each other (Section 2.3.9) and to our model results (Section 3.2.1), we similarly indicate uncertainty using confidence intervals based on sample sizes.

#### 5.1.2 How Much Did Accuracy Improve Through Calibration?

Because we did not have access to most of the validation datasets when calibration began, we do not have a full picture of how much accuracy improved through calibration. For residential, we did track the progress of QOIs over time; the figures presented in Section 4.1.5 provide insights into how QOI error changes over time. Notably, some model changes caused QOI errors to worsen. In some cases this was because a bug was introduced, but in most cases, we believe it meant that the model had been accurate for the wrong reason, and prompted us to continue refining the model.

### **Residential EIA Accuracy Improvement**

For residential, we had access to some EIA and LRD when calibration began, so we can examine those comparisons before and after the calibration effort. The comparisons of ResStock to EIA-861 residential electricity sales by utility and state and EIA-176 residential gas sales by state are shown before and after calibration in Figure 365. Before calibration, ResStock was overestimating annual utility-scale electricity sales with the slope of the linear fit at 1.32 and an  $R^2$  of 0.95, with significant errors for several utilities including Entergy Texas, Pacific Gas & Electric, and Southern California Edison. After calibration, including the effect of the output correction model, the slope was reduced to 1.05, and  $R^2$  improved to 0.99 with all major outliers corrected. Before calibration ResStock had significant errors in state electric retail sales for some states. After calibration, ResStock significantly increased the  $R^2$  from 0.92 to 0.99. Similarly, the estimation of residential natural gas by state was improved from an underestimation with slope of 0.75 to 0.92; the  $R^2$  improved from 0.95 to 0.96. A correction model has not been developed for natural gas, though this is planned future work.

### **Residential LRD Accuracy Improvement**

There were four residential LRD regions that were present in both the original and final sets of LRD (see Section 2.3.6). Note that the before and after comparison years are different, so the LRD reflects changes in weather and number of customers. Looking at the before and after comparisons for these four LRD regions (see Figure 366) is quite useful, because it can tell us how much accuracy improved in regions that were not a focus on the regional calibration efforts. For example, before calibration, ResStock overestimated peaks for Baltimore Gas and Electric (BGE) in Maryland by 68% (summer) and 61% (winter). After calibration, the overestimation was reduced to 10% and 22%, respectively (keeping in mind the uncertainty present in LRD). The improvement for ERCOT, covering most of Texas, is similar. The summer peak error in PG&E California did not see as much improvement, remaining at 45% after calibration, though the absolute error was significantly reduced. It is notable how much the shape of winter load improved through calibration across all four locations, with the peaks becoming much more diversified.

### **Residential End-Use Dataset Accuracy Improvement**

We did not perform comprehensive before and after comparisons against the residential end-use datasets. We show comparisons between pre-calibration ResStock output and the end-use datasets for occupant-driven loads (Figures 15–24). The validation of the new residential stochastic occupant behavior model shows the improvement resulting from that change, which is primarily visible in smoothing out the pre-calibration spikiness caused by coincident schedules (Figure 45). Before and after comparisons in Appendix E show these improvements in variation for all major occupant-driven end uses and hot water draws.

### **Commercial Accuracy Improvement**

For commercial, there was not as much data available for comparison before calibration began, and the empirical data we compared against changed over time as we iterated on meter metadata matching, misclassification, and outlier removal, which makes before/after comparisons less meaningful. We can show one example of commercial before/after comparisons for the Fort Collins AMI data. Figure 367 shows the before and after comparisons for two building types: full service restaurant and small offices. Before calibration, ComStock overestimated peaks for full service restaurants by 30% (summer) and 80% (winter), relative to the middle of the AMI data's 80% confidence interval. After calibration, the overestimation was reduced to 10% and 30%, respectively. The improvement in small offices load was more modest, with 40% (summer) and 30% (winter) underestimates reduced to 30% and 15% underestimates, respectively.

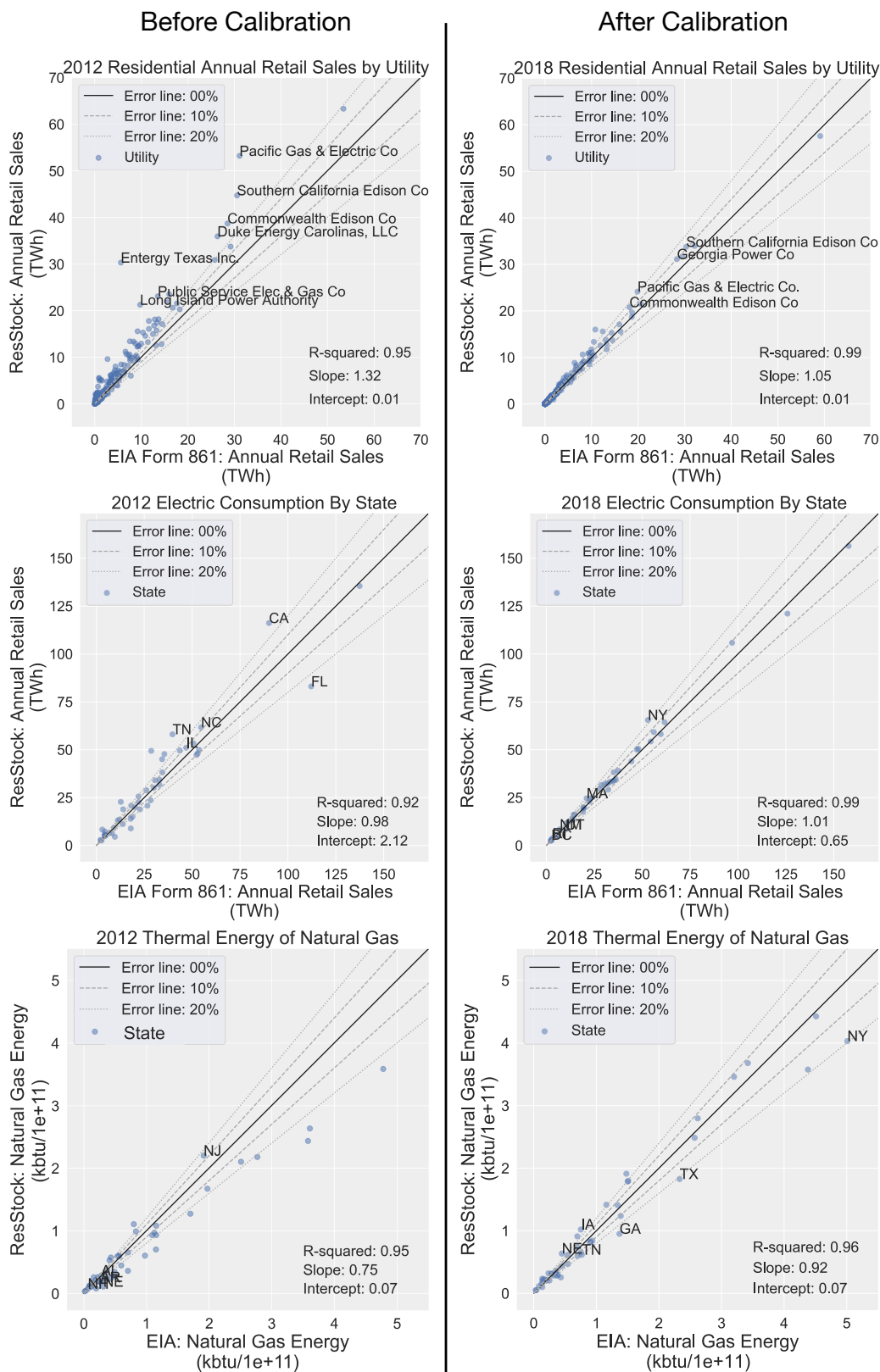


Figure 365. Comparisons of ResStock annual results to EIA Form 861 utility retail sales (top), EIA Form 861 retail sales by state (middle), and natural gas energy by state (bottom) before and after calibration

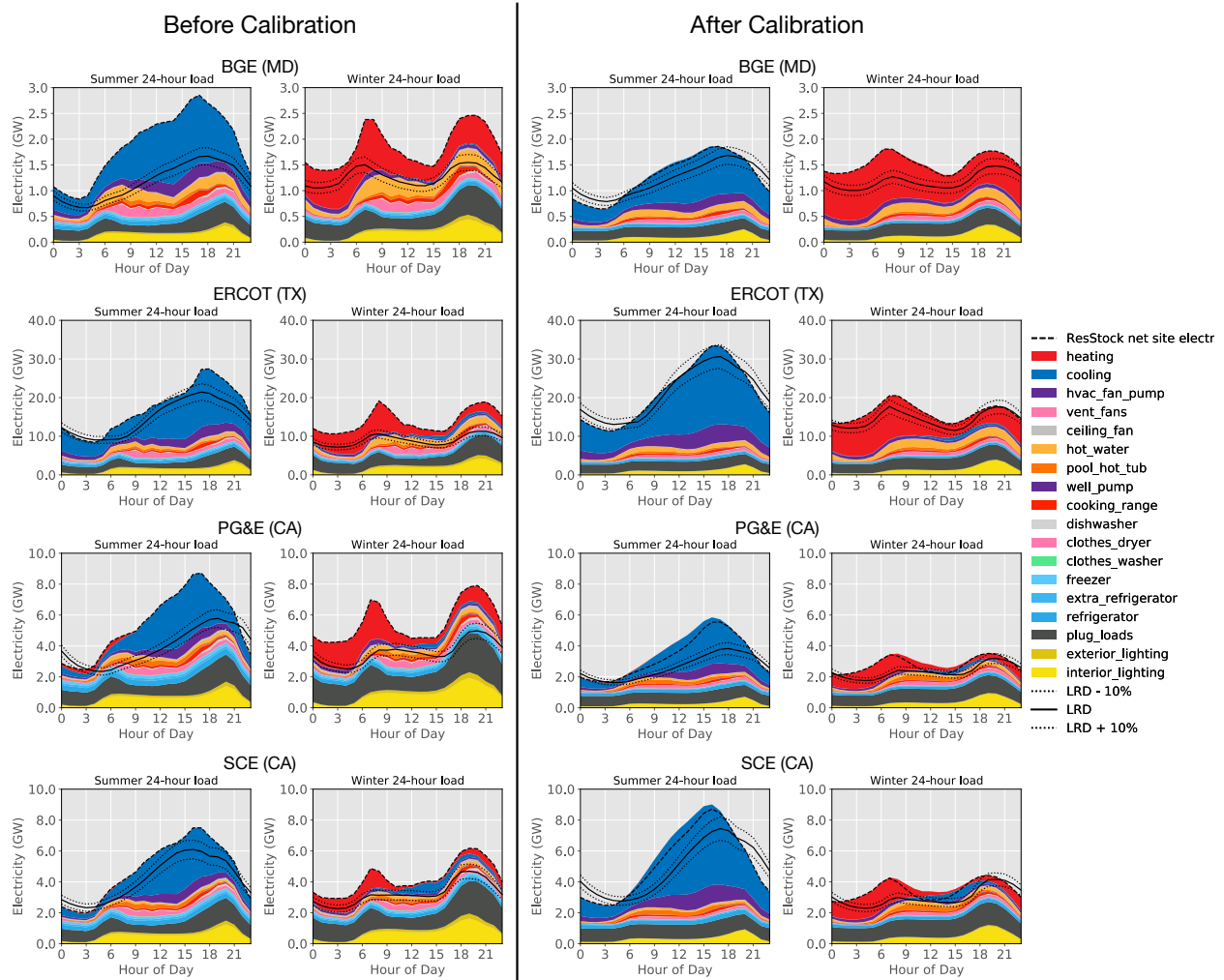


Figure 366. Before and after calibration: ResStock comparisons to LRD average daily profiles

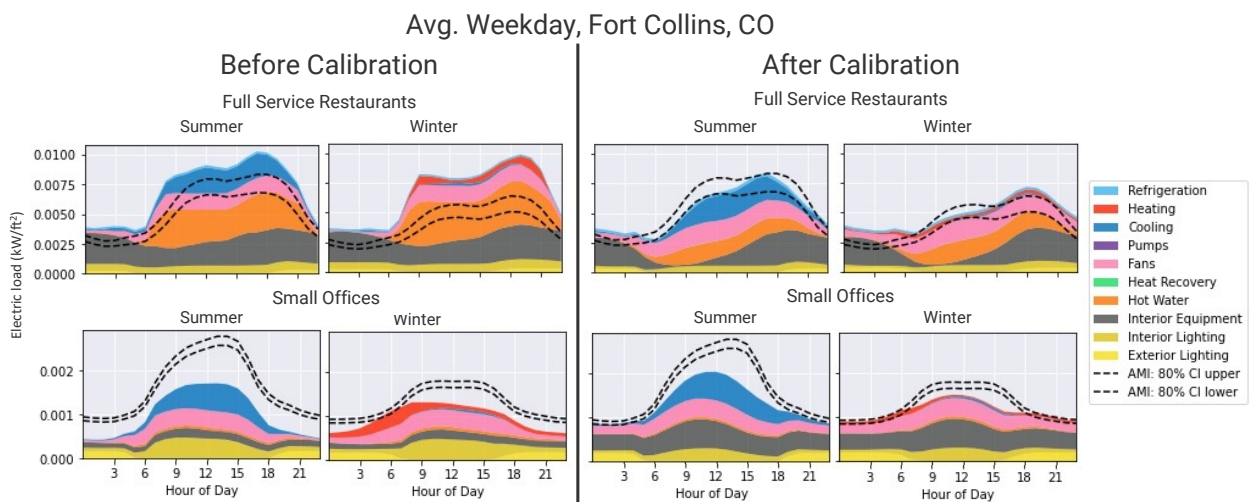


Figure 367. Before and after calibration: ComStock comparisons to Fort Collins AMI average daily profiles



### 5.1.3 How Much Uncertainty Is There in the Results?

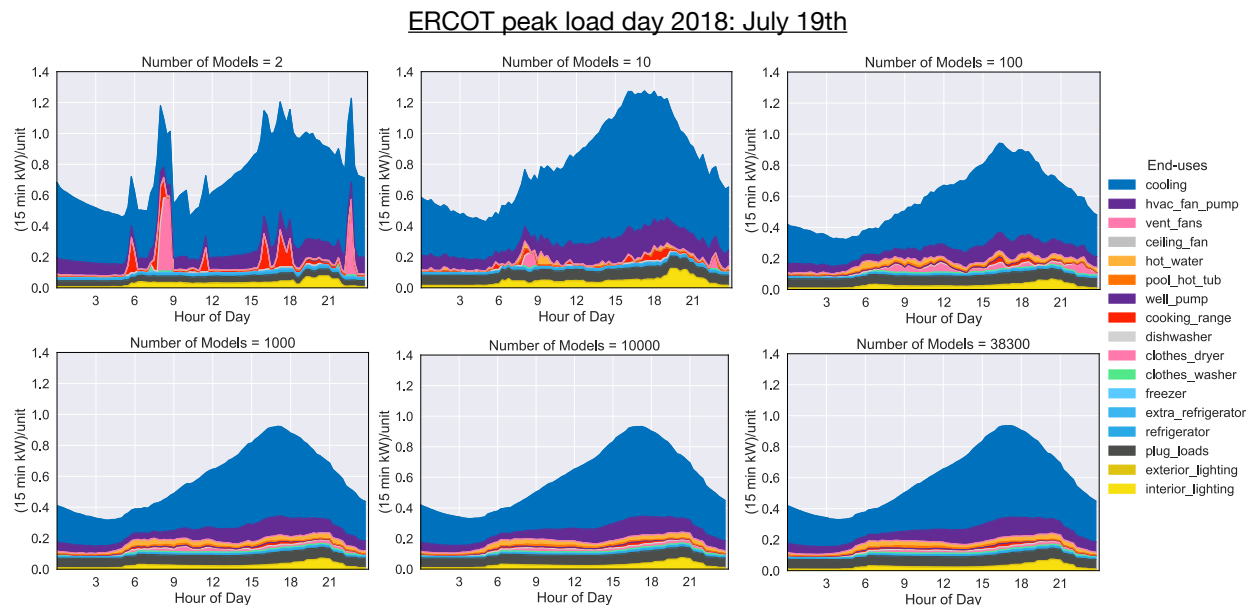
In addition to validation accuracy, dataset users should understand the uncertainty of the EULP data produced by the ResStock and ComStock models. As introduced in Section 2.5, we believe that the only significant sources of uncertainty are (1) the uncertainty in stock-level model input parameters, and (2) uncertainty due to insufficient number of ResStock or ComStock samples.

#### *Uncertainty in Model Input Parameters*

Our approach to propagating input parameter uncertainty ranges is covered in Section 2.5. The results of this process are presented in Section 4.4. In general, the uncertainty in QOIs is relatively small, compared to the uncertainty in the empirical validation data. For example, the uncertainty in seasonal or top 10 day average peak magnitude is typically in the 3–9% range for ResStock and 4–11% for ComStock, depending on the season, building type, and region. The daily minimum base load magnitude has slightly lower uncertainty, with values in the 1–4% range for ResStock, and 3–6% for ComStock. Annual energy uncertainty is in the 3–6% range for ResStock and 3–8% for ComStock. Peak time uncertainty is typically 45–90 minutes for ResStock and 30–90 minutes for ComStock, depending on the season, building type, and region.

#### *Uncertainty Due to Model Sample Size*

Both ComStock and ResStock represent the commercial and residential building stocks with conditionally dependent distributions of various building characteristics. The timeseries loads are produced by sampling from this probabilistic space of building characteristics. A single sample or a handful of samples are not enough to fully describe the diversity of building characteristics and occupant behavior in a given location. Figure 368 shows how the number of samples affects the modeled aggregate 15-minute load of the residential building stock on a peak load day in the ERCOT region (Texas). With only two sampled models, the occupant driven loads from cooking and clothes drying dominate the load shape. As the number of models increases, the stochastic nature of the occupant driven loads eventually smooths out, until there is almost no visible difference between 10,000 samples and the 38,300 samples (the number of models typically run during a national scale ResStock simulation). As a result, a sufficient number of models are needed to accurately estimate the timeseries end-use loads from ResStock and ComStock.



**Figure 368. Visual convergence of ERCOT end-use loads by the number of models used to construct the end-use timeseries**

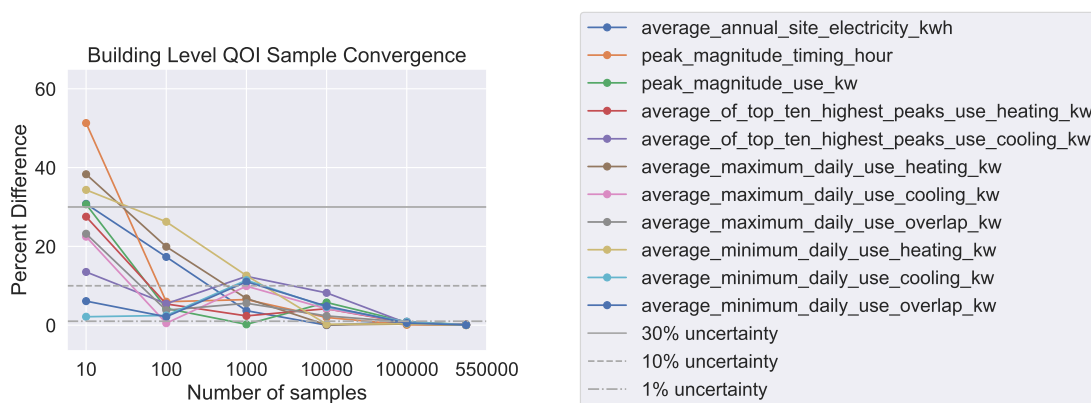
#### *Sample Size Recommendations for Users*

Convergence analysis is a technique to quantify the accuracy of estimating a QOI with a given number of samples. The same problem is solved with increasing number of samples and the QOI is calculated for each set of samples. Figure 369 shows how the various QOIs vary with increasing sample size, with the percent error calculated relative to the QOI with 1,000,000 samples. With smaller sample sizes, any of the QOIs can have a large deviation from



the expected converged QOI values. As the number of samples increases, the deviation from the expected estimate decreases. The rate at which this uncertainty due to sample size decreases is based on the sampling method used to sample the conditionally dependent distributions that characterize the building stock. The sampling method for ComStock uses Sobol' sequences, while ResStock uses deterministic quota-based sampling with random assignment of non-correlated parameters. Both these methods strategically choose the samples to obtain a better sample of the probabilistic space compared to simple random sampling (Monte Carlo sampling).

From Figure 369, 1,000 samples result in about 15% uncertainty or less for all the QOIs. At about 10,000 samples the uncertainty drops to around 10%. When interacting with the published data of this project, we recommend a minimum of 1,000 samples be used to reduce uncertainty due to sample size. If there are less than 1,000 samples in a query of the data or downloaded in an aggregate timeseries data file, then we recommend combining geographies nearby the area of interest to increase the sample size until the minimum recommended number of samples is obtained. Sampling uncertainty for annual site energy is less than 15% across all sample sizes, so if only annual (as opposed to peak) metrics are of interest, then fewer samples may be necessary.



**Figure 369. Sample convergence plot for the magnitude building level QOIs. The percent error relative to result with 1,000,000 samples is shown. Using 1,000 or more samples ensures that sampling uncertainty is less than 15% across all metrics.**

#### 5.1.4 How Does the EULP Dataset Compare to Previously Available EULPs?

As discussed in Section 2.1 of Mims Frick et al. (2019), most existing EULP sources are limited to one region and most do not cover all building types and end uses. Here we discuss how the EULP dataset presented in this report compares to two publicly available and widely used sources with national coverage.

##### EPRI Load Shape Library

EPRI conducts a wide range of load research on behalf of their members. For years, they have hosted a publicly available load data repository of end-use load profiles (*EPRI Load Shape Library* 2021). This Load Shape Library (LSL)<sup>TM</sup> 8.0 collects end-use, whole-premise, and technology measure load shapes from a variety of sources dating back to 2000, and makes them available to users via a web interface. Some data are only available to EPRI members. The publicly available end-use load shapes on the Load Shape Library 8.0 are quite different from the EULP dataset because they are only shapes and do not include magnitudes or building weights. Additionally, they are not 8,760 hourly profiles, but rather 24-hour profiles for six day types. The documentation states that “Users should treat the LSL data as a sample reference. Confidence and precision levels of the data are unknown.”

##### Commercial and Residential Hourly Load Profiles for all TMY3 Locations in the United States

This widely used dataset hosted on OpenEI.org (2021) was created around 2012 as a byproduct of various analyses of solar photovoltaics and solar water heating (Ong, Campbell, and Clark 2012; Ong et al. 2013). The commercial load profiles are simply the 16 ASHRAE 90.1-2004 DOE Commercial Prototype Models simulated in all (approximately 1,000) TMY3 locations, with building insulation levels changing based on ASHRAE 90.1-2004 requirements in each climate zone. The residential load profiles are simply five EnergyPlus models (one per climate region) representing 2009 IECC construction single-family detached homes simulated in all TMY3 locations.

The biggest limitation of both the commercial and residential datasets is that they are limited to one construction vintage that does not represent the overall existing building stock. They also lack any diversity in building characteristics and operation. The residential dataset includes a HIGH building load profile that was intended to provide a rough approximation of older home vintages, but it combines poor thermal insulation with larger house size, tighter thermostat setpoints, and less efficient HVAC equipment. Conversely, the LOW building combines excellent thermal insulation with smaller house size, wider thermostat setpoints, and more efficient HVAC equipment. Therefore, they likely represent very small fractions of the housing stock. Another limitation of the residential dataset is that in each of the five climate zones, there is only one type of heating and one type of water heating (e.g., all-electric homes are in the hot-humid zone only). Finally, the profiles do not include building counts or weights, so they cannot be used to understand aggregate load impacts.

## 5.2 Conclusion

This report documents the technical work completed through the three-year *End-Use Load Profiles for the U.S. Building Stock* project funded by the U.S. Department of Energy (DOE) in fiscal years 2019–2021. The work culminated in the release of a publicly available dataset of simulated end-use load profiles (EULPs) for the stock of residential and commercial buildings in the contiguous United States.

Historical and contemporary efforts to develop EULPs for particular regions have used direct submetering of a statistically representative sample of buildings. Applying such an approach to the contiguous United States would have cost an order of magnitude higher than the already significant budget for this project; would have likely been limited in coverage of building types and end uses; and, most importantly, would not have resulted in calibrated models that enable future what-if analyses of scenarios involving energy efficiency, electrification, demand flexibility, or changes in climate. Our hybrid approach—combining a wide range of empirical data with detailed physics-simulation building stock models—produced EULPs covering all major commercial and residential building types and end uses, for all locations of the contiguous United States, and the calibrated building models enable what-if scenario analyses. At the same time, our national approach would not have been possible without the foresight of organizations that have invested in regional end-use submetering load research. Our work was also made possible by the ratepayer- and taxpayer-funded investments in advanced metering infrastructure over the past decade. We have fully documented our hybrid approach methodology in Section 2 and model calibration updates in Section 3.

Although our hybrid approach was similar to some previous examples that used tens or hundreds of building energy models (Itron 2006; Baroiant et al. 2019), the scale of our application was unprecedented, both in terms of empirical data gathered—2.3 million meters worth of hourly data from 11 utilities—and in terms of the granularity of building stock simulation—900,000 building energy models representing 58 billion ft<sup>2</sup> of commercial buildings and 133 million residential dwelling units.

Beyond the scale of application, we developed novel approaches for both the sensitivity analysis and uncertainty quantification of the ComStock and ResStock models (Sections 3.1 and 2.5). The sensitivity analyses involved random forest regressions trained on the inputs and outputs of the high-granularity building stock models. This novel approach to sensitivity analysis—which did not require additional model runs—was possible because with thousands of samples, the stock models already cover much of the variation of their parameter spaces. In contrast, our uncertainty quantification methodology required thousands of model runs to propagate uncertainty in model inputs through to model outputs. To accomplish this, we developed timeseries surrogate models for the EnergyPlus simulations and applied them at an unprecedented scale using high-performance computing, so that millions of model realizations could be conducted to evaluate uncertainty for the quantities of interest. Although model input uncertainty is important, one conclusion from this project is that there can be just as much uncertainty in “ground truth” load data as in modeled results, whether that is because of sampling bias (Figure 175), or because of uncertainties in building type and floor area metadata (*Biases in Commercial AMI Data*).

The final product of the project is calibrated and validated 15-minute resolution load profiles for all major residential and commercial building types and end uses, across all climate regions in the United States. It is worth emphasizing that “validated” is not the colloquial meaning that suggests the outputs perfectly match the available empirical data; it means that model accuracy was evaluated for the quantities of interest and reported so that EULP data users know what level of confidence they should have when putting them to use. Similarly, “calibrated” does not mean that model error was minimized through automated or manual tuning of inputs, which is prone to overfitting (in

other words, getting the right answer for the wrong reason). Instead, the objective of calibration was to make model improvements that reduce model error, but only when supported by data. Model improvements were typically in the form of increased accuracy, diversity, or resolution of input parameter distributions, as thoroughly documented in Section 3. We did develop a data-driven residential output correction model to reduce some of the remaining model error, as documented in Section 3.2.10.

The resulting version 1.0 EULP dataset is available in three formats—(1) via a web viewer, (2) as downloadable spreadsheet files, and (3) in a detailed format that can be queried with big data tools—to maximize accessibility for different types of users. We also made the OpenStudio model input files available for building energy modelers to use for their own analyses. Utility planners, consultants, regulators, state energy offices, researchers, and building owners are now able to use these resources, along with tools such as Berkeley Lab’s forthcoming Time-Sensitive Value Calculator, to estimate the value of energy efficiency, demand response, and other distributed energy resources. Such analysis can be used to guide utility resource and distribution system planning, research and development prioritization, and state and local energy planning and regulation.

The EULP dataset can be updated over time by incorporating new years of input data (weather, census, real estate, time use, EIA surveys, and so on) into the models as they become available, giving the load models longevity beyond this calibration effort as well as longevity beyond EULPs developed through a pure end-use submetering approach. Model outputs can continue to be compared to available empirical data, such as monthly EIA sales data and load research data, to understand how well the models perform in future years. Additionally, the calibrated models can be a foundation to develop *end-use savings shapes* that describe the difference in energy consumption between a baseline building and a building with an energy saving or flexibility measure applied.

Our nation, states, cities, and utilities are embarking on an ambitious transition to a 100% clean energy economy by 2050. The EULP dataset and calibrated building stock models can play a key role by helping us understand, with more accuracy than ever before, how buildings and their occupants interact with the national electricity system, and the role that high-performing, energy-efficient, and demand-flexible buildings can play in an equitable transition to a decarbonized, affordable, and reliable energy system.

## References

- Abramson, D. L., Barry; Lung-Sing Wong; Herman. 2006. “Service Life Data from an Interactive Web-Based Owning and Operating Cost Database”. *ASHRAE Transactions* 112, no. 1 (): 81–92.
- ASHRAE. 2014. *ASHRAE Guideline 14–2014, Measurement of Energy, Demand, and Water Savings*.
- . 2021. *ASHRAE Owning and Operating Cost Database*. <http://weblegacy.ashrae.org/publicdatabase/summary.asp>. Accessed: 2021-09-05. ASHRAE.
- Barbour, E., R. Ciraulo, V. Nubbe, S. Robinson, and J. Stanley. Expected publication 2021. *NFRC Commercial Fenestration Market Study*. Tech. rep. Guidehouse Inc.
- Baroiant, S., J. Barnes, D. Chapman, S. Keates, and J. Phung. 2019. *California Investor-Owned Utility Electricity Load Shapes: Final Project Report*. Tech. rep. Prepared by ADM Associates, Inc. for the California Energy Commission.
- Bertagnolio, S. 2012. “Evidence-Based Model Calibration For Efficient Building Energy Services”. PhD thesis, University of Liège.
- Bianchi, C., L. Zhang, D. Goldwasser, A. Parker, and H. Horsey. 2020. “Modeling occupancy-driven building loads for large and diversified building stocks through the use of parametric schedules”. *Applied Energy* 276:115470.
- Bickel, S., E. Phan-Gruber, and S. Christie. 2013. *Residential Windows and Window Coverings: A Detailed View of the Installed Base and User Behavior*. Tech. rep. Prepared for Building Technologies Office, U.S. Department of Energy, by D&R International.
- Big Ladder Software, L. 2021. *EnergyPlus weather file (EPW) data dictionary*. Version 9.5. <https://bigladdersoftware.com/epx/docs/9-5/auxiliary-programs/energyplus-weather-file-epw-data-dictionary.html>.
- Breiman, L. 2011. “Random Forests”. *Machine Learning* 45 (1): 5–32. ISSN: 1573-0565. doi:<https://doi.org/10.1023/A:1010933404324>. <https://link.springer.com/article/10.1023/A:1010933404324>.
- Buccitelli, N., C. Elliott, S. Schober, and M. Yamada. 2017. “2015 U.S. Lighting Market Characterization”. *Washington DC*.
- Building America Partnership for Improved Residential Construction. 2021. *Phased Deep Retrofits Data Portal*. <http://www.infomonitors.com/pdr/>. Accessed: 2021-08-26.
- Building Performance Database (BPD)*. <https://buildings.lbl.gov/cbs/bpd>. Accessed: 2021-08-26.
- Butzbaugh, J., L. Sandahl, and M. Baechler. 2017. “US HPWH Market Transformation: where we’ve been and where to go next”. *Energy Efficiency in Domestic Appliances and Lighting (EEDAL’17)*: 922.
- Chan, W. R., J. Joh, and M. H. Sherman. 2012. *Air Leakage of US Homes: Regression Analysis and Improvements from Retrofit*. Tech. rep. Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States).
- Chaudhary, G., J. New, J. Sanyal, P. Im, Z. O’Neill, and V. Garg. 2016. “Evaluation of “Autotune” calibration against manual calibration of building energy models”. *Applied Energy* 182:115–134. ISSN: 0306-2619. doi:<https://doi.org/10.1016/j.apenergy.2016.08.073>. <https://www.sciencedirect.com/science/article/pii/S030626191631159X>.
- Chen, J., R. Adhikari, E. J. H. Wilson, J. Robertson, B. Polly, and O. Olawale. 2022. “Stochastic simulation of occupant-driven energy use in a bottom-up residential building stock model”. In preparation.
- Chen, J., B. Polly, and E. J. H. Wilson. 2022. “A Review of Bottom-Up Stochastic Simulation Approaches for Residential Building Electricity Consumption”. In preparation.
- Christensen, C, S Horowitz, T Givler, A Courtney, and G Barker. 2005. “BEopt: Software for Identifying Optimal Building Designs on the Path to Zero Net Energy; Preprint” (). <https://www.osti.gov/biblio/15016068>.
- Coakley, D., P. Raftery, and M. Keane. 2014. “A review of methods to match building energy simulation models to measured data”. *Renewable and Sustainable Energy Reviews* 37 (Supplement C): 123 –141. ISSN: 1364-0321. doi:<https://doi.org/10.1016/j.rser.2014.05.007>. <http://www.sciencedirect.com/science/article/pii/S1364032114003232>.
- Cochran, J., P. Denholm, M. Mooney, D. Steinberg, E. Hale, G. Heath, B. Palmintier, B. Sigrin, D. Keyser, D. McCamey, et al. 2021. *The Los Angeles 100% Renewable Energy Study (LA100)*. Tech. rep. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Code Status Maps: Commercial Energy Code Adoption*. <http://bcapcodes.org/code-status/>. Building Codes Assistance Project.
- ComEd. 2018. *ComEd Anonymous Data Service*. <https://www.comed.com/SmartEnergy/InnovationTechnology/pages/anonymousdataservice>.
- Commercial and Residential Hourly Load Profiles for all TMY3 Locations in the United States*. 2021. <https://data.openepi.org/submissions/153>.

- Commercial Demand Module of the National Energy Modeling System: Model Documentation*. 2017. Tech. rep. U.S. Energy Information Administration.
- Crawley, D., et al. 2001. “EnergyPlus: Creating a New-Generation Building Energy Simulation Program”. *Energy and Buildings* 33 (): 319–331. doi:10.1016/S0378-7788(00)00114-6.
- Culp, T., S. Widder, and K. Cort. 2015. *Thermal and Optical Properties of Low-E Storm Windows and Panels*. Tech. rep. Pacific Northwest National Laboratory.
- Cutler, D., J. Winkler, N. Kruis, C. Christensen, and M. Brandemuehl. 2013. “Improved Modeling of Residential Air Conditioners and Heat Pumps for Energy Calculations” (). doi:10.2172/1219902. <https://www.osti.gov/biblio/1219902>.
- Dahlhausen, M. 2020. “Determining Measurement Requirements for Whole Building Energy Model Calibration”. PhD thesis, University of Maryland.
- Das, S., E. Wilson, and E. Williams. 2021. “The impact of behavioral and geographic heterogeneity on residential-sector carbon abatement costs”. *Energy and Buildings* 231:110611.
- Database for Energy Efficient Resources*. 2021. <http://deeresources.com/>. Accessed: 2021-09-05. California Public Utilities Commission.
- Davila, C. C. 2017. “Building Archetype Calibration for Effective Urban Building Energy Modeling”. PhD thesis, Massachusetts Institute of Technology, Department of Architecture.
- DeWitt, P. E., C. CaraDonna, A. LeBar, M. Dahlhausen, E. Present, R. Adhikari, and A. Parker. 2022. “Misclassification and Outlier Detection of Buildings in a Linked Utility and Real-Estate Data Set”. In preparation.
- DNV GL. 2021. *2019 California Residential Appliance Saturation Study (RASS)*. <https://www.energy.ca.gov/data-reports/surveys/2019-residential-appliance-saturation-study>.
- Egerter, A., and M. Campbell. 2020. *Prefabricated Zero Energy Retrofit Technologies: A Market Assessment*. Tech. rep. Rocky Mountain Institute.
- Elevate. 2017. *Commonwealth Edison’s Anonymous Data Service: A Review and Recommendations*. Tech. rep. Elevate.
- Energy Efficiency Policy Manual Version 6 For Post-2018 Programs*. 2020. Tech. rep. California Public Utilities Commission.
- ENERGY STAR. 2020a. *ENERGY STAR® Certified Room Air Conditioners*. [https://www.energystar.gov/partner\\_resources/products\\_partner\\_resources/brand\\_owner\\_resources/unit\\_shipment\\_data/archives](https://www.energystar.gov/partner_resources/products_partner_resources/brand_owner_resources/unit_shipment_data/archives). Accessed: 2020-07-02.
- . 2020b. *ENERGY STAR® Program Requirements Product Specification for Room Air Conditioners - Eligibility Criteria, Version 3.0*. [https://www.energystar.gov/sites/default/files/specs/private/ENERGY\\_STAR\\_Final\\_Version\\_3\\_Room\\_Air\\_Conditioner\\_Specification.pdf](https://www.energystar.gov/sites/default/files/specs/private/ENERGY_STAR_Final_Version_3_Room_Air_Conditioner_Specification.pdf). Accessed: 2020-06-01.
- . 2020c. *ENERGY STAR® Program Requirements Product Specification for Room Air Conditioners - Eligibility Criteria, Version 4.0*. <https://www.energystar.gov/sites/default/files/ENERGY%20STAR%20Final%20Version%204.0%20Room%20Air%20Conditioners%20Specification.pdf>. Accessed: 2020-06-01.
- . 2020d. *ENERGY STAR® Unit Shipment and Sales Data Archives*. <https://data.energystar.gov/Active-Specifications/ENERGY-STAR-Certified-Room-Air-Conditioners/rg68-9xmm/data>. Accessed: 2020-07-02.
- EPRI Load Shape Library*. 2021. Electric Power Research Institute, Inc. <https://loadshape.epri.com/>.
- ERCOT. 2021. *Backcasted (Actual) Load Profiles - Historical*. <http://www.ercot.com/mktinfo/loadprofile/alp/>.
- Fenaughty, K., D. Parker, and E. Martin. 2017. “Phased Deep Retrofit Project: Real-Time Measurement of Energy End-uses and Retrofit Opportunities”.
- Fennell, P. J., et al. 2021. “Challenges and Lessons Learned in Applying Sensitivity Analysis to Building Stock Energy Models”. In *17th IBPSA International Conference and Exhibition, Building Simulation 2021*.
- FERC. 2007. *Form No. 714 - Annual Electric Balancing Authority Area and Planning Area Report*. <https://www.ferc.gov/industries-data/electric/general-information/electric-industry-forms/form-no-714-annual-electric/data>. Accessed: 2021-08-26.
- Frick, N., and M. Pigman. 2022. *End-Use Load Profiles for the U.S. Building Stock: Applications and Opportunities*. Tech. rep. In preparation.
- Frick, N. M. 2019. “End Use Load Profile Inventory”.
- Frick, N. M., T. Eckman, and C. A. Goldman. 2017. *Time-varying value of electric energy efficiency*. Tech. rep.
- Garrett, A., and J. R. New. 2016. *Suitability of ASHRAE guideline 14 metrics for calibration*. Tech. rep. Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States).



- Giuntella, O., and F. Mazzonna. 2019. "Sunset time and the economic effects of social jetlag: evidence from US time zone borders". *Journal of Health Economics* 65:210–226. ISSN: 0167-6296. doi:<https://doi.org/10.1016/j.jhealeco.2019.03.007>. <https://www.sciencedirect.com/science/article/pii/S0167629618309718>.
- Guide for the Verification and Validation of Computational Fluid Dynamics Simulations*. 1998. Tech. rep. AIAA (American Institute for Aeronautics and Astronautics).
- Hale, E., et al. 2021. "The Los Angeles 100% Renewable Energy Study (LA100) Chapter 3: Electricity Demand Projections" (). doi:10.2172/1774871. <https://www.nrel.gov/docs/fy21osti/79444-3.pdf>.
- Hiller, C. 2000. "Determining equipment service life". *ASHRAE Journal* 42 (): 48–50+52.
- Ho, J., J. Becker, M. Brown, P. Brown, I. Chernyakhovskiy, S. Cohen, W. Cole, S. Corcoran, K. Eurek, W. Frazier, et al. 2021. *Regional Energy Deployment System (ReEDS) Model Documentation: Version 2020*. Tech. rep. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Horel, J., M. Splitt, L. Dunn, J. Pechmann, B. White, C. Ciliberti, S. Lazarus, J. Slemmer, D. Zaff, and J. Burks. 2002. "Mesowest: Cooperative mesonets in the western United States". *Bulletin of the American Meteorological Society* 83 (2): 211–226.
- Horowitz, S., J. Robertson, J. Maguire, U. O. of Energy Efficiency, and R. Energy. 2019. *OpenStudio® Energy Rating Index (ERI) Workflow*. doi:10.11578/dc.20190228.1. <https://www.osti.gov/biblio/1497076>.
- Horry County. 2018. *IMAGINE 2040 Comprehensive Plan, Chapter 2*. \url{<https://www.horrycounty.org/portals/0/Docs/planningandzoning/Imagine2040/Chapter%20%20-%20Population%2010.1.2018.pdf>}.
- Huang, Y. J. 2018. "How Much Should We Trust Computer Simulated Weather Data?" In *2018 Building Performance Analysis Conference and SimBuild*. Unpublished; proceedings listed at <http://toc.proceedings.com/53732webtoc.pdf>. ASHRAE and IBPSA-USA.
- Ingraham, C. 2019. "How living on the wrong side of a time zone can be hazardous to your health". *The Washington Post*.
- Itron. 2006. *2006 California Commercial End-Use Survey (CEUS)*. Tech. rep.
- Judkoff, R., and J. Neymark. 2013. "Twenty Years On!: Updating the IEA BESTEST Building Thermal Fabric Test Cases for ASHRAE Standard 140" (). doi:10.2172/1220110. <https://www.osti.gov/biblio/1220110>.
- Judkoff, R., B. Polly, M. Bianchi, and J. Neymark. 2010. "Building Energy Simulation Test for Existing Homes (BESTEST-EX); Phase 1 Test Procedure: Building Thermal Fabric Cases". doi:10.2172/988600. <https://www.osti.gov/biblio/988600>.
- Judkoff, R., and J. Neymark. 1995. *International Energy Agency building energy simulation test (BESTEST) and diagnostic method*. Tech. rep. National Renewable Energy Lab.
- . 2006. *Model validation and testing: The methodological foundation of ASHRAE Standard 140*. Tech. rep. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Judkoff, R., B. Polly, and J. Neymark. 2016. "A Method to Test Model Calibration Techniques". In *2016 ACEEE Summer Study on Energy Efficiency in Buildings*.
- Judkoff, R., D. Wortman, and J. Burch. 1982. *Empirical validation of building energy-analysis simulation programs: a status report*. Tech. rep. Solar Energy Research Inst., Golden, CO (USA).
- Judkoff, R., et al. 2017. "Airside HVAC BESTEST: HVAC Air-Distribution System Model Test Cases for ASHRAE Standard 140" (). <https://www.osti.gov/biblio/1379463>.
- Kankiewicz, A., and G. Novotny. 2015. *How Misuse of Solar Resource Datasets is Reducing Solar Industry Profits*. Tech. rep. Clean Power Research.
- KEMA. 2009. *End-Use Load Data Update Project Final Report, Phase 1: Cataloguing Available End-Use and Efficiency Measure Load Data*. Tech. rep.
- KEMA. 2010. *2009 California Residential Appliance Saturation Study (RASS)*. <https://www.energy.ca.gov/data-reports/surveys/2019-residential-appliance-saturation-study/2009-and-2003-residential>.
- KEMA-XENERGY. 2004. *2003 California Residential Appliance Saturation Study (RASS)*. <https://www.energy.ca.gov/data-reports/surveys/2019-residential-appliance-saturation-study/2009-and-2003-residential>.
- Labs, K., J. Carmody, R. Sterling, L. Shen, Y. J. Huang, and D. Parker. 1988. *Building Foundation Design Handbook*. Tech. rep. ORNL.
- Langevin, J., J. L. Reyna, S. Ebrahimigharehbaghi, N. Sandberg, P. Fennell, C. Nägeli, J. Laverge, M. Delghust, É. Mata, M. Van Hove, et al. 2020. "Developing a common approach for classifying building stock energy models". *Renewable and Sustainable Energy Reviews* 133:110276.

- Larson, B., L. Gilman, R. Davis, M. Logsdon, J. Usan, B. Hannas, D. Baylon, P. Storm, V. Mugford, and N. Kvaltine. 2014. *2011 Residential Building Stock Assessment: Metering Study*. <https://neea.org/resources/2011-rbsa-metering-study>. Prepared by Ecotope Inc. for the Northwest Energy Efficiency Alliance.
- LBNL. 2020. *Home Energy Saver & Score: Engineering Documentation: Heating and Cooling Equipment Efficiencies*. <http://hes-documentation.lbl.gov/calculation-methodology/calculation-of-energy-consumption/heating-and-cooling-calculation/heating-and-cooling-equipment/heating-and-cooling-equipment-efficiencies>. Accessed: 2020-06-01.
- Li, H., Z. Wang, T. Hong, A. Parker, and M. Neukomm. 2021a. “Characterizing patterns and variability of building electric load profiles in time and frequency domains”. *Applied Energy* 291:116721. ISSN: 0306-2619. doi:<https://doi.org/10.1016/j.apenergy.2021.116721>. <https://www.sciencedirect.com/science/article/pii/S0306261921002397>.
- . 2021b. “Characterizing patterns and variability of building electric load profiles in time and frequency domains”. *Applied Energy* 291:116721. ISSN: 0306-2619. doi:<https://doi.org/10.1016/j.apenergy.2021.116721>. <https://www.sciencedirect.com/science/article/pii/S0306261921002397>.
- Mai, T. T., P. Jadun, J. S. Logan, C. A. McMillan, M. Muratori, D. C. Steinberg, L. J. Vimmerstedt, B. Haley, R. Jones, and B. Nelson. 2018. *Electrification futures study: Scenarios of electric technology adoption and power consumption for the United States*. Tech. rep. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Meier, A., L. Rainer, A. Daken, T. Ueno, M. Pritoni, and D. Baldewicz. 2019. “What can connected thermostats tell us about American heating and cooling habits?”
- Mills, E. 2008. “The Home Energy Saver: Documentation of Calculation Methodology, Input Data, and Infrastructure, LBNL-51938”. Berkeley, CA: Lawrence Berkeley National Laboratory.
- Mims Frick, N., E. J. H. Wilson, J. Reyna, A. S. Parker, E. K. Present, J. Kim, T. Hong, H. Li, and T. Eckman. 2019. *End-Use Load Profiles for the U.S. Building Stock: Market Needs, Use Cases, and Data Gaps*. Tech. rep. Golden. doi:10.2172/1576489.
- Muehleisen, R. T., and J. Bergerson. 2016. “Bayesian Calibration - What, Why and How”. In *Proceedings of the 4th International High Performance Buildings Conference at Purdue, July 11-14, 2016*.
- National Oceanic and Atmospheric Administration. 2021. *Climate at a glance. National time series. Cooling degree days*. [https://www.ncdc.noaa.gov/cag/national/time-series/110/cdd/12/12/1895-2021?base\\_prd=true&begbaseyear=1901&endbaseyear=2000](https://www.ncdc.noaa.gov/cag/national/time-series/110/cdd/12/12/1895-2021?base_prd=true&begbaseyear=1901&endbaseyear=2000).
- National Research Council. 2012. *Assessing the reliability of complex models: mathematical and statistical foundations of verification, validation, and uncertainty quantification*. National Academies Press.
- Navigant. 2014. *2014 Commercial Building Stock Assessment: Final Report*. <https://neea.org/resources/2014-cbsa-final-report>. Prepared by Navigant Consulting for the Northwest Energy Efficiency Alliance.
- . 2018. *RES 1 Baseline Load Shape Study*. Tech. rep. Prepared for the Electric and Gas Program Administrators of Massachusetts.
- NEEA. 2021a. *Northwest End Use Load Research Project - Commercial Energy Metering Study*. <https://neea.org/data/nw-end-use-load-research-project/>.
- . 2021b. *Northwest End Use Load Research Project - Home Energy Metering Study*. <https://neea.org/data/nw-end-use-load-research-project/>.
- . 2021c. *Residential Building Stock Assessment Website*. <https://neea.org/data/residential-building-stock-assessment>.
- New, J. R., J. Sanyal, M. Bhandari, and S. Shrestha. 2012. “Autotune e+ building energy models”. *Proceedings of SimBuild 5* (1): 270–278.
- NREL. 2021a. *NSRDB: National Solar Radiation Database – About – U.S. Data*. <https://nsrdb.nrel.gov/about/u-s-data.html>. Accessed: 2021-09-05.
- . 2021b. *OpenStudio-HPXML*. Version 1.2.0. <https://github.com/NREL/OpenStudio-HPXML>.
- NREL, LBNL, ORNL, and PNNL. 2021. *OpenStudio-Standards*. Version 0.2.14. <https://github.com/NREL/openstudio-standards>.
- Ong, S., C. Campbell, and N. Clark. 2012. *Impacts of regional electricity prices and building type on the economics of commercial photovoltaic systems*. Tech. rep. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Ong, S., N. Clark, P. Denholm, and R. Margolis. 2013. *Breakeven prices for photovoltaics on supermarkets in the United States*. Tech. rep. National Renewable Energy Lab.(NREL), Golden, CO (United States).



- Parker, D., K. Sutherland, D. Chasar, J. Montemurno, B. Amos, and J. Kono. 2016. *Phased Retrofits in Existing Homes in Florida Phase I: Shallow and Deep Retrofits*. Tech. rep. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Pecan Street Inc. 2021. *Pecan Street Dataport*. <https://www.pecanstreet.org/dataport/>. Accessed: 2021-08-26.
- Polly, B., N. Kruis, and D. Roberts. 2011. *Assessing and improving the accuracy of energy analysis for residential buildings*. Tech. rep. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Present, E., E. Wilson, C. CaraDonna, J. Kim, R. Adhikari, A. C. McCreery, E. Titus, and N. Frick. 2020. "Putting Our Industry's Data to Work: A Case Study of Large-Scale Data Aggregation". In *2020 ACEEE Summer Study on Energy Efficiency in Buildings*.
- Price, P. 2010. *Methods for analyzing electric load shape and its variability*. Tech. rep. Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States).
- Puckett, C. D., L. Molander, C. Holmes, K. Gomatom, R. Jackson, and M. Starke. 2014. *End-Use Data Development: It's Time to Better Understand our Customers*. [http://nilmworkshop.org/2016/proceedings/Sponsor\\_DNV-GL.pdf](http://nilmworkshop.org/2016/proceedings/Sponsor_DNV-GL.pdf).
- Reddy, T. A. 2006a. "Literature review on calibration of building energy simulation programs: uses, problems, procedures, uncertainty, and tools". *ASHRAE transactions* 112:226.
- Reddy, T. A. 2006b. "Literature Review on Calibration of Building Energy Simulation Programs: Uses, Problems, Procedures, Uncertainty and Tools". *ASHRAE Transactions* 112, no. 1 ().
- Regional Technical Forum. 2013. *SEEM RBSA Calibration, Phase II: Electric Heating Energy Adjustments due to Supplemental Heat, Program Eligibility, and Related Factors*. Tech. rep. RTF Staff Technical Report.
- Reid, M. 2020. *MatthewReid854/reliability: v0.5.1*. Version v0.5.1. doi:10.5281/zenodo.3938000. \url{https://doi.org/10.5281/zenodo.3938000}.
- Reinhart, C. F., and C. C. Davila. 2016. "Urban building energy modeling—A review of a nascent field". *Building and Environment* 97:196–202.
- Reliability terminology*. 2021. <https://www.ti.com/support-quality/reliability/reliability-terminology.html>. Accessed: 2021-09-05. Texas Instruments.
- Residential and Commercial Sector Energy Code Adoption and Compliance Rates*. 2017. Tech. rep. Prepared by ICF International L.L.C for the U.S. Energy Information Administration. U.S. Energy Information Administration.
- RESNET. 2019. *ANSI/RESNET/ICC 301-2019 Standard for the Calculation and Labeling of the Energy Performance of Dwelling and Sleeping Units using an Energy Rating Index*. [http://www.resnet.us/wp-content/uploads/archive/resblog/2019/01/ANSIRESNETICC301-2019\\_vf1.23.19.pdf](http://www.resnet.us/wp-content/uploads/archive/resblog/2019/01/ANSIRESNETICC301-2019_vf1.23.19.pdf). Accessed: 2021-09-08.
- Richman, E. E., E. M. Rauch, J. N. Phillips, K. A. Petty, J. Knappek, and P. L. Lopez-Rangel. 2008. *National Commercial Construction Characteristics and Compliance with Building Energy Codes: 1999-2007*. Tech. rep. Pacific Northwest National Lab.(PNNL), Richland, WA (United States).
- Robertson, J., B. Polly, and J. Collis. 2013. *Evaluation of automated model calibration techniques for residential building energy simulation*. Tech. rep. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Roth, A., D. Goldwasser, and A. Parker. 2016. "There's a Measure for That!" *Energy and Buildings* 117 (). ISSN: 0378-7788. doi:10.1016/j.enbuild.2015.09.056. <https://www.osti.gov/biblio/1245363>.
- Ruggles, S., S. Flood, S. Foster, R. Goeken, J. Pacas, M. Schouweiler, and M. Sobek. 2021. *IPUMS USA: Version 11.0 [dataset]*. <https://doi.org/10.18128/D010.V11.0>. Minneapolis, MN.
- (SEIA), S. E. I. A. 2020. "Solar market insight report 2020 year review". <https://www.seia.org/research-resources/solar-market-insight-report-2020-year-review>.
- Skumatz, L. A. 2012. "What Makes a Good EUL? Analysis of Existing Estimates and Implications for New Protocols for Estimated Useful Lifetimes (EULs)". In *2012 International Energy Program Evaluation Conference, Rome, Italy*.
- Smith, A., N. Lott, and R. Vose. 2011. "The integrated surface database: Recent developments and partnerships". *Bulletin of the American Meteorological Society* 92 (6): 704–708.
- Taylor, Z. T., Y. Xie, C. D. Burleyson, N. Voisin, and I. Kraucunas. 2019. "A multi-scale calibration approach for process-oriented aggregated building energy demand models". *Energy and Buildings* 191:82–94. ISSN: 0378-7788. doi:<https://doi.org/10.1016/j.enbuild.2019.02.018>. <https://www.sciencedirect.com/science/article/pii/S0378778818330950>.
- Titus, E., and C. McChalicher. 2021. *NEEP Regional End Use Load Profile Data Inventory and Needs Assessment*. Tech. rep. Northeast Energy Efficiency Partnerships.

- U.S. Census Bureau. 2017. *2017 American Housing Survey Public Use File*. <https://www.census.gov/programs-surveys/ahs.html>.
- U.S. Department of Homeland Security. 2021. *Homeland Infrastructure Foundation-Level Data (HIFLD) Portal*. <https://hifld-geoplatform.opendata.arcgis.com/>. Accessed: 2021-08-26.
- U.S. Energy Information Administration. 2007. *Form EIA-861M (formerly EIA-826) detailed data*. <https://www.eia.gov/electricity/data/>. Accessed: 2021-08-26.
- . 2012. *2012 Commercial Building Energy Consumption Survey (CBECS)*. <http://www.eia.doe.gov/emeu/cbecs/>. Accessed: 2021-09-08.
  - . 2013. *2009 Residential Energy Consumption Survey (RECS) Microdata*. <https://www.eia.gov/consumption/residential/data/2009/>. Accessed: 2021-08-26.
  - . 2017. *Residential Energy Consumption Survey (RECS): Using the 2015 microdata file to compute estimates and standard errors (RSEs)*. Tech. rep. U.S. Energy Information Administration. <https://www.eia.gov/consumption/residential/data/2015/pdf/microdata.pdf>.
  - . 2018a. *2015 Consumption and Expenditures Technical Documentation Summary*. Tech. rep. U.S. Energy Information Administration.
  - . 2018b. *2015 Residential Energy Consumption Survey (RECS) Microdata*. <https://www.eia.gov/consumption/residential/data/2015/>. Accessed: 2021-08-26.
  - . 2018c. *Form EIA-861M (formerly EIA-826) detailed data*. <https://www.eia.gov/electricity/data/eia861m/>. Accessed: 2021-08-26.
  - . 2021a. Personal communication with EIA staff.
  - . 2021b. *Annual Electric Power Industry Report, Form EIA-861 detailed data files*. <https://www.eia.gov/electricity/data/eia861/>. Accessed: 2021-08-26.
  - . 2021c. *Form EIA-861M (formerly EIA-826) detailed data*. <https://www.eia.gov/electricity/data/eia861m/>. Accessed: 2021-08-26.
  - . 2021d. *Heat Content of Natural Gas Consumed*. [https://www.eia.gov/dnav/ng/ng\\_cons\\_heat\\_a\\_EPG0\\_VGTH\\_btucf\\_a.htm](https://www.eia.gov/dnav/ng/ng_cons_heat_a_EPG0_VGTH_btucf_a.htm).
  - . 2021e. *Natural Gas Data*. <https://www.eia.gov/naturalgas/data.php>. Accessed: 2021-08-26.
  - . 2021f. *Residential Energy Consumption Survey (RECS)*. <https://www.eia.gov/consumption/residential/index.php>. Accessed: 2021-08-26.
- Valovcin, S, A. Hering, B Polly, and M Heaney. 2014. “A statistical approach for post-processing residential building energy simulation output”. *Energy and buildings* 85:165–179.
- Walker, I., R. Chan, and M. Spears. 2020. *Residential Diagnostics Database (ResDB)*. Residential Building Systems Group, Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory. <https://resdb.lbl.gov/>.
- Wilcox, S., and W. Marion. 2008. *Users manual for tmy3 data sets (revised)*. Tech. rep. NREL/TP-581-43156. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Wilson, E., C. Engebrecht-Metzger, S. Horowitz, and R. Hendron. 2014. “2014 Building America House Simulation Protocols” (). doi:10.2172/1126820. <https://www.osti.gov/biblio/1126820>.
- Wilson, E., C. Christensen, S. Horowitz, and H. Horsey. 2016. “A high-granularity approach to modeling energy consumption and savings potential in the US residential building stock”. *Proceedings of SimBuild* 6 (1).
- Wilson, E. J. H., C. B. Christensen, S. G. Horowitz, J. J. Robertson, and J. B. Maguire. 2017. *Energy efficiency potential in the US single-family housing stock*. Tech. rep. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Winkler, J., C. Booten, D. Christensen, and J. Tomerlin. 2012. *Laboratory Performance Testing of Residential Window Mounted Air Conditioners*. Tech. rep. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Xu, K. 2012. “Assessing the Minimum Instrumentation to Well Tune Existing Medium Sized Office Building Energy Models”. PhD thesis, The Pennsylvania State University.
- Yamada, M., J. Penning, S. Schober, K. Lee, and C. Elliott. 2019. *Energy savings forecast of solid-state lighting in general illumination applications*. Tech. rep. Prepared by Navigant Consulting for the U.S. Department of Energy. U.S. Department of Energy.
- Yang, Z., and B. Becerik-Gerber. 2015. “A model calibration framework for simultaneous multi-level building energy simulation”. *Applied Energy* 149 (Supplement C): 415–431. ISSN: 0306-2619. doi:<https://doi.org/10.1016/j.apenergy.2015.03.048>. <http://www.sciencedirect.com/science/article/pii/S0306261915003323>.

Zhang, L., et al. 2021. “High-resolution hourly surrogate modeling framework for physics-based large-scale building stock modeling”. *Sustainable Cities and Society* 75 (April): 103292. ISSN: 22106707. doi:10.1016/j.scs.2021.103292. <https://doi.org/10.1016/j.scs.2021.103292>.

## Appendix A List of Technical Advisory Group Members and Organizations

Name	Organization	Name	Organization
Charles Alonge	NYISO	Jim Leverette	Southern Company
Ayad Al-Shaikh	CalTF	Jessica Lin	Oracle
Jen Amann	ACEEE	Angela Long	PGE
Kausar Ashraf	Onatrio ISO	Kimberly Lukasiak	NV Energy
Jamie Barber	GA PSC	Ross Macwhinney	City of New York
Cyrus Bhedwar	SEEA	Cecily McChalicher	NEEP
Mark Bielecki	Center for Sustainable Energy	Bill McNary	EIA
Stephen Bird	Clarkson University	Pasi Miettinen	Sagewell
Michael Bishop	SolarReviews	Erik Miller	AES Indiana
Danielle Bond	C Power	Katherine Mitchell	Autodesk
Brad Borum	IURC	Claire Miziolek	Energy Solutions
Ali Bozorgi	ICF	Sarah Mullkoff	MI PSC
Kristen Brown	Electron	Mike Myser	Energy Platforms
Chris Burgess	MEEA	Paulomi (Lucy) Nandy	MEEA
Dave Chassin	SLAC	Shinvani Nathoo	ISO Ontario
Rebecca Chen	Ontario ISO	Chris Neme	Energy Futures Group
David Clement	NEEA	Clay Nesler	WRI
Sue Coakley	NEEP	Laura Ortega	CPS Energy
Barry Coflan	NEEP	Abhijeet Pande	TRC Solutions
Erin Cosgrove	NEEP	Dave Parsons	HI PUC
Timothy Costa	ISO New England	Bob Pauley	IURC
Matt Cox	Greenlink Analytics	David Podorson	Xcel Energy
Ron Domitrovic	EPRI	Susan Powers	Clarkson University
Paul Donohoo-Vallett	DOE	Curt Puckett	DNV
Tom Eckman	LBNL	Bob Ramirez	DNV
Tony Faris	BPA	Barbara Richards	Southern Company
Jamie Fine	Environmental Defense Fund	Rachel Scheu	Elevate Energy
Michael Fink	VEIC	Scott Schuetter	Slipstream
Ellen Franconi	PNNL	Prasenjit Shil	Ameren
Adam Gerza	Energy Toolbase	Rodney Sobin	NASEO
Krish Gomatam	EPRI	Justin Spencer	Apex Analytics
Benjamin Griffiths	Massachusetts	Robert Stephenson	VEIC
Mike Hamilton	Seattle City Light	Kenji Takahashi	Synapse
Alex Hofmann	APPA	Greg Thomas	PSD Consulting
Chris Holmes	EPRI	Elizabeth Titus	NEEP
Meg Howard	MA Clean Energy Center	JJ Vandette	VEIC
David Jacobson	Jacobs Energy	Puja Vohra	Slipstream
Bryan Jungers	E Source	Valerie von Schramm	CPS Energy
Steven Keates	ADM (now Truckee Donner PUD)	Dave Walker	MI PSC
Phillip Kelsven	BPA	Robert Weber	BPA
Sami Khawaja	Cadmus	Adam Wehmann	VEIC
Ben King	DOE (now Rhodium Group)	Bob Willen	Ameren
Kurtis Kolnowski	AEG	Cody Williams	Xcel Energy
Peter Langbein	PJM	Dan Williams	Measurable
Will Lange	WaterFurnace International	Craig Williamson	DNV
Greg Lawson	EIA	Dan York	ACEEE
Tuan Le	CPS	Henry Yoshimura	ISO New England

## Appendix B Roles of Authors and Non-Author Contributors

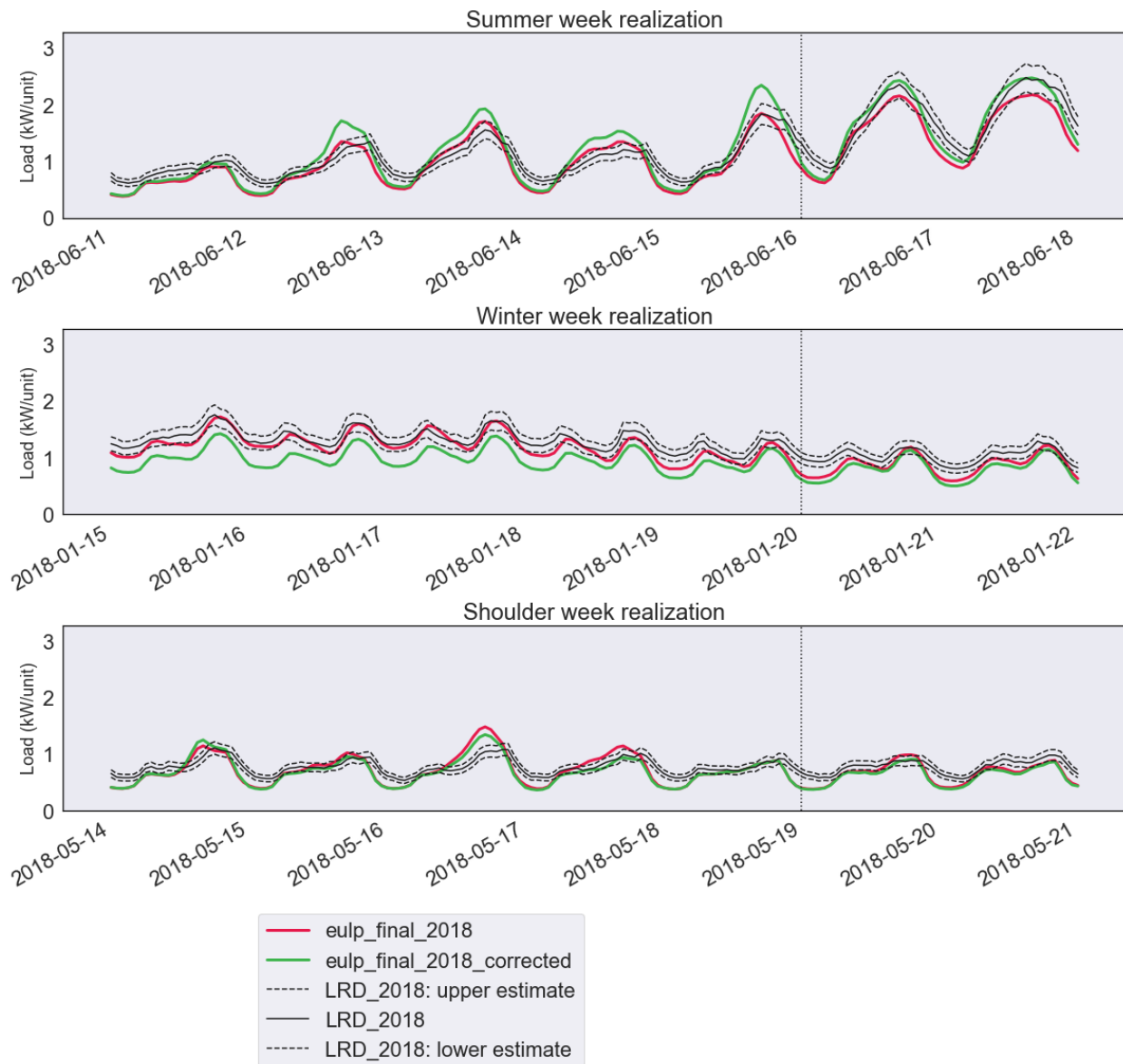
Unless otherwise specified, the listed authors and non-author contributors are affiliated with NREL at the time of publication.

Role or Team	Lead	Team Members
Principal Investigator	Eric J.H. Wilson	
Co-Principal Investigators	Andrew Parker, Natalie Mims Frick (LBNL)	
Project Management	Kim Trenbath, Aileen Blair	
Scientific Communication Lead	Janet L. Reyna	
Subcontract Technical Monitor	Lieko Earle	Andrew Parker, Eric J.H. Wilson
Supervision Review	David Roberts	
DOE Technical Monitor	Monica Neukomm (DOE)	
Data Acquisition	Elaina Present	Christopher CaraDonna, Janghyun Kim, Rajendra Adhikari, Noel Merket
Data Processing	Lixi Liu	Rajendra Adhikari, Noel Merket, Janghyun Kim, Elaina Present, Carlo Bianchi, Chris CaraDonna, Philip White
Calibration - Residential	Anthony Fontanini	Eric Wilson, Rajendra Adhikari, Elaina Present, Lixi Liu, Andrew Speake, Oluwatobi Adekanye (former NREL), Liang Zhang, Janet Reyna, Joseph Robertson, Scott Horowitz, Prateek Munankarmi
Correction Model - Residential	Rajendra Adhikari	Anthony Fontanini, Eric Wilson
Calibration - Commercial	Matthew Dahlhausen	Amy LeBar, Chris CaraDonna, Janghyun Kim, Marlana Praprost, Andrew Parker, Peter DeWitt, Elaina Present, Lixi Liu, Rajendra Adhikari, Jonathan Gonzalez (former NREL), Eric Bonnema, Liang Zhang
Uncertainty Quantification	Janet Reyna	Liang Zhang, Matthew Reynolds, Siby Plaththotam (ANL), Dalton Jones (former NREL), Peter DeWitt, Rawad El Kontar, Noel Merket, Kevin Sayers, Anthony Fontanini, Ralph T. Muehleisen (ANL), Qi Li (former ANL), Xinshuo Yang (former NREL)
Stochastic Occupancy - Residential	Eric Wilson	Rajendra Adhikari, Anthony Fontanini, Jianli Chen (former NREL), Ben Polly, Opeoluwa Wonuola Olawale
Stochastic Occupancy - Commercial	Tianzhen Hong (LBNL)	Zhe Wang (LBNL), Han Li (LBNL), Andrew Parker, Jonathan Gonzalez (former NREL), Mahmoud Alahmad (UNL)
Weather Data	Carlo Bianchi	Matthew Dahlhausen, Henry Horsey, Eric Bonnema, Anthony Fontanini
Repository Management	Joseph Robertson, Rajendra Adhikari	
Data Viewer Leads	Noel Merket, Henry Horsey	

## Appendix C Supplemental Residential AMI Comparisons

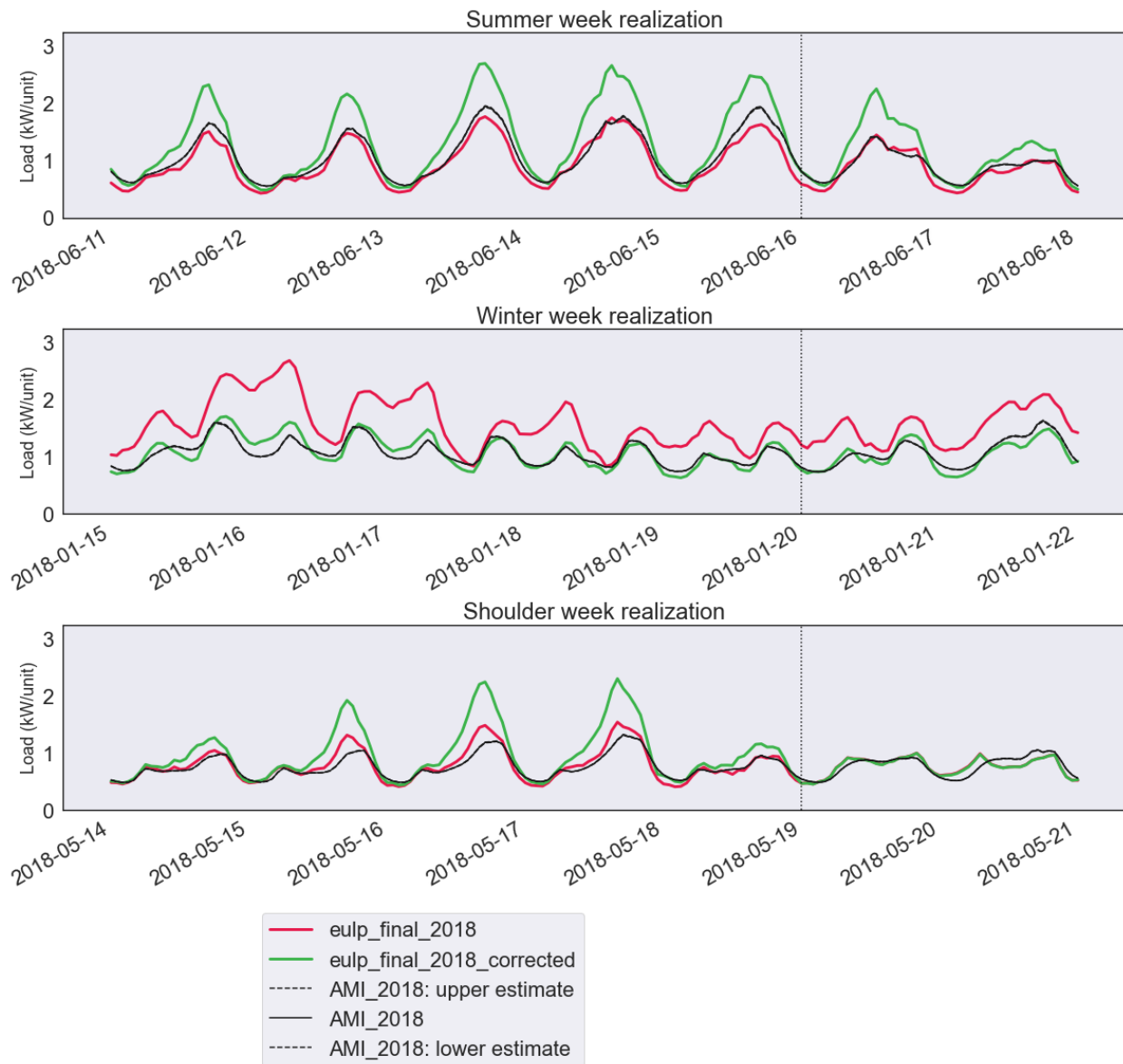
This appendix contains comparisons of ResStock modeled load to residential AMI data. Figures 370–377 show comparisons for an example week in each season. Figures 378–385 show comparisons for a year of daily sums of electricity use per unit. The daily sum plots show seasonal dynamics instead of diurnal dynamics.

ResStock load is shown before and after the EIA-861M-based heating and cooling correction model described in Section 3.2.10 is applied. The correction model is driven by statewide monthly consumption differences, so sometimes it improves the fit with AMI data (VEIC in summer; Fort Collins, Seattle, EPB in winter) and sometimes it worsens the fit (Fort Collins, Seattle, EPB in summer; ComEd, Tallahassee, Cherryland in winter; Fort Collins in the shoulder season:).



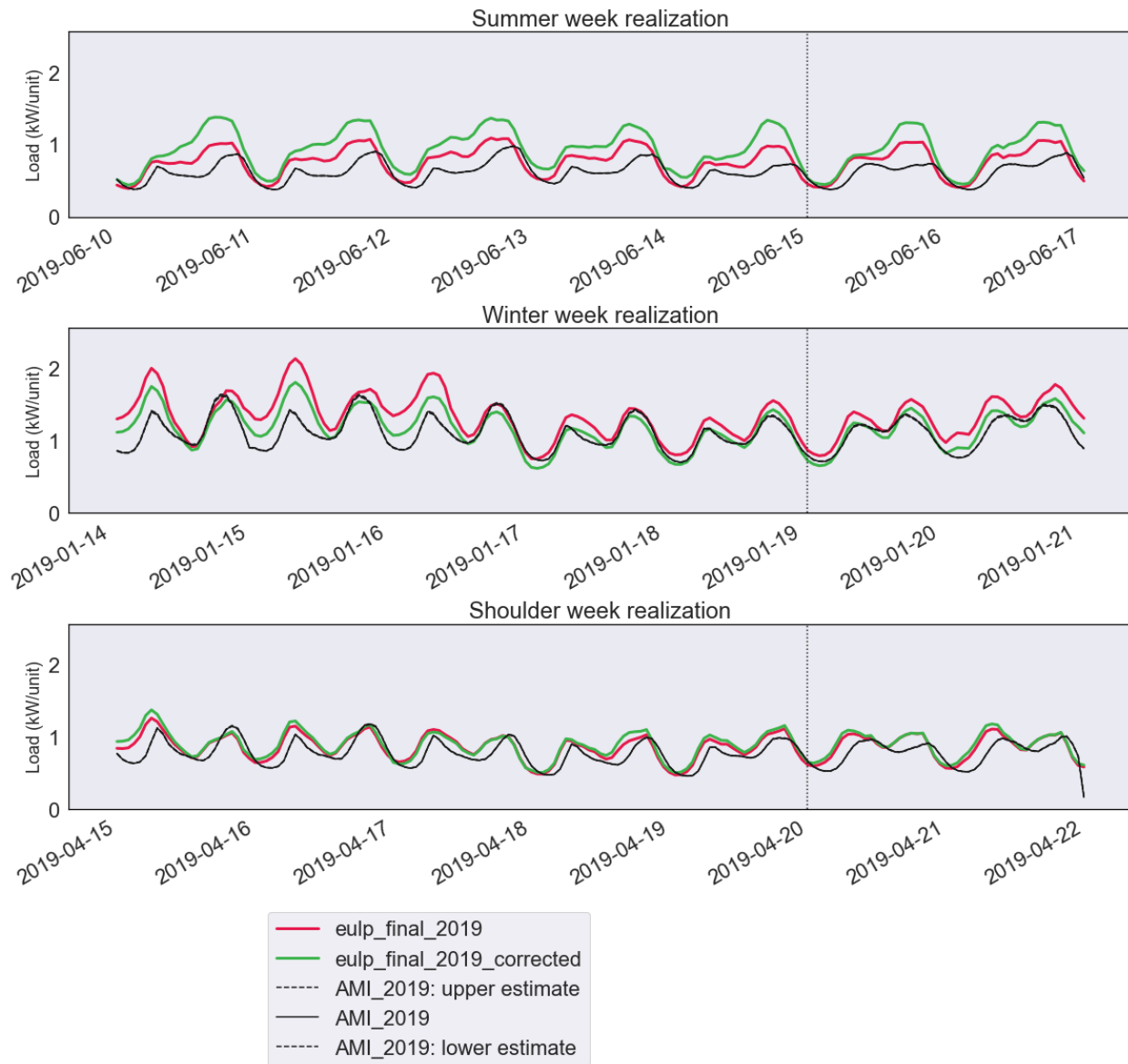
**Figure 370. Comparison of ResStock modeled load to ComEd residential AMI data for an example week in each season. ResStock load (“eulp\_final”) is shown before and after the EIA-861M-based heating and cooling correction model is applied (see Section 3.2.10). The correction model is driven by statewide monthly consumption differences; sometimes it improves the fit with AMI data and sometimes it worsens the fit.**



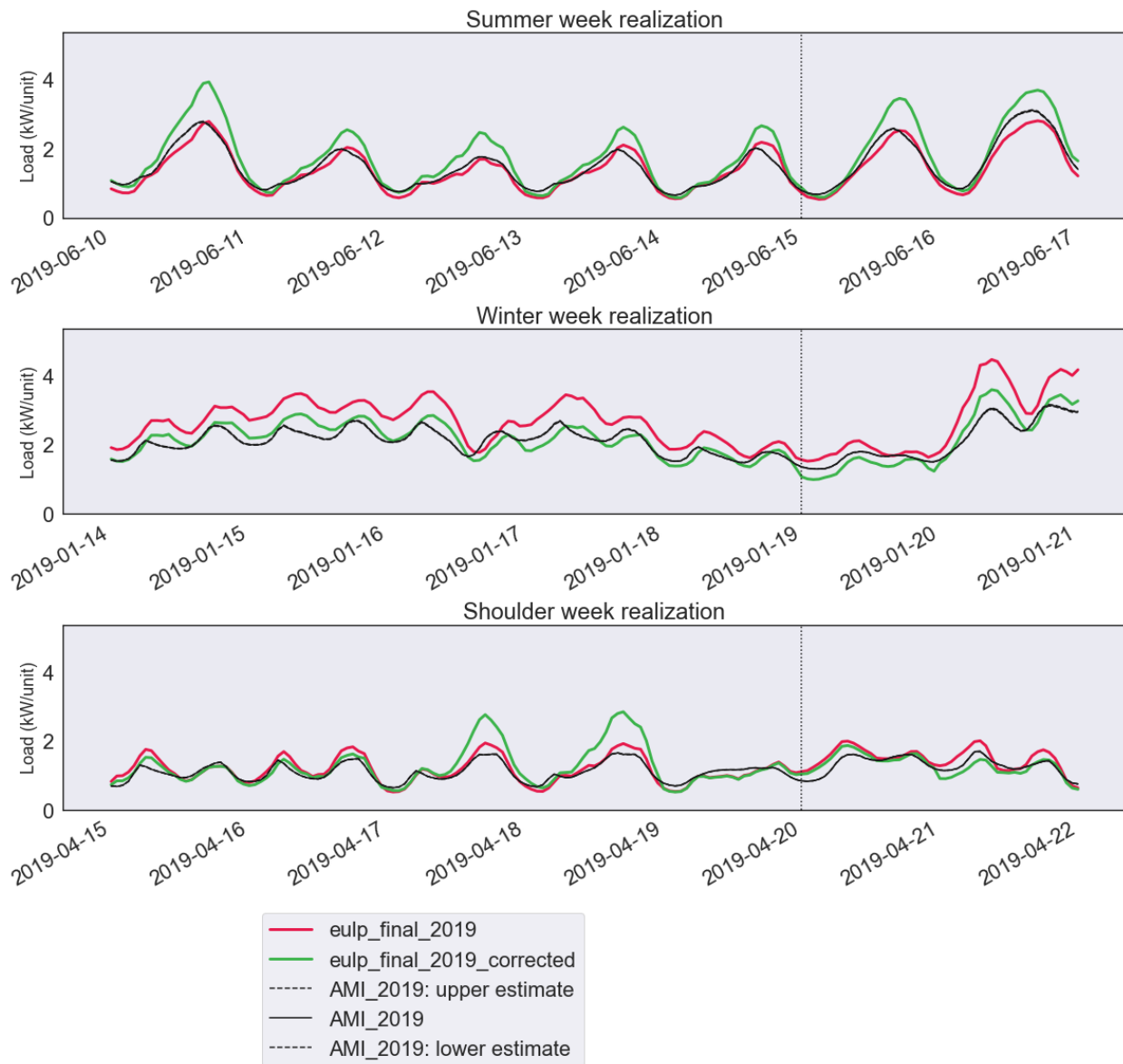


**Figure 371. Comparison of ResStock modeled load to City of Fort Collins residential AMI data for an example week in each season. ResStock load (“eulp\_final”) is shown before and after the EIA-861M-based heating and cooling correction model is applied (see Section 3.2.10). The correction model is driven by statewide monthly consumption differences; sometimes it improves the fit with AMI data and sometimes it worsens the fit.**

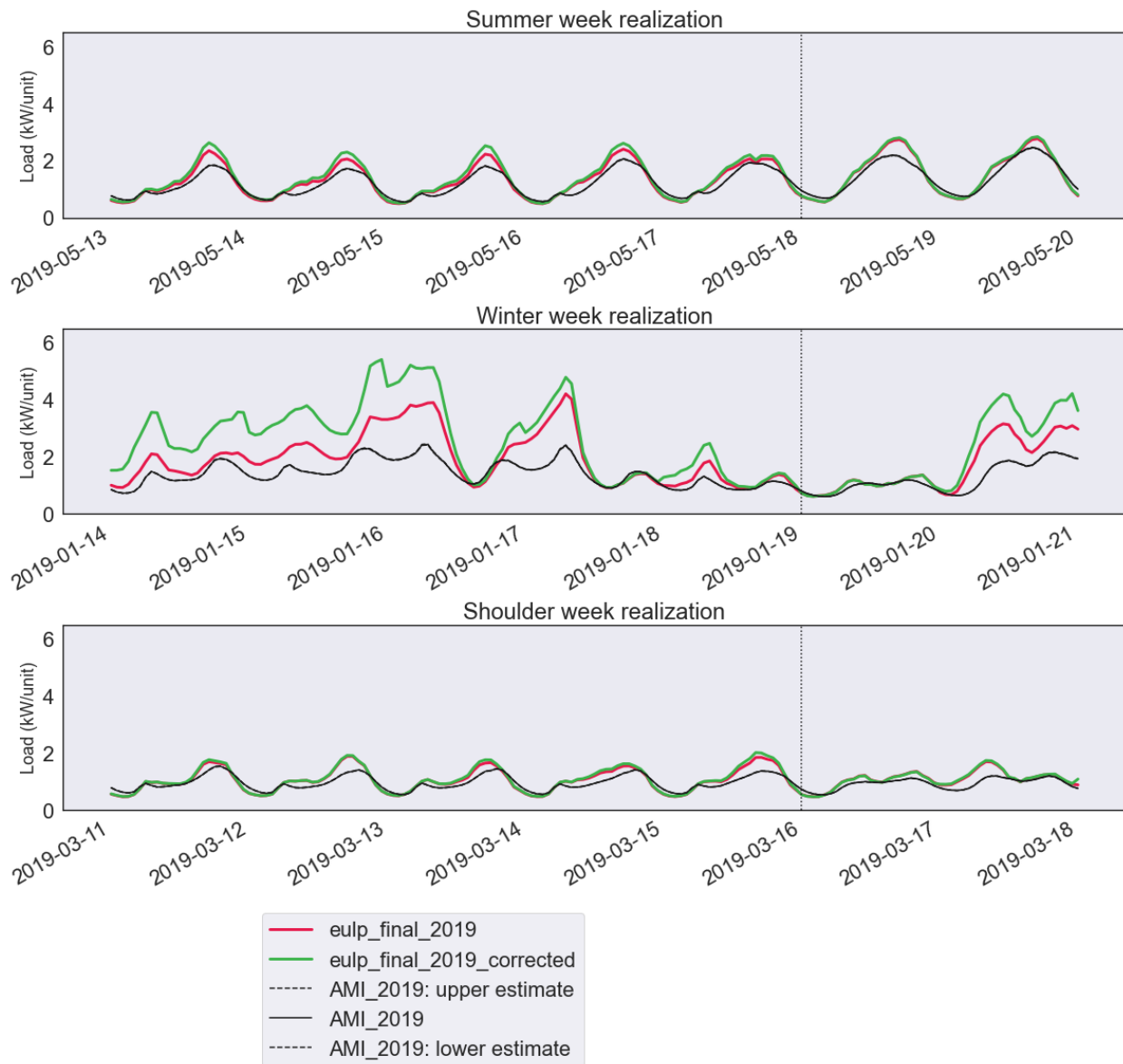




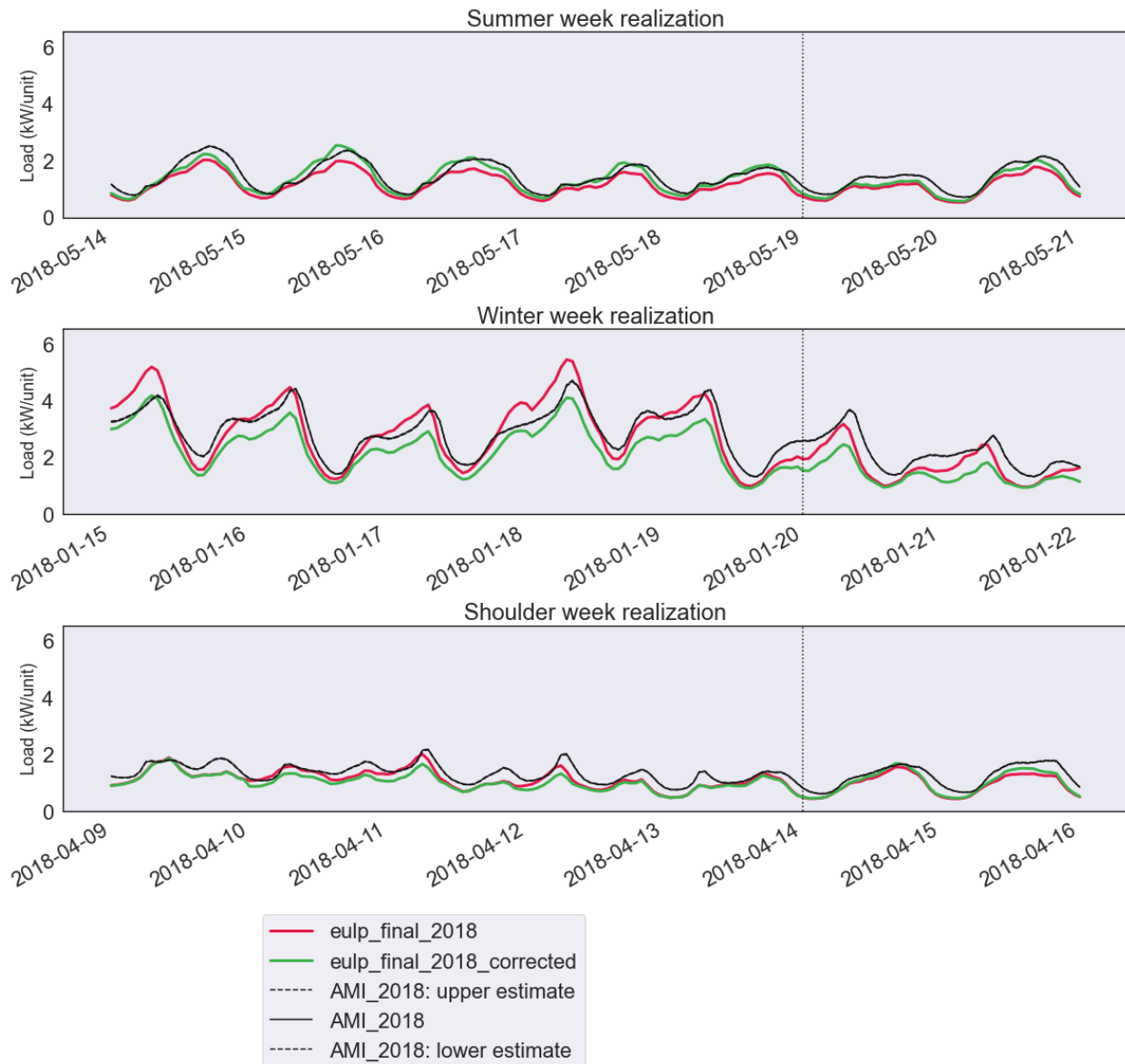
**Figure 372. Comparison of ResStock modeled load to Seattle City Light residential AMI data for an example week in each season. ResStock load (“eulp\_final”) is shown before and after the EIA-861M-based heating and cooling correction model is applied (see Section 3.2.10). The correction model is driven by statewide monthly consumption differences; sometimes it improves the fit with AMI data and sometimes it worsens the fit.**



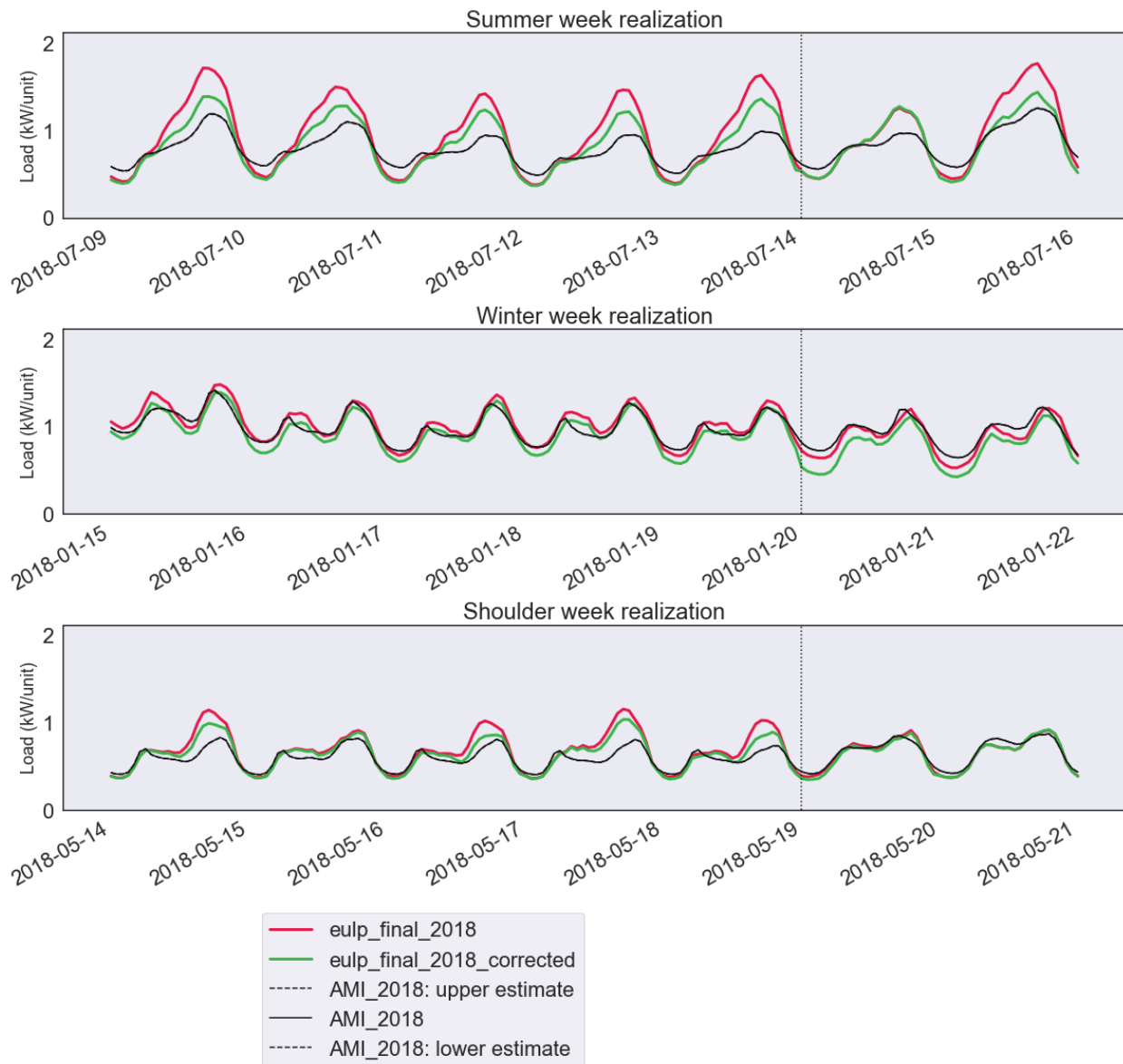
**Figure 373. Comparison of ResStock modeled load to EPB Chattanooga residential AMI data for an example week in each season. ResStock load (“eulp\_final”) is shown before and after the EIA-861M-based heating and cooling correction model is applied (see Section 3.2.10). The correction model is driven by statewide monthly consumption differences; sometimes it improves the fit with AMI data and sometimes it worsens the fit.**



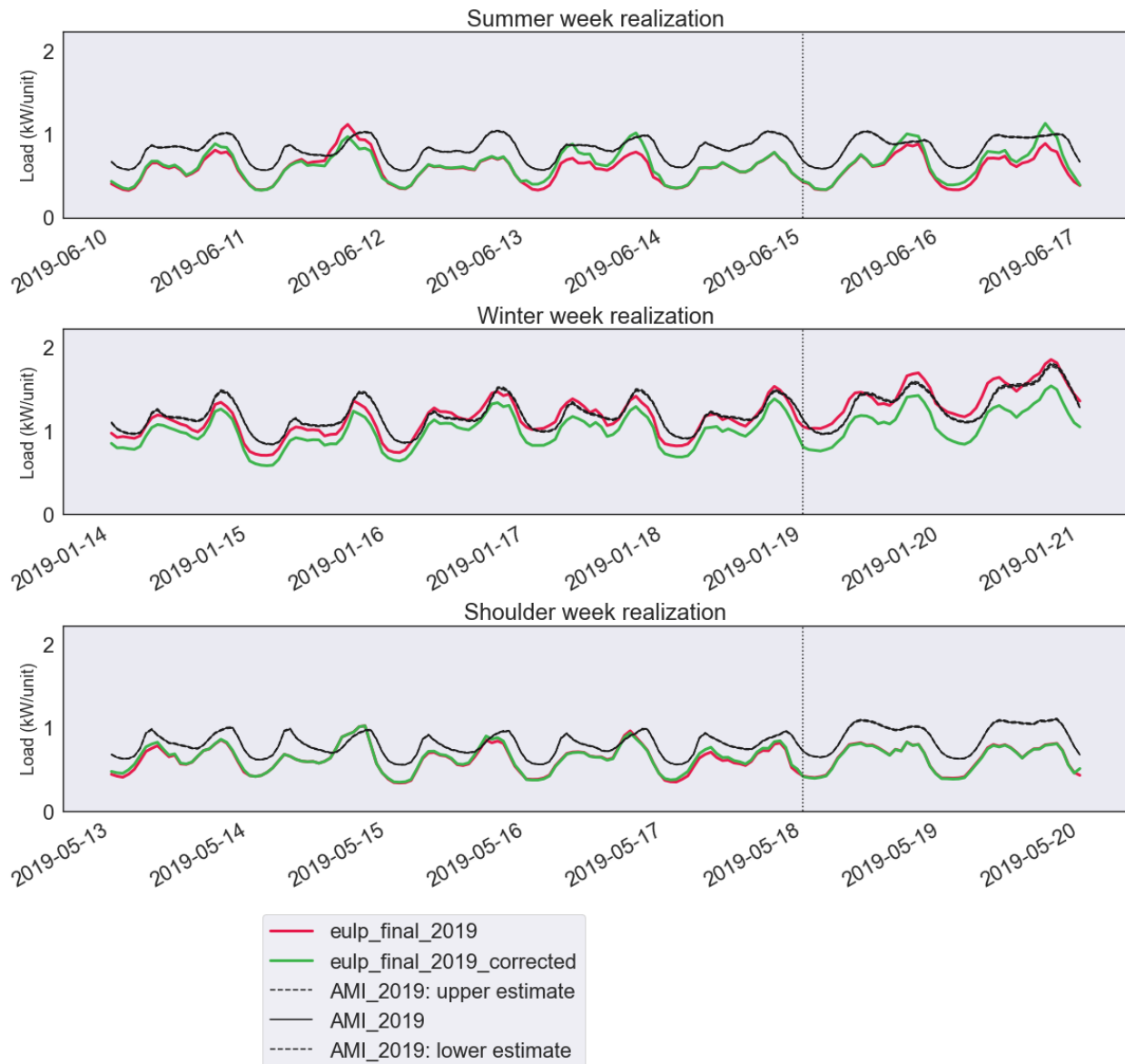
**Figure 374. Comparison of ResStock modeled load to City of Tallahassee residential AMI data for an example week in each season. ResStock load (“eulp\_final”) is shown before and after the EIA-861M-based heating and cooling correction model is applied (see Section 3.2.10). The correction model is driven by statewide monthly consumption differences; sometimes it improves the fit with AMI data and sometimes it worsens the fit.**



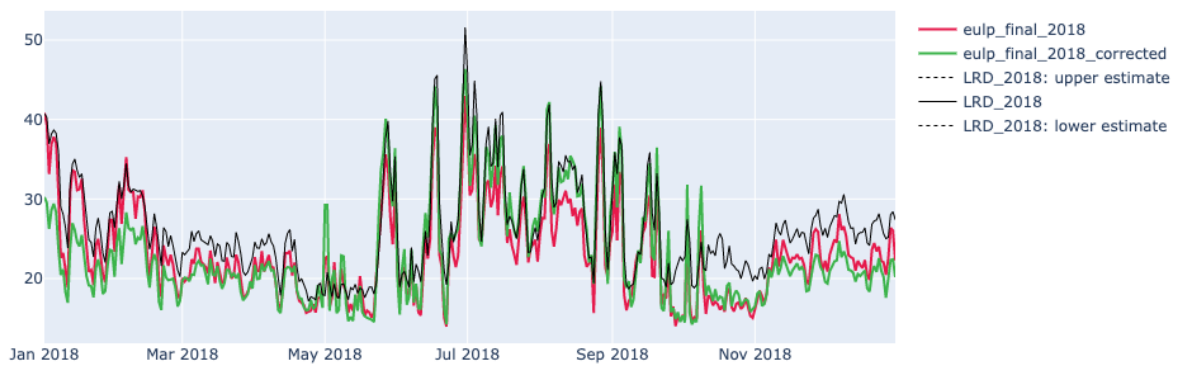
**Figure 375. Comparison of ResStock modeled load to Horry Electric Cooperative residential AMI data for an example week in each season. ResStock load (“eulp\_final”) is shown before and after the EIA-861M-based heating and cooling correction model is applied (see Section 3.2.10). The correction model is driven by statewide monthly consumption differences; sometimes it improves the fit with AMI data and sometimes it worsens the fit.**



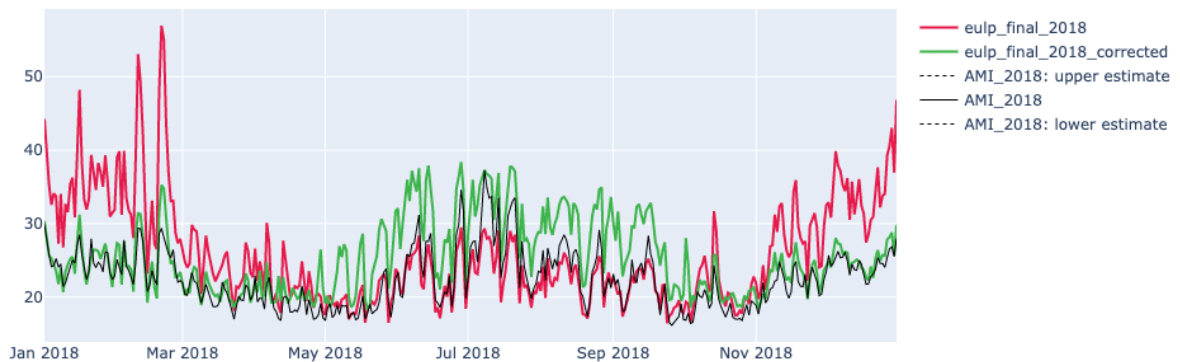
**Figure 376. Comparison of ResStock modeled load to VEIC residential AMI data for an example week in each season. ResStock load (“eulp\_final”) is shown before and after the EIA-861M-based heating and cooling correction model is applied (see Section 3.2.10). The correction model is driven by statewide monthly consumption differences; sometimes it improves the fit with AMI data and sometimes it worsens the fit.**



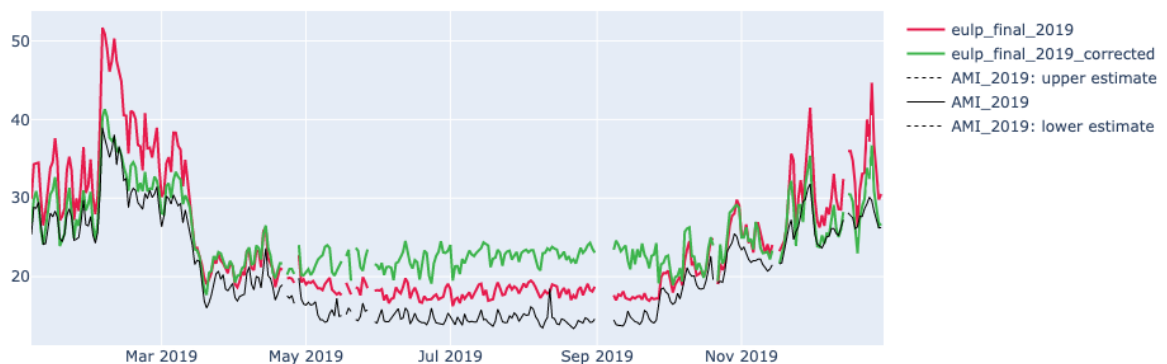
**Figure 377. Comparison of ResStock modeled load to Cherryland Electric Co-op residential AMI data for an example week in each season. ResStock load (“eulp\_final”) is shown before and after the EIA-861M-based heating and cooling correction model is applied (see Section 3.2.10). The correction model is driven by statewide monthly consumption differences; sometimes it improves the fit with AMI data and sometimes it worsens the fit.**



**Figure 378.** Comparison of ResStock modeled load to ComEd residential AMI data for daily sums of electricity per unit (kWh/unit/day). ResStock load (“eulp\_final”) is shown before and after the EIA-861M-based heating and cooling correction model is applied (see Section 3.2.10). The correction model is driven by statewide monthly consumption differences; sometimes it improves the fit with AMI data and sometimes it worsens the fit.

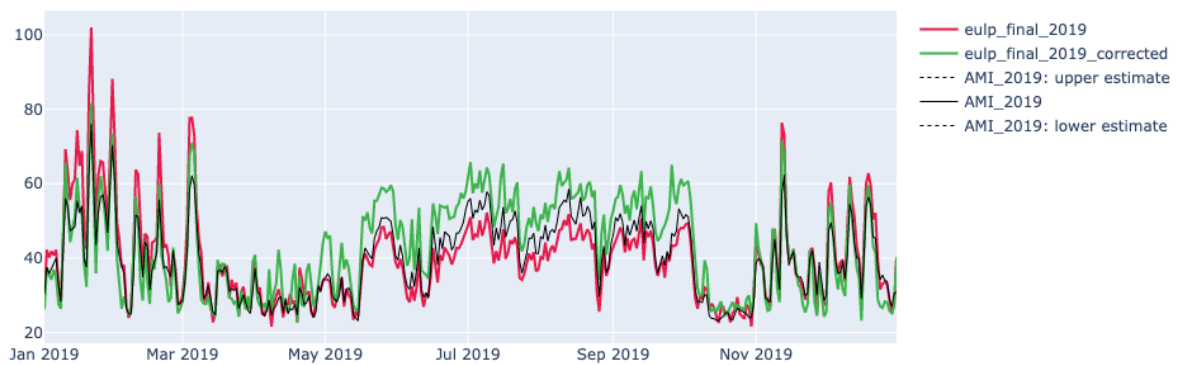


**Figure 379.** Comparison of ResStock modeled load to City of Fort Collins residential AMI data for daily sums of electricity per unit (kWh/unit/day). ResStock load (“eulp\_final”) is shown before and after the EIA-861M-based heating and cooling correction model is applied (see Section 3.2.10). The correction model is driven by statewide monthly consumption differences; sometimes it improves the fit with AMI data and sometimes it worsens the fit.

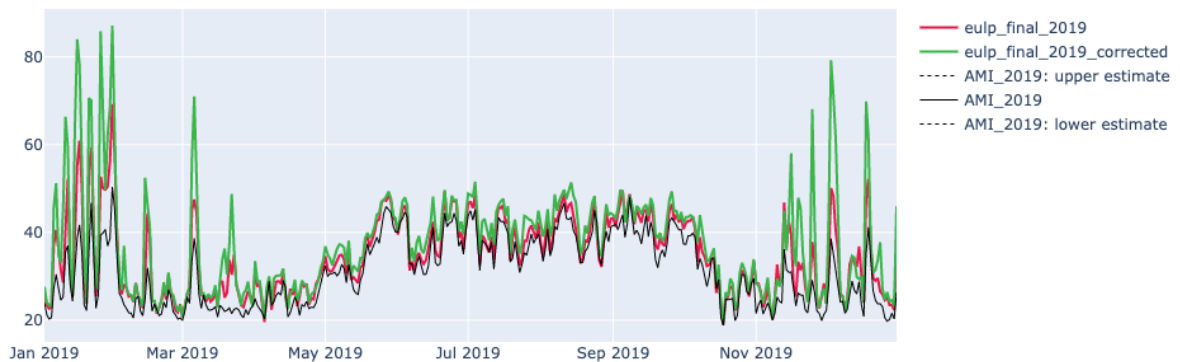


**Figure 380.** Comparison of ResStock modeled load to Seattle City Light residential AMI data for daily sums of electricity per unit (kWh/unit/day). ResStock load (“eulp\_final”) is shown before and after the EIA-861M-based heating and cooling correction model is applied (see Section 3.2.10). The correction model is driven by statewide monthly consumption differences; sometimes it improves the fit with AMI data and sometimes it worsens the fit.

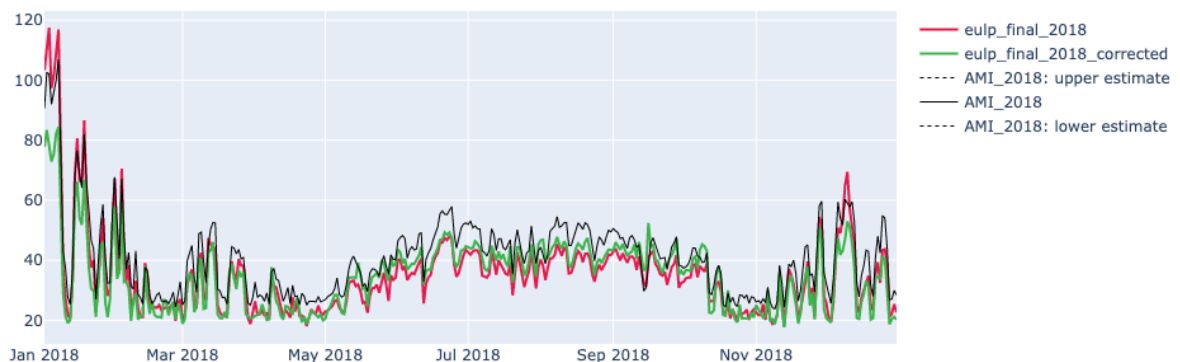




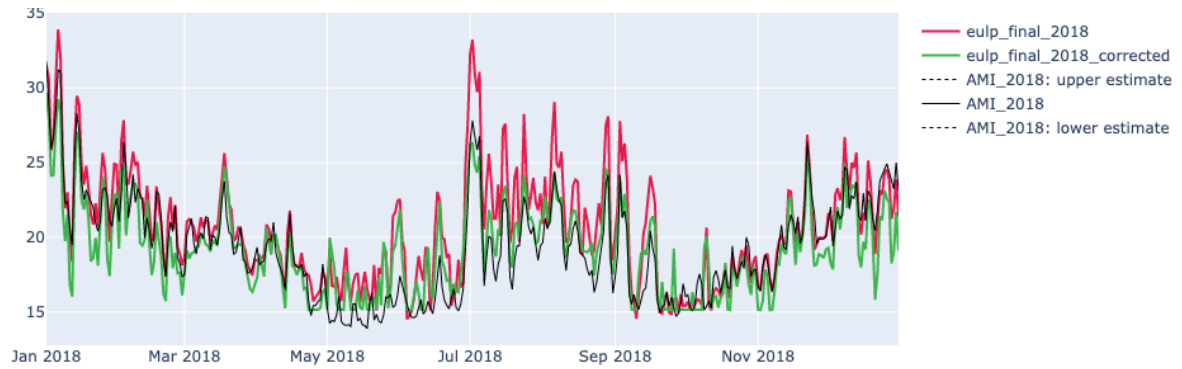
**Figure 381.** Comparison of ResStock modeled load to EPB Chattanooga residential AMI data for daily sums of electricity per unit (kWh/unit/day). ResStock load (“eulp\_final”) is shown before and after the EIA-861M-based heating and cooling correction model is applied (see Section 3.2.10). The correction model is driven by statewide monthly consumption differences; sometimes it improves the fit with AMI data and sometimes it worsens the fit.



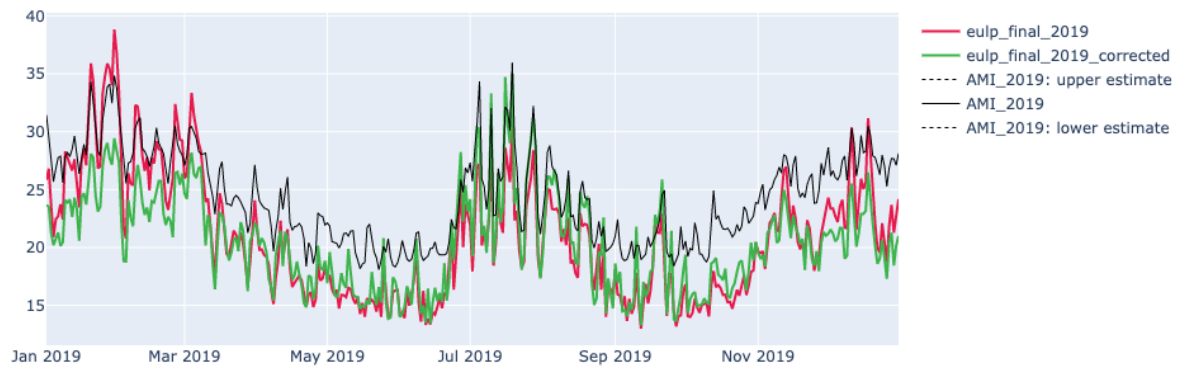
**Figure 382.** Comparison of ResStock modeled load to City of Tallahassee residential AMI data for daily sums of electricity per unit (kWh/unit/day). ResStock load (“eulp\_final”) is shown before and after the EIA-861M-based heating and cooling correction model is applied (see Section 3.2.10). The correction model is driven by statewide monthly consumption differences; sometimes it improves the fit with AMI data and sometimes it worsens the fit.



**Figure 383.** Comparison of ResStock modeled load to Horry Electric Cooperative residential AMI data for daily sums of electricity per unit (kWh/unit/day). ResStock load (“eulp\_final”) is shown before and after the EIA-861M-based heating and cooling correction model is applied (see Section 3.2.10). The correction model is driven by statewide monthly consumption differences; sometimes it improves the fit with AMI data and sometimes it worsens the fit.



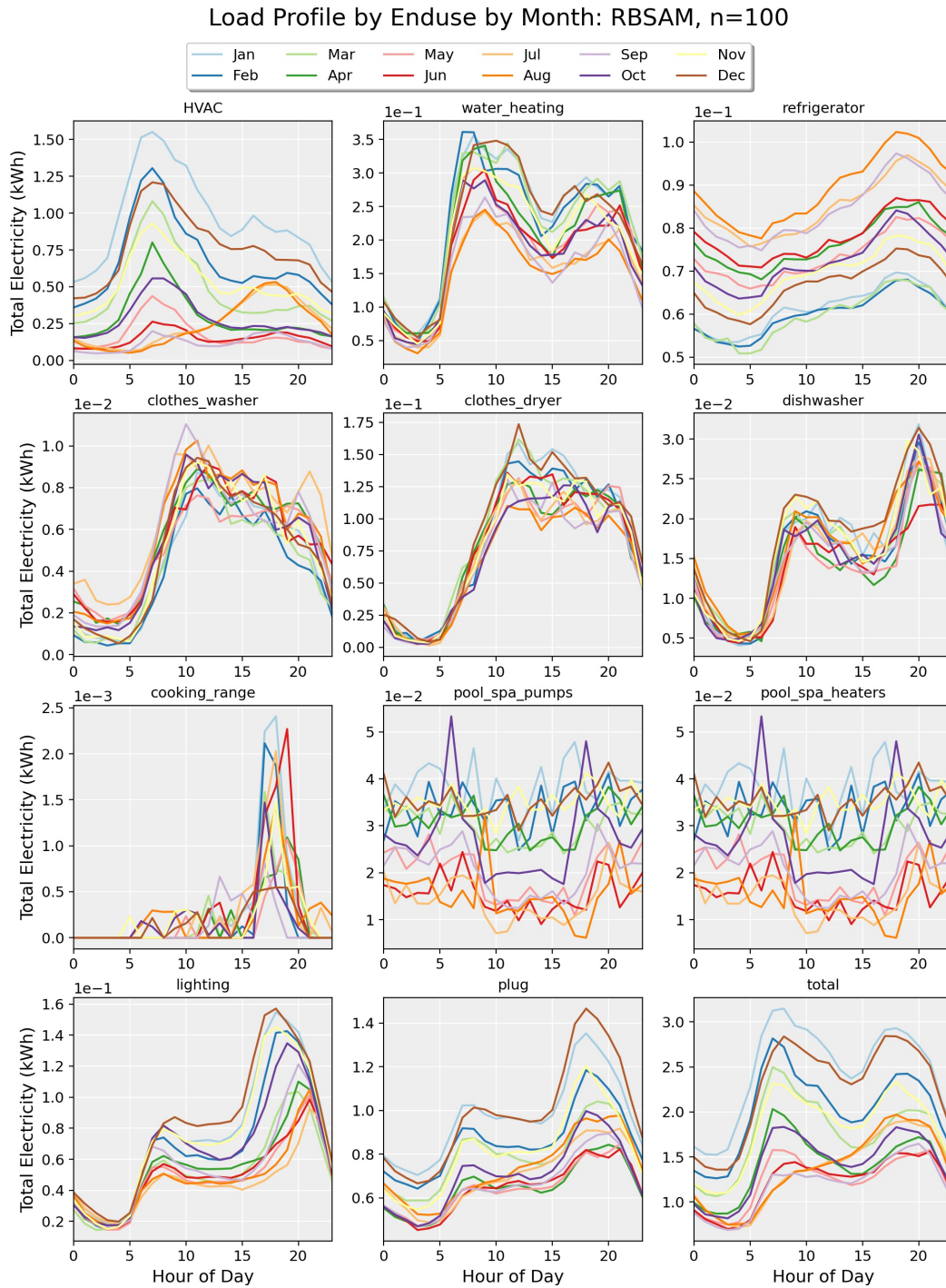
**Figure 384.** Comparison of ResStock modeled load to VEIC residential AMI data for daily sums of electricity per unit (kWh/unit/day). ResStock load (“eulp\_final”) is shown before and after the EIA-861M-based heating and cooling correction model is applied (see Section 3.2.10). The correction model is driven by statewide monthly consumption differences; sometimes it improves the fit with AMI data and sometimes it worsens the fit.

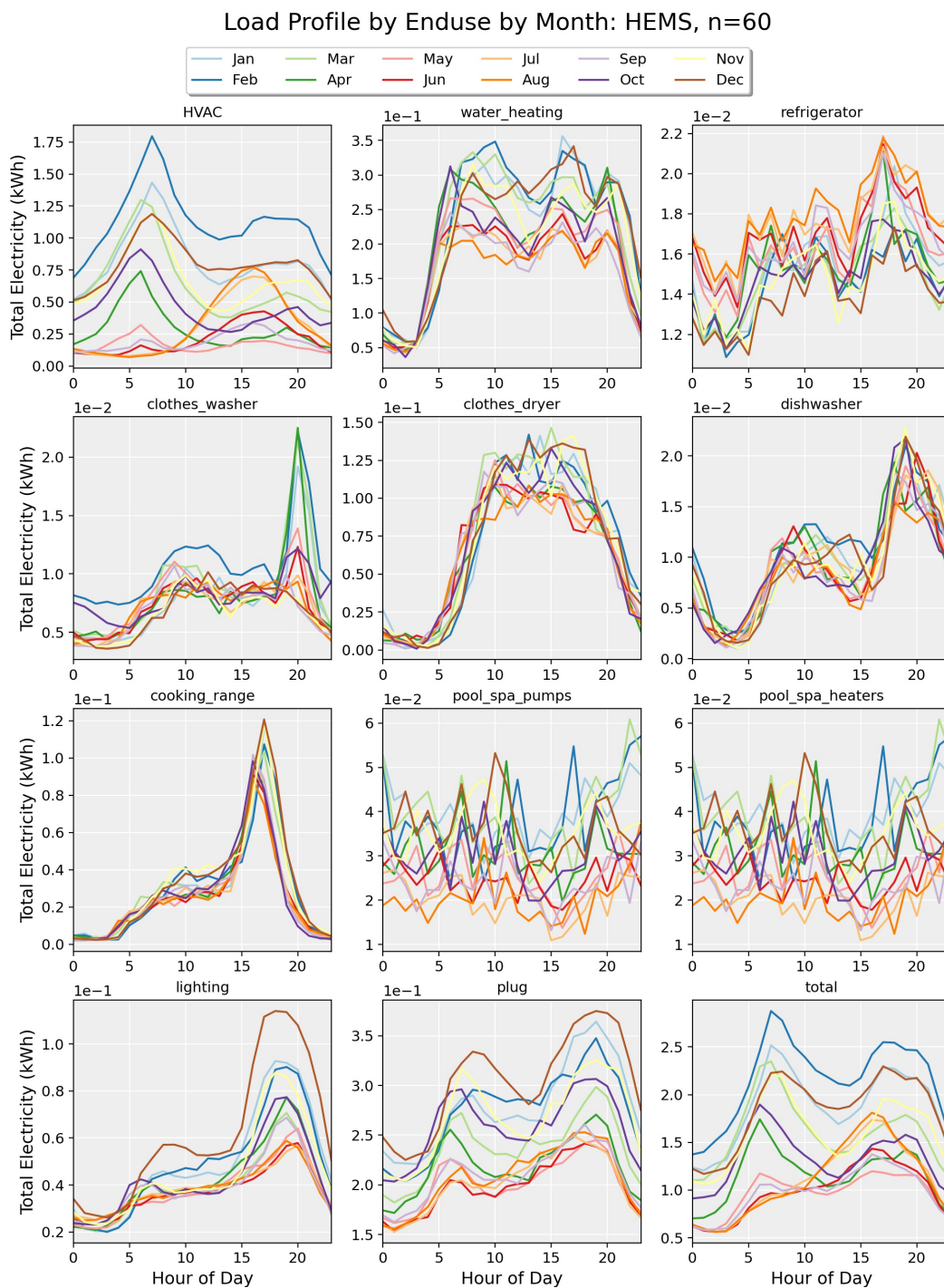


**Figure 385.** Comparison of ResStock modeled load to Cherryland Electric Co-op residential AMI data for daily sums of electricity per unit (kWh/unit/day). ResStock load (“eulp\_final”) is shown before and after the EIA-861M-based heating and cooling correction model is applied (see Section 3.2.10). The correction model is driven by statewide monthly consumption differences; sometimes it improves the fit with AMI data and sometimes it worsens the fit.

## Appendix D Supplemental Residential End-Use Dataset Analysis

This appendix contains supplemental figures from analysis of the five residential end use datasets. Each figure shows average 24-hour profiles for each month for each end use category measured. One can see that some end uses do not vary much from month to month (e.g., washer, dryer, dishwasher). Other end uses exhibit a lot of variation from month to month in either magnitude (e.g., refrigerator) or both shape and magnitude (e.g., pool pumps, lighting).







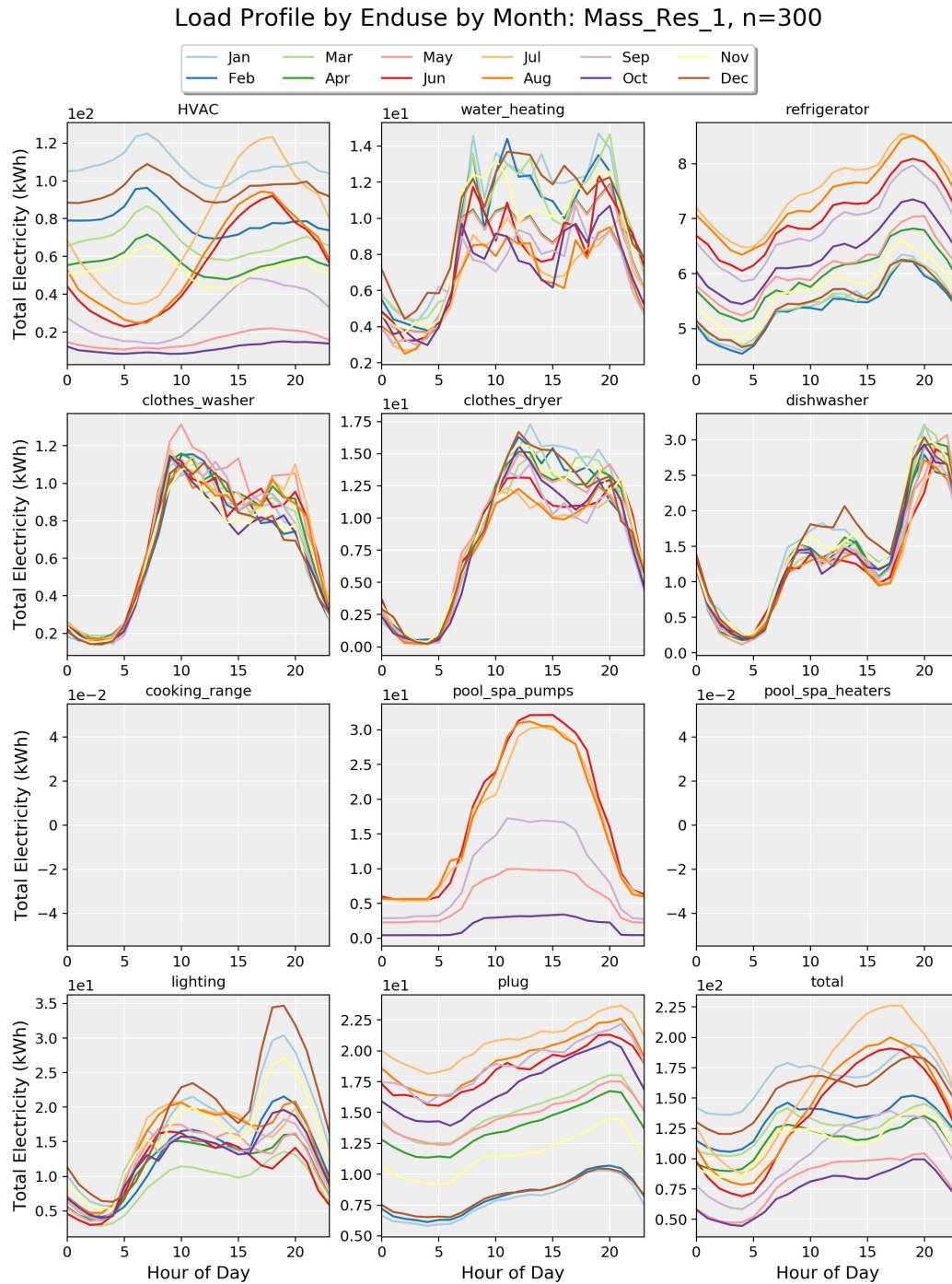


Figure 388. Submetered EULPs by month—Mass. RES 1

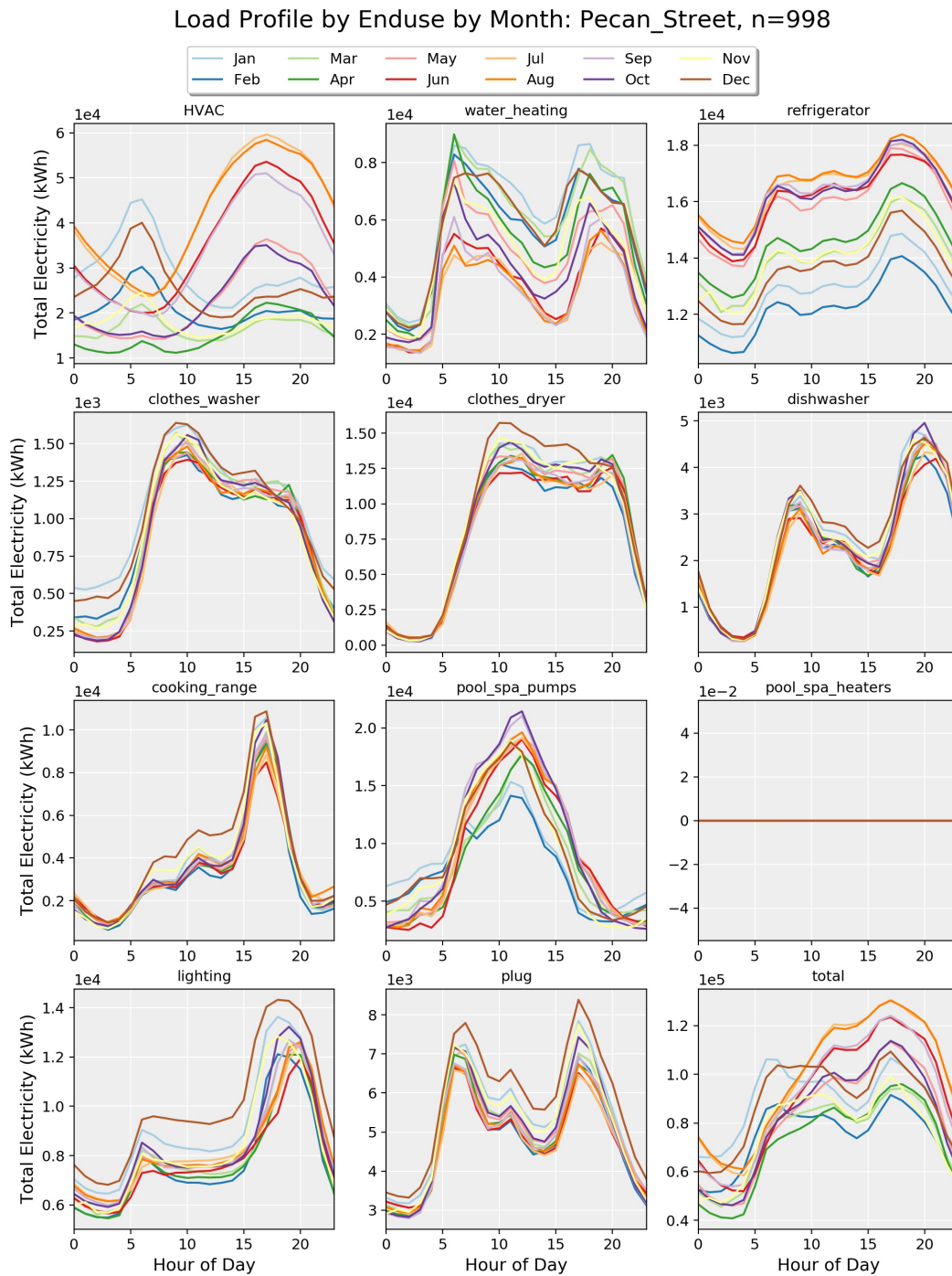
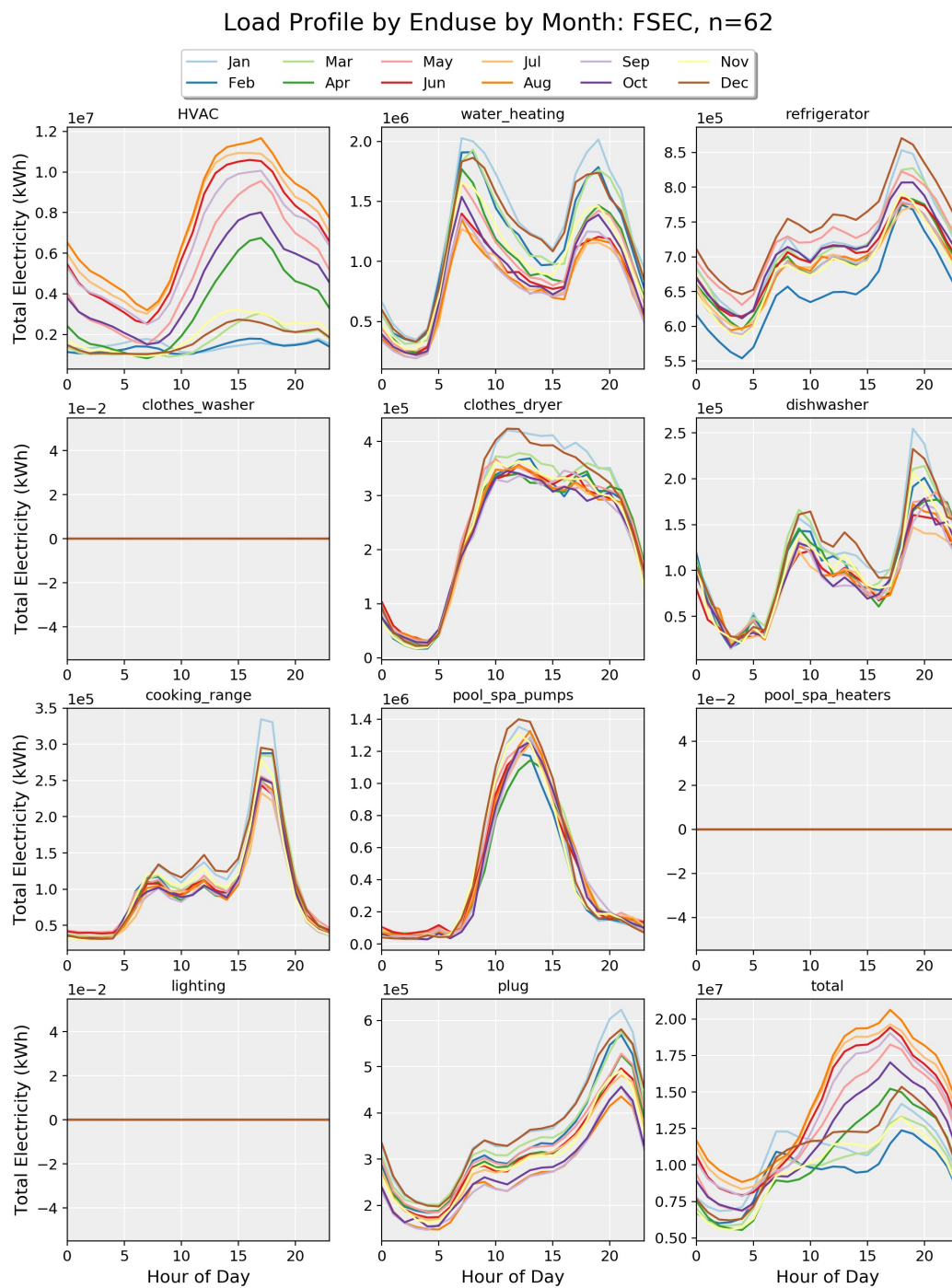


Figure 389. Submetered EULPs by month—Pecan Street

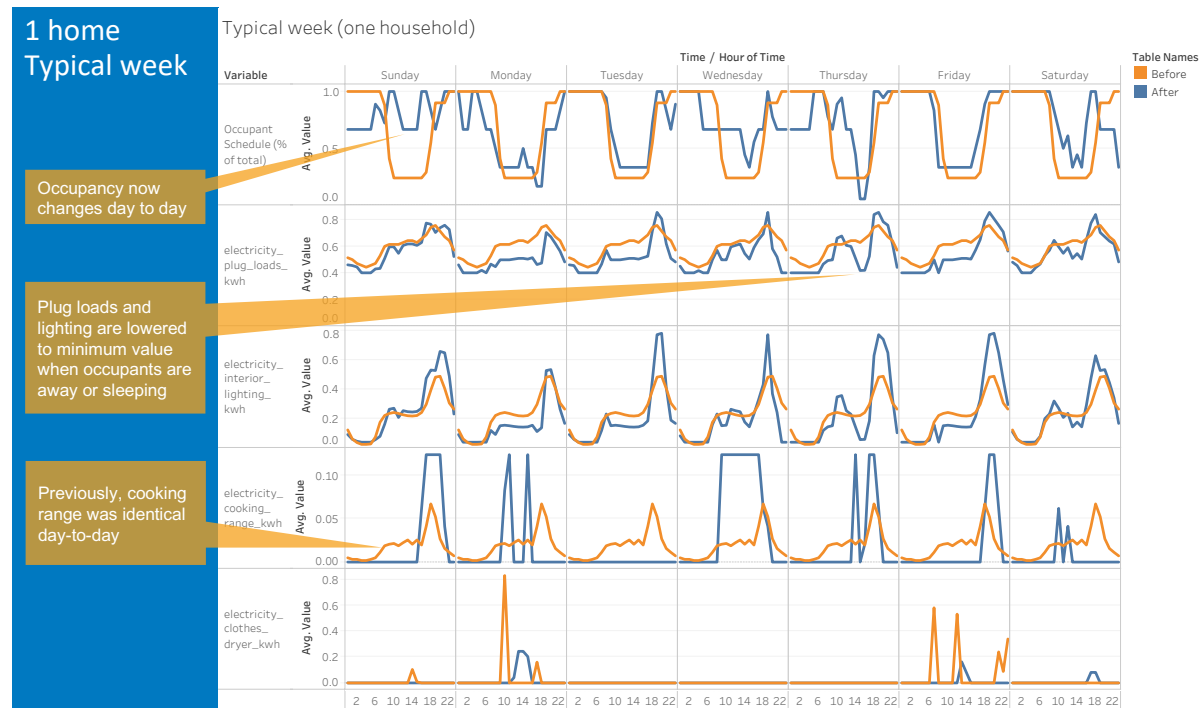




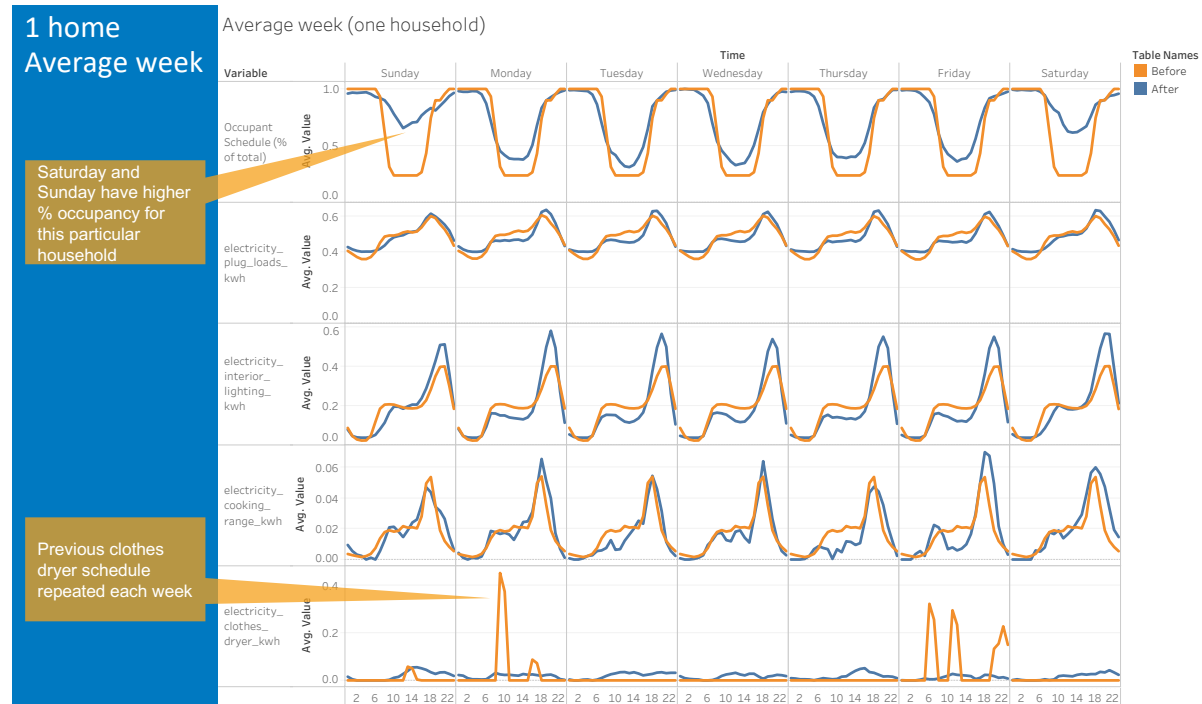
**Figure 390. Submetered EULPs by month—FSEC**

## **Appendix E Supplemental Figures for Residential Stochastic Occupant-Driven Loads**

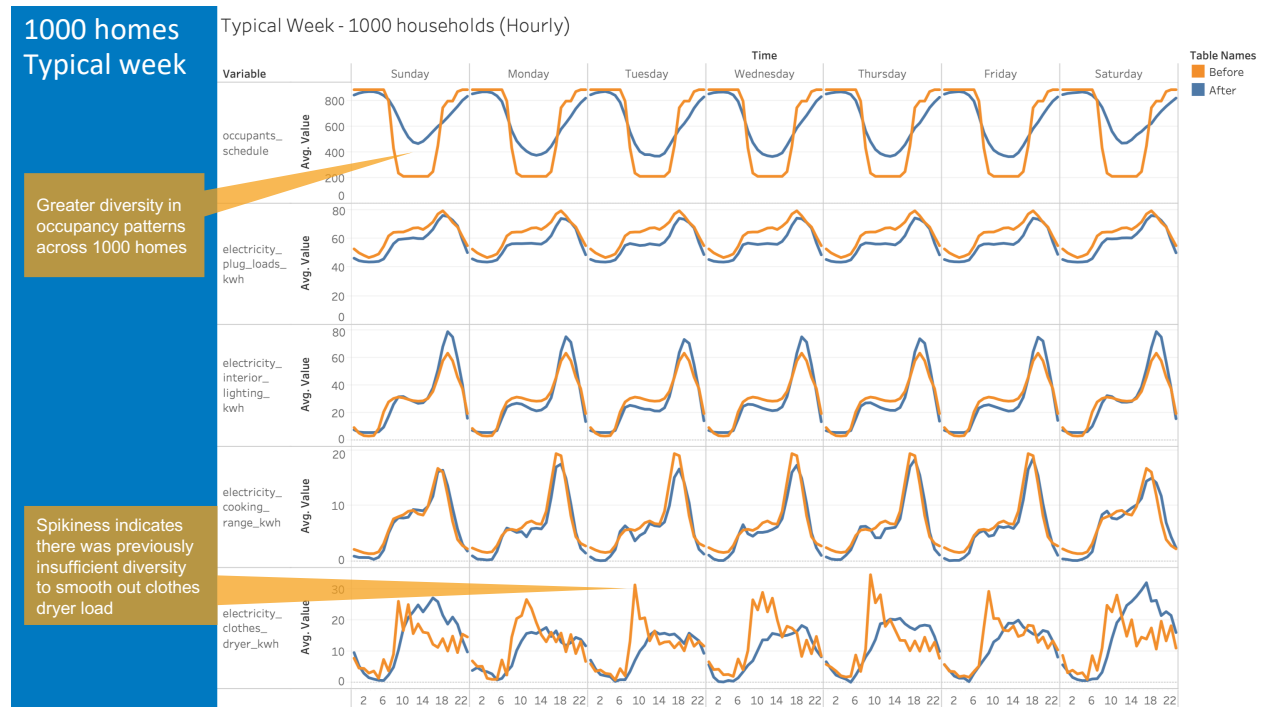
This appendix contains supplemental figures showing the impact of the new residential stochastic schedule generator. They show both typical and average weeks for one and 1,000 homes for several electric end uses and five domestic hot water schedules.



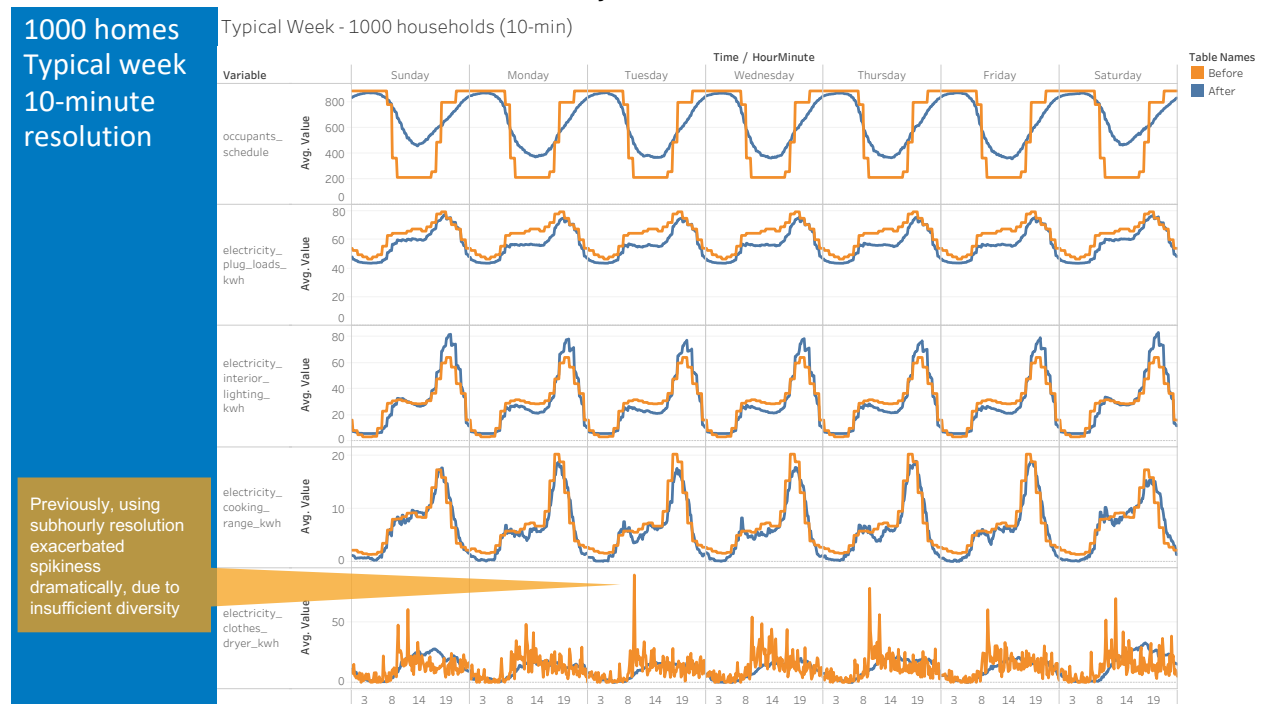
**Figure 391. Impact of new residential schedule generator on selected EULPs: one home, typical week**



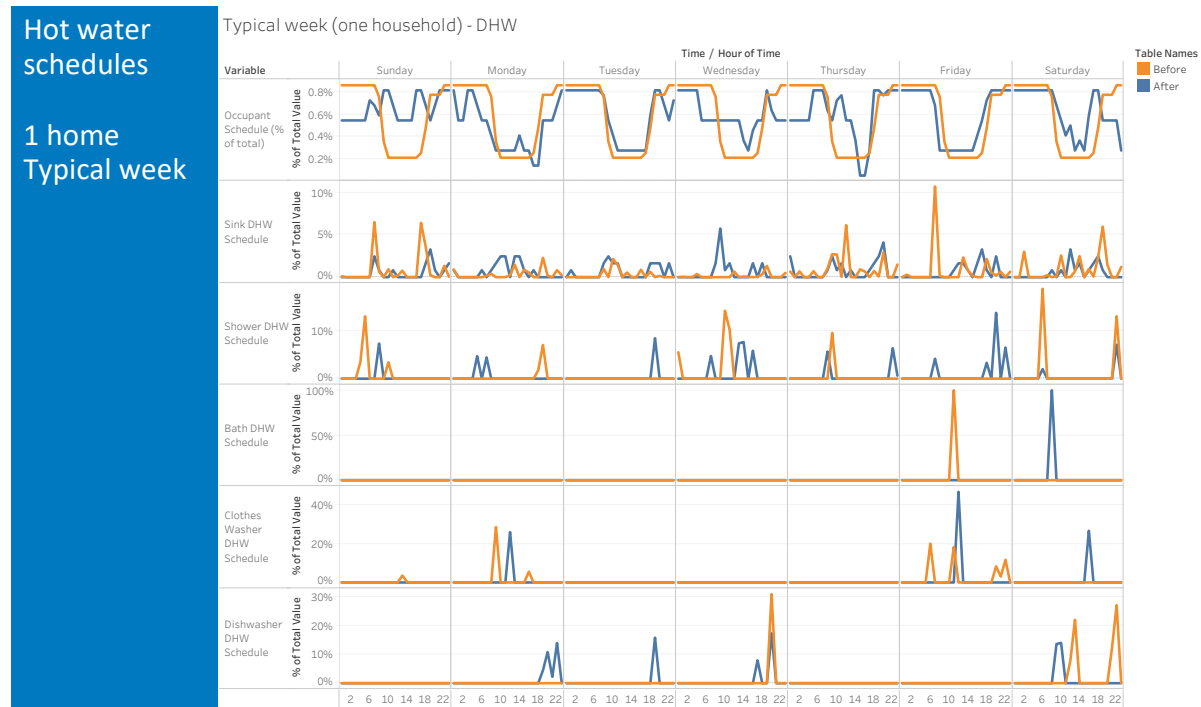
**Figure 392. Impact of new residential schedule generator on selected EULPs: one home, average week**



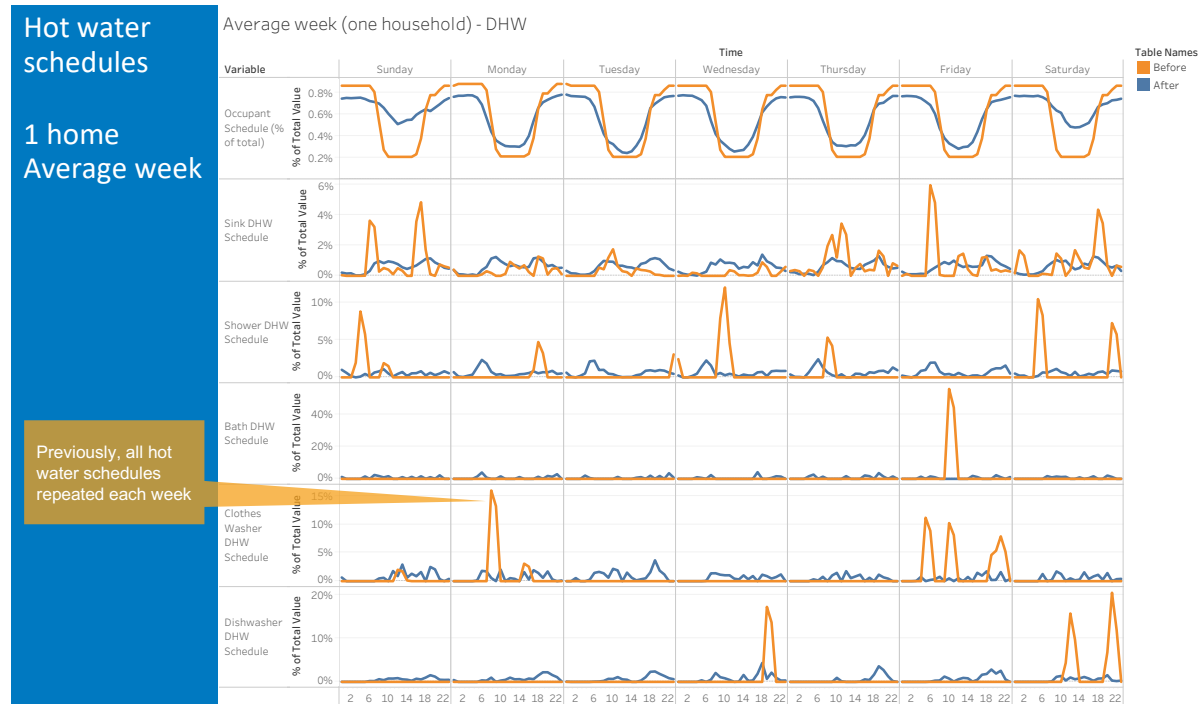
**Figure 393. Impact of new residential schedule generator on selected EULPs: 1,000 homes, typical week, hourly resolution**



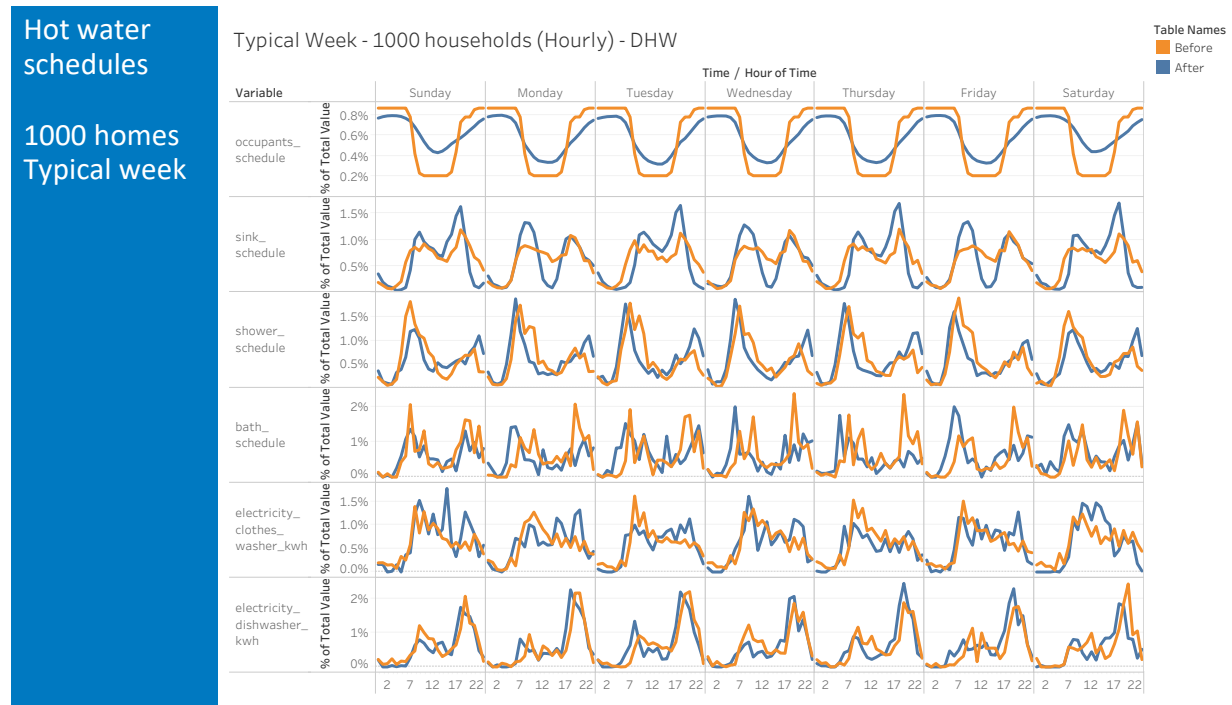
**Figure 394. Impact of new residential schedule generator on selected EULPs: 1,000 homes, typical week, 10-minute resolution**



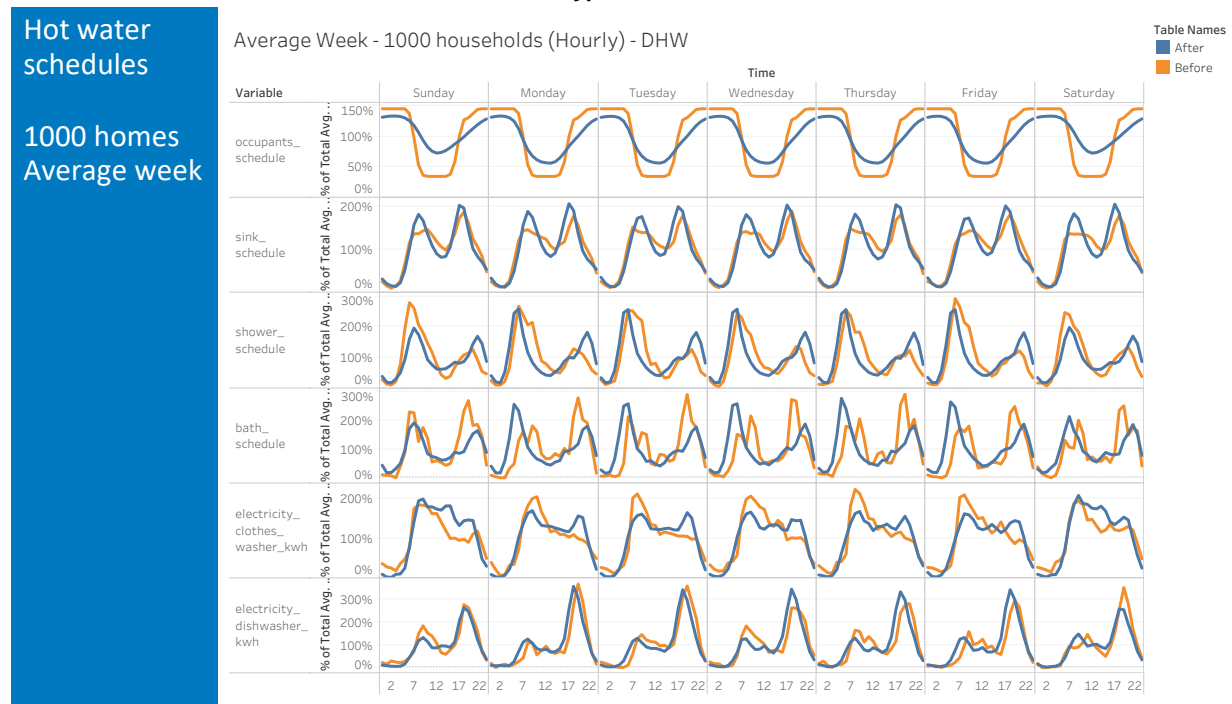
**Figure 395. Impact of new residential schedule generator on domestic hot water schedules: one home, typical week**



**Figure 396. Impact of new residential schedule generator on domestic hot water schedules: one home, average week**



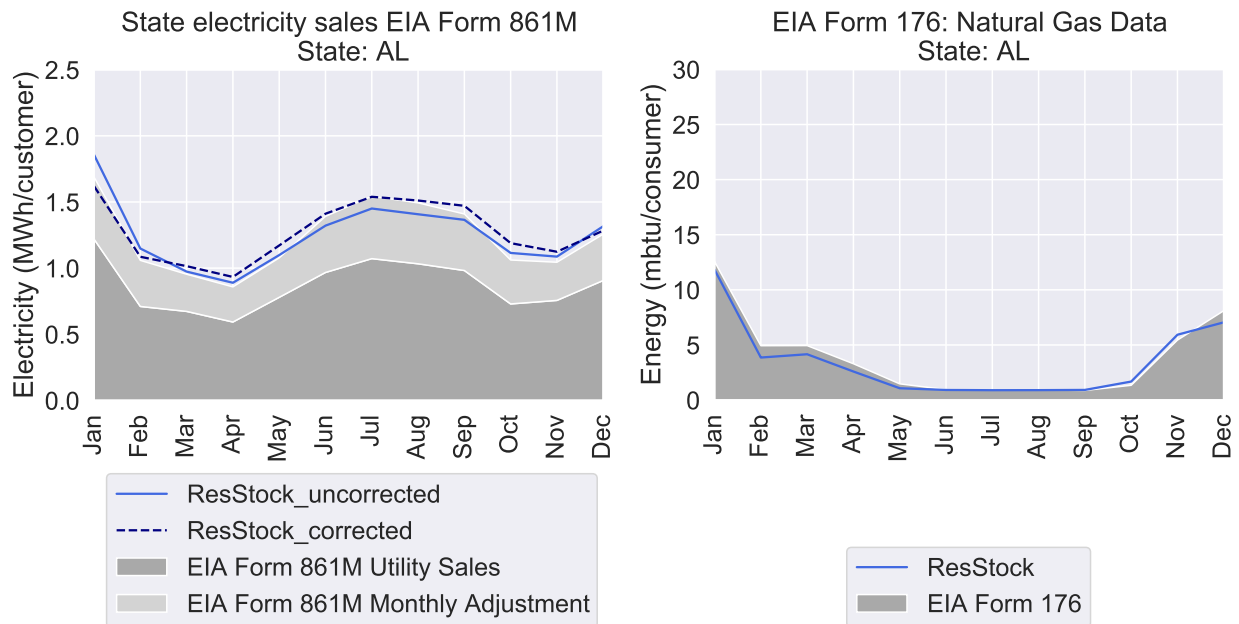
**Figure 397. Impact of new residential schedule generator on domestic hot water schedules: 1,000 homes, typical week**



**Figure 398. Impact of new residential schedule generator on domestic hot water schedules: 1,000 homes, average week**

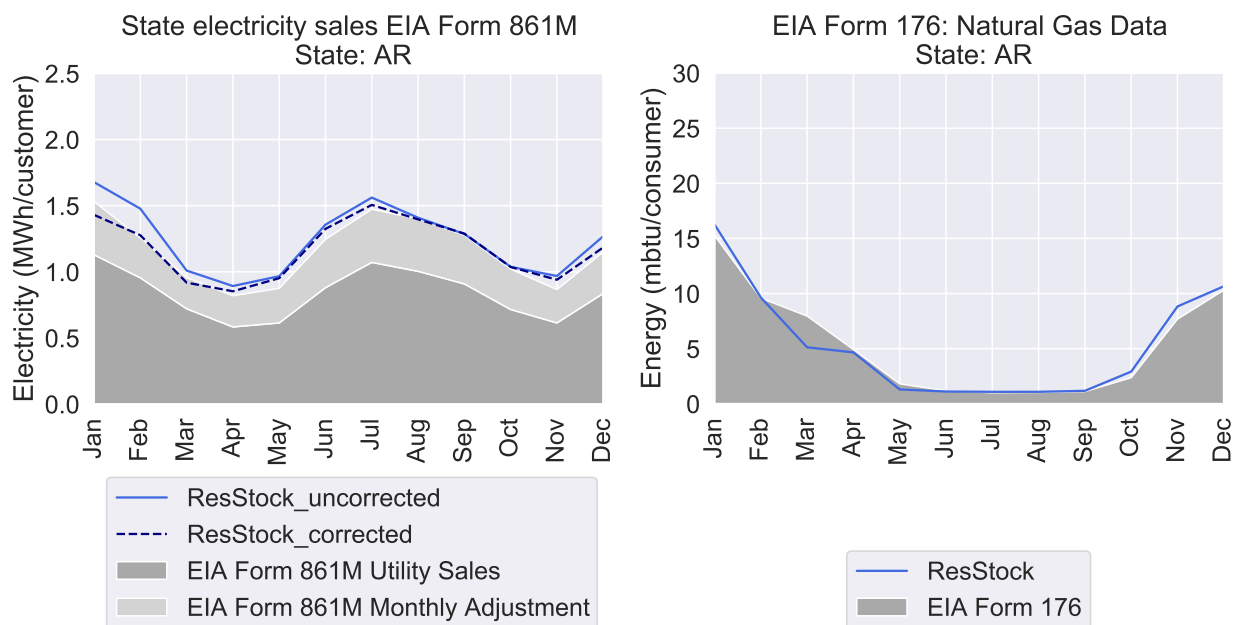
## Appendix F Residential EIA Monthly Electricity and Natural Gas Consumption Comparisons by State

This appendix contains comparisons of ResStock to EIA monthly electricity and monthly natural gas data for each state. The electricity comparisons contain data from 2018 EIA Form 861M. The natural gas comparisons contain data from 2018 data from EIA Form 176. The data displayed for EIA form 861M contains the total utility reported monthly sales by state and the monthly state adjustment supplied by EIA.

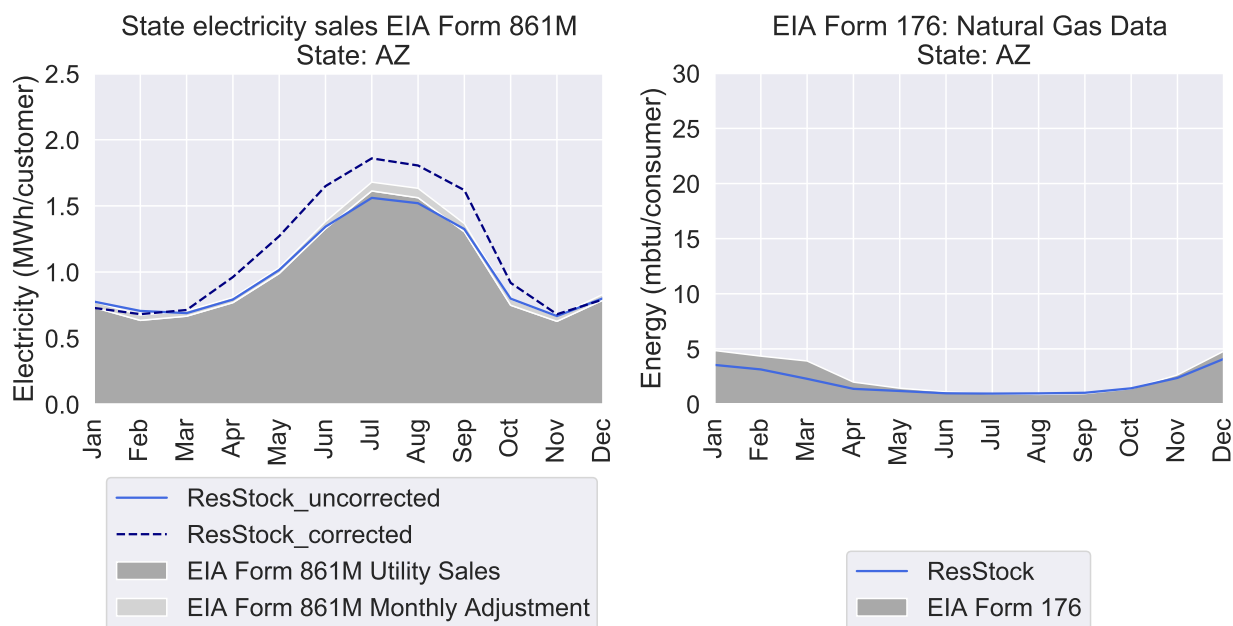


**Figure 399. ResStock AL monthly electric sales before and after correction compared to 2018 AL sales reported in EIA Form 861M (left). ResStock AL monthly natural gas energy compared to 2018 AL natural gas energy reported in EIA Form 176 (right).**

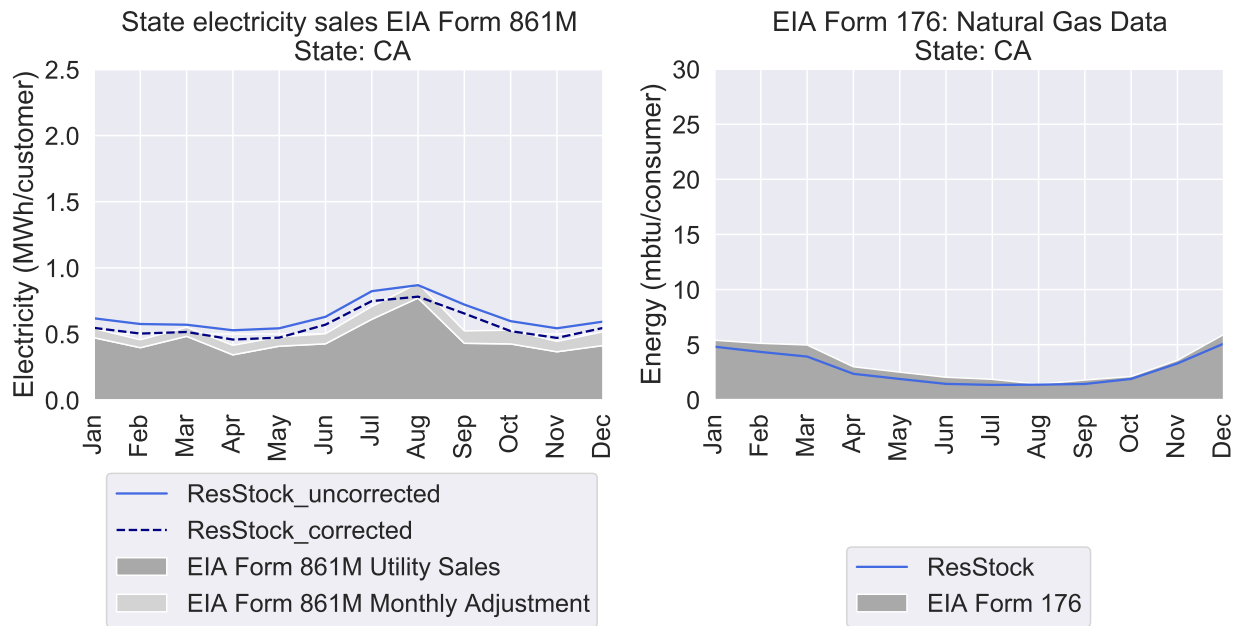




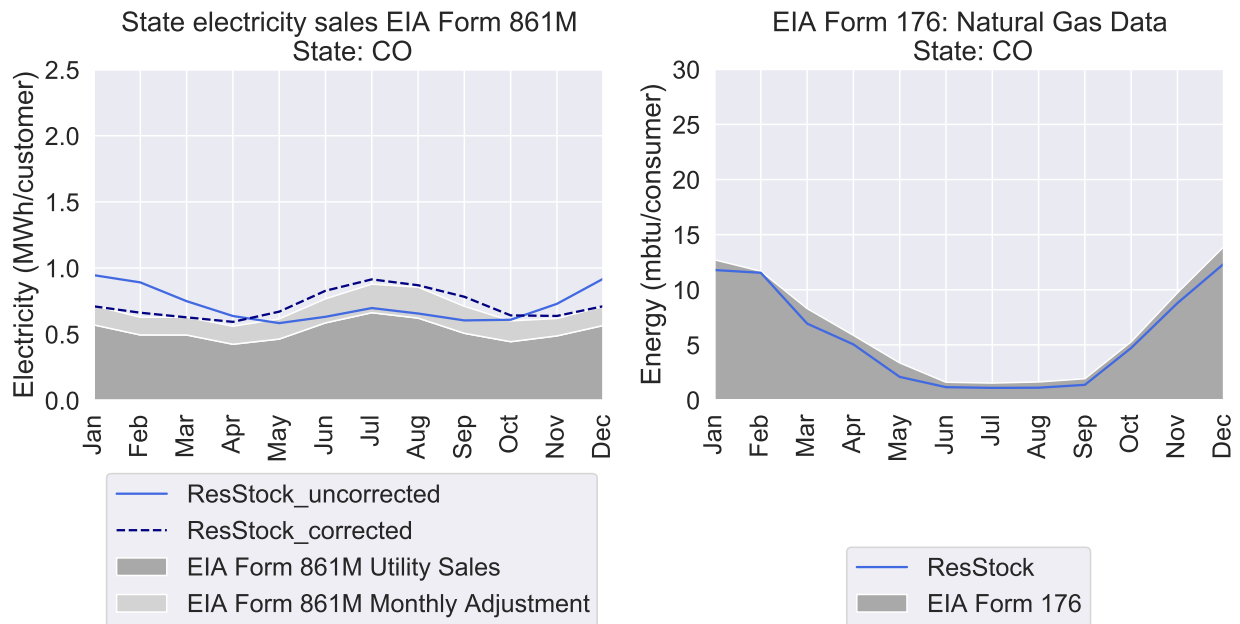
**Figure 400. ResStock AR monthly electric sales before and after correction compared to 2018 AR sales reported in EIA Form 861M (left). ResStock AR monthly natural gas energy compared to 2018 AR natural gas energy reported in EIA Form 176 (right).**



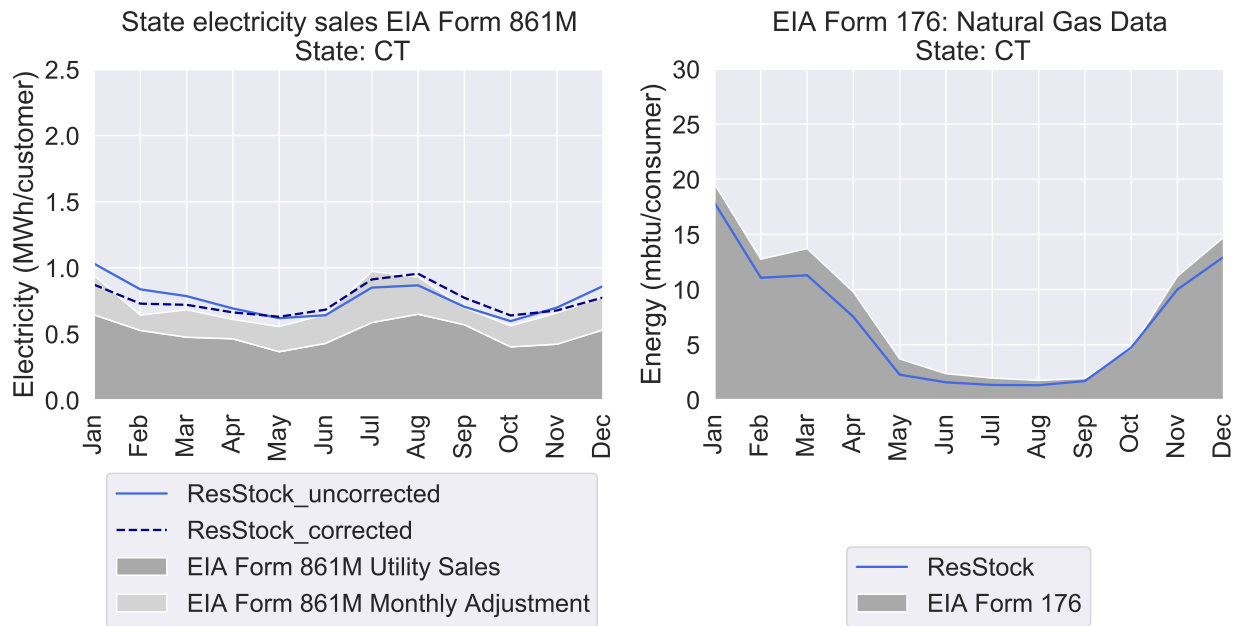
**Figure 401. ResStock AZ monthly electric sales before and after correction compared to 2018 AZ sales reported in EIA Form 861M (left). ResStock AZ monthly natural gas energy compared to 2018 AZ natural gas energy reported in EIA Form 176 (right).**



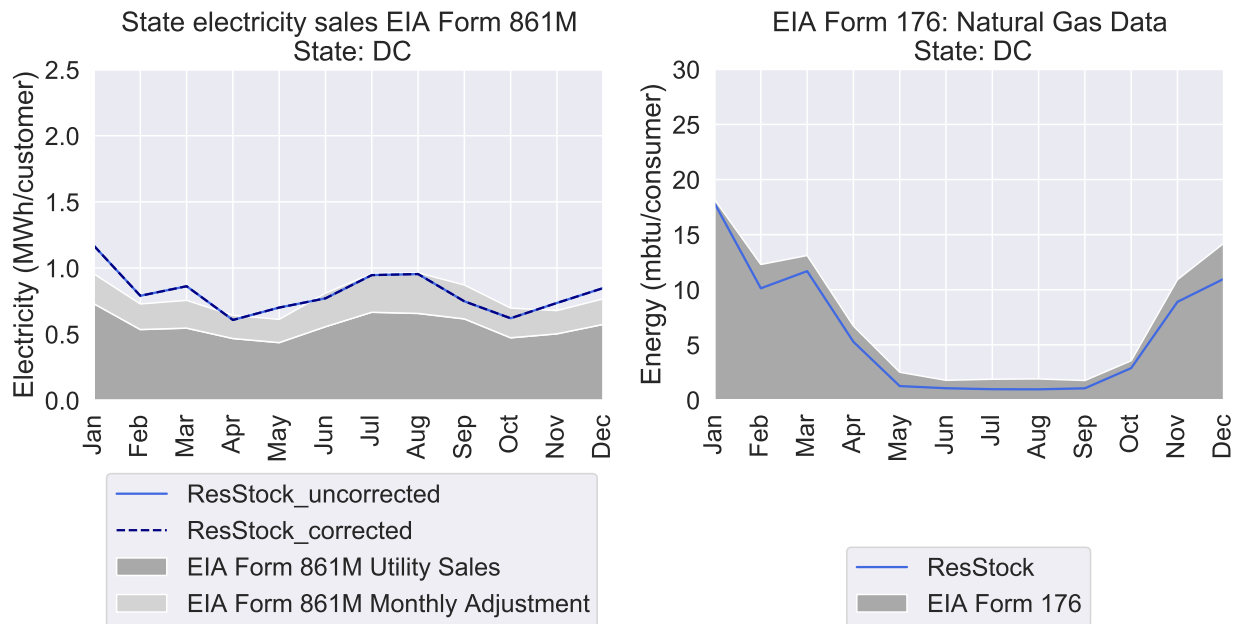
**Figure 402. ResStock CA monthly electric sales before and after correction compared to 2018 CA sales reported in EIA Form 861M (left). ResStock CA monthly natural gas energy compared to 2018 CA natural gas energy reported in EIA Form 176 (right).**



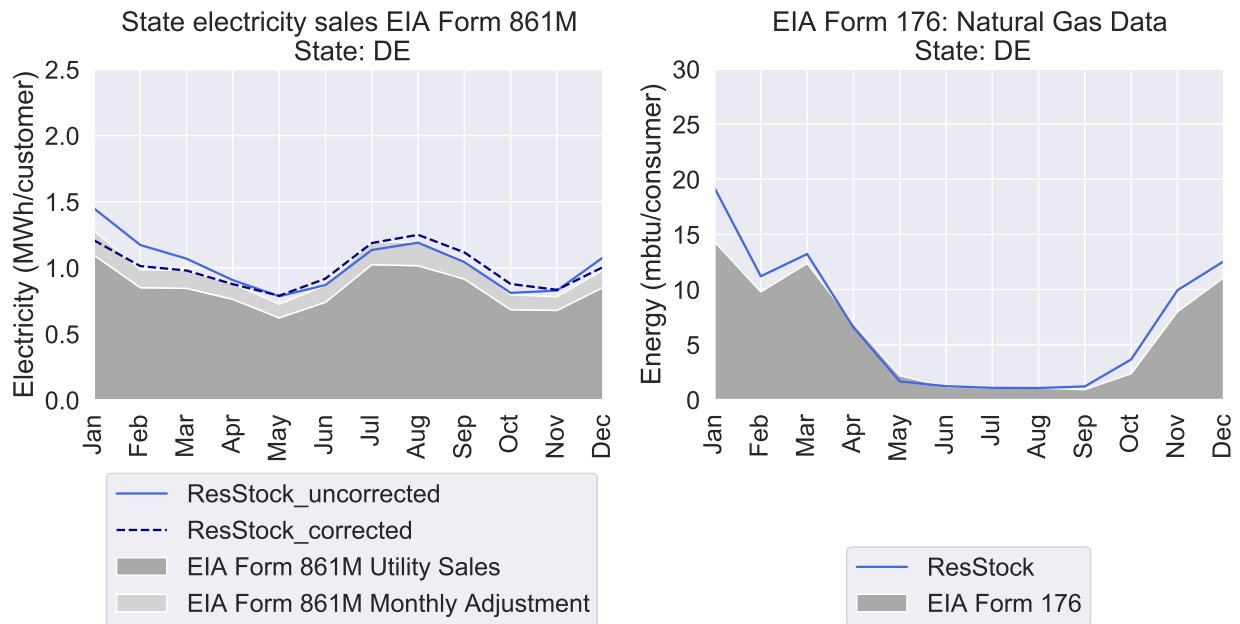
**Figure 403. ResStock CO monthly electric sales before and after correction compared to 2018 CO sales reported in EIA Form 861M (left). ResStock CO monthly natural gas energy compared to 2018 CO natural gas energy reported in EIA Form 176 (right).**



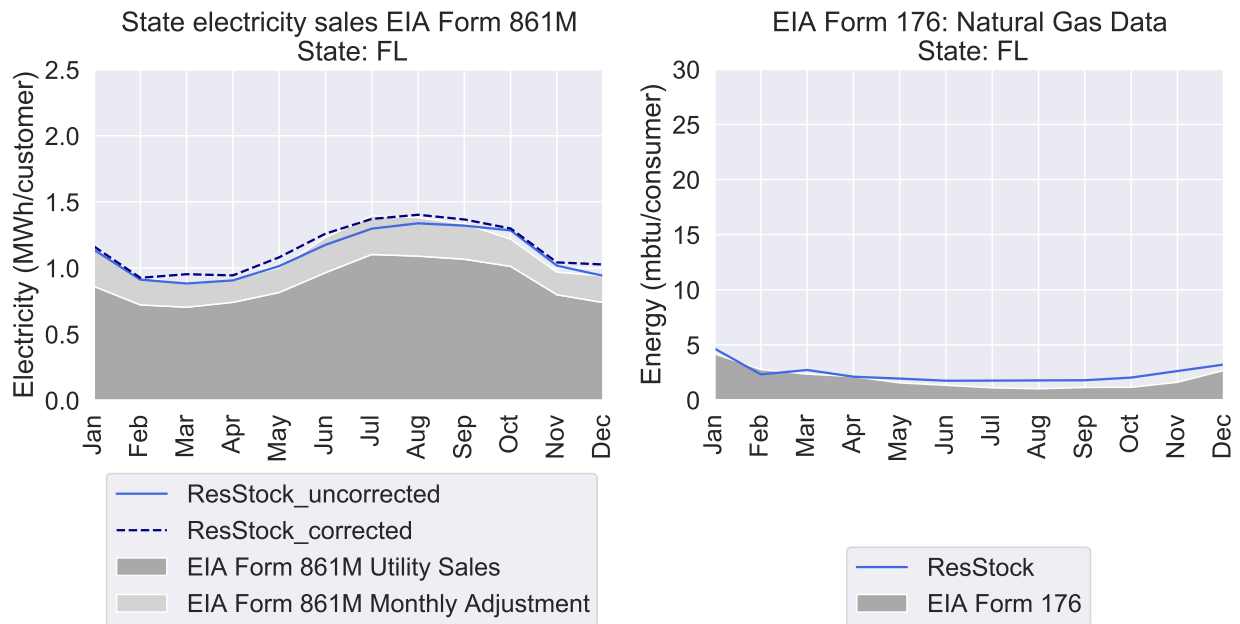
**Figure 404. ResStock CT monthly electric sales before and after correction compared to 2018 CT sales reported in EIA Form 861M (left). ResStock CT monthly natural gas energy compared to 2018 CT natural gas energy reported in EIA Form 176 (right).**



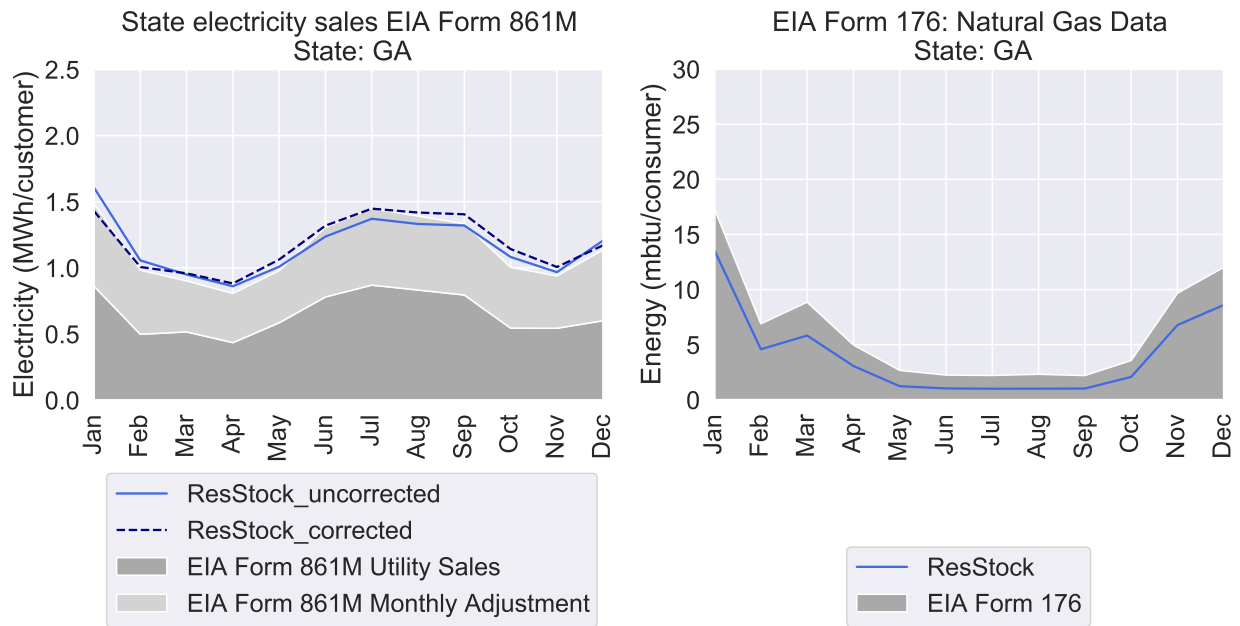
**Figure 405. ResStock DC monthly electric sales before and after correction compared to 2018 DC sales reported in EIA Form 861M (left). ResStock DC monthly natural gas energy compared to 2018 DC natural gas energy reported in EIA Form 176 (right).**



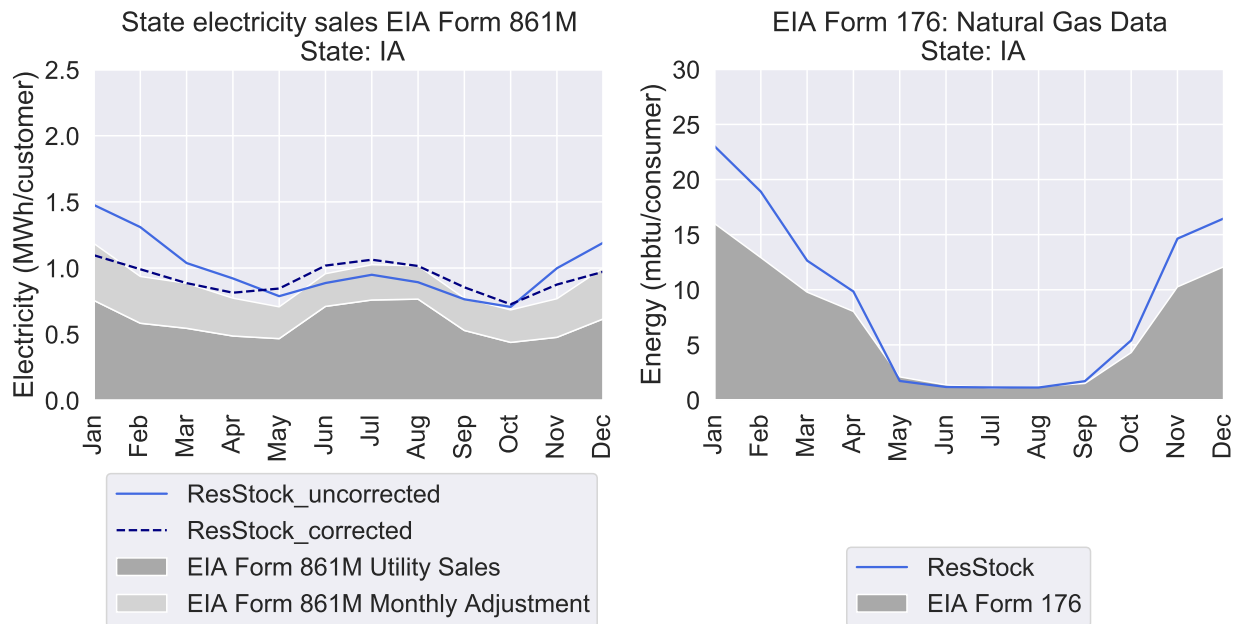
**Figure 406. ResStock DE monthly electric sales before and after correction compared to 2018 DE sales reported in EIA Form 861M (left). ResStock DE monthly natural gas energy compared to 2018 DE natural gas energy reported in EIA Form 176 (right).**



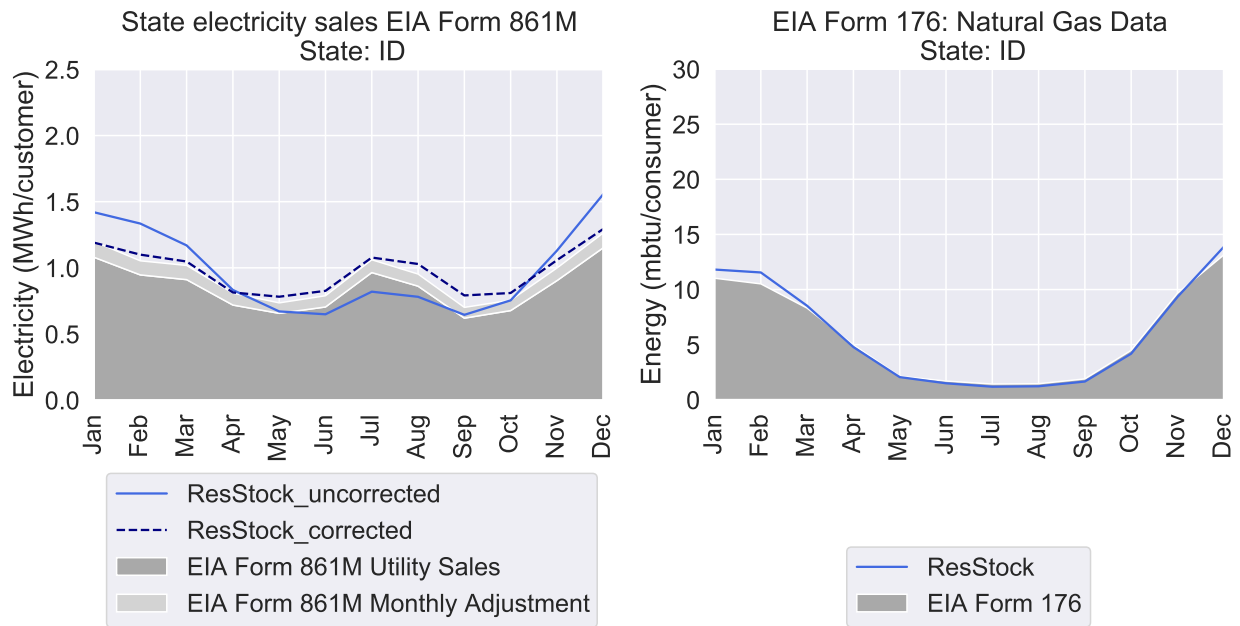
**Figure 407. ResStock FL monthly electric sales before and after correction compared to 2018 FL sales reported in EIA Form 861M (left). ResStock FL monthly natural gas energy compared to 2018 FL natural gas energy reported in EIA Form 176 (right).**



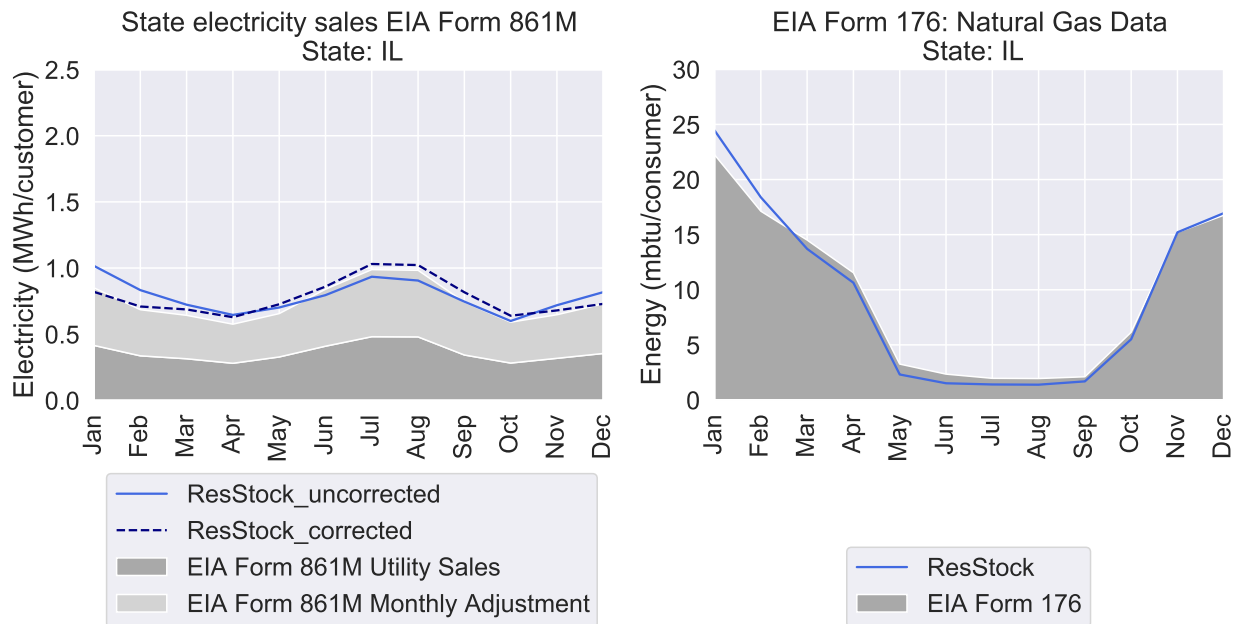
**Figure 408. ResStock GA monthly electric sales before and after correction compared to 2018 GA sales reported in EIA Form 861M (left). ResStock GA monthly natural gas energy compared to 2018 GA natural gas energy reported in EIA Form 176 (right).**



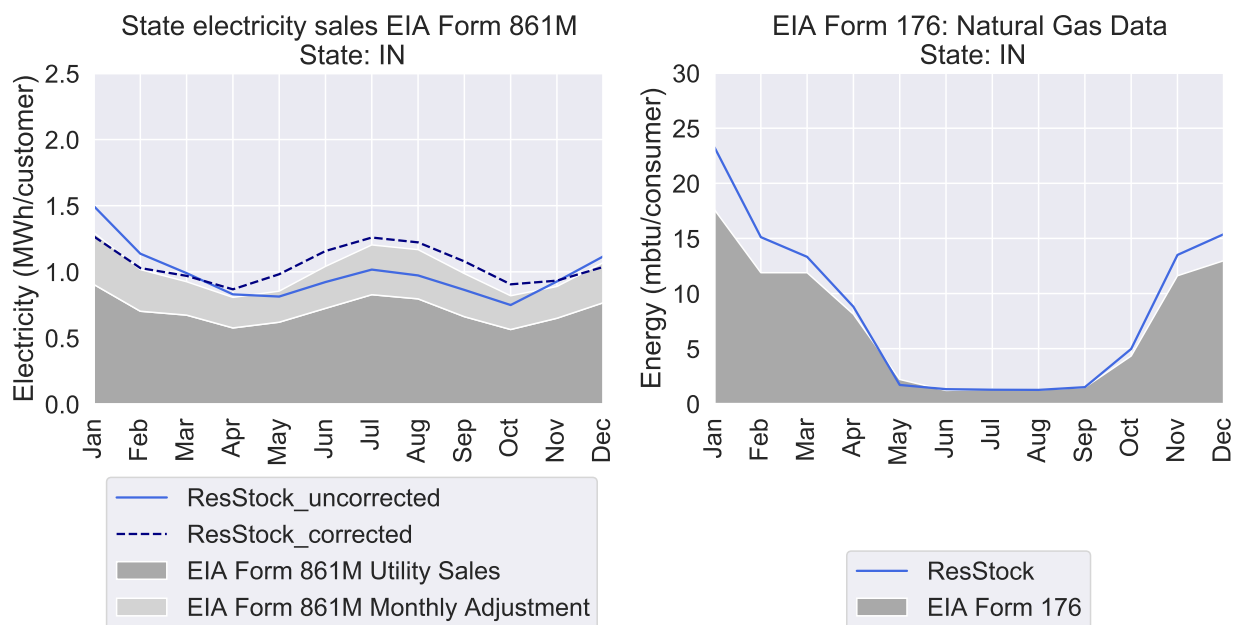
**Figure 409. ResStock IA monthly electric sales before and after correction compared to 2018 IA sales reported in EIA Form 861M (left). ResStock IA monthly natural gas energy compared to 2018 IA natural gas energy reported in EIA Form 176 (right).**



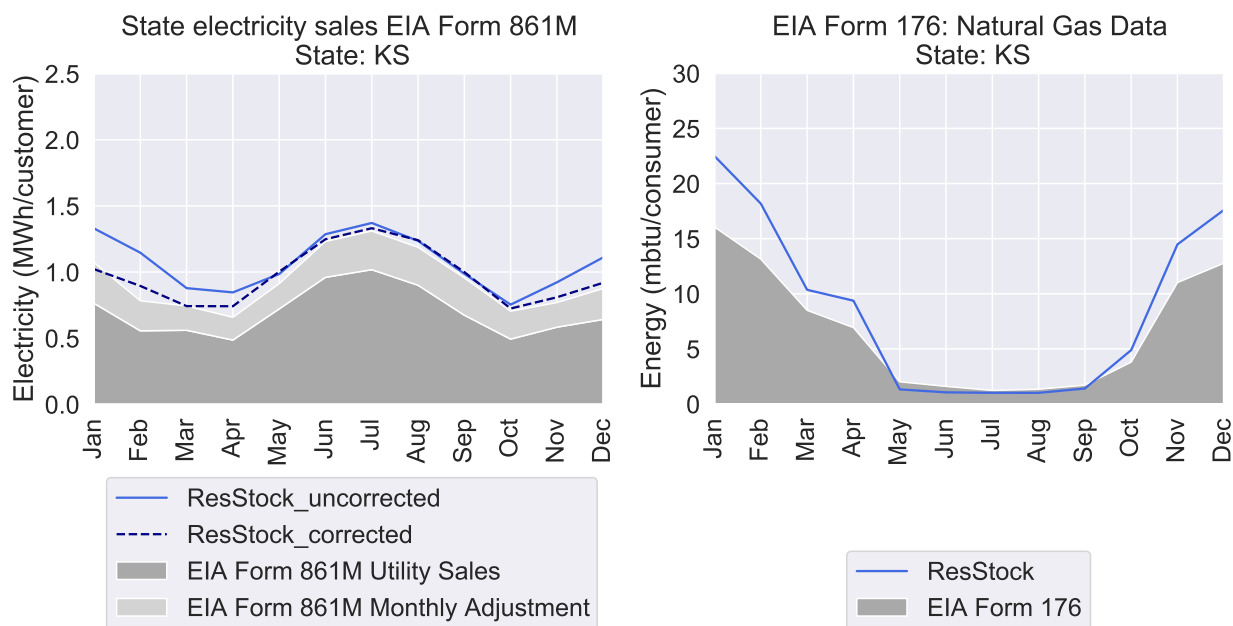
**Figure 410. ResStock ID monthly electric sales before and after correction compared to 2018 ID sales reported in EIA Form 861M (left). ResStock ID monthly natural gas energy compared to 2018 ID natural gas energy reported in EIA Form 176 (right).**



**Figure 411. ResStock IL monthly electric sales before and after correction compared to 2018 IL sales reported in EIA Form 861M (left). ResStock IL monthly natural gas energy compared to 2018 IL natural gas energy reported in EIA Form 176 (right).**

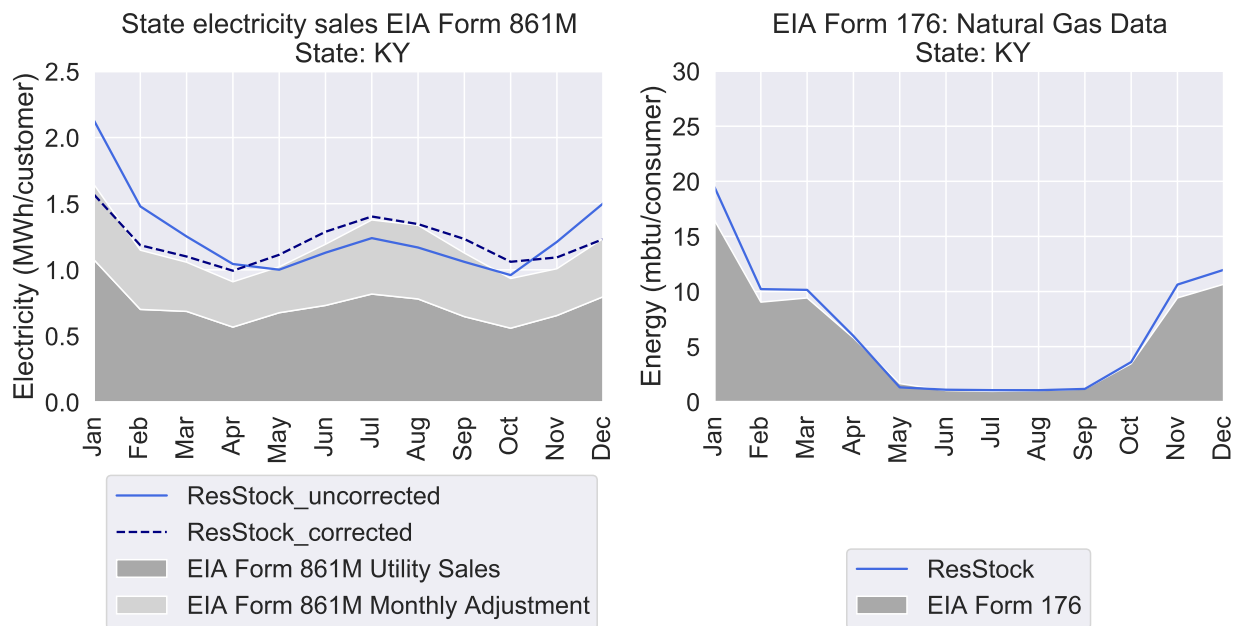


**Figure 412. ResStock IN monthly electric sales before and after correction compared to 2018 IN sales reported in EIA Form 861M (left). ResStock IN monthly natural gas energy compared to 2018 IN natural gas energy reported in EIA Form 176 (right).**

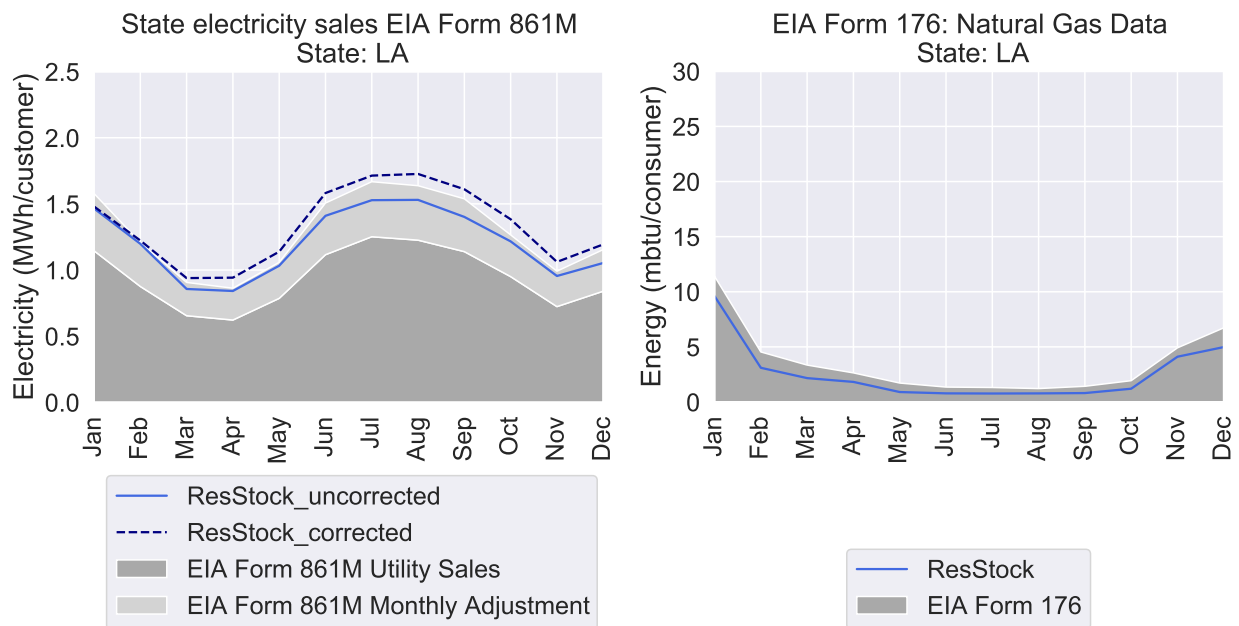


**Figure 413. ResStock KS monthly electric sales before and after correction compared to 2018 KS sales reported in EIA Form 861M (left). ResStock KS monthly natural gas energy compared to 2018 KS natural gas energy reported in EIA Form 176 (right).**

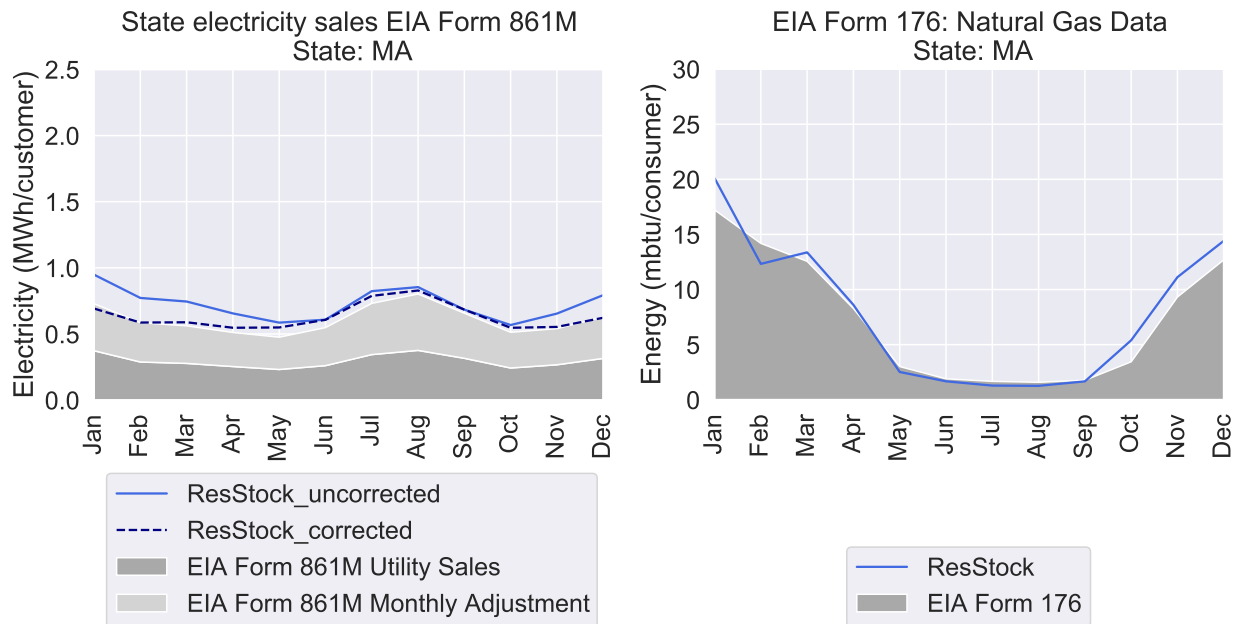




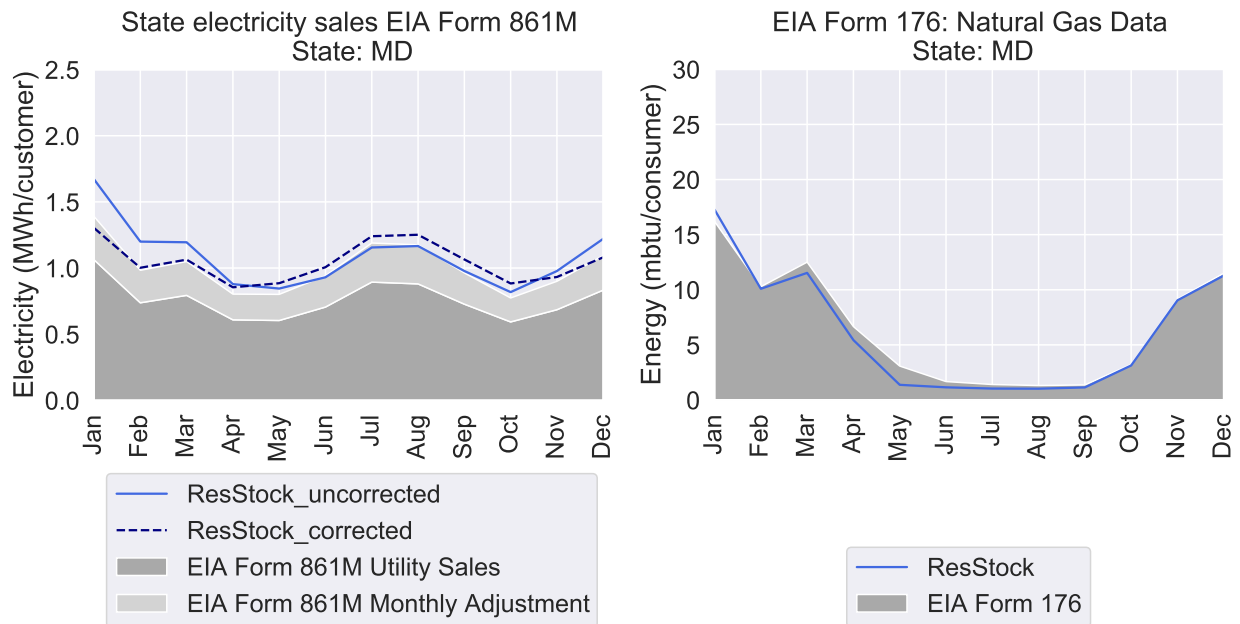
**Figure 414. ResStock KY monthly electric sales before and after correction compared to 2018 KY sales reported in EIA Form 861M (left). ResStock KY monthly natural gas energy compared to 2018 KY natural gas energy reported in EIA Form 176 (right).**



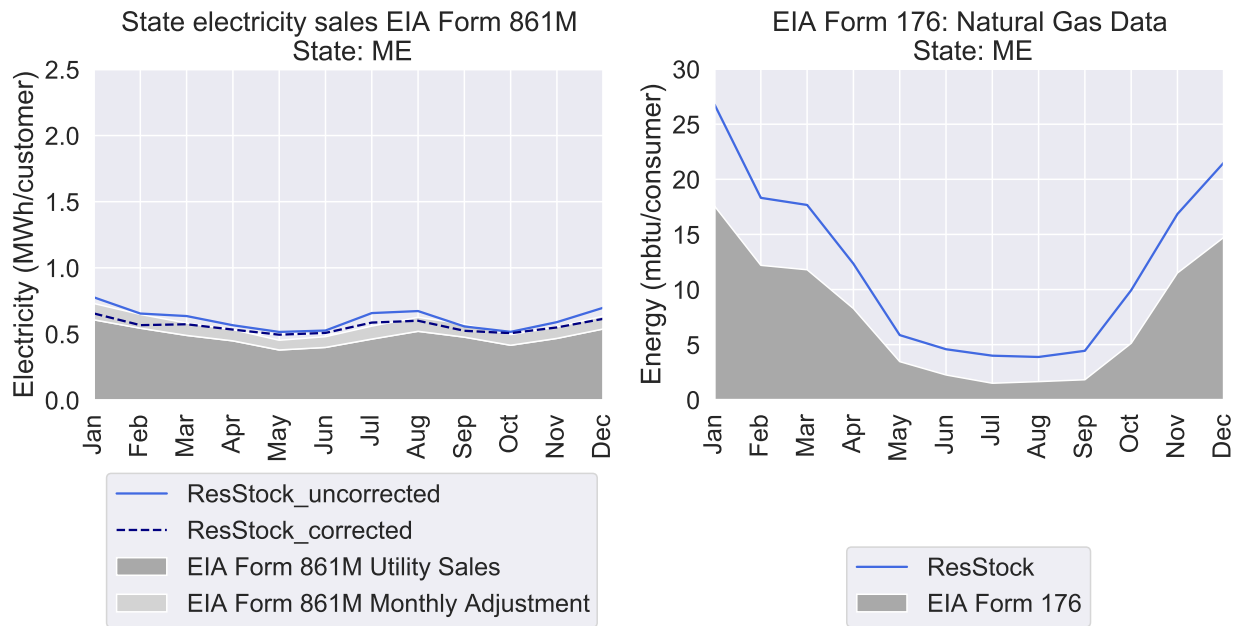
**Figure 415. ResStock LA monthly electric sales before and after correction compared to 2018 LA sales reported in EIA Form 861M (left). ResStock LA monthly natural gas energy compared to 2018 LA natural gas energy reported in EIA Form 176 (right).**



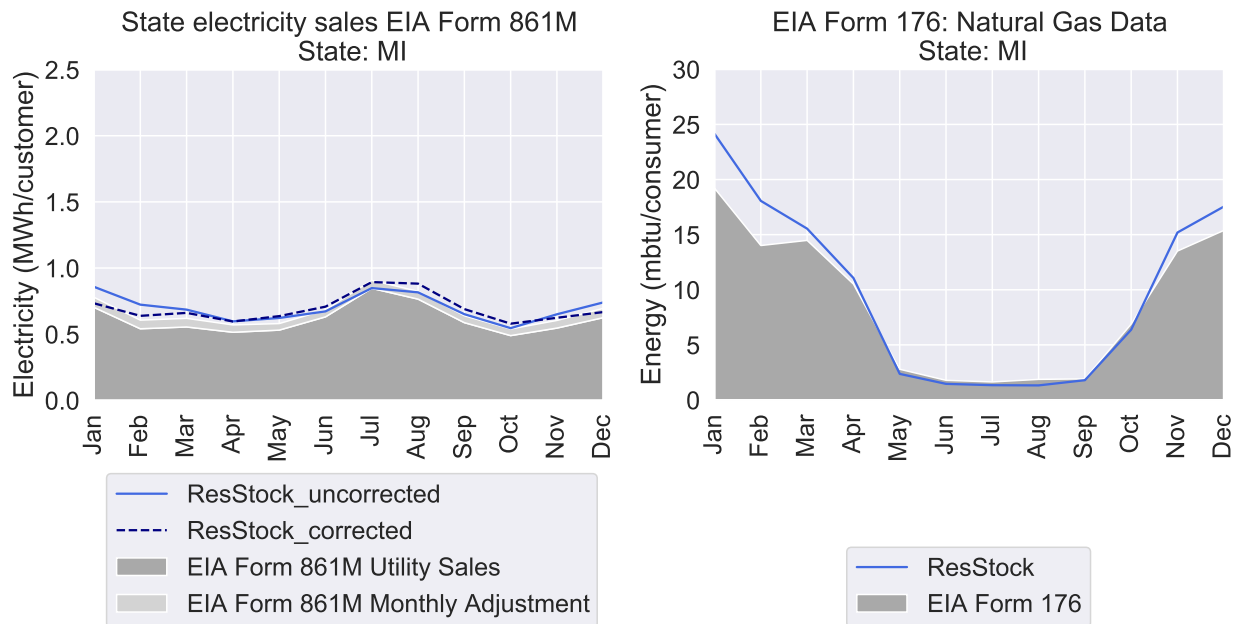
**Figure 416. ResStock MA monthly electric sales before and after correction compared to 2018 MA sales reported in EIA Form 861M (left). ResStock MA monthly natural gas energy compared to 2018 MA natural gas energy reported in EIA Form 176 (right).**



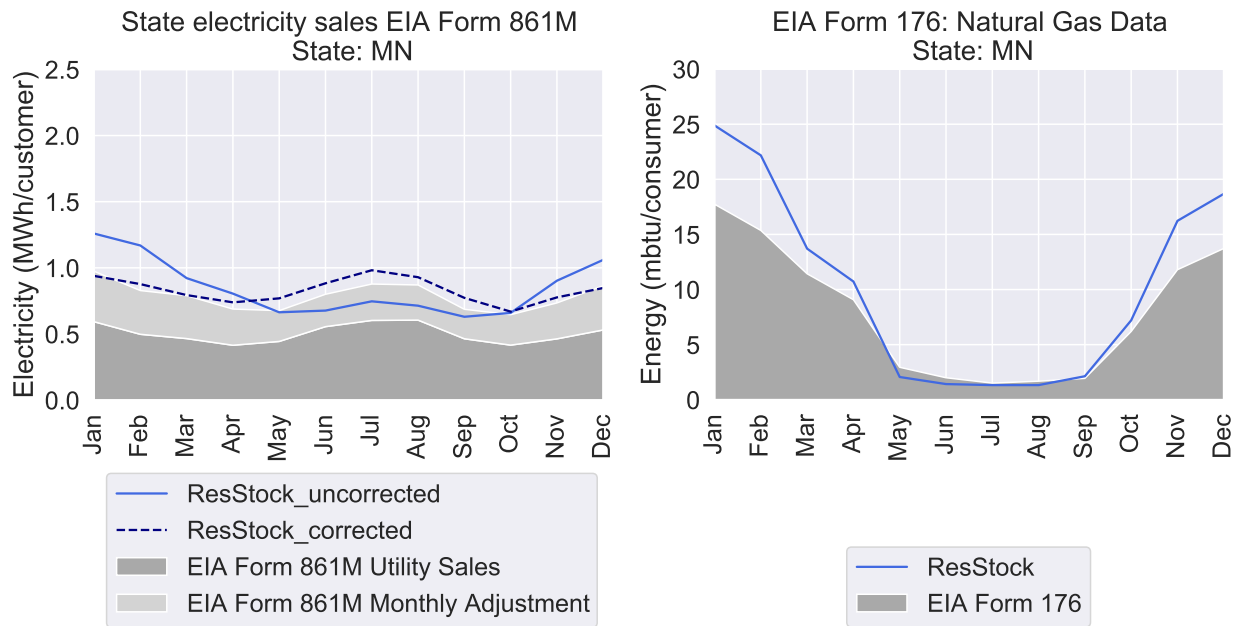
**Figure 417. ResStock MD monthly electric sales before and after correction compared to 2018 MD sales reported in EIA Form 861M (left). ResStock MD monthly natural gas energy compared to 2018 MD natural gas energy reported in EIA Form 176 (right).**



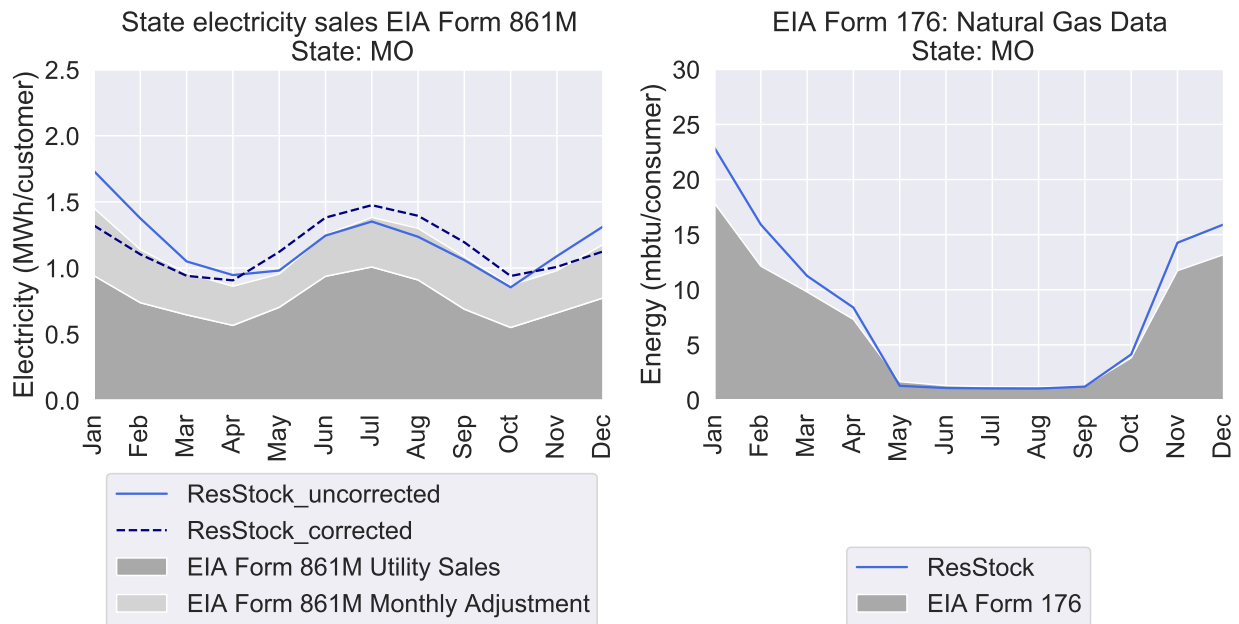
**Figure 418. ResStock ME monthly electric sales before and after correction compared to 2018 ME sales reported in EIA Form 861M (left). ResStock ME monthly natural gas energy compared to 2018 ME natural gas energy reported in EIA Form 176 (right).**



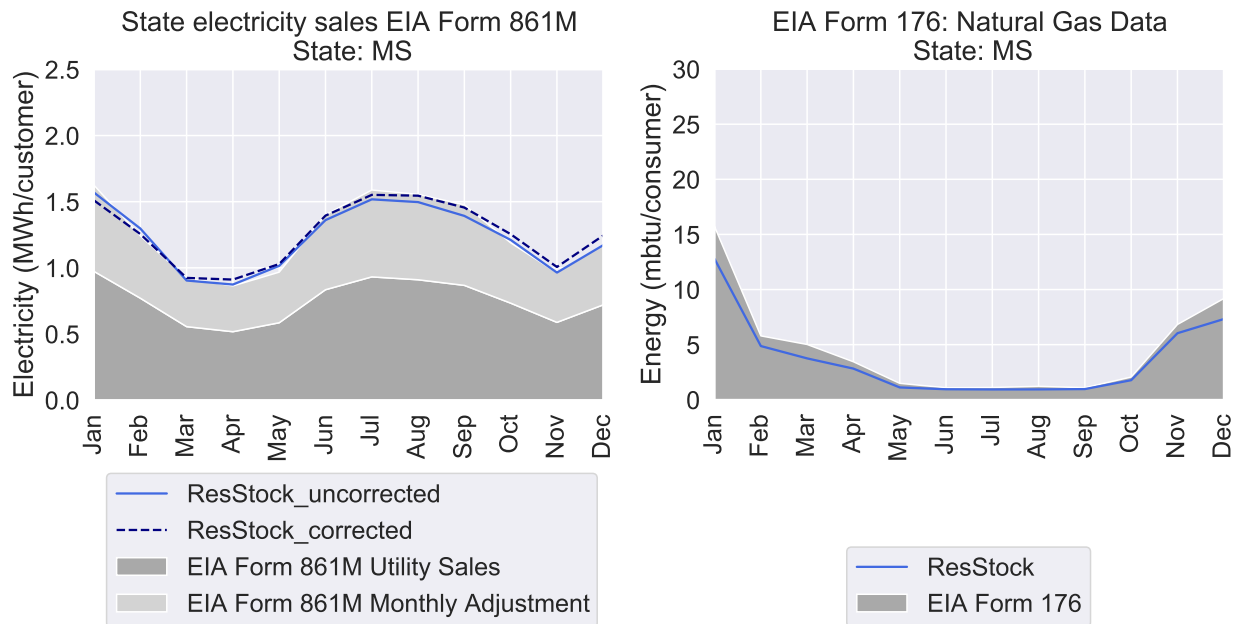
**Figure 419. ResStock MI monthly electric sales before and after correction compared to 2018 MI sales reported in EIA Form 861M (left). ResStock MI monthly natural gas energy compared to 2018 MI natural gas energy reported in EIA Form 176 (right).**



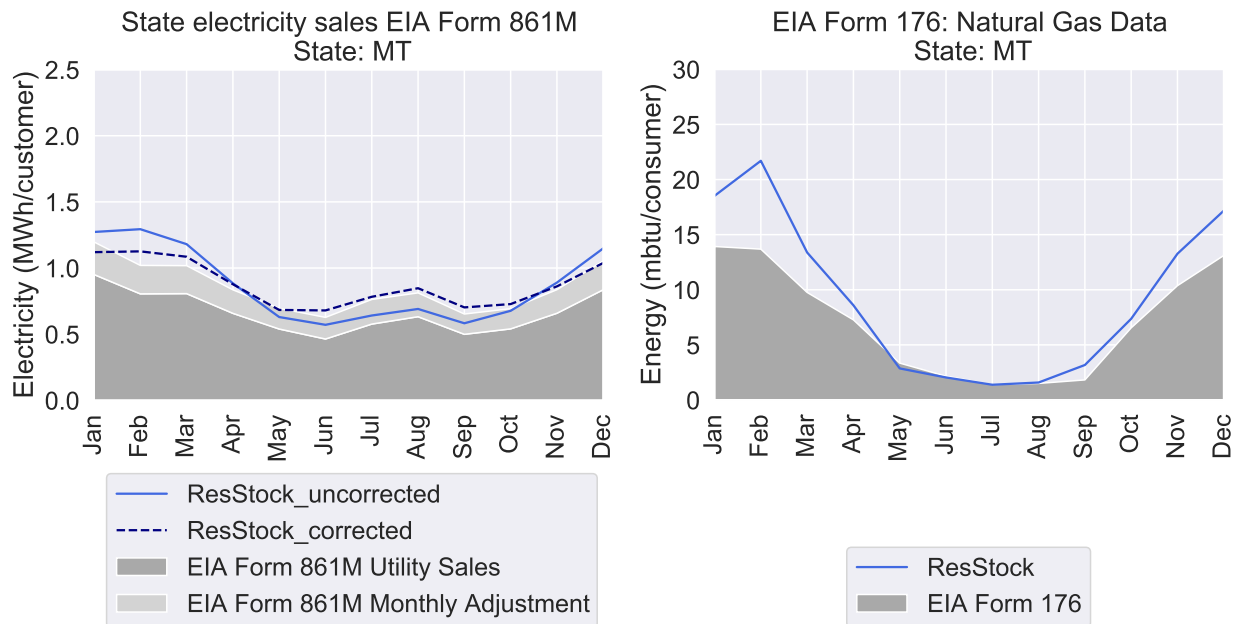
**Figure 420. ResStock MN monthly electric sales before and after correction compared to 2018 MN sales reported in EIA Form 861M (left). ResStock MN monthly natural gas energy compared to 2018 MN natural gas energy reported in EIA Form 176 (right).**



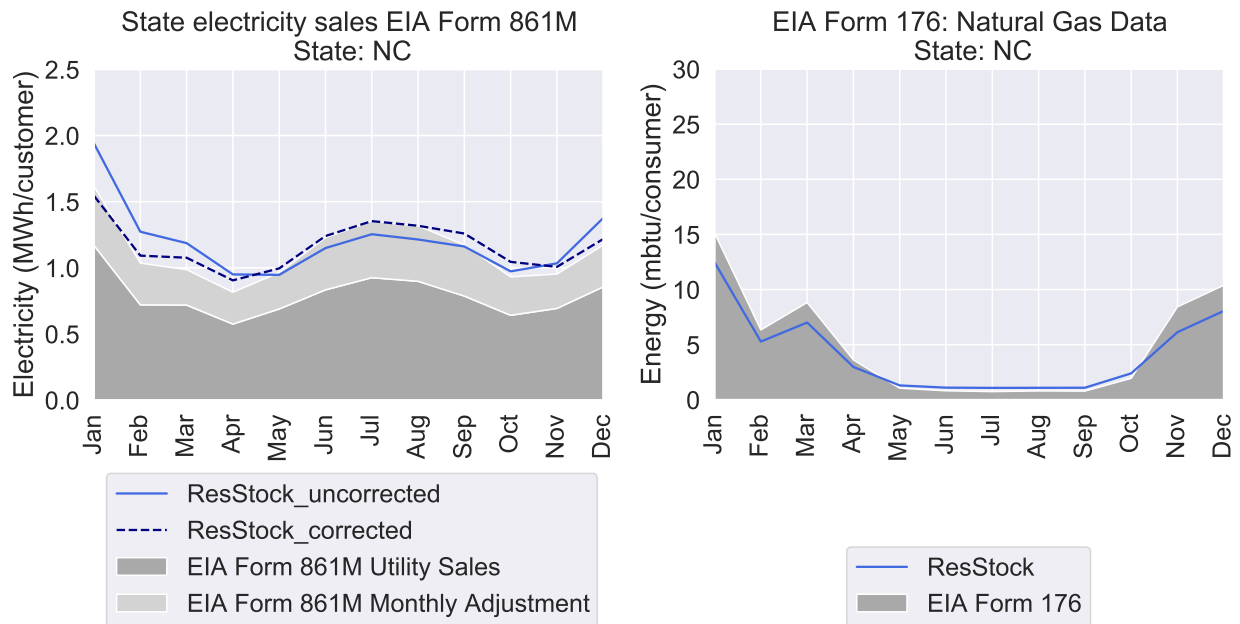
**Figure 421. ResStock MO monthly electric sales before and after correction compared to 2018 MO sales reported in EIA Form 861M (left). ResStock MO monthly natural gas energy compared to 2018 MO natural gas energy reported in EIA Form 176 (right).**



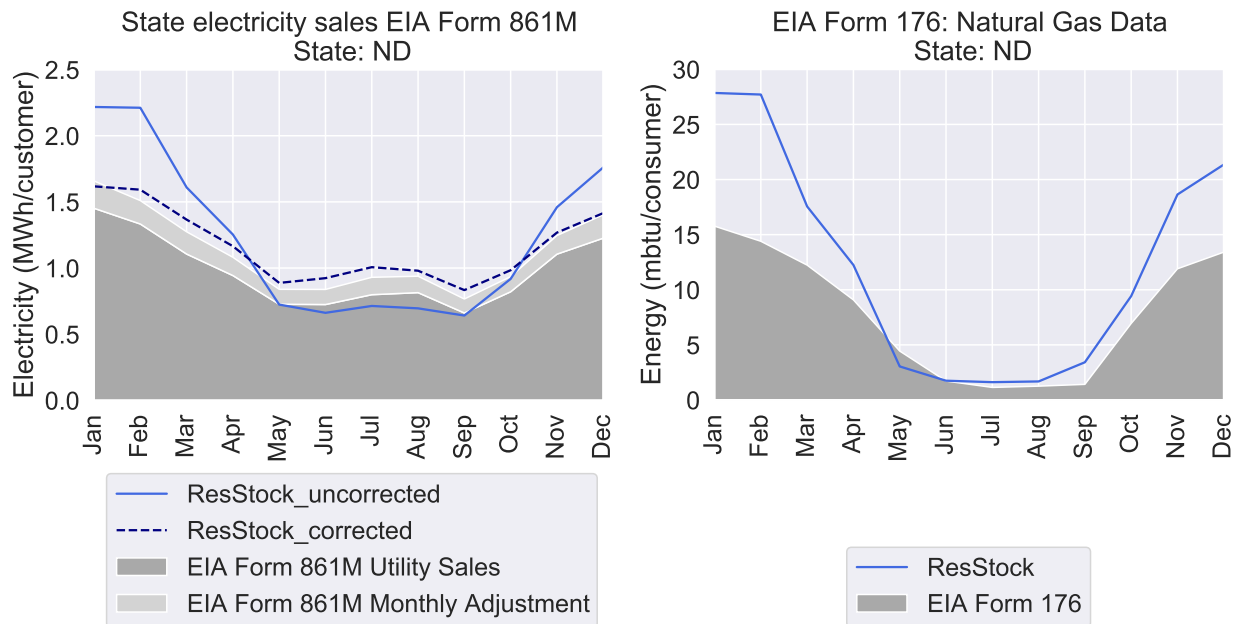
**Figure 422. ResStock MS monthly electric sales before and after correction compared to 2018 MS sales reported in EIA Form 861M (left). ResStock MS monthly natural gas energy compared to 2018 MS natural gas energy reported in EIA Form 176 (right).**



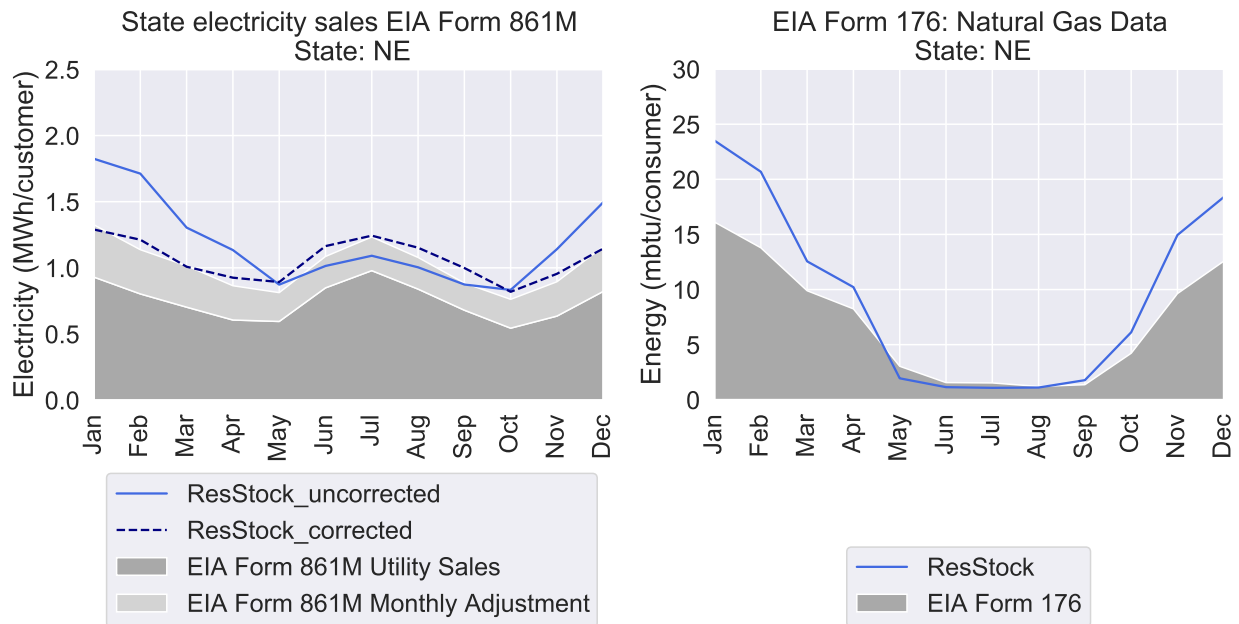
**Figure 423. ResStock MT monthly electric sales before and after correction compared to 2018 MT sales reported in EIA Form 861M (left). ResStock MT monthly natural gas energy compared to 2018 MT natural gas energy reported in EIA Form 176 (right).**



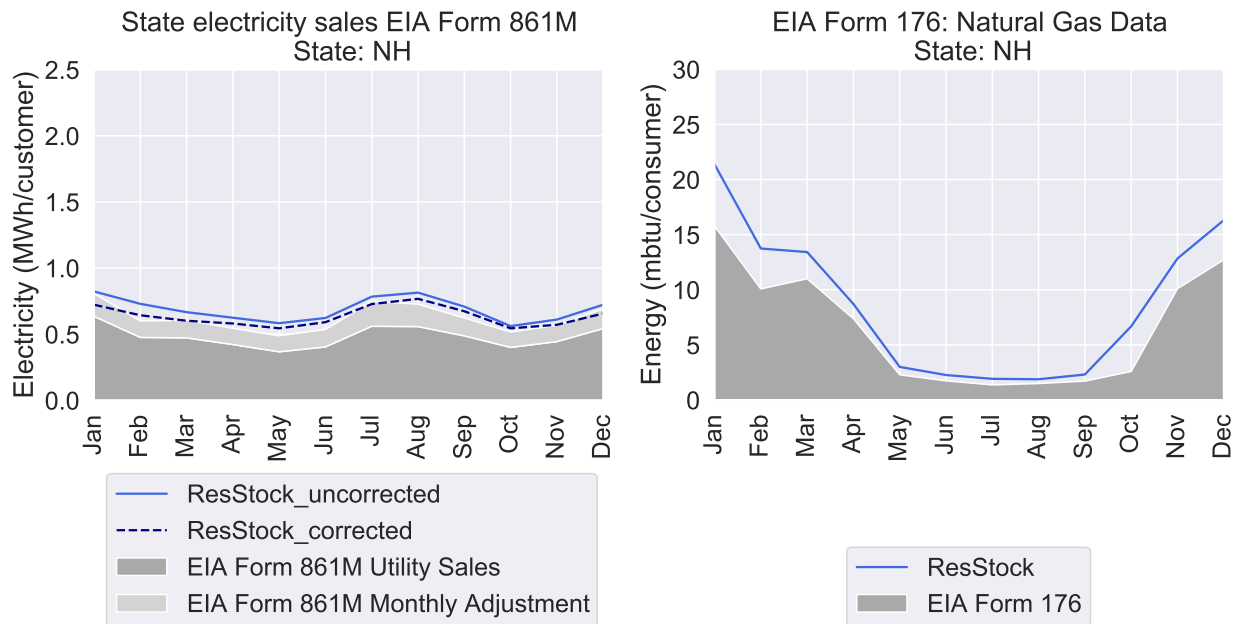
**Figure 424. ResStock NC monthly electric sales before and after correction compared to 2018 NC sales reported in EIA Form 861M (left). ResStock NC monthly natural gas energy compared to 2018 NC natural gas energy reported in EIA Form 176 (right).**



**Figure 425. ResStock ND monthly electric sales before and after correction compared to 2018 ND sales reported in EIA Form 861M (left). ResStock ND monthly natural gas energy compared to 2018 ND natural gas energy reported in EIA Form 176 (right).**

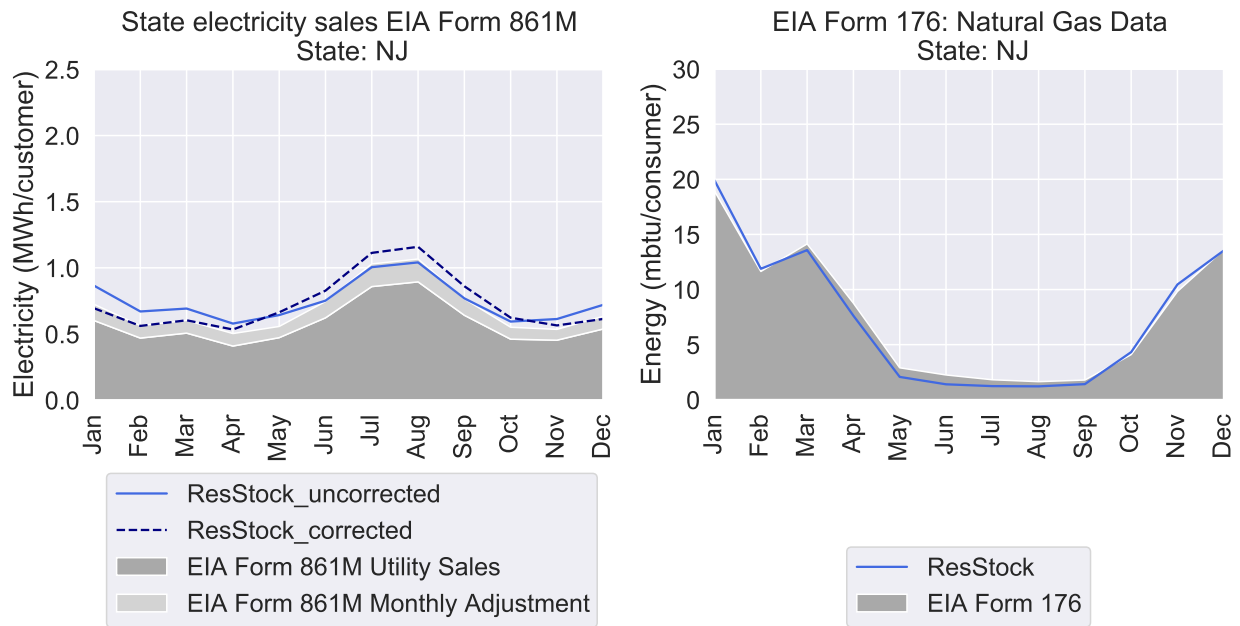


**Figure 426. ResStock NE monthly electric sales before and after correction compared to 2018 NE sales reported in EIA Form 861M (left). ResStock NE monthly natural gas energy compared to 2018 NE natural gas energy reported in EIA Form 176 (right).**

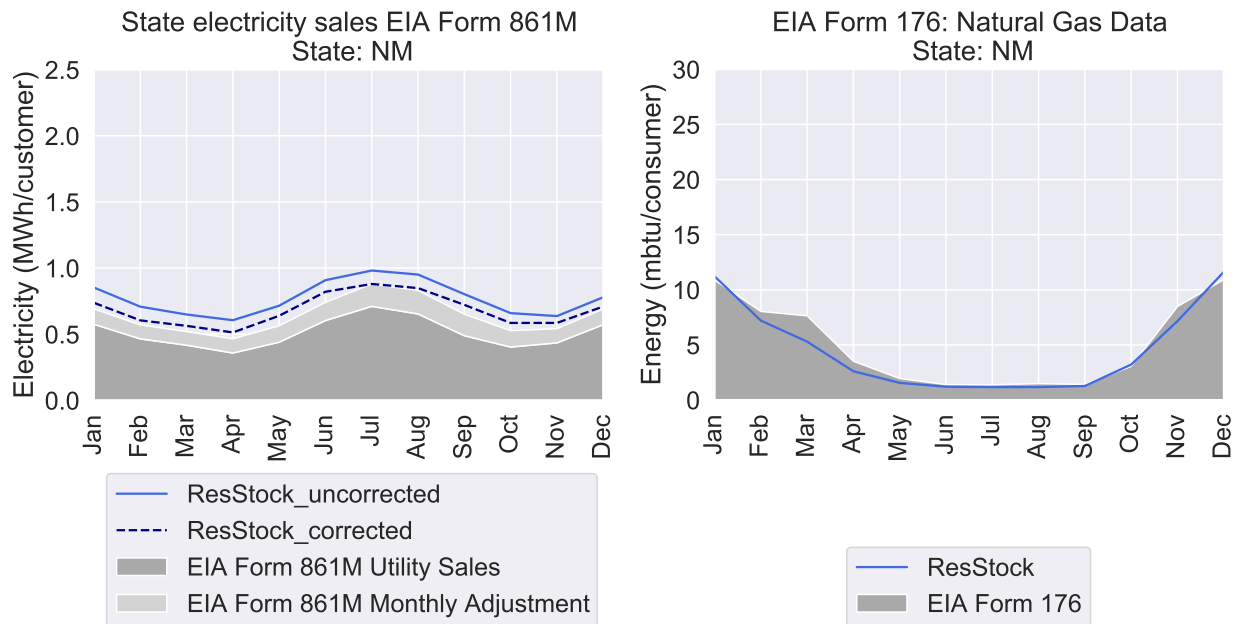


**Figure 427. ResStock NH monthly electric sales before and after correction compared to 2018 NH sales reported in EIA Form 861M (left). ResStock NH monthly natural gas energy compared to 2018 NH natural gas energy reported in EIA Form 176 (right).**

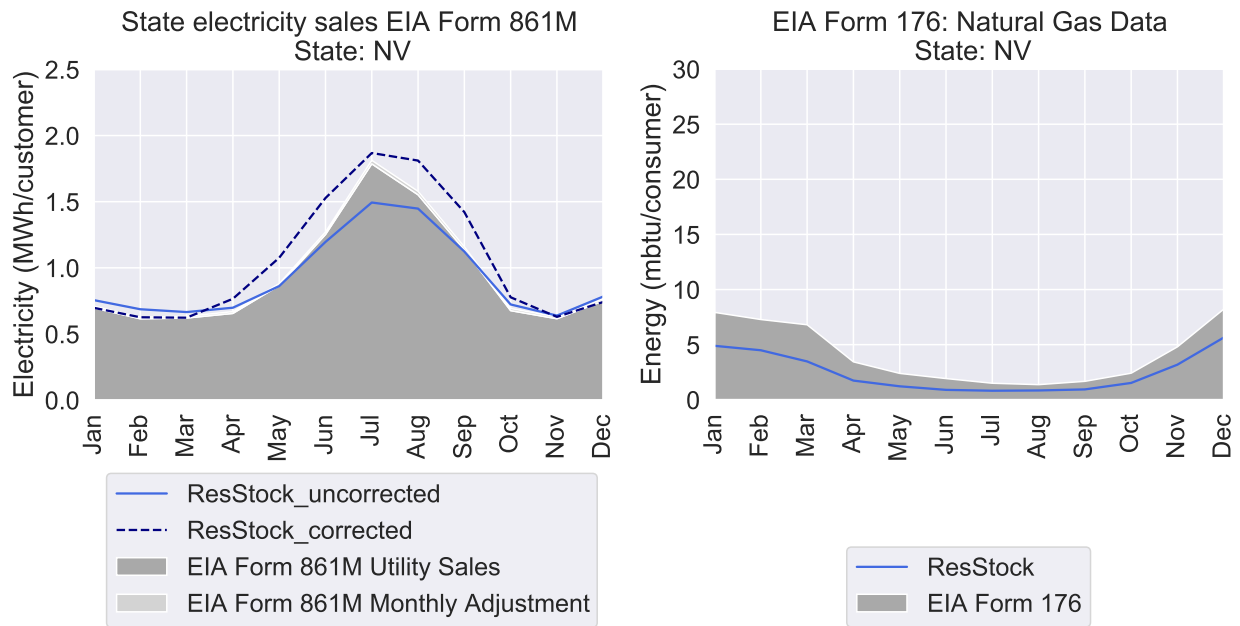




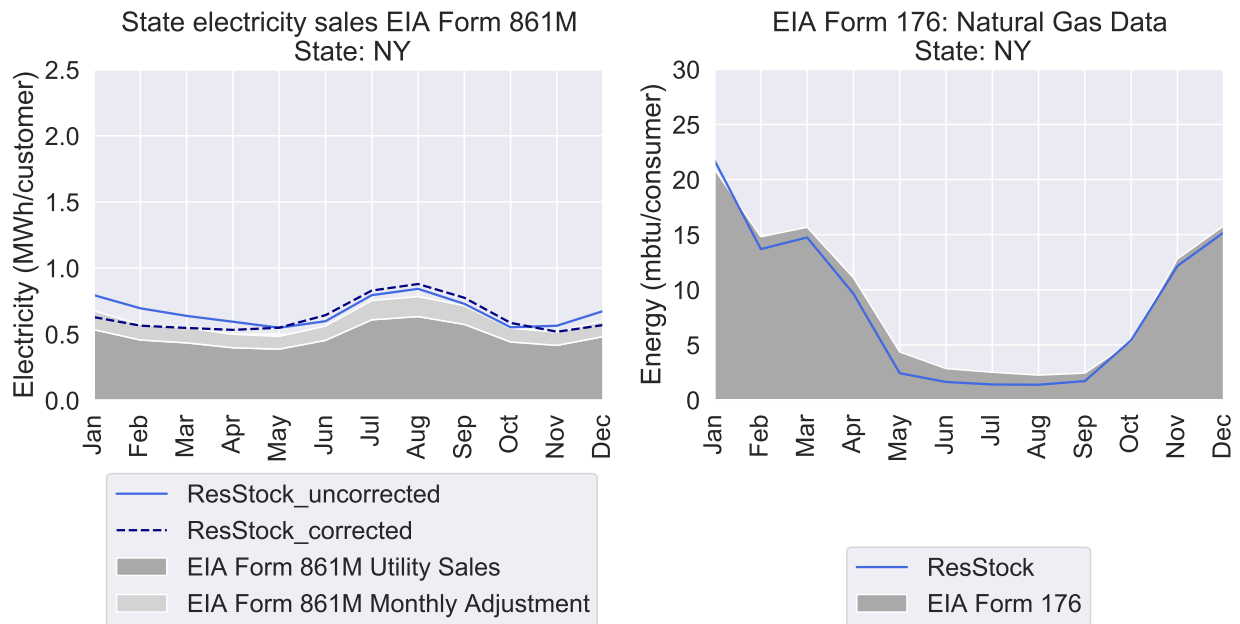
**Figure 428. ResStock NJ monthly electric sales before and after correction compared to 2018 NJ sales reported in EIA Form 861M (left). ResStock NJ monthly natural gas energy compared to 2018 NJ natural gas energy reported in EIA Form 176 (right).**



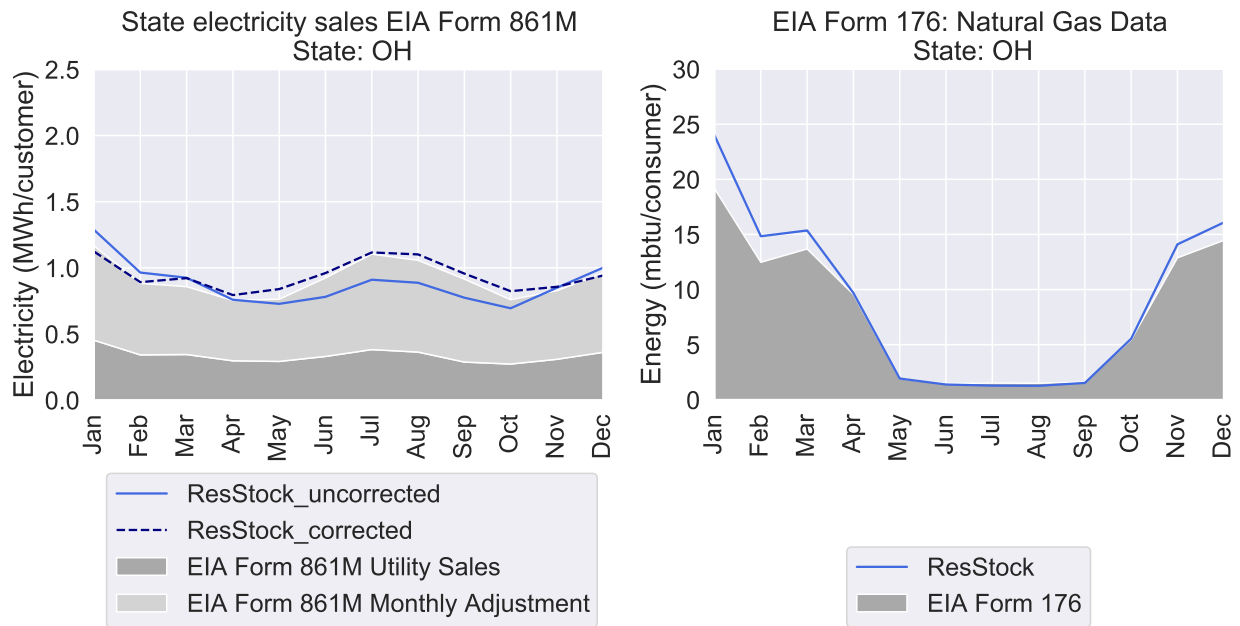
**Figure 429. ResStock NM monthly electric sales before and after correction compared to 2018 NM sales reported in EIA Form 861M (left). ResStock NM monthly natural gas energy compared to 2018 NM natural gas energy reported in EIA Form 176 (right).**



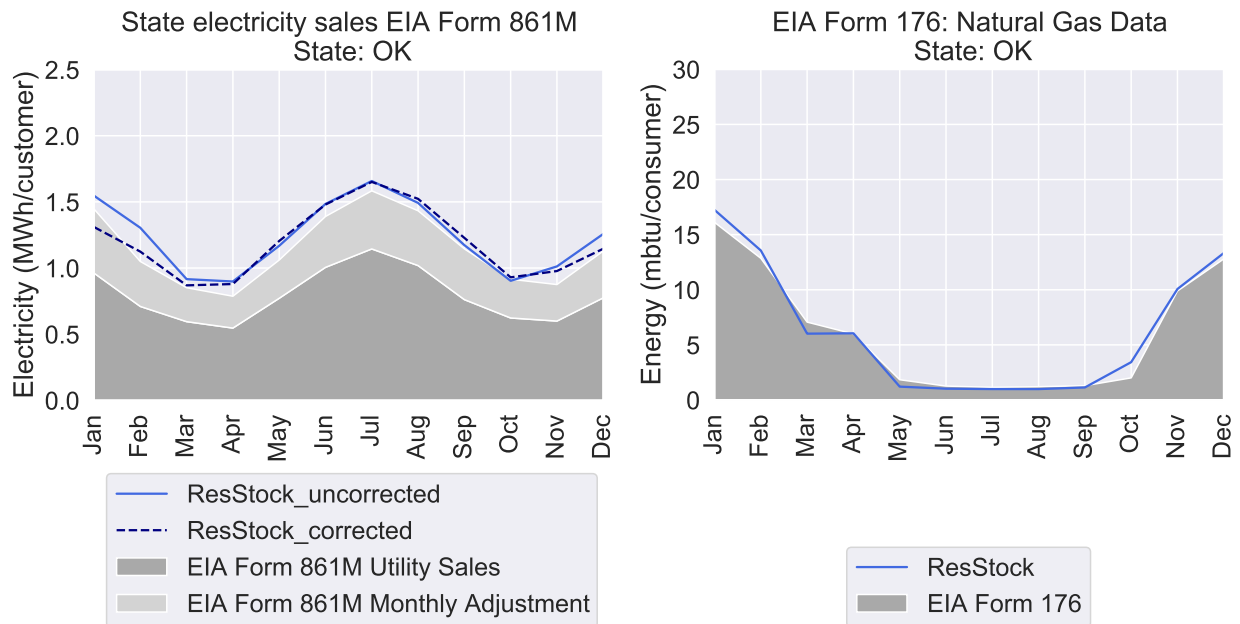
**Figure 430. ResStock NV monthly electric sales before and after correction compared to 2018 NV sales reported in EIA Form 861M (left). ResStock NV monthly natural gas energy compared to 2018 NV natural gas energy reported in EIA Form 176 (right).**



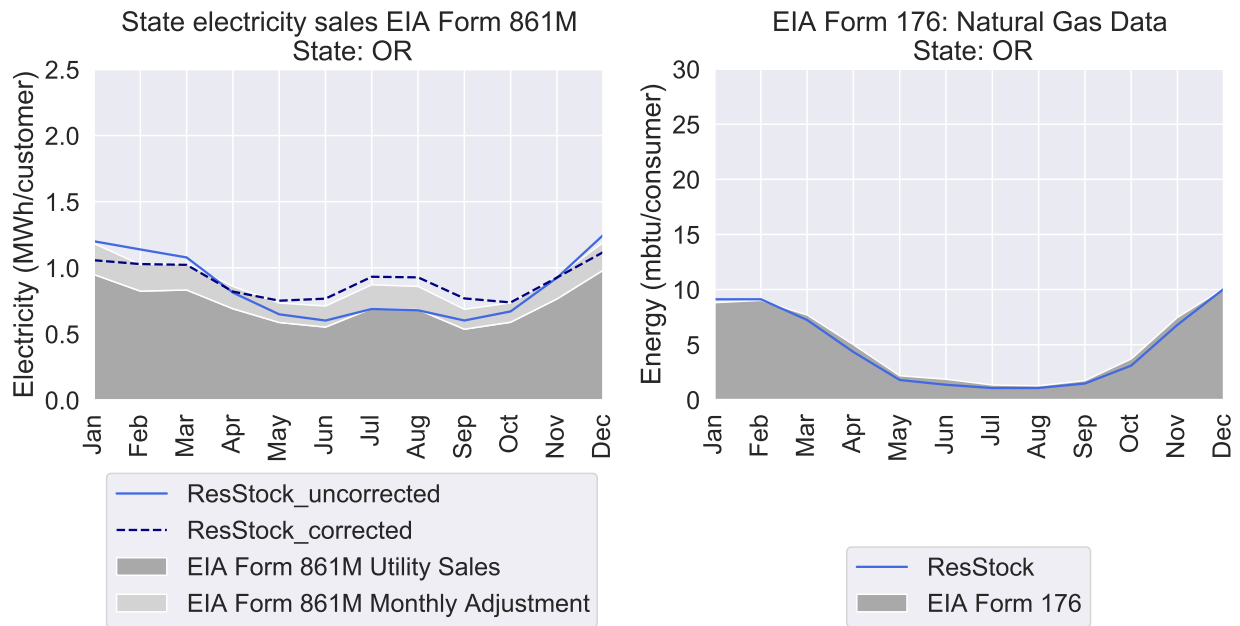
**Figure 431. ResStock NY monthly electric sales before and after correction compared to 2018 NY sales reported in EIA Form 861M (left). ResStock NY monthly natural gas energy compared to 2018 NY natural gas energy reported in EIA Form 176 (right).**



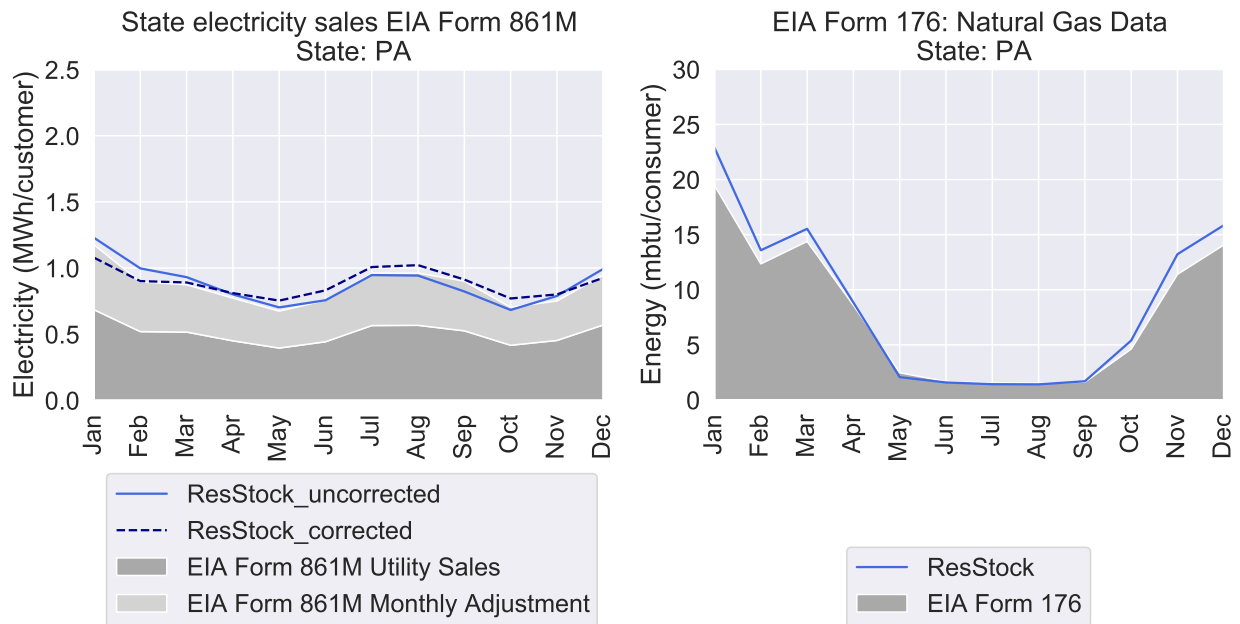
**Figure 432. ResStock OH monthly electric sales before and after correction compared to 2018 OH sales reported in EIA Form 861M (left). ResStock OH monthly natural gas energy compared to 2018 OH natural gas energy reported in EIA Form 176 (right).**



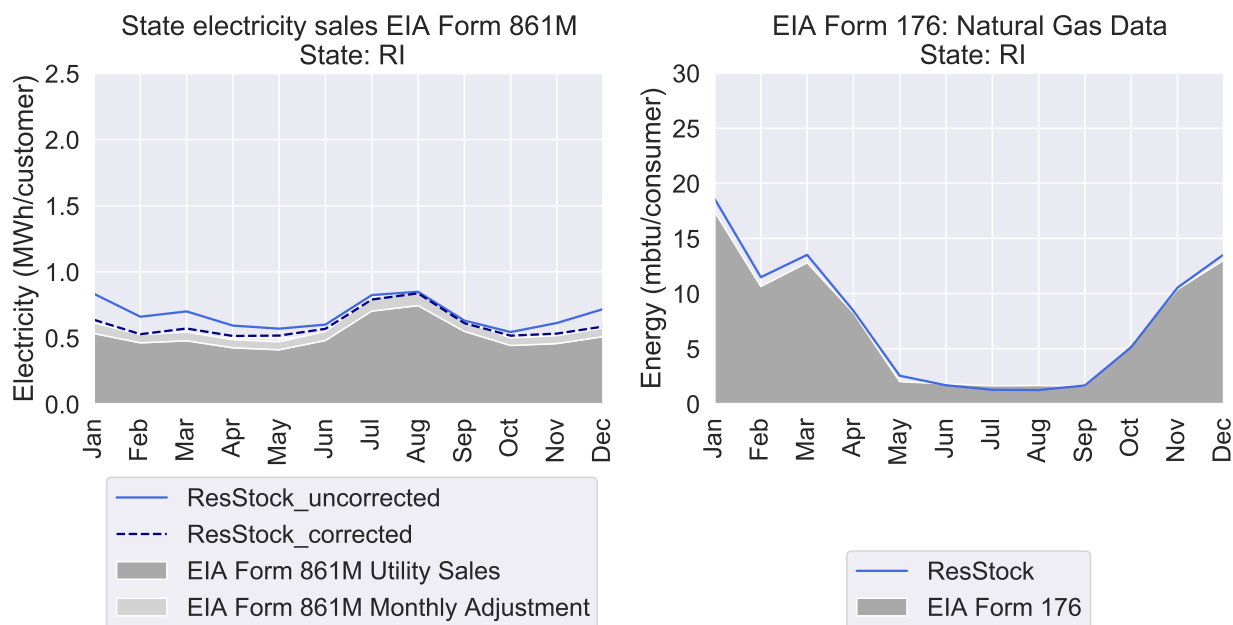
**Figure 433. ResStock OK monthly electric sales before and after correction compared to 2018 OK sales reported in EIA Form 861M (left). ResStock OK monthly natural gas energy compared to 2018 OK natural gas energy reported in EIA Form 176 (right).**



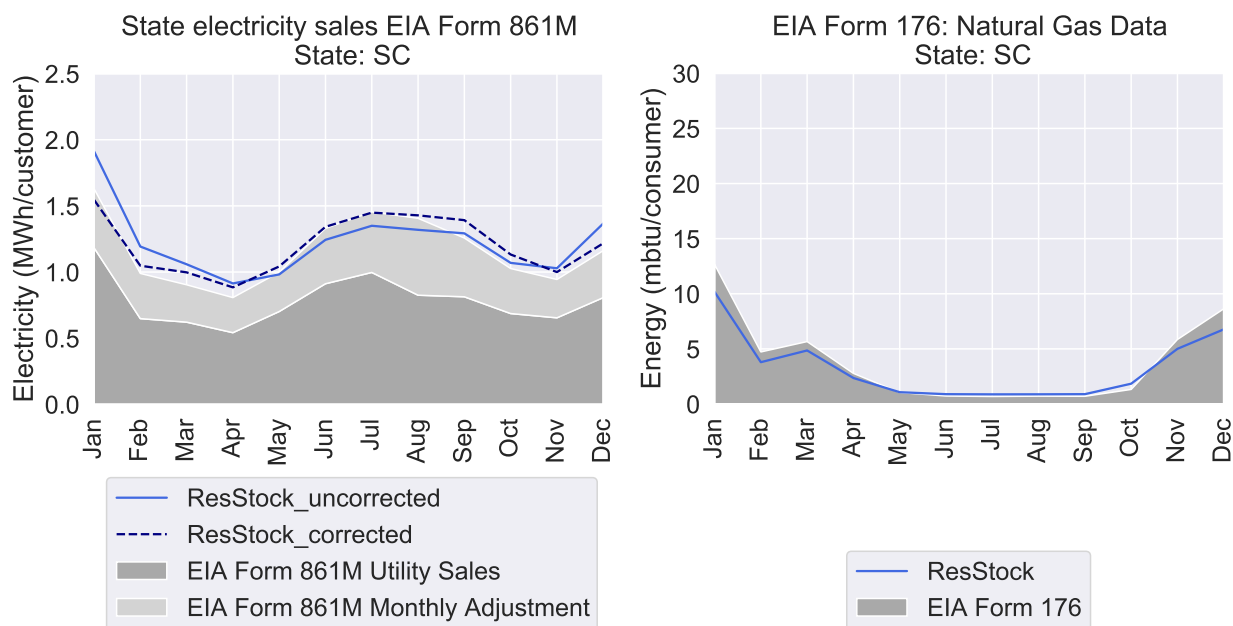
**Figure 434. ResStock OR monthly electric sales before and after correction compared to 2018 OR sales reported in EIA Form 861M (left). ResStock OR monthly natural gas energy compared to 2018 OR natural gas energy reported in EIA Form 176 (right).**



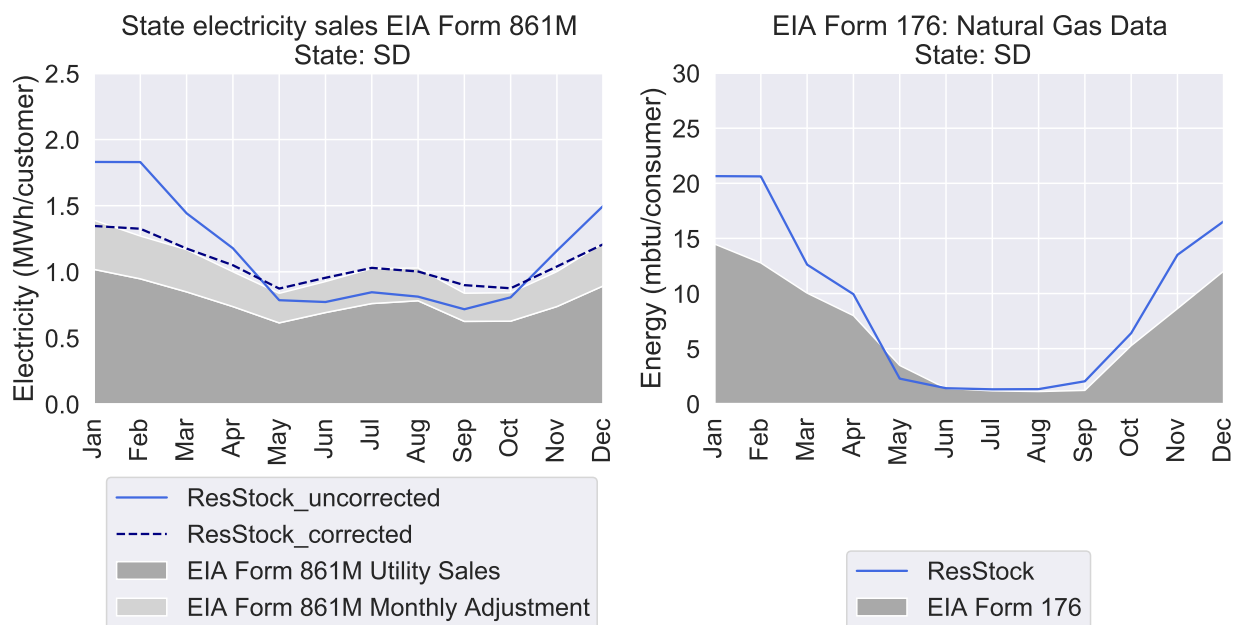
**Figure 435. ResStock PA monthly electric sales before and after correction compared to 2018 PA sales reported in EIA Form 861M (left). ResStock PA monthly natural gas energy compared to 2018 PA natural gas energy reported in EIA Form 176 (right).**



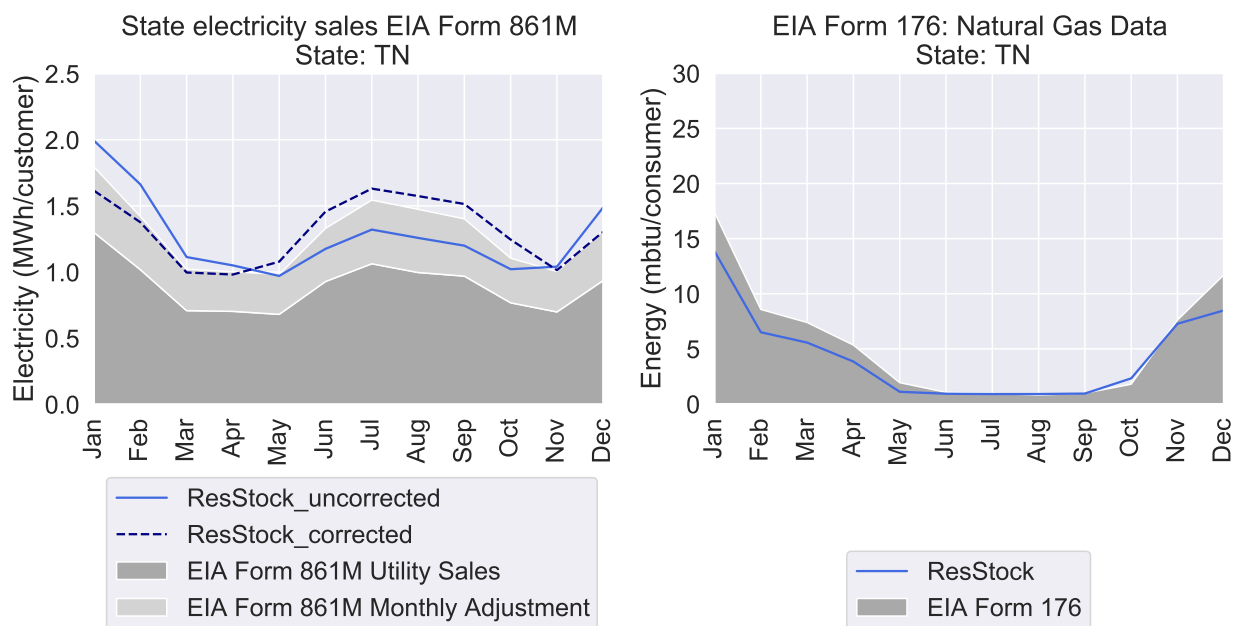
**Figure 436. ResStock RI monthly electric sales before and after correction compared to 2018 RI sales reported in EIA Form 861M (left). ResStock RI monthly natural gas energy compared to 2018 RI natural gas energy reported in EIA Form 176 (right).**



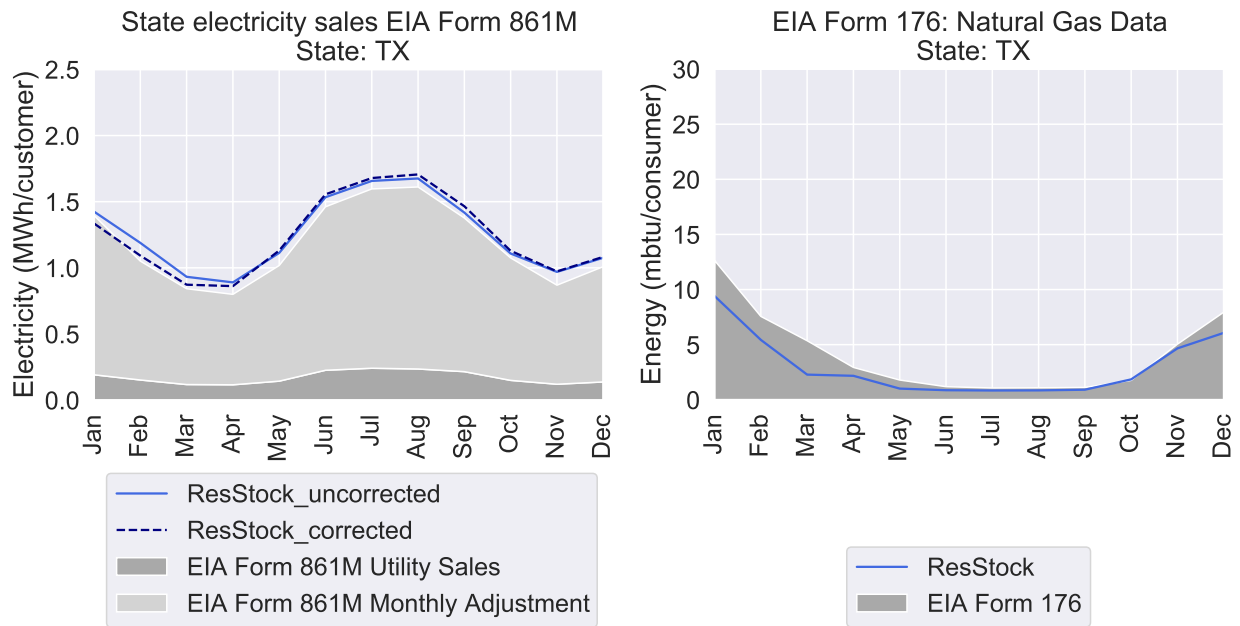
**Figure 437. ResStock SC monthly electric sales before and after correction compared to 2018 SC sales reported in EIA Form 861M (left). ResStock SC monthly natural gas energy compared to 2018 SC natural gas energy reported in EIA Form 176 (right).**



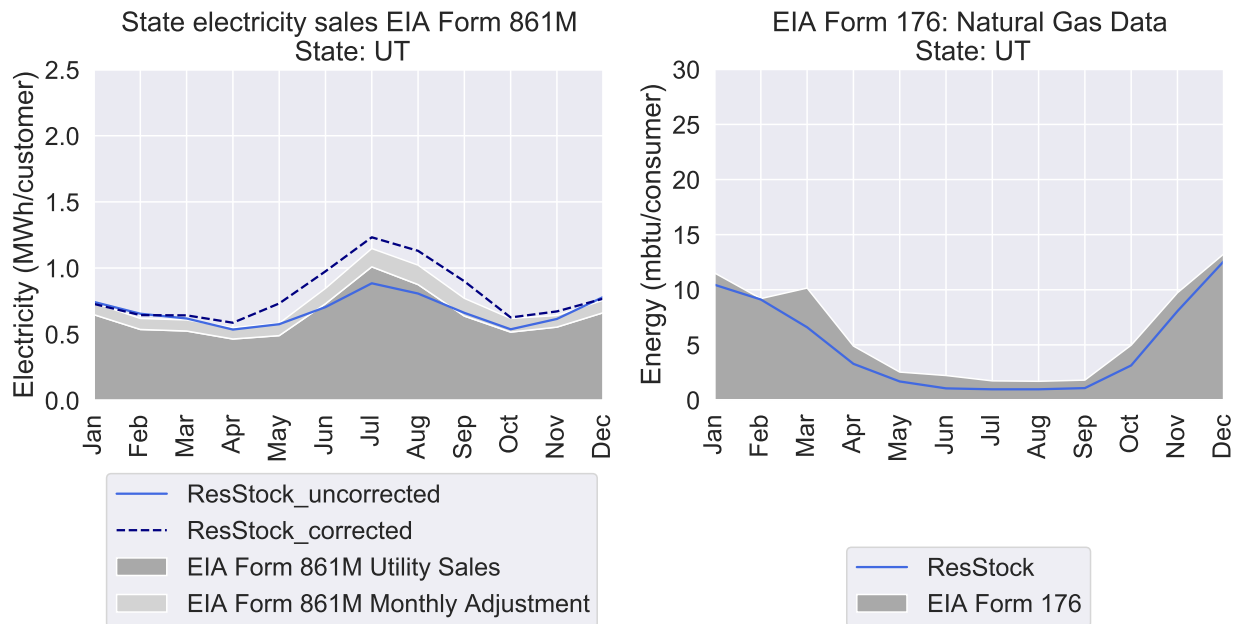
**Figure 438. ResStock SD monthly electric sales before and after correction compared to 2018 SD sales reported in EIA Form 861M (left). ResStock SD monthly natural gas energy compared to 2018 SD natural gas energy reported in EIA Form 176 (right).**



**Figure 439. ResStock TN monthly electric sales before and after correction compared to 2018 TN sales reported in EIA Form 861M (left). ResStock TN monthly natural gas energy compared to 2018 TN natural gas energy reported in EIA Form 176 (right).**

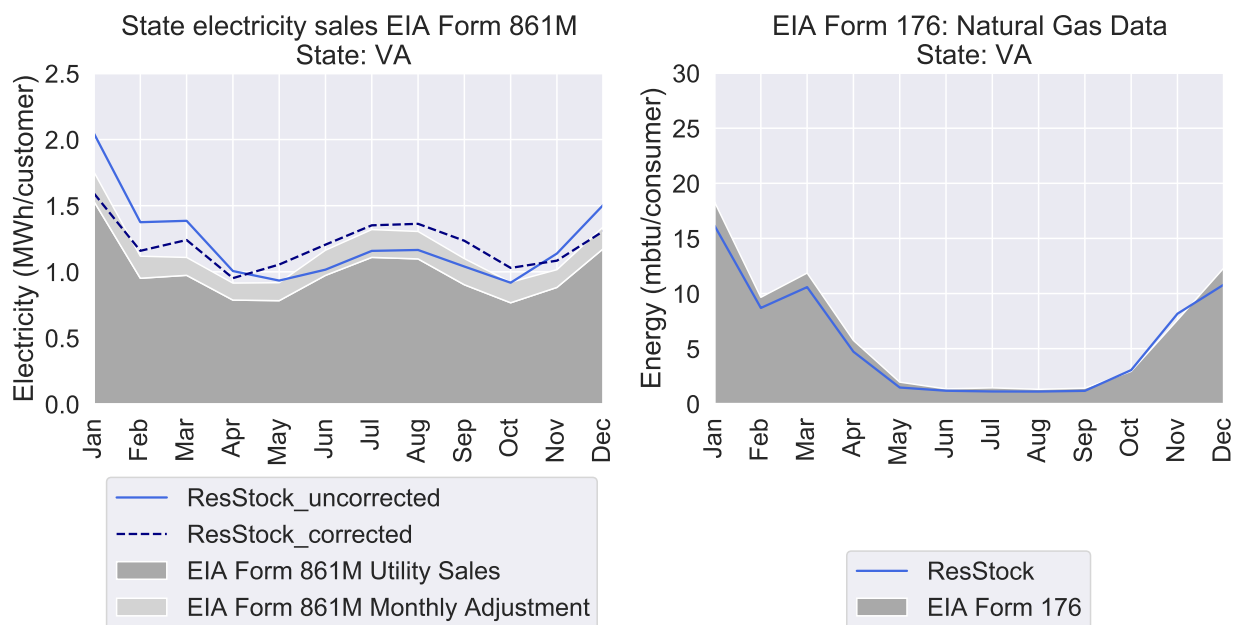


**Figure 440. ResStock TX monthly electric sales before and after correction compared to 2018 TX sales reported in EIA Form 861M (left). ResStock TX monthly natural gas energy compared to 2018 TX natural gas energy reported in EIA Form 176 (right).**

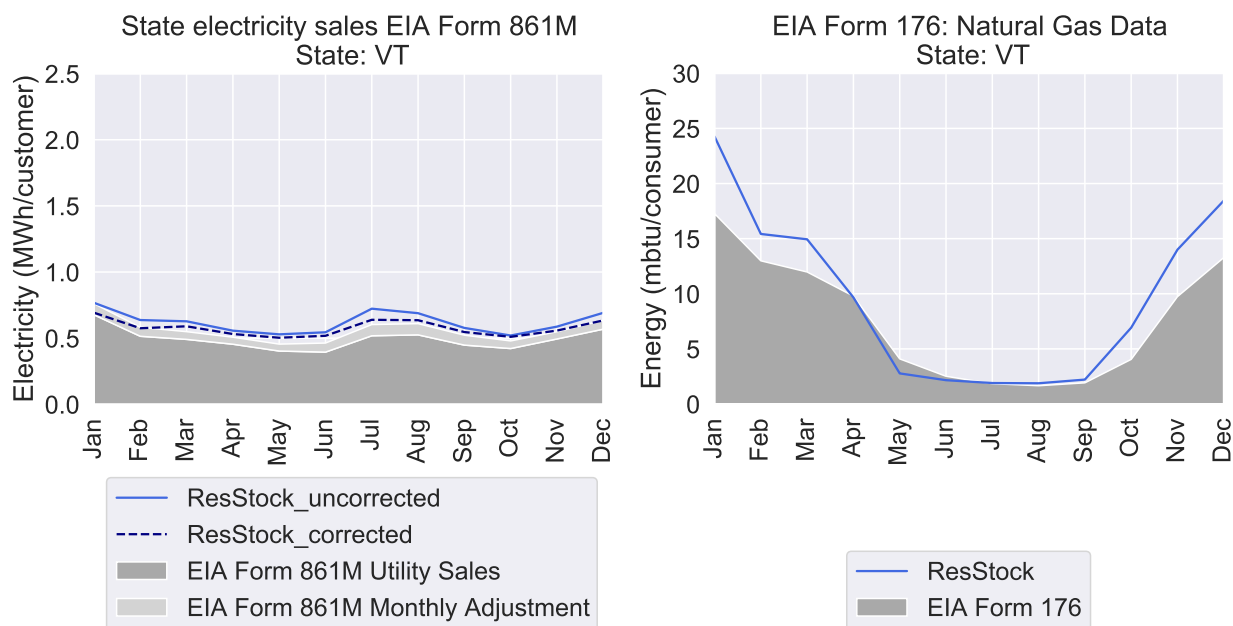


**Figure 441. ResStock UT monthly electric sales before and after correction compared to 2018 UT sales reported in EIA Form 861M (left). ResStock UT monthly natural gas energy compared to 2018 UT natural gas energy reported in EIA Form 176 (right).**

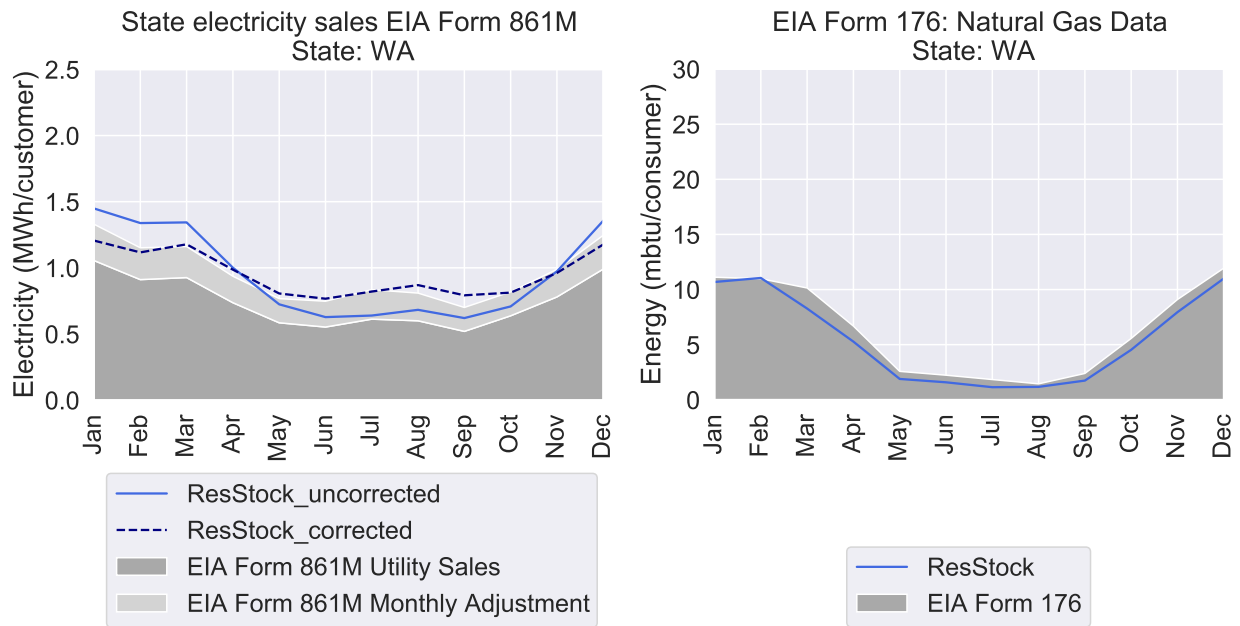




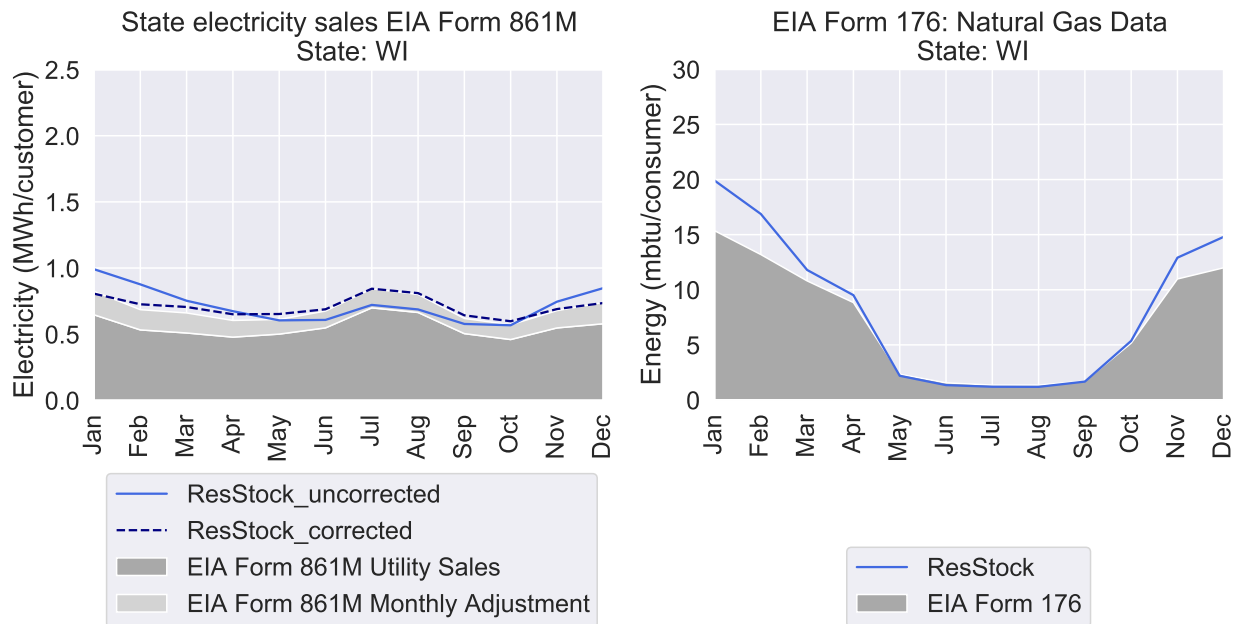
**Figure 442. ResStock VA monthly electric sales before and after correction compared to 2018 VA sales reported in EIA Form 861M (left). ResStock VA monthly natural gas energy compared to 2018 VA natural gas energy reported in EIA Form 176 (right).**



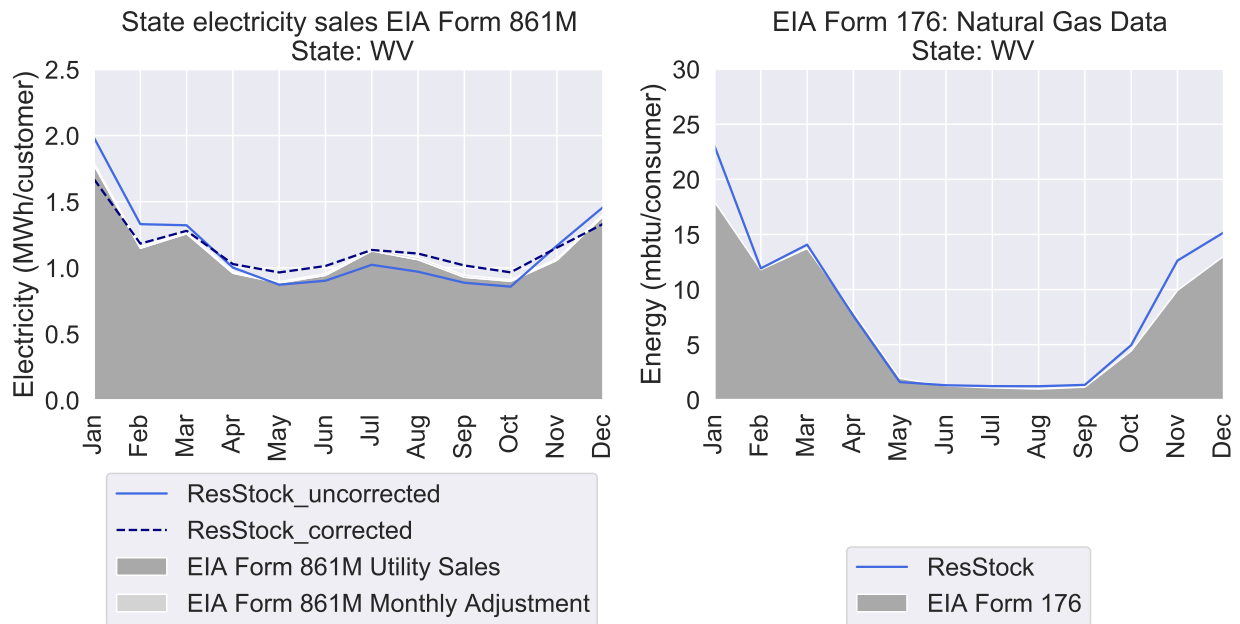
**Figure 443. ResStock VT monthly electric sales before and after correction compared to 2018 VT sales reported in EIA Form 861M (left). ResStock VT monthly natural gas energy compared to 2018 VT natural gas energy reported in EIA Form 176 (right).**



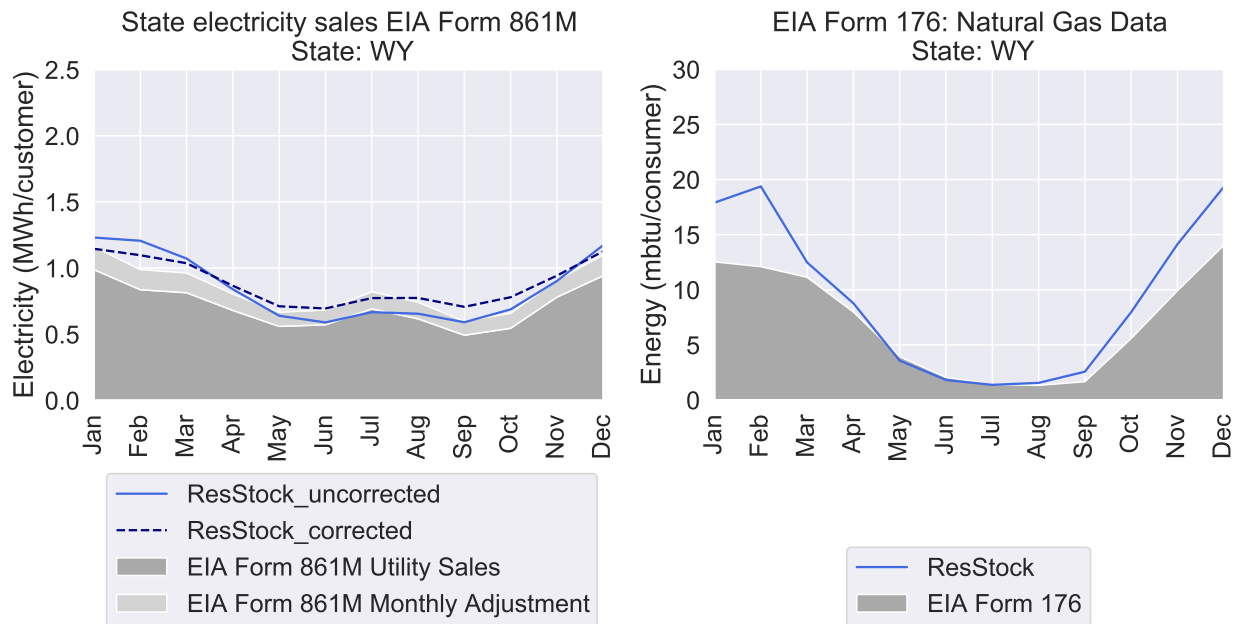
**Figure 444. ResStock WA monthly electric sales before and after correction compared to 2018 WA sales reported in EIA Form 861M (left). ResStock WA monthly natural gas energy compared to 2018 WA natural gas energy reported in EIA Form 176 (right).**



**Figure 445. ResStock WI monthly electric sales before and after correction compared to 2018 WI sales reported in EIA Form 861M (left). ResStock WI monthly natural gas energy compared to 2018 WI natural gas energy reported in EIA Form 176 (right).**



**Figure 446. ResStock WV monthly electric sales before and after correction compared to 2018 WV sales reported in EIA Form 861M (left). ResStock WV monthly natural gas energy compared to 2018 WV natural gas energy reported in EIA Form 176 (right).**



**Figure 447. ResStock WY monthly electric sales before and after correction compared to 2018 WY sales reported in EIA Form 861M (left). ResStock WY monthly natural gas energy compared to 2018 WY natural gas energy reported in EIA Form 176 (right).**



