PNNL-27687

Pacific Northwest NATIONAL LABORATORY

Proudly Operated by Battelle Since 1965

# The Building ENergy Demand (BEND) Model: A Comprehensive Composite Building Energy Modeling Framework

## July 2018


YL Xie
CD Burleyson
I Kraucunas
ZT Taylor
N Voisin



U.S. DEPARTMENT OF ENERGY

Prepared for the U.S. Department of Energy under Contract DE-AC05-76RL01830

PACIFIC NORTHWEST NATIONAL LABORATORY
*operated by*
BATTELLE
*for the*
UNITED STATES DEPARTMENT OF ENERGY
*under Contract DE-AC05-76RL01830*

**Printed in the United States of America**

**Available to DOE and DOE contractors from**
**the Office of Scientific and Technical Information,**
**P.O. Box 62, Oak Ridge, TN 37831-0062**
www.osti.gov
**ph: (865) 576-8401**
**fox: (865) 576-5728**
**email: reports@osti.gov**

**Available to the public from the National Technical Information Service**
**5301 Shawnee Rd., Alexandria, VA 22312**
**ph: (800) 553-NTIS (6847)**
**or (703) 605-6000**
**email: info@ntis.gov**
**Online ordering: http://www.ntis.gov**

PNNL-27687

# The Building ENergy Demand (BEND) Model: A Comprehensive Composite Building Energy Modeling Framework

YL Xie
CD Burleyson
I Kraucunas
ZT Taylor
N Voisin

July 2018

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99352

# Executive Summary

The Building ENergy Demand (BEND) model was developed at Pacific Northwest National Laboratory (PNNL) to simulate climate-dependent hourly building energy demand with the ability to disaggregate down to or aggregate up to any geographic area of interest, including counties, states, electric utility service territories, and census regions. The BEND modeling framework is designed to study the effects of climate change, economics, technology advancement, and various mitigation policies on aggregate building energy demand. To make the process computationally feasible, BEND reduces granularity in building characteristics and weather forcing while maintaining enough detail to minimize systematic differences (biases) between BEND results and data from the Energy Information Administration (EIA).

Based on the Commercial Building Energy Consumption Survey (CBECS) and Residential Energy Consumption Survey (RECS) databases, BEND establishes a set of *representative buildings* with predefined sets of building types, sizes, and vintages. These *representative buildings* are defined for each of the four U.S. census regions and the five climate zones used by the EIA for building surveys (hereafter "EIA climate zones"). BEND's *representative buildings* are based on a statistical analysis of the CBECS and RECS databases that infers detailed parameters such as construction type, thermal properties, internal loads, equipment and mechanical systems, and operation schedules from the high-level descriptors in those databases (building type, vintage, size, and location). The *representative buildings* are simplified and consolidated into a set of *prototype buildings* for the purpose of simulating building energy demand.

Combining the *prototype buildings* with weather files from selected locations and years, BEND creates a set of *simulation-ready building models* to be run using the EnergyPlus simulation engine (EnergyPlus). The base spatial unit of each BEND simulation is a subregion that is the intersection of an International Energy Conservation Code climate zone, a census region, and an EIA climate zone; analysis for each spatial unit is influenced by the weather file BEND selects to represent the climate-similar area. The major simulation outputs of interest to BEND are hourly energy consumption values for each simulated year, possibly for multiple fuel types. Based on building-stock weights derived from the CBECS/RECS data, the hourly output streams are aggregated to a single, 8760-hour load profile for each of the major building categories—residential and commercial—with the possibility of segregating by fuel types.

To correct for biases inherent in the use of a limited suite of *representative buildings*, BEND is calibrated against observational targets of total annual building energy consumption and hourly load profiles (including peak demand) obtained for the decision-relevant geographic scales of interest (e.g., total electric load data for individual balancing authorities available from EIA). Based on available data such as population distribution, BEND is able to aggregate both up and down scales, allowing the study of building energy demand over a range of spatial scales (e.g., states or counties within the Western Electricity Coordinating Council, Eastern Interconnection, or the entire U.S.) and/or subsets of building types (e.g., office buildings with particular size ranges).

The BEND modeling framework is not a stand-alone piece of software, but rather a workflow that integrates multiple components and processes:

- the CBECS and RECS databases
- a series of specification/mapping tables for building characteristics coupled with regression models and regression coefficient tables for the designation of building components and mechanical systems based on high-level building characteristics
- calculation of building weights based on their fraction of the total number of buildings
- generation of building model templates in eXtensible Markup Language (XML) format and macros

- creation of building model input files for EnergyPlus
- large-scale invocation and management of EnergyPlus runs.

These components are integrated through a series of utility programs that can be customized to fit a given analysis.

This report documents the concept, methodology, and workflow of the BEND modeling platform with an application over the western U.S. as a case study.

# Acronyms and Abbreviations

| | |
|---|---|
| AHU | air handling unit |
| BA | balancing authority |
| BEND | Building ENergy Demand (model) |
| CBECS | Commercial Building Energy Consumption Survey |
| CDD | cooling degree days |
| CR | census region |
| CZ | climate zone |
| DX | direct expansion |
| EIA | U.S. Energy Information Administration |
| EIC | Eastern Interconnection |
| EnergyPlus | EnergyPlus simulation engine |
| FCU | fan coil unit |
| FEDS | Facility Energy Decision System |
| HDD | heating degree days |
| HVAC | heating, ventilating, and air conditioning |
| IECC | International Energy Code Council |
| MSU | minimum spatial unit |
| NREL | National Renewable Energy Laboratory |
| PNNL | Pacific Northwest National Laboratory |
| RECS | Residential Energy Consumption Survey |
| WECC | Western Electricity Coordinating Council |
| XML | eXtensible Markup Language |

# Contents

# Figures



# Tables

# 1.0 Introduction

A majority of available large-spatial-scale (e.g., cities, balancing authorities [BAs], or states) load forecasting models are based on statistical analysis (e.g., multiple regression or neural nets) of historical loads. These models have become very accurate for short-term load forecasting based on weather forecasts. However, this type of model offers little opportunity to explore the effects of new building codes, urban expansion, technology shocks, and/or changes in population or population distribution. The Building ENergy Demand (BEND) model is intended to facilitate these more detailed analyses to enhance planning for future instantiations of the electricity grid.

According to the U.S. Energy Information Administration (EIA 2018), building energy accounts for 40% of the total U.S. energy consumption and 74% of the total U.S. electricity consumption in 2010. Space heating and cooling currently account for a large share of the overall energy demand for buildings—49% in residential buildings and 44% in commercial buildings. Therefore, numerous studies that aim to project future loads have focused on building energy demand, and in particular how climate change is expected to affect building energy systems through, for example, decreased heating and increased cooling loads. A variety of approaches have been used to investigate the effect of climate on building energy demand. Broadly, these can be categorized into statistical approaches, which establish relationships between building energy demand and climate based on historical data, and modeling approaches, where building stocks are represented through building models and building energy demand is simulated using physically based models. The BEND modeling framework developed by Pacific Northwest National Laboratory (PNNL) is one such simulation tool. The BEND model, originally developed as a component model of the Platform for Regional Integrated Modeling and Analysis (PRIMA) initiative at PNNL (Kraucunas et al. 2015), simulates climate-dependent hourly building energy demands. The prototype model was used in the Dirks et al. (2015) publication which focused on the eastern U.S.

BEND uses EnergyPlus as the simulation engine. EnergyPlus is a U.S. Department of Energy whole-building energy simulation model that simulates electric load in individual buildings (DOE 2013). EnergyPlus has been used to explore the impact of various building codes and project the change in building demand due to changes in temperature. EnergyPlus simulates individual buildings, each of which must be described in detail. Evaluating building energy at decision-relevant scales that may be used for planning purposes therefore requires a rather complex process of specifying a large number of prototype buildings and aggregating simulation results to the desired scales. This complexity is valuable when using a modeling approach rather than a statistical approach because detailed building models can take into account the great diversity of buildings that exist in any given spatial domain and would be affected in different ways by changes in climate. In BEND, a large number of detailed building models are specified to represent actual building stocks with high spatial resolution. The BEND modeling framework is not a stand-alone program file, but rather a workflow that integrates multiple components and processes:

- The Commercial Building Energy Consumption Survey (CBECS) and Residential Energy Consumption Survey (RECS) databases (EIA 2003a, EIA 2009a)
- A series of specification/mapping tables for building characteristics coupled with regression models and regression coefficient tables for the designation of building components and mechanical systems based on building characteristics
- Calculation of building weights as a portion of the total number of buildings
- Generation of building model templates in eXtensible Markup Language (XML) format and macros
- Creation of building model input files for the EnergyPlus software
- Large-scale invocation and management of EnergyPlus runs.

These components are integrated through a series of utility programs that can be customized to fit a given analysis.

This report documents the concept, methodology, and workflow of the BEND modeling platform with an application over the western U.S. as a case study.

Based on the CBECS and RECS databases, BEND creates a set of *representative buildings* based on grouping the surveyed buildings into predefined sets based on building types, sizes, and vintages. These *representative buildings* are defined for each of the four U.S. census regions and the five climate zones used by the U.S. Energy Information Administration (EIA) for building surveys (hereafter "EIA climate zones"). First, BEND generates the *representative buildings* based on a statistical analysis of the CBECS and RECS databases; then it derives detailed descriptions for parameters such as construction, thermal properties, internal load, equipment and mechanical systems, operation schedules; finally it creates a set of *prototype buildings* for the purpose of simulating building energy demand. Combining the *prototype buildings* with weather files from selected locations, BEND creates a set of *simulation-ready building models* to be run using the EnergyPlus simulation engine. The base spatial unit of each BEND simulation is the subregion defined by the intersection between the International Energy Code Council (IECC) climate zone and the census region/EIA climate zone, influenced by the weather file BEND selects to represent a climate-similar area. The major outputs of BEND are annual hourly energy consumption for both residential and commercial buildings. The calibration of BEND is performed on decision-relevant geographic scales from which observational targets of total annual building energy consumption and hourly load profiles can be obtained. BEND is also able to provide more-detailed energy demand information for components such as fuel types (e.g., electricity, natural gas, or oil) or end use (e.g., cooling, heating, lighting, fans, pumps, or service water). By being able to aggregate both up and down scale, BEND is able to simulate building energy demand over a range of spatial scales (e.g., states or counties in the Western Electricity Coordinating Council [WECC], Eastern Interconnection [EIC], or the entire U.S.) and/or subsets of buildings (e.g., office buildings with particular size ranges). The BEND modeling framework facilitates the evaluation of the effects of climate change, economics, technology advancement, and various mitigation policies on aggregate building energy demand.

# 2.0 Overview of the BEND Workflow

The BEND modeling framework is a series of steps for simulating aggregate building energy demand in a geographic region of interest, with the U.S. Department of Energy EnergyPlus building model as the primary simulation element. The key pre-simulation task of BEND is strategically designing building models representative of the observed building stock within a given region. Those building models are combined with either observed or predicted weather information to produce EnergyPlus simulations for each of the representative building models. Finally, after the EnergyPlus runs, BEND aggregates the outputs to decision-relevant scales for calibration and analysis.

To make the process computationally feasible, BEND reduces granularity in building characteristics and weather forcing while maintaining enough detail to minimize systematic differences (biases) between BEND results and data from the EIA.

The *representative buildings*, a representative sample of CBECS and RECS data, are a statistical representation of the total building stock in a given geographical area. They have statistically similar physical and thermal characteristics, functions and activities, operations, and heating/cooling systems. Representative buildings have associated weights that account for the fraction of the total number of buildings they represent.

Four high-level specifications of the *representative buildings* (type, vintage, size, and location) are converted into a much larger number of detailed parameters in the *prototype buildings*.

The combination of *prototype buildings* and weather files for a given region creates *simulation-ready building models*. A single EnergyPlus simulation is conducted for each of the *simulation-ready building models*.

An overview of the BEND workflow is given in Figure 1. Several key points are illustrated by the figure:

- Building stock information is drawn from recent editions of EIA's CBECS (EIA 2003a) and RECS (EIA 2009a)
- After commercial and residential building stocks are characterized separately, the two sets are subjected to a similar series of steps to create *simulation-ready building models* (Steps 1–3)
- The resulting set of EnergyPlus simulations is conducted on a supercomputer (Step 4)
- The initial aggregation of results is done at the level of balancing authorities (BAs) due to the availability of historical data for calibration (Step 5)
- Final aggregation is done by first downscaling to the county level based on population data, then upscaling to the decision-relevant level of interest (Step 6).

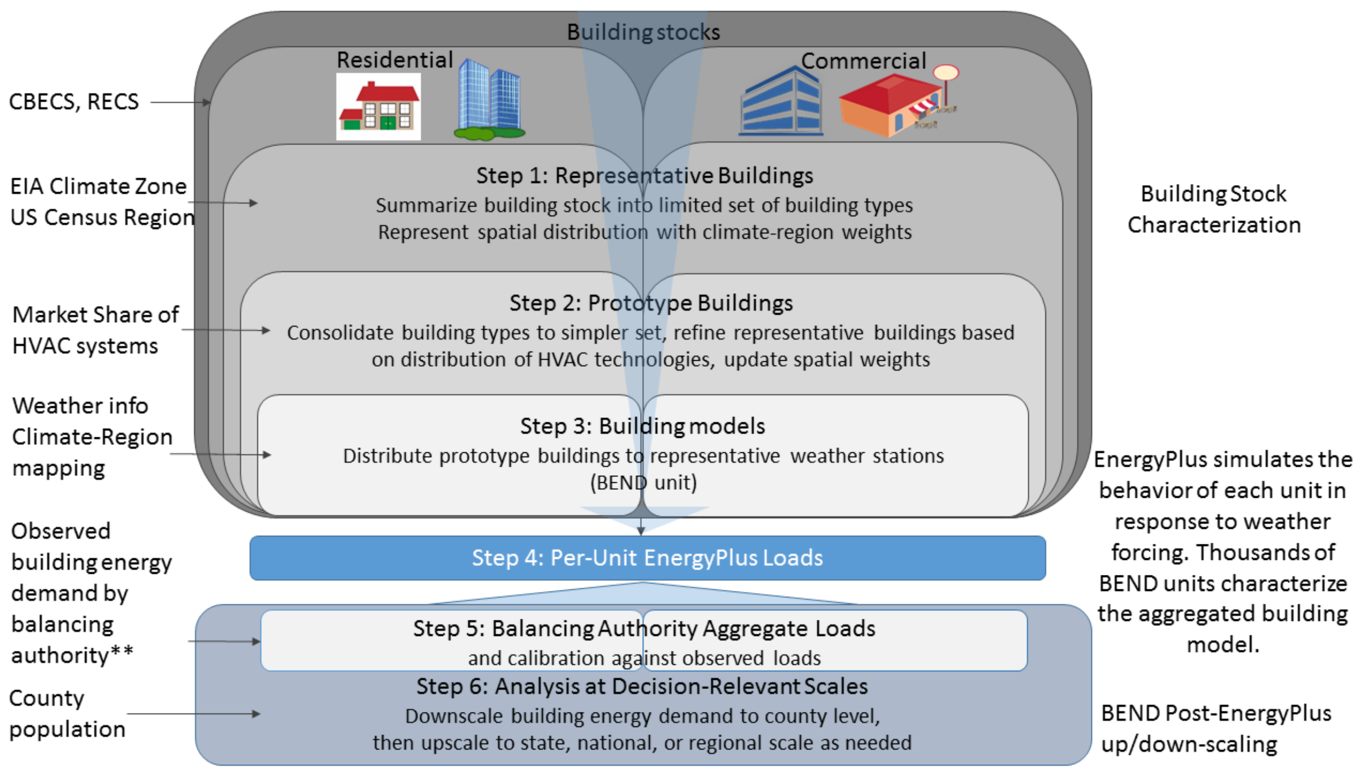


**Figure 1**. Overview of the BEND workflow

This section contains an overview of the tasks for the preprocessing, simulation, and post-processing steps. Section 3.0 gives detailed descriptions of the various methods used to carry out each task in BEND.

## 2.1 Preprocessing Tasks

In preparation for the EnergyPlus simulations, the BEND workflow includes (a) developing a set of statistically *representative buildin*gs for a specific geographic or climatic region of interest, including possibly the entire U.S.; (b) estimating a weight for each of the *representative buildings*, which carries information on the total number of similar buildings each *representative building* represents; (c) selecting base *prototype building* models that specify design characteristics for each of the *representative buildings*; (d) specifying detailed construction, thermal properties, and mechanical systems for the *prototype building* models to generate simulation inputs in EnergyPlus format based on four high-level features of the *representative buildings*; and (e) selecting a set of weather stations to represent the region of interest and combining each *prototype building* with an appropriate weather station to create *simulation-ready building models* for input to EnergyPlus.

### 2.1.1 Developing a Set of Representative Buildings

The building stock for a region of interest comprises an unknown number of unique buildings that vary in purpose, size, shape, age, composition, and many other features. The EIA conducts periodic surveys of the building stock with the goal of understanding the number of buildings, their characteristics, and the amount of energy they use. The CBECS and RECS are national surveys that sample thousands and tens of thousands of buildings, respectively. They include information on the spatial distribution of buildings, the distribution of buildings across service types, construction, age, and size. More details of the CBECS and RECS data sets are given in Section 3.1.1 (Drawing Building Stocks from CBECS and RECS). Because it is not computationally feasible to simulate all of the buildings captured by the surveys, BEND uses the CBECS and RECS data sets to develop a set of *representative buildings* to simulate.

BEND uses U.S. census regions and EIA climate zones, which are the base spatial segmentations in both CBECS and RECS, to characterize residential and commercial buildings within a given region of interest (Figure 2 and Figure 3). There are 15 unique combinations of census region and EIA climate zone in the continental U.S. (hereafter CR/CZs; Figure 4). For each of these 15 CR/CZs, BEND identifies a subset of *representative buildings* based on key features of the buildings within the region. BEND then uses this set of *representative buildings* as a statistically representative sample of the total building stock within the region. More details on the creation of each set of *representative buildings* are given in Section 3.1.2 (Developing a Set of Representative Buildings).

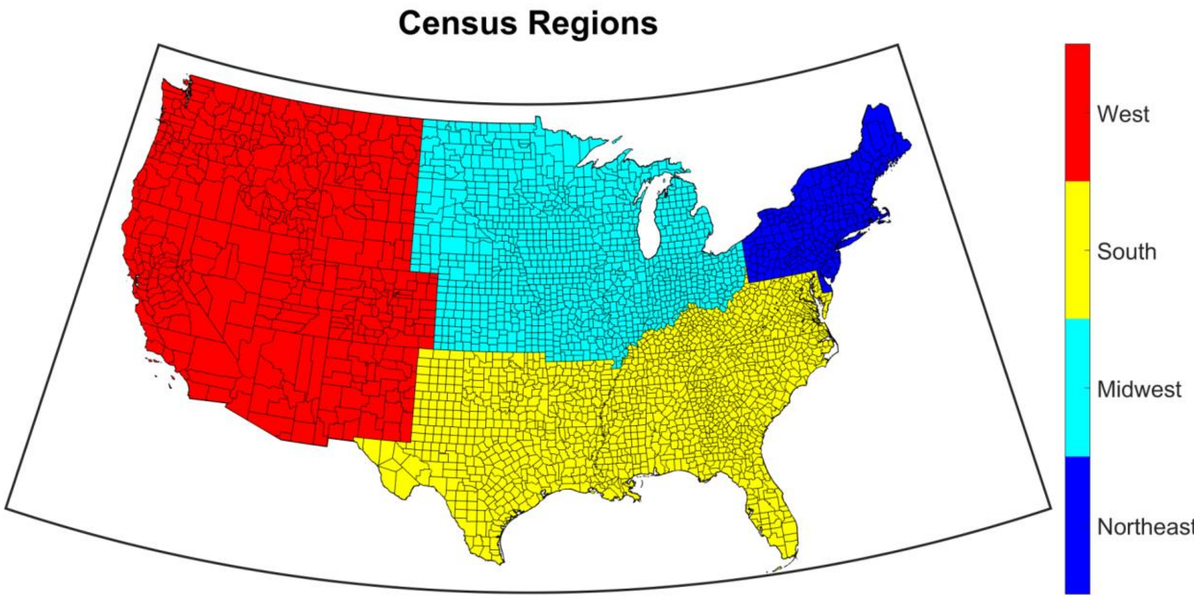


**Figure 2**. CBECS and RECS census regions

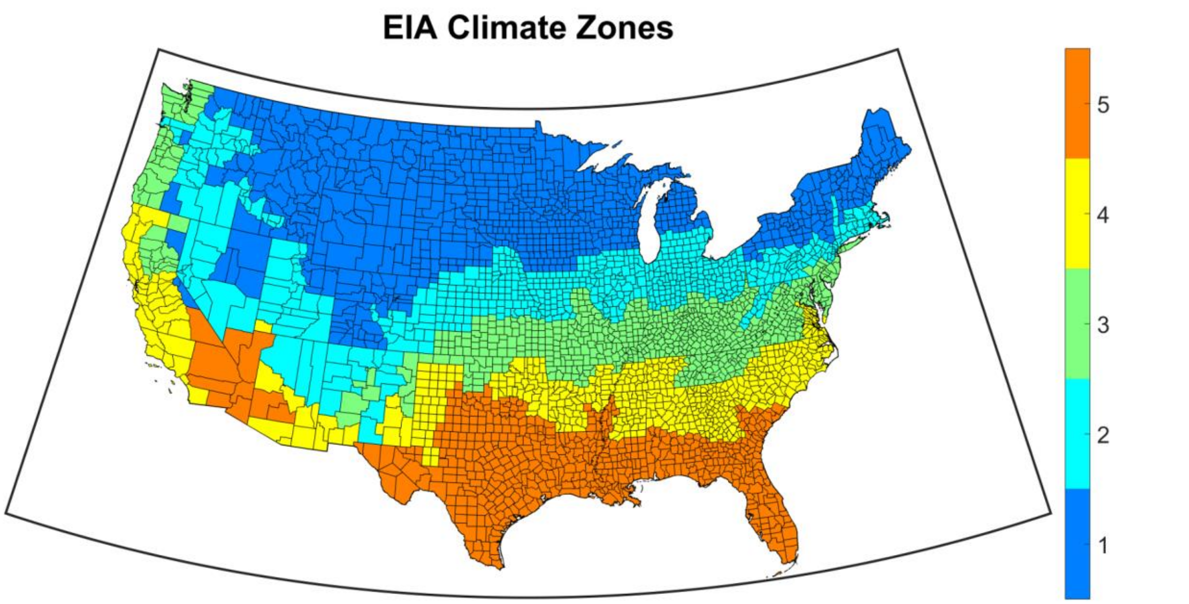


**Figure 3**. EIA climate zones

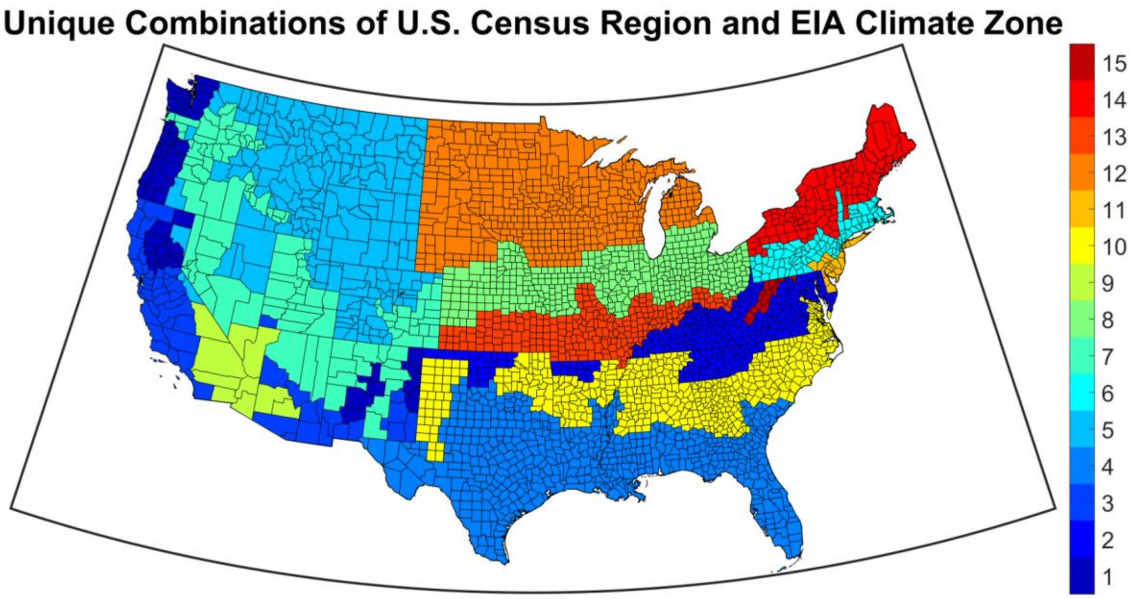


**Figure 4**. Unique combinations of CBECS and RECS census region and EIA climate zone (CR/CZs)

Each building surveyed in CBECS and RECS has an associated weight that denotes the number of buildings that the surveyed building represents. Because the *representative buildings* in BEND are designed to reflect the total building stock, these weights must be distributed from the total set of all buildings sampled down to the subset of *representative buildings*. In some cases, the *representative buildings* reflect multiple types of buildings surveyed, so the weights of those building types must be aggregated. Section 3.1.3 (Estimating the Initial Weights for the Set of Representative Buildings) discusses the method BEND uses to estimate the weights for each of the *representative buildings*.

### 2.1.2 Developing a Set of Prototype Buildings

The *representative buildings* are a statistical representation of the total building stock. Their characteristics reflect the key features of buildings in a given geographical area. For example, buildings in Montana will generally have more insulation than buildings in Arizona. In order to use the EnergyPlus simulation engine, BEND must generate more-detailed specifications for each of the *representative buildings*. For example, specifications of construction, thermal properties, internal loads, and equipment use are needed as input to EnergyPlus. To this end, BEND uses a set of *prototype buildings* that are detailed characterizations of the set of *representative buildings.* There is one base *prototype building* template for each of the *representative buildings*. This prototype building template has a geometrical skeleton that is specific for the building type and is scalable across the range of possible sizes in a given region. The detailed parameters for construction (e.g., wall type, roof type, and floor type), thermal properties (e.g., insulation of the building envelopes and U-factor and solar heat gain coefficient[1] of the windows), internal load (e.g., people and lighting densities), and the age and performance of the heating, ventilating, and air conditioning (HVAC) systems depend on the four high-level properties of the representative building (type, vintage, size, and location).

[1] The U-Factor measures the rate of heat transfer and tells how well the window insulates. The Solar Heat Gain Coefficient (SHGC) measures the fraction of solar energy transmitted and tells how well the product blocks heat caused by sunlight.
https://www.energy.gov/energysaver/design/windows-doors-and-skylights/energy-performance-ratings-windows-doors-and

The conversion from the set of *representative buildings* to the set of *prototype buildin*gs involves two key transformations. First, the four high-level properties (type, vintage, size, and location) of the *representative buildings* are converted into a much larger number of detailed parameters in the *prototype buildings*. Obviously there is variation in these parameters even after constraining the high-level features. For example, two commercial buildings of similar sizes that were built in the same year and county may have different HVAC systems. Because of this, the set of the *representative buildings* is subdivided into a larger set of *prototype buildings* to account for this variability. Borrowing the example above about different HVAC systems, when multiple HVAC systems are used for a single type of *representative building*, the *representative building* is subdivided into multiple *prototype buildings* that have distinct HVAC systems, but retain all other building properties. During this partitioning, the weight carried by each *representative building* is redistributed among the multiple *prototype buildings* according to the relative market share of each possible HVAC system. The conversion from the set of *representative buildings* to the set of *prototype buildings* is discussed in more detail in Section 3.1.4 (Developing a Set of Prototype Buildings).

### 2.1.3 Developing a Set of Simulation-Ready Building Models

In a given EnergyPlus run for a unique *prototype building,* the driving factor for changes in building energy demand is the weather. This weather can either be observed (i.e., based on historical observations) or predicted (i.e., taken from a climate model). In both cases, the weather from a single location has to be associated with a *prototype building* for EnergyPlus. Previous studies (e.g., NREL 2011; Hong et al. 2013; PNNL 2015) have based weather file selection for building simulations on IECC climate zones (Figure 5). The IECC climate zones, which are designed to support building energy construction codes, segregate the U.S. into geographical regions of similar climate. IECC climate zone boundaries are constrained to fall along county lines, which means they incorporate information other than climate, such as population and even historical precedent for certain building code requirements, and also that they ignore some important climatic considerations in favor of convenient geographical definitions that favor highly populated areas. For BEND’s purposes, the IECC climate zones’ consideration of multiple climate variables—temperature, humidity, rainfall, solar radiation, etc.—and the influence of population distribution on their boundaries make them an advantageous framework for choosing weather stations with which to drive the EnergyPlus simulations.

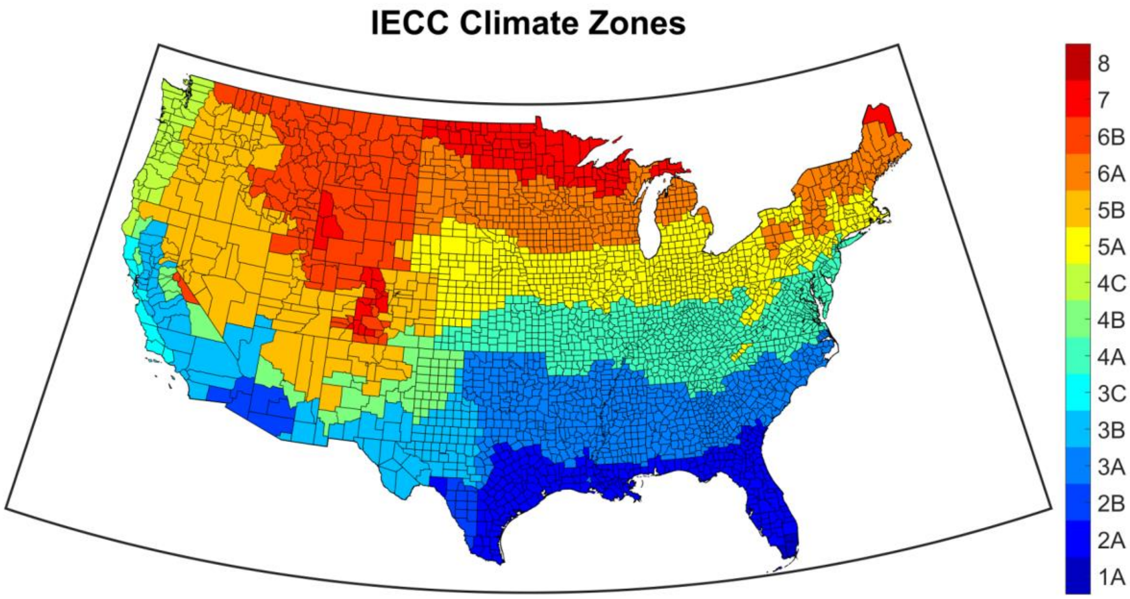


**Figure 5**. IECC climate zones

Using the IECC climate zones to determine which weather stations to use, and using the CR/CZs to determine the building characteristics, the intersection of IECC climate zones and CR/CZs defines the base spatial unit for BEND. At least one unique weather file is required for each of these base spatial units. In an ideal scenario, higher spatial granularity can be achieved if more than one weather station is used to capture the weather variability within each base spatial unit. The sensitivity to the number of weather stations utilized is evaluated in Burleyson et al. (2017).

The binding of weather files with *prototype buildings* for a given region leads to a set of unique *simulation-ready building models*. Because the spatial unit of the *prototype buildings* (CR/CZs; Figure 4) can be distinct from the spatial unit of the *simulation-ready building models* (the intersection of a CR/CZ and an IECC climate zone), each *prototype building* may be associated with more than one weather station. As with the conversion between the *representative buildings* set and the *prototype buildings* set, the weights carried by the *prototype buildings* need to be redistributed to the *simulation-ready building models* set. BEND uses county-level populations to carry out this redistribution. More details about the creation of the set of *simulation-ready building models* and the weight redistribution are given in Section 3.1.5 (Assigning Weather Stations and Developing a Set of Simulation-Ready Building Models).

## 2.2 EnergyPlus Runs

Once the set of *simulation-ready building models* has been generated, a set of EnergyPlus input files (i.e., .idf files) can then be derived. In the current version of BEND, PNNL's GParm (General PAraMetrics) tool first generates a set of XML files, and then uses a utility program developed by the National Renewable Energy Laboratory (NREL) to convert the .xml files into .idf files. After this step, there is one .idf file for each of the *simulation-ready building models*. At this point, a single EnergyPlus simulation is then conducted for each of the *simulation-ready building models*, using the .idf files to initialize the simulation. The generation of .idf files and mechanics of running of EnergyPlus many thousands of times is discussed in more detail in Section 3.2 (EnergyPlus Runs).

## 2.3 Post-Processing Tasks

Generally speaking, the base spatial units of BEND are not the most useful or insightful for forecasting or analyzing building energy demand for most users. For example, an individual BA may use BEND to forecast building energy demand under a changing climate, and thus would be interested in the aggregate demand within the bounds of their particular BA. Other questions of interest might be relevant at the state level. These are referred to as "decision-relevant spatial scales." Examples of decision-relevant spatial scales include BAs, the nodes/zones of a production cost model (e.g., PROMOD or PLEXOS), or larger-scale utility entities such as the WECC or the EIC. To provide estimates of building energy demand at decision-relevant spatial scales, BEND has the ability to upscale or downscale the simulation results by aggregating or disaggregating the simulated building energy demand across or to multiple base spatial units. BEND uses county-level populations to determine the weights for converting between larger and smaller spatial scales (in either direction). While the decision-relevant spatial scale will vary depending on the research question at hand, the up/downscaling scheme itself is generic. The scheme is discussed in more detail in Section 3.3.1 (Aggregation of the EnergyPlus Runs).

# 3.0 Detailed Methodology

This section describes in more detail each step of the BEND workflow as well as the various sources of data called upon along the way.

## 3.1 Preprocessing Tasks

### 3.1.1 Drawing Building Stocks from CBECS and RECS

CBECS comprises information on the physical characteristics of commercial buildings, building use and occupancy patterns, equipment use, and energy-conservation features and practices. By definition, commercial buildings are all buildings in which at least half of the floor space is used for a purpose that is not residential, industrial, or agricultural. The current CBECS represents data from calendar year 2012.[1] Like CBECS, RECS is also administered by the EIA. The information collected by RECS includes energy characteristics of each housing unit, energy usage patterns, and household demographics. The current RECS was conducted in 2015.[2]

CBECS and RECS include data from thousands and tens of thousands of sample buildings, respectively. There is a weight included for each sample in the survey, which is the reciprocal of the probability of the building being selected into the sample. The weights are adjusted to account for non-response bias. The adjusted weights denote the total number of the buildings that each sampled building represents; therefore, the adjusted weights can be used to obtain a national or regional estimate of the total number of buildings that are represented in the surveys.

The current version of BEND is based on the 2003 version of CBECS (hereafter CBECS03) and the 2009 version of RECS (hereafter RECS09). This means that BEND is not currently based on the most up-to-date estimates of the building stock in the U.S. Future development of the model may include an update of the base CBECS and RECS data sets. CBECS03 consists of 5,215 survey records representing

[1] EIA – Energy Information Administration. *Commercial Buildings Energy Consumption Survey*. https://www.eia.gov/consumption/commercial/about.php.
[2] EIA – Energy Information Administration. *Residential Buildings Energy Consumption Survey*. https://www.eia.gov/consumption/residential/about.php.

4,859,000 commercial buildings with a total estimated floor space of 71,658,000,000 square feet. RECS09 was based on 12,083 household samples, which represent 113,600,000 residential buildings with a total estimated floor space of 239,583,000,000 square feet.

Geographically, both the CBECS and RECS surveys are organized by U.S. census regions (Figure 2) and EIA climate zones (Figure 3), which are based on historical heating degree days and cooling degree days (HDD and CDD, respectively). There are 15 unique intersections between census regions and EIA climate zones (CR/CZs; Figure 4). These 15 CR/CZs are the geographic and climate areas in which the survey samples define BEND's *representative buildings* set.

The descriptions, meanings, and value ranges of all of the variables in CBECS03 and RECS09 are documented in EIA 2003b and EIA 2009b, respectively. Variables important to the setup of BEND include census region (REGION8 in CBECS03 or REGIONC in RECS09) as shown in Figure 2 and Table 1, EIA climate zone (CLIMAT in CBECS03 or AIA_Zone in RECS09) as shown in Figure 3 and Table 2, and building type (PBA8 in CBECS03 or TYPEHUQ in RECS09) as shown in Table 3 and Table 4 for CBECS03 and RECS09, respectively.

In total, there are more than 60 (CBECS03) and 930 (RECS09) variables in the surveys that describe the characteristics of the buildings sampled. As described in Section 2.1.1 (Developing a Set of Representative Buildings), a subset of these variables is used to characterize the important features of the buildings in the *representative buildings* set (Table 5). This subset is building type (PBA in CBECS03 or TYPEHUQ in RECS09), square footage (SQFT8 in CBECS03 or TOTSQFT in RECS09), construction year (YRCON8 in CBECS03 or YEARMADE in RECS09), and the associated weights of each sample (ADJWT8 in CBECS03 or NWEIGHT in RECS09).

**Table 1**. Census regions for CBECS03 and RECS09

| Label | REGION8 (CBECS03) \| REGIONC (RECS09) |
|---|---|
| 1 | Northeast |
| 2 | Midwest |
| 3 | South |
| 4 | West |

**Table 2**. EIA climate zones for CBECS03 and RECS09

| Label | CLIMAT (CBECS03) \| AIA_Zone (RECS09) |
|---|---|
| 1 | CDD < 2000 and HDD > 7000 |
| 2 | CDD < 2000 and 5500 < HDD < 7000 |
| 3 | CDD < 2000 and 4000 < HDD < 5499 |
| 4 | CDD < 2000 and HDD < 4000 |
| 5 | CDD ≥ 2000 and HDD < 4000 |

**Table 3**. Building types in CBECS03

| Label | PBA8 |
|---|---|
| 1 | Vacant |
| 2 | Office |
| 4 | Laboratory |
| 5 | Nonrefrigerated Warehouse |
| 6 | Food Sales |
| 7 | Public Order and Safety |
| 8 | Outpatient Health Care |
| 11 | Refrigerated Warehouse |
| 12 | Religious Worship |
| 13 | Public Assembly |
| 14 | Education |
| 15 | Food Service |
| 16 | Inpatient Health Care |
| 17 | Nursing |
| 18 | Lodging |
| 23 | Strip Shopping Mall |
| 24 | Enclosed Mall |
| 25 | Retail Other than Mall |
| 26 | Service |
| 91 | Other |

**Table 4**. Building types in RECS09

| Label | TYPEHUQ |
|---|---|
| 1 | Mobile Home |
| 2 | Single-Family Detached |
| 3 | Single-Family Attached |
| 4 | Apartment in Building with 2–4 Units |
| 5 | Apartment in Building with 5+ Units |

**Table 5**. Variables used to characterize *representative buildings*

| Description | CBECS03 | RECS09 |
|---|---|---|
| Building type | PBA | TYPEHUQ |
| Square footage | SQFT8 | TOTSQFT |
| Year of construction | YRCON8 | YEARMADE |
| Final full sample building weight | ADJWT8 | NWEIGHT |

### 3.1.2 Developing a Set of Representative Buildings

As described in Section 2.1.1 (Developing a Set of Representative Buildings), not only is it computationally impractical to simulate all of the thousands of buildings in the CBECS and RECS samples, but doing so would also require many difficult-to-justify inferences about the large number of building parameters required to run EnergyPlus for each individual building. Instead, BEND is designed to group buildings in CBECS and RECS based on common characteristics. BEND uses an approach similar to the so-called “most likely parameter-value” approach developed in the Facility Energy Decision System (FEDS; PNNL 2014) to develop sets of *representative buildings* for each of the 15 CR/CZs.

BEND uses a three high-level features (type, vintage, and size) to build a limited number of categories or bins within each CR/CZ. Using these features, BEND groups the buildings in the surveys into smaller subsets of *representative buildings*. The key assumption underpinning this is that buildings with the same four high-level characteristics (type, vintage, size, and location) will have statistically similar physical and thermal characteristics, functions and activities, operations, and heating/cooling systems. Hence, they will likely demand similar amounts of energy. Under this assumption, all of the various building samples that go into a given *representative building* may be simulated with the same base model without a significant loss of information.

In the current version of BEND there are five residential and 11 commercial building types. These building types are described in Table 6, which also shows the mapping between the BEND building type and the CBECS/RECS building types. These building types are roughly consistent with the nonmilitary FEDS building types. For mall buildings, a limited amount of information was collected in CBECS03, and as a result there are many blank fields in the data. Thus mall buildings are excluded from the current version of BEND. In addition to malls, surveyed building types of vacant, nursing, religious worship, and laboratory were excluded from the current version of BEND for similar reasons.

BEND uses seven vintage categories to represent the broad range of building ages and six size bins to capture the wide size distribution, as shown in Table 7 and Table 8, respectively. Universal size bins are defined for the commercial buildings and building-type-specific size bins are used for residential buildings.

All of the various combinations of high-level features (i.e., 11 commercial and five residential building types, seven vintage categories, and six size bins) leads to a set of 672 *representative buildings* ([11 + 5] × 7 × 6) in each of the 15 CR/CZs. Overall, BEND is based on a set of 10,080 *representative buildings* for the entire U.S. (i.e., 15 × 672; Table 9). Constructing sets of *representative buildings* for each of the 15 CR/CZs ensures that BEND has sufficient spatial granularity to support regional building energy studies.

The set of *representative buildings* is currently a static statistical representation of the building stock for the entire U.S. as well as the various subregions within in it. However, this set of *representative buildings* could be revised or replaced if new versions of CBECS or RECS are released, if new definitions are used for the building types, vintage categories, and size bins, or if new methods were used to derive the initial weights. Furthermore, the set of *representative buildings* could also be modified to reflect projected economic changes, technology innovation, population growth or redistribution, or climate change. This makes BEND a powerful tool for understanding the dynamic and potentially nonlinear response of building energy demand to any number of societally relevant future scenarios. However, without those changes, the set of the *representative buildings* can be considered a fixed and task-independent representation of the U.S. building stock. The set of *representative buildings* in BEND can be stored in a table with rows for the different *representative buildings* and columns denoting the four high-level features (type, vintage, size, and location).

**Table 6**. BEND building type mapped to CBECS and RECS building type

| CBECS03 \| RECS09 Index | Building Activities | BEND Index | BEND Type | XML Type |
|---|---|---|---|---|
| 1 | Vacant | – | – | – |
| 2 | Office | 8 | Office | Office |
| 4 | Laboratory | – | – | – |
| 5 | Nonrefrigerated warehouse | 10 | Warehouse & Storage | Warehouse |
| 6 | Food sales | 3 | Food Sales | Grocery |
| 7 | Public order and safety | 9 | Public Order/Safety | Police Station |
| 8 | Outpatient health care | 5 | Health Care | Health Care |
| 11 | Refrigerated warehouse | – | – | – |
| 12 | Religious worship | – | – | – |
| 13 | Public assembly | 1 | Assembly | Convention Center |
| 14 | Education | 2 | Education | Education:Secondary |
| 15 | Food service | 4 | Food Service | Dining:Family |
| 16 | Inpatient health care | 5 | Health Care | Health Care |
| 17 | Nursing | – | – | – |
| 18 | Lodging | 6 | Lodging | Motel |
| 23 | Strip shopping mall | – | – | – |
| 24 | Enclosed mall | – | – | – |
| 25 | Retail other than mall | 7 | Mercantile and Service | Retail |
| 26 | Service | 7 | Mercantile and Service | – |
| 91 | Other | 11 | Other | All Other |
| 1 | Mobile home | 16 | Mobile Homes | Residential |
| 2 | Single-family detached | 12 | Single Family Detached | Residential |
| 3 | Single-family attached | 13 | Single Family Attached | Residential |
| 4 | Apartment in building with 2–4 units | 14 | 2 to 4 Unit Buildings | Multi-Family Housing |
| 5 | Apartment in building with 5+ units | 15 | 5+ Unit Buildings | Multi-Family Housing |

**Table 7**. BEND building vintage categories

| Vintage Category | Vintages |
|---|---|
| 1 | <1946 |
| 2 | 1946–1960 |
| 3 | 1961–1973 |
| 4 | 1974–1979 |
| 5 | 1980–1986 |
| 6 | 1987–1996 |
| 7 | 1997–2003 |

**Table 8**. BEND size bins

| Size Bins | CBECS (all types) | RECS Single Family Detached | RECS Single Family Attached | RECS Apartment 2–4 Units | RECS Apartment 5+ Units | RECS Mobile |
|---|---|---|---|---|---|---|
| 1 | <5000 | <1350 | <1200 | <2100 | <5570 | <700 |
| 2 | 5000–10,000 | 1350–1800 | 1200–1600 | 2100–2700 | 5570–8050 | 700–850 |
| 3 | 10,000–25,000 | 1800–2300 | 1600–2000 | 2700–3300 | 8050–11,900 | 850–1000 |
| 4 | 25,000–50,000 | 2300–2800 | 2000–2600 | 3300–4500 | 11,900–21,400 | 1000–1200 |
| 5 | 50,000–100,000 | 2800–3650 | 2600–3600 | 4500–6900 | 21,400–53,400 | 1200–1450 |
| 6 | >100,000 | >3650 | >3600 | >6900 | >53,400 | >1450 |

**Table 9**. Number of *representative buildings* in BEND for each CR/CZ

| | | EIA Climate Zone | | | | |
|---|---|---|---|---|---|---|
| Census Region | | 1 | 2 | 3 | 4 | 5 |
| 1 | Northeast | 672 | 672 | 672 | – | – |
| 2 | Midwest | 672 | 672 | 672 | – | – |
| 3 | South | – | 672 | 672 | 672 | 672 |
| 4 | West | 672 | 672 | 672 | 672 | 672 |

### 3.1.3 Estimating the Initial Weights for the Set of Representative Buildings

The CBECS and RECS data sets include weights for the sampled buildings that reflect how many total buildings each sample represents. Those weights need to be translated to the 672 *representative buildings* used in BEND. In an ideal scenario there would be a one-to-one match. However, in many cases there are either zero or more than one surveyed buildings associated with each of the 672 *representative buildings* in each of the 15 CR/CZs. Weights need to be derived for each of the *representative buildings,* and the sum of the weights should match the sum of the weights in CBECS and RECS. BEND uses a combination of smoothing and interpolating the weights from CBECS and RECS to estimate the number of buildings and the floor space for each of the 672 *representative buildings*.

The smoothing and interpolation approach assumes that the distribution of surveyed buildings in a given CR is independent among EIA climate zones, building types, vintage bins, and size bins:

- The fraction of the total square footage and of the total number of buildings are first calculated for the above geographic and building characteristics
- The product of these individual fractions provides the overall fraction of the combination of building types, size bins, and vintage bins in a given CR/CZ
- The product of these fractions and the total square footage or the total number of buildings in the CR provides the number of buildings and total square footage of each combination.

The number of buildings is used as the initial weight. Together, the number of buildings and the square footage can be used to calculate an average square footage per building for each combination. This value may be useful in the future to calibrate the square footage.

Equation 1 calculates the overall number of buildings and total square footage in each CR.

$$\begin{cases} Wgt^{CR} = \sum_{i=1}^{5}\sum_{j=1}^{16}\sum_{k=1}^{7}\sum_{l=1}^{6} Wgt_{i,j,k,l}^{CR} \\ Sf^{CR} = \sum_{i=1}^{5}\sum_{j=1}^{16}\sum_{k=1}^{7}\sum_{l=1}^{6} Sf_{i,j,k,l}^{CR} \end{cases} \tag{1}$$

where *CR* denotes census region, *Wgt* and *Sf* are the weight and square footage of the surveyed building, respectively, and indices *i*, *j*, *k*, and *l* denote EIA climate zones, building types, building vintage bins, and building size bins, respectively.

By assuming independence among building types, vintages, and sizes, the overall number of buildings and square footage for each combination of EIA climate zone, building type, vintage, and size in each of the census regions are calculated using Equations 2-5, respectively.

EIA climate zone

$$\begin{cases} Wgt_{CZ}^{CR} = \sum_{j=1}^{16}\sum_{k=1}^{7}\sum_{l=1}^{6} Wgt_{j,k,l}^{CR-CZ} \\ Sf_{CZ}^{CR} = \sum_{j=1}^{16}\sum_{k=1}^{7}\sum_{l=1}^{6} Sf_{j,k,l}^{CR-CZ} \end{cases} \quad \text{where } CZ = 1,\ldots,5 \tag{2}$$

Type

$$\begin{cases} Wgt_{Bldg}^{CR} = \sum_{i=1}^{5}\sum_{k=1}^{7}\sum_{l=1}^{6} Wgt_{i,k,l}^{CR-Bldg} \\ Sf_{Bldg}^{CR} = \sum_{i=1}^{5}\sum_{k=1}^{7}\sum_{l=1}^{6} Sf_{i,k,l}^{CR-Bldg} \end{cases} \quad \text{where } Bldg = 1,\ldots,16 \tag{3}$$

Vintage

$$\begin{cases} Wgt_{Vintage}^{CR} = \sum_{i=1}^{5}\sum_{j=1}^{16}\sum_{l=1}^{6} Wgt_{i,j,l}^{CR-Vintage} \\ Sf_{Vintage}^{CR} = \sum_{i=1}^{5}\sum_{j=1}^{16}\sum_{l=1}^{6} Sf_{i,j,l}^{CR-Vintage} \end{cases} \quad \text{where } Vintage = 1,\ldots,7 \tag{4}$$

Size

$$\begin{cases} Wgt_{Size}^{CR} = \sum_{i=1}^{5}\sum_{j=1}^{16}\sum_{k=1}^{7} Wgt_{i,j,k}^{CR-Size} \\ Sf_{Size}^{CR} = \sum_{i=1}^{5}\sum_{j=1}^{16}\sum_{k=1}^{7} Sf_{i,j,k}^{CR-Size} \end{cases} \quad \text{where } Size = 1,\ldots,6 \tag{5}$$

The relative number or square footage of buildings that belong to a given EIA climate zone, building type, vintage, or size bin within a given census region can be calculated using Equations 6-9, respectively:

EIA climate zone
$$\begin{cases} F\_Wgt_{CZ}^{CR} = Wgt_{CZ}^{CR}/Wgt^{CR} \\ F\_Sf_{CZ}^{CR} = Sf_{CZ}^{CR}/Sf^{CR} \end{cases} \quad \text{where } CZ = 1,\ldots,5 \tag{6}$$

Type
$$\begin{cases} F\_Wgt_{Bldg}^{CR} = Wgt_{Bldg}^{CR}/Wgt^{CR} \\ F\_Sf_{Bldg}^{CR} = Sf_{Bldg}^{CR}/Sf^{CR} \end{cases} \quad \text{where } Bldg = 1,\ldots,16 \tag{7}$$

Vintage
$$\begin{cases} F\_Wgt_{Vintage}^{CR} = Wgt_{Vintage}^{CR}/Wgt^{CR} \\ F\_Sf_{Vintage}^{CR} = Sf_{Vintage}^{CR}/Sf^{CR} \end{cases} \quad \text{where } Vintage = 1,\ldots,7 \tag{8}$$

Size
$$\begin{cases} F\_Wgt_{Size}^{CR} = Wgt_{Size}^{CR}/Wgt^{CR} \\ F\_Sf_{Size}^{CR} = Sf_{Size}^{CR}/Sf^{CR} \end{cases} \quad \text{where } Size = 1,\ldots,6 \tag{9}$$

The overall relative number or square footage of buildings characterized by a given combination of high-level features (type, vintage, size, and location) can be calculated using Equation 10.

$$\begin{cases} F\_Wgt_{CZ-Bldg-Vinatge-Size}^{CR} = F\_Wgt_{CZ}^{CR} \times F\_Wgt_{Bldg}^{CR} \times F\_Wgt_{Vintage}^{CR} \times F\_Wgt_{Size}^{CR} \\ F\_Sf_{CZ-Bldg-Vinatge-Size}^{CR} = F\_Sf_{CZ}^{CR} \times F\_Sf_{Bldg}^{CR} \times F\_Sf_{Vintage}^{CR} \times F\_Sf_{Size}^{CR} \end{cases} \tag{10}$$

By scaling by the total number or square footage of all buildings in each CR, the total number or the total square footage of the buildings characterized by a given combination of high-level features (type, vintage, size, and location), which are the set of *representative buildings*, can be calculated using Equation 11.

$$\begin{cases} Wgt_{CZ-Bldg-Vinatge-Size}^{CR} = F\_Wgt_{CZ-Bldg-Vinatge-Size}^{CR} \times Wgt^{CR} \\ Sf_{CZ-Bldg-Vinatge-Size}^{CR} = F\_Sf_{CZ-Bldg-Vinatge-Size}^{CR} \times Sf^{CR} \end{cases} \tag{11}$$

We use data from the WECC as an example to clarify the steps for deriving the initial weights for the set of representative buildings. The WECC lies entirely within a single census region (CR 4) and spans all five EIA climate zones (CZs), so in total there are 1 × 5 × 672 = 3360 *representative buildings* in BEND that are used to represent the building stock in the WECC. An R script (1_wgt_calc_from_database.R) is used to analyze the CBECS03 and RECS09 databases to derive the initial weights for the 3360 *representative buildings* in the WECC.

### 3.1.4 Developing a Set of Prototype Buildings

The *representative buildings* in BEND are characterized by a relatively small number of high-level building and geographic/climatic features. More-detailed specifications are needed in order to prepare the buildings for simulation using EnergyPlus. For example, EnergyPlus requires detailed construction and thermal properties, internal loads and equipment use, and HVAC system parameters. The set of *representative buildings* are expanded by adding the necessary details and creating a set of *prototype buildings*. Each of the 672 *representative buildings* in a given CR/CZ is matched to a *prototype building* template based on the building type. The *prototype building* templates have a geometrical skeleton that is scalable according to the size bin of the *representative building*. The detailed parameters for construction (e.g., wall type, roof type, and floor type), thermal properties (e.g., insulation of the building envelopes and U-factor and solar heat gain coefficient of the windows), internal load (e.g., people and lighting densities), and the age and performance of the HVAC systems are inferred from the four high-level properties of the *representative buildings* (type, vintage, size, and location).

For a given *representative building* with a given type-vintage-size combination in a given CR/CZ, there are naturally large variations in construction types, thermal properties, internal loads and equipment, HVAC systems, and schedules. For simplicity, the current version of BEND does not account for variations in most of the parameters. Instead, for most parameters it simplifies things by specifying the most likely set of parameters for a given *representative building*. However, BEND does account for variations in HVAC systems. During the expansion between *representative buildings* and *prototype buildings*, each prototype building is given a unique HVAC system while retaining all other physical, thermal, mechanical, and operational properties.

BEND uses paired heating and cooling technologies to represent the range of possible HVAC systems in use. Table 10 and Table 11 give details about the different heating and cooling technologies considered in BEND. A larger number of the HVAC pairs were constructed based on the heating and cooling technologies. BEND uses 33 HVAC pairs defined for the 11 different types of commercial buildings (Table 12). A different list of heating technologies, cooling technologies, and possible HVAC pairs were defined for each of the five different types of residential buildings (not shown).

**Table 10**. Heating technologies

| Heating Technology # | Heating Technology |
|---|---|
| 1 | Furnace |
| 2 | Natural draft boiler w/ AHU |
| 3 | Natural draft boiler w/ FCU |
| 4 | Natural draft boiler w/ radiator |
| 5 | Mechanical draft boiler w/ AHU |
| 6 | Mechanical draft boiler w/ FCU |
| 7 | Mechanical draft boiler w/ radiator |
| 8 | Purchased heat w/ AHU |
| 9 | Purchased heat boiler w/ FCU |
| 10 | Purchased heat w/ radiator |
| 11 | Air-source heat pump |
| 12 | Ground-coupled heat pumps |
| 13 | Gas-driven heat pumps |
| 14 | Individual space heater |
| 15 | No heating |
| 16 | Infer heating technology from high-level properties |

AHU is air handling unit and FCU is fan coil unit

**Table 11.** Cooling technologies

| Cooling Technology # | Cooling Technology |
|---|---|
| 1 | Air cooled reciprocating chiller w/ AHU |
| 2 | Air cooled reciprocating chiller w/ FCU |
| 3 | Water cooled reciprocating chiller w/ AHU |
| 4 | Water cooled reciprocating chiller w/ FCU |
| 5 | Water cooled centrifugal chiller w/ AHU |
| 6 | Water cooled centrifugal chiller w/ FCU |
| 7 | Absorption chiller w/ AHU |
| 8 | Absorption chiller w/ FCU |
| 9 | Purchased cooling w/ AHU |
| 10 | Purchased cooling w/ FCU |
| 11 | Gas engine driven chiller w/ AHU |
| 12 | Gas engine driven chiller w/ FCU |
| 13 | Air-source heat pump |
| 14 | Water-source heat pump |
| 15 | Gas-driven heat pump |
| 16 | Air-source DX |
| 17 | Room A/C |
| 18 | No cooling |
| 19 | Infer cooling technology from high-level properties |

DX is direct expansion

**Table 12**. HVAC pairs for BEND commercial buildings

| System Name | Heating | Cooling |
|---|---|---|
| Direct fired heat with DX | Electric heat | Electric cooling |
| Direct fired heat with DX | Gas heat | Electric cooling |
| Direct fired heat with DX | Liquid heat | Electric cooling |
| Direct fired heat with chiller | Electric heat | Electric cooling |
| Direct fired heat with chiller | Gas heat | Electric cooling |
| Direct fired heat with chiller | Liquid heat | Electric cooling |
| Direct fired heat with chiller | Gas heat | Other cooling |
| Direct fired heat with no cooling | Electric heat | No cooling |
| Direct fired heat with no cooling | Gas heat | No cooling |
| Direct fired heat with no cooling | Liquid heat | No cooling |
| No heating and No cooling | No heat | No cooling |
| Heat pumps and heat pumps | Electric heat | Electric cooling |
| Distr. steam/hot water and distr. chilled water | Distr. heat | Distr. cooling |
| Indiv. space heaters and individual room A/C | Electric heat | Electric cooling |
| Indiv. space heaters and individual room A/C | Gas heat | Electric cooling |
| Distr. steam/hot water and central chillers | Distr. heat | Electric cooling |
| Distr. steam/hot water and central chillers | Distr. heat | Other cooling |
| Indiv. space heaters and no cooling | Electric heat | No cooling |
| Indiv. space heaters and no cooling | Gas heat | No cooling |
| Indiv. space heaters and no cooling | Liquid heat | No cooling |
| Boilers and individual room A/C | Electric heat | Electric cooling |
| Boilers and individual room A/C | Gas heat | Electric cooling |
| Boilers and individual room A/C | Liquid heat | Electric cooling |
| No heating and DX | No heat | Electric cooling |
| Direct fired and distr. chilled water | Electric heat | Electric cooling |
| Direct fired and distr. chilled water | Gas heat | Electric cooling |
| Direct fired and distr. chilled water | Liquid heat | Electric cooling |
| Distr. steam/hot water and no cooling | Distr. heat | No cooling |
| Distr. steam/hot water and DX | Distr. heat | Electric cooling |
| Indiv. space heaters and distr. chilled water | Electric heat | Electric cooling |
| No heating and central chillers | No heat | Electric cooling |
| No heating and distr. chilled water | No heat | Distr. cooling |
| No heating and individual room A/C | No heat | Electric cooling |

After identifying the various possible combinations of HVAC systems, the next step is to estimate how often each combination occurs across the building stock. BEND uses a series of regression models, one model for each of the possible HVAC systems, to estimate the market share of each of the possible HVAC systems for a given *representative building*. The regression estimates are treated as probabilities and are used to rank the relative market shares of the possible HVAC systems. The regression coefficients are documented in an Excel spreadsheet (SetBuilderAll.xlsx). Rather than using all possible HVAC systems, BEND takes a subset of the most frequently used systems. Starting from 0% market share, BEND incrementally adds more and more HVAC systems (starting from the most likely and going down the list) until more than 60% of the total market share of possible HVAC systems is accounted for. The set of *N* HVAC systems that add up to more than 60% of the total market shares is then used and a *prototype building* is created for each of the *N* systems for a given *representative building*. The exact number (*N*) and specifications of the HVAC systems vary with each *representative building*. The market shares of the *N* selected HVAC systems for each *representative building* are renormalized to sum to 1 in order to provide the relative frequency of each HVAC system in each *representative building*.

For example, for the *representative building* characterized by the type, vintage, and size combinations of 1-1-1 in CR 1 and CZ 1, there are 33 possible combinations of heating and cooling technologies estimated from the regression models for commercial buildings. Two types of HVAC systems dominate the market: HVAC pair 2 had 47% of the total market shares and HVAC pair 9 had another 14.2%. The combination of those two HVAC pairs exceeds the 60% threshold and thus only those two types are retained. Their renormalized fractions are 0.769 (HVAC pair 2) and 0.233 (HVAC pair 9). Therefore the *representative building* 1-1-1 in CR 1 and CZ1 becomes two *prototype buildings*, one of which uses HVAC pair 2 and the other HVAC pair 9. Recall that all parameters other than heating and cooling technologies of the two prototype buildings are the same. The initial weight of this *representative building* (i.e., the number of buildings that it represents) is then redistributed into portions of 76.9% (HVAC pair 2) and 23.3% (HVAC pair 9) for the two *prototype buildings*.

Starting from the original 672 *representative buildings* in each CR/CZ (Table 9), the expansion to account for multiple HVAC systems results in a much larger set of *prototype buildings* (Table 13). The exact number of *prototype buildings* varies by CR/CZ, but in each CR/CZ there are at least twice as many *prototype buildings* as *representative buildings*. Using the WECC as an example (CR 4), the 672 *representative buildings* are converted into 9682 *prototype buildings* distributed among the 5 CZs. The initial weights associated with the 672 *representative buildings* are redistributed to the *prototype buildings* according to their relative frequencies.

**Table 13**. Number of *prototype buildings* in BEND for each CR/CZ

| | | EIA Climate Zone | | | | |
|---|---|---|---|---|---|---|
| Census Region | | 1 | 2 | 3 | 4 | 5 |
| 1 | Northeast | 2174 | 2008 | 2201 | – | – |
| 2 | Midwest | 1517 | 1373 | 1537 | – | – |
| 3 | South | – | 1956 | 2087 | 1879 | 1693 |
| 4 | West | 2042 | 1869 | 2046 | 1896 | 1729 |

Regression relationships similar to the one described above for designation of the most common HVAC systems are available for inferring the values of many other detailed parameters from the four high-level building characteristics (type, vintage, size, and location). The current version of BEND ignores the variation of all parameters other than HVAC system. Instead, BEND uses the regression relationships to identify a single—the most likely—value for each of the parameters for each *representative building.* For example, consider the wall type of commercial buildings. There are five dominant wall types: engineered metal, masonry on steel, masonry on masonry frame, masonry on wood frame, and siding on wood frame. Regression models were established to relate the high-level building features to each of the five dominant wall types based on CBECS03. BEND applies these regression models to each commercial *representative building*, and the wall type with the largest regression weight is selected as the representative wall type for that particular *representative building*. All *prototype buildings* based on this *representative building* will use that one type of wall. Once the wall type is selected, the insulation levels of the wall (rinsul) and its sheathing (rsheath), both of which are required to run EnergyPlus, are retrieved from Table 14 based on the size of the *representative building* and which CR/CZ it is in.

**Table 14**. Insulation of wall construction

| Size Bin → | 1, 2, 3 | | 4 | | 5 | | 6 | | 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | rinsul | rsheath | rinsul | rsheath | rinsul | rsheath | rinsul | rsheath | rinsul | rsheath | EIA climate_zone |
| Siding on Wood Frame | 7 | 0.62 | 11 | 2.62 | 19 | 2.62 | 19 | 2.62 | 19 | 2.62 | 1 |
| | 7 | 0.62 | 11 | 2.62 | 19 | 2.62 | 19 | 2.62 | 19 | 2.62 | 2 |
| | 7 | 0.62 | 11 | 2.62 | 19 | 2.62 | 19 | 2.62 | 19 | 2.62 | 3 |
| | 7 | 0.62 | 11 | 2.62 | 19 | 2.62 | 19 | 2.62 | 19 | 2.62 | 4 |
| | 7 | 0.62 | 11 | 2.62 | 19 | 2.62 | 19 | 2.62 | 19 | 2.62 | 5 |
| Masonry on Wood Frame | 7 | 0.62 | 11 | 2.62 | 19 | 2.62 | 19 | 2.62 | 19 | 2.62 | 1 |
| | 7 | 0.62 | 11 | 2.62 | 19 | 2.62 | 19 | 2.62 | 19 | 2.62 | 2 |
| | 7 | 0.62 | 11 | 2.62 | 19 | 2.62 | 19 | 2.62 | 19 | 2.62 | 3 |
| | 7 | 0.62 | 11 | 2.62 | 19 | 2.62 | 19 | 2.62 | 19 | 2.62 | 4 |
| | 7 | 0.62 | 11 | 2.62 | 19 | 2.62 | 19 | 2.62 | 19 | 2.62 | 5 |
| Masonry on Masonry Frame | 3.2 | 0 | 3.2 | 0 | 3.2 | 0 | 6.4 | 0 | 13 | 0 | 1 |
| | 3.2 | 0 | 3.2 | 0 | 3.2 | 0 | 4.8 | 0 | 11 | 0 | 2 |
| | 3.2 | 0 | 3.2 | 0 | 3.2 | 0 | 4.8 | 0 | 11 | 0 | 3 |
| | 3.2 | 0 | 3.2 | 0 | 3.2 | 0 | 4.8 | 0 | 11 | 0 | 4 |
| | 3.2 | 0 | 3.2 | 0 | 3.2 | 0 | 4.8 | 0 | 11 | 0 | 5 |
| Masonry on Steel | 7 | 0 | 11 | 0 | 11 | 0 | 19 | 0 | 19 | 0 | 1 |
| | 7 | 0 | 11 | 0 | 11 | 0 | 11 | 0 | 19 | 0 | 2 |
| | 7 | 0 | 11 | 0 | 11 | 0 | 11 | 0 | 13 | 0 | 3 |
| | 7 | 0 | 11 | 0 | 11 | 0 | 11 | 0 | 13 | 0 | 4 |
| | 7 | 0 | 11 | 0 | 11 | 0 | 11 | 0 | 13 | 0 | 5 |
| Engineered Metal | 5.39 | 0 | 6.63 | 0 | 8.79 | 0 | 12.8 | 0 | 19 | 0 | 1 |
| | 5.39 | 0 | 6.63 | 0 | 8.79 | 0 | 11.2 | 0 | 13 | 0 | 2 |
| | 5.39 | 0 | 6.63 | 0 | 8.79 | 0 | 9.6 | 0 | 9.6 | 0 | 3 |
| | 5.39 | 0 | 6.63 | 0 | 8.79 | 0 | 9.6 | 0 | 9.6 | 0 | 4 |
| | 5.39 | 0 | 6.63 | 0 | 8.79 | 0 | 9.6 | 0 | 9.6 | 0 | 5 |

Many other detailed parameters needed for EnergyPlus are derived using a similar process. For each *representative building* with four high-level characteristics (type, vintage, size, and location), the values of more than 500 parameters are defined in each of the associated *prototype buildings*. Building on the WECC example (CR 4), the parameters for each of the 9582 *prototype buildings* are stored in a parameter file with 9582 rows and 500+ columns, with each column containing the parameter values needed to run EnergyPlus. The definitions of the predefined building types and parameters of the building construction, equipment, and mechanical systems along with the regression coefficients for the *representative buildings* characterized by given high-level features are documented in an Excel spreadsheet (InputWorksheet.xlms).

In the current version of BEND, the set of *prototype buildings* is considered to be static. However, if the set of *representative buildings* were updated, or if existing building construction types, equipment, lighting, or mechanical systems were redefined or improved, it is possible to modify the *prototype buildings* by taking a market-share approach to other parameters in addition to HVAC systems. For example, the current predefined HVAC systems may not reflect changes in HVAC system popularity or include any developments in HVAC technology in the future. The regression approach used for designating the detailed building parameters could also be enhanced. While all of these are possible in BEND, the current set of *prototype buildings* can be considered a fixed and task-independent set of detailed buildings that are representative of the current total U.S. building stock.

### 3.1.5 Assigning Weather Stations and Developing a Set of Simulation-Ready Building Models

For a given *prototype building* with fixed parameters, a time series of varying weather conditions is the driving force of variability in the EnergyPlus simulation. Energy usage is simulated using forcing from EnergyPlus Weather (.epw) files that contain an hourly time series of observed or predicted meteorological variables (e.g., temperature, humidity, or solar radiation). These forcing files are the primary mechanism by which the model responds to changes in weather and they represent the physical linkage between climate and building energy demand. The selection of weather data sets or locations to force the model is a key component of the simulation design. The binding of a *prototype building* and an .epw weather file leads to a unique *simulation-ready building model*. The development of these *simulation-ready building models* is the final preprocessing step for BEND.

The energy-related climate regions adopted by the IECC and the American Society of Heating, Refrigerating, and Air-Conditioning Engineers are widely used in building energy simulations. The IECC climate zones, which were developed by PNNL in the early 2000s, are based on an analysis of 4,775 weather stations across the U.S. as well as widely accepted classifications of world climate regimes that have been applied in a variety of different disciplines. The IECC climate zones divide the U.S. into eight primarily temperature-oriented zones and three moisture regimes (moist [A], dry [B], and marine [C]; Figure 5). Although there are theoretically $8 \times 3 = 24$ possible IECC zone/regime combinations, only 15 occur in the U.S. For clarity these are referred to as IECC climate zones, which are similar to but do not necessary align directly with the EIA climate zones used to distinguish building characteristics in the CBECS and RECS data sets.

The set of *representative buildings,* and subsequently the set of *prototype buildings,* are defined for census regions and EIA climate zones (CR/CZs). BEND uses the IECC climate zones to identify the best weather stations for a given region. Associating weather stations with specific *prototype buildings* is facilitated by overlaying the IECC climate zones on the CR/CZs. Every unique intersection of these overlaid data sets represents a unique combination of building characteristics and weather (Figure 6). These small regions constitute the minimum spatial unit (MSU) in the BEND model.

**Weather Station and CR/CZ Combinations**

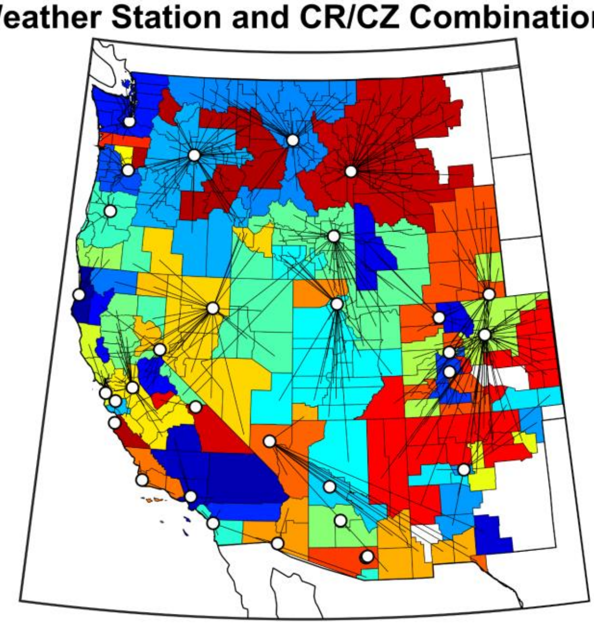

**Figure 6**. Example mapping between weather stations (4 per IECC climate zone) and CR/CZs

The selection of both the number and location of weather files for building energy simulations is a broad topic of research. Ideally, if there are no issues with the availability of weather data or restrictions on computational resources, one should use as many weather stations as possible in order to most completely capture the weather within each IECC climate zone. In the continental U.S., there are on the order of 1,000 possible weather stations. As an alternative to using data from weather stations, which are heterogeneously distributed point measurements, one could use gridded weather data from an atmospheric model or reanalysis. Many atmospheric models or reanalyses can produce hourly estimates of the weather conditions on a regular latitude-longitude grid with 1/8° spatial resolution (~12 km × 12 km). These gridded weather datasets contain over 65,000 pseudo-stations for the continental U.S. In reality, both the availability of the pertinent weather data and limitations on computational resources restrict the total number of stations that can be practically used. Given these restrictions, the choice of number and location of weather stations to use is a critical, user-specific exercise.

The WECC's eight distinct IECC climate zones suggest that a minimum of eight weather stations are needed. Burleyson et al. (2017) investigated biases in temperature and simulated building energy demand when different numbers of weather stations, ranging from eight (one from each IECC climate zone in the WECC) to roughly 150 (all available weather stations in the WECC), were used. They found that the minimum set of eight weather stations leads to significant biases in both temperature and simulated building energy demand. Both biases were reduced when using more than one weather station per IECC climate zone. However, the degree to which the bias is reduced was nonlinear. Total biases decreased significantly when the quantity of weather stations was increased from one to four per IECC climate zone, but decreased much more slowly when using more than four weather stations per IECC climate zone. Based on this result, the current version of BEND, which is focused on simulating energy demand in the WECC, uses four weather stations per IECC climate zone for a total of 32 weather stations across the WECC (Figure 6). The specific weather stations used in each IECC climate zone in the WECC are given in

Table 15 and in each EIA climate zone in Table 16. Note that after these stations were selected, two of the stations were dropped due to issues with data availability. The selected weather stations were reorganized by CR/CZ to match the spatial scale of the *prototype buildings*. Each *prototype building* in a given CR/CZ is simulated using forcing from all of the weather stations in that CR/CZ, and the total output is aggregated based on the methods described in Section 3.3.1. The binding of each *prototype building* with a unique weather file leads to a *simulation-ready building model*. For the current setup of BEND, there are 90,307 *simulation-ready building models* for the WECC based on the choice of four weather stations per IECC climate zone (Table 17).

It should be pointed out that while the set of *representative buildings* and the set of *prototype buildings* are generic in the current version of BEND, the creation of the set of *simulation-ready building models* is task specific. As described above, the *simulation-ready building models* depend on the region of interest and the choice of the quantity and locations of the weather stations used to force the *prototype models*. The total of 90,307 *simulation-ready building models* described above and listed in Table 17 is specific to a WECC study using four stations per IECC climate zone to force the simulation. If a different number or different locations of the weather stations were used (e.g., the 65,000 pseudo-stations), the number of *simulation-ready building models* will change. The total number is not linearly dependent on the number of weather stations used.

**Table 15**. The WMO[(a)] identifier for the four weather stations in each of the eight IECC climate zones in the WECC

| | IECC Climate Zone | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 2B | 3B | 3C | 4B | 4C | 5B | 6B | 7 |
| 1 | 699604 | 722880 | 724940 | 723723 | 725945 | 724666 | 725640 | 724676 |
| 2 | 722780 | 722926 | 722895 | 723650 | 726930 | 725755 | 725785 | 724677 |
| 3 | 722745 | 723865 | 724915 | 724800 | 726985 | 725830 | 726797 | 725715 |
| 4 | 722740[(b)] | 724920 | 724945 | 725846 | 727938 | 727845 | 727730 | 724673[(b)] |

a. WMO is the World Meteorological Organization
b. Station not used because of poor data availability

**Table 16**. The WMO identifier for the 30 weather stations in the WECC mapped to EIA climate zones

| EIA CZ1 | EIA CZ2 | EIA CZ3 | EIA CZ4 | EIA CZ5 |
|---|---|---|---|---|
| 724666 | 723650 | 723650 | 722745 | 699604 |
| 724676 | 723865 | 724666 | 722880 | 722780 |
| 724677 | 724666 | 724920 | 722895 | 722880 |
| 725640 | 724676 | 725830 | 722926 | 722926 |
| 725715 | 725640 | 725945 | 723723 | 723865 |
| 725755 | 725755 | 726930 | 723865 | 724800 |
| 725785 | 725785 | 726985 | 724800 | – |
| 725830 | 725830 | 727938 | 724915 | – |
| 726797 | 726985 | – | 724920 | – |
| 727730 | 727845 | – | 724940 | – |
| 727845 | 727938 | – | 724945 | – |
| – | – | – | 725830 | – |
| – | – | – | 725846 | – |
| – | – | – | 725945 | – |

**Table 17**. Total number of *simulation-ready building models* in the WECC for each EIA climate zone

| | | EIA Climate Zone | | | | | |
|---|---|---|---|---|---|---|---|
| Census Region | | 1 | 2 | 3 | 4 | 5 | Total |
| 4 | West | 22,462 | 20,559 | 16,368 | 26,544 | 10,374 | 96,307 |

## 3.2 EnergyPlus Runs

Once the set of *prototype buildings* is determined, a set of EnergyPlus input files (i.e., .idf files) can then be derived. The generation of the set of .idf files is a two-step process in BEND. First, the parameter file describing the detailed characteristics of the *prototype buildings* and a template file describing the high-level characteristics are processed by a PNNL-developed utility program called GParm (General PARaMetrics). The template file, which for BEND is in extensible markup language (XML) format, contains a full description of the building but with its detailed characteristics left out. The characteristics are given as replaceable tags. The parameter file contains the values that the replaceable tags take on with one *prototype building* per row. GParm reads each row of the parameter file, which consists of values of the entire set of parameters for a particular *prototype building*, and uses the values to replace the replaceable tags in the template file. The output of this process is a set of .xml files, with the number of .xml files equal to the number of *prototype buildings*. The next step is to use an NREL-developed XML processor to convert the set of .xml files into a corresponding set of EnergyPlus model input files (i.e., .idf files). There is one .idf file for each of the *prototype buildings*. Finally, after each *prototype building* is paired with one or more weather stations to produce the set of *simulation-ready building models*, an EnergyPlus simulation is conducted for each *simulation-ready building model*. The output of each EnergyPlus simulation is an annual hourly time series of total electrical energy consumption and, if desired, the energy consumption for individual fuel types or end uses.

## 3.3 Post-Processing Tasks

### 3.3.1 Aggregation of the EnergyPlus Runs

The EnergyPlus simulation results correspond to the MSUs of BEND. In each BEND MSU, there is a large number of *simulation-ready building models* representing the mix of the possible buildings in that region. The simulations of the *simulation-ready building models* in each MSU are aggregated into separate residential and commercial building energy demand profiles. This aggregation is straightforward because each *simulation-ready building model* has an associated weight derived from CBECS and RECS.

As the BEND MSUs are not generally the spatial scale of interest the results for this first aggregation step must be transformed—either upscaled to larger regional units or downscaled to smaller regional units. The spatial regions of interest will be at a decision-relevant scale such as BAs, nodes/zones of a production cost model (e.g., PROMOD or PLEXOS), regional administration authorities (e.g., the WECC or EIC), states, or even the entire U.S. The scaling process is task dependent, but the scaling scheme is generic. Usually both the MSUs and the decision-relevant scale of interest involve geographical regions larger than individual counties, so counties are used as a common base for scaling the energy simulation results. Therefore, the aggregate residential and commercial profiles in each MSU are first downscaled (disaggregated) to the county level and then upscaled (aggregated) to the scale of interest using population weighting in both directions.

As an example, consider the use of BEND to estimate the building energy demand for each of the 29 BAs in the WECC to describe the process of aggregating and scaling the raw EnergyPlus simulation results to residential and commercial demand profiles at each of the BAs. Equation 12 starts with the outcome of the 96,307 *simulation-ready building models* that have been simulated for the WECC. There is a unique weight associated with each *simulation-ready building model*. Spatially, these 96,307 *simulation-ready building models* are distributed over the 50 MSUs in BEND for the WECC (Figure 6). The annual hourly load profiles of the *simulation-ready building models* that fall into each of these 50 MSUs are aggregated by applying the weights. The aggregation is conducted separately on commercial and residential buildings. The aggregation is a simple arithmetic weighted sum and the results of the aggregation process are 50 pairs of commercial and residential annual hourly load profiles, one for each of the 50 MSUs in the WECC.

$$\left[\begin{array}{l}\text{8,760}\\ \text{annual}\\ \text{hourly load}\\ \text{profiles for}\\ \text{each of the}\\ \text{96,307}\\ \text{building}\\ \text{models}\end{array}\right]_{8,760}^{96,307} \times \left[\begin{array}{l}\text{Diagonal}\\ \text{matrix with}\\ \text{weights of}\\ \text{the 96,307}\\ \text{building}\\ \text{models in}\\ \text{the}\\ \text{diagonal}\\ \text{elements}\end{array}\right]_{96,307}^{96,307} \times \left[\begin{array}{l}\text{Source}\\ \text{matrix with}\\ \text{“1” in the}\\ \text{element}\\ \text{when a}\\ \text{building}\\ \text{model is}\\ \text{located in}\\ \text{one of the}\\ \text{50 spatial}\\ \text{intersections}\end{array}\right]_{96,307}^{50} = \left[\begin{array}{l}\text{8,760 annual}\\ \text{hourly load}\\ \text{profiles at}\\ \text{each of the}\\ \text{50 spatial}\\ \text{intersections}\end{array}\right]_{8,760}^{50} \qquad (12)$$

Because the county serves as the base for the scaling process, the 50 pairs of commercial and residential annual hourly load profiles in each MSU are scaled down to the county level first. There are 388 counties in the WECC and for each county there is a designated EIA climate zone, an IECC climate zone, and a weather location used for simulation, which collectively give an unambiguous mapping between the MSUs and counties. Because population has a major effect on energy demand, the populations of the counties are used to create a transformation matrix to downscale the aggregated commercial and residential annual hourly load profiles in each MSU into a pair of commercial and residential annual hourly load profiles for each county (Equation 13).

$$ {}_{8,760}\begin{bmatrix} \text{8,760 annual hourly load profiles at each of the 50 spatial intersections} \end{bmatrix}^{50} \times {}_{50}\begin{bmatrix} \text{Population fractions of the county in each of the 50 spatial intersections} \end{bmatrix}^{388} = {}_{8,760}\begin{bmatrix} \text{8,760 annual hourly load profiles for each of the 388 counties in WECC} \end{bmatrix}^{388} \tag{13} $$

The resulting county-level results can be similarly upscaled to the decision-relevant scale of interest. In this example there are 29 BAs in the WECC (Equation 14).

$$ {}_{8,760}\begin{bmatrix} \text{8,760 annual hourly load profiles for each of the 388 counties in WECC} \end{bmatrix}^{388} \times {}_{388}\begin{bmatrix} \text{Population fractions of each county in the 29 BAs in WECC} \end{bmatrix}^{29} = {}_{8,760}\begin{bmatrix} \text{8,760 annual hourly load profiles for each of the 29 BAs in WECC} \end{bmatrix}^{29} \tag{14} $$

The final outcome of this downscaling and upscaling steps is an 8,760 hour time series of residential building energy demand and commercial building energy demand in each of the 29 BAs in the WECC.

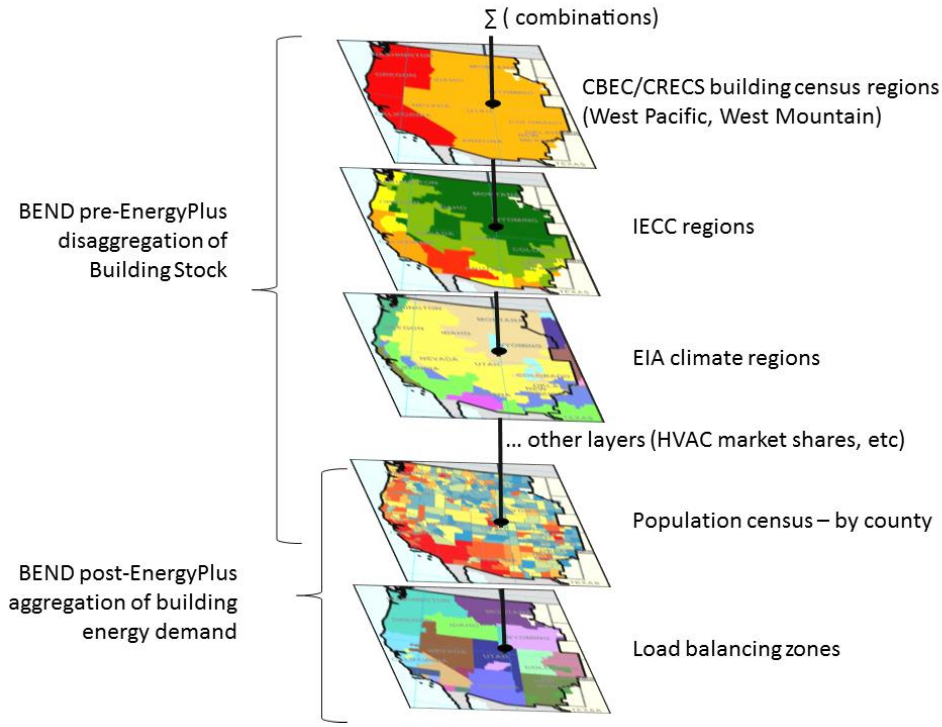


**Figure 7**. Graphical depiction of the various layers involved in the disaggregation and aggregation of building loads in BEND

### 3.3.2 Bias, Bias Correction, and Calibration

The annual hourly load profiles at spatial scales of interest, being built up from numerous individual building models, will not exactly match the real (reported) loads available from the EIA. Indeed, for a multitude of reasons, systematic differences (biases) can be expected. Biases result from the many approximations used in developing BEND's load profiles, including but not limited to:

- BEND's CBECS/RECS-based approximation of the building stock
- BEND's segregation of building types, sizes, and vintages into limited category bins and its simplification of all buildings into a set of *representative buildings*
- The limited set of detailed building characteristics used to approximate buildings within the established type, vintage categories, and size bins and within climate regions
- The limited treatment of HVAC system types and their distribution within the other BEND categories and regions
- The limited set of weather stations used to drive EnergyPlus simulations
- The population-weighted rescaling of results across spatial scales.

These and other differences result in biases in BEND's building energy demand estimates, which may be systematic with respect to location, building type, season, or time of day. BEND's estimates are therefore subjected to a calibration process to apply a high-level correction for those systematic biases. The calibration process is complex and is described in more detail in a paper currently under review. This section contains a simplified summary of the calibration process.

BEND's annual hourly energy demand profiles for residential and commercial buildings are scaled to the level of BAs and then compared against residential and commercial loads derived from EIA data reported by those BAs. However, a direct comparison requires that EIA-reported hourly profiles be segregated into residential and commercial portions, a process that itself introduces some uncertainty. Based on a statistical analysis of the differences between BEND's profiles and the EIA-derived profiles, adjustment factors that vary by month and hour (time of day) are developed. These adjustment factors, which are simple multipliers applied to BEND's profiles, are used to adjust any further BEND results from specific related analyses (e.g., using various future climate scenarios, population changes, or technology improvements).

# 4.0 Summary and Future Improvements

## 4.1 Summary

Based on CBECS and RECS databases, BEND creates a set of *representative buildings* with predefined sets of building types, vintages, sizes, and locations. BEND constructs the set of *representative buildings* based on a statistical analysis of the CBECS and RECS databases. That analysis is used to derive detailed descriptions for the building parameters such as construction, thermal properties, internal load, equipment and mechanical systems, and operation schedules and to create a set of *prototype buildings* to be simulated by EnergyPlus. BEND selects weather forcing based on climate zones used by the IECC. Within each intersection of IECC climate zone and CR/CZ, one or more weather files is chosen to represent the weather forcing to be used in the simulation of the *prototype buildings.* The binding of weather files with *prototype building* creates a set of *simulation-ready building models*. The spatial region influenced by the weather file constitutes the MSU in BEND. The outcomes of the EnergyPlus simulations are annual hourly energy consumption values for each of the *simulation-ready building models*. In each of the MSUs, the energy demand profiles of the individual *simulation-ready building models* are aggregated into a pair of residential and commercial hourly time series. Through population weighting, the pair of BEND residential and commercial energy demand time series can be scaled down or up to different spatial scales (e.g., counties, states, or BAs), which allows it to provide annual hourly building energy consumption (the sum of residential and commercial) scaled to decision-relevant geographic scales. From these hourly profiles the total annual building energy consumption and peak electricity demand can be obtained. BEND can also provide quantities other than total energy consumption and peak demand, such as energy demand by fuel type (i.e., electricity, natural gas, or oil), by end use (i.e., cooling, heating, lighting, fans, pumps, or service water), by subregion of a spatial region (e.g., BAs within the WECC or EIC) and/or by subset of buildings (e.g., only large office buildings).

## 4.2 Future Improvements

This section contains an incomplete list of possible future improvements to BEND.

- The current BEND framework uses data from CBECS03 and RECS09. Newer versions of CBECS and RECS exist and BEND could be updated based on the new versions.

- The definitions of building types, vintages, and sizes used in creating the *representative buildings* may be revisited, especially on the new version of CBECS and RECS. Currently omitted building types may be restored and new building types may be introduced.

- The current BEND framework uses ordinary linear regression to capture the relationships between most of the building parameters and the four high-level building characteristics (type, vintage, size, and location). The estimates of the linear regression are treated as probabilities. This may be modified using more theoretically sound statistical approaches.

- The algorithm that calculates building weights for the *representative buildings* could be enhanced.

- The algorithm through which building construction, thermal properties, internal load, schedule, and HVAC systems are designated could be upgraded.

- Currently, BEND only considers the variation of HVAC systems and ignores the variation of all other construction types, equipment, and schedules. This could be improved by introducing variation of more parameters.

- The predefined construction types (e.g., wall types and insulation) could be revisited.

- Templates for the various buildings could be refined.

- The system for identifying the appropriate weather forcing could be upgraded.

# 5.0 List of Utility Programs and Spreadsheets

**Set_Builder_all.xlms**: Excel spreadsheet with a macro to designate the number and types of the HVAC systems for the set of *representative buildings*. This process expands the set of *representative buildings* into a set of *prototype buildings* with the consideration of variation of the HVAC systems.

**InputWorksheet.xlms**: Excel spreadsheet with a large number of regression coefficient tables and regression estimate tables as lookup tables to designate all detailed building parameters (e.g., wall, roof, floor type and insulation, lighting power density, occupant density, HVAC, and schedules) for the set of *representative buildings*. This process converts the few high-level building and geographic/climatic characteristics (type, size, vintage, and location) into a large number of parameters for detailed description of the buildings. The result of this process is table-format parameter files that will be used to generate EnergyPlus .idf files.

**EPXMLPreproc**: A utility program developed by NREL that is used to generate XML files based on the parameter files and the building template files in the XML format.

**GParm**: A "General PARaMetrics" utility program developed by PNNL that reads each row of the parameter file, which consists of values of the entire set of parameters for a particular *prototype building* and uses the values to replace the replaceable tags in the template file. The output of this process is a set of .xml files, with the number of .xml files equal to the number of *prototype buildings*.

**1_wgt_calc_from_database.R**: An R script that is used to apply the BEND building type, vintage, and size definition on the CBECS and RECS databases to derive building weights for the set of *representative buildings*. Smoothing and interpolation are used to derive the building weights assuming the distributions of building type, size, and vintage are independent in a given CR/CZ.

**2_wgt_extend_to_HVAC_fractions.R**: An R script that takes the weights of the *representative buildings* and the fractions of the selected HVAC systems comprising more than 60% of the market shares to expand the set of *representative buildings* into a set of *prototype buildings* recalculated by designating one or more HVAC systems for each of the *representative buildings.* The weights of the *representative buildings* are redistributed proportionally to the weights of the set of *prototype buildings*.

**3_wgt_extend_to_WMO.R**: An R script that distributes the weight of a *prototype building* into the weights of one or more *simulation-ready building models* when the *prototype building* is paired with one or more weather files. The original weights carried by the *prototype buildings* need to be redistributed to the *simulation-ready building models*. Population is used for the weight distribution. County is treated as the common base for the mapping among CR/CZs and IECC climate zones. Once a decision has been made on how many and which weather stations will be used to represent the weather in the intersections, weather stations are assigned to the county within which the weight redistribution will be performed based on the population.

# 6.0 Acknowledgements

The prototype BEND model was conceived of and designed by James Dirks, Will Gorrissen, John Hathaway, Daniel Skorski, Michael Scott, Trenton Pulsipher, Maoyi Huang, Ying Liu, and Jennie Rice – all of whom were at PNNL at the time the model was developed. The current BEND team is greatly appreciative of their efforts. The initial development BEND was supported by the Laboratory Directed Research and Development Program at Pacific Northwest National Laboratory. Ongoing development and deployment of BEND was supported by the U.S. Department of Energy, Office of Science, as part of research in the Multi-Sector Dynamics, Earth and Environmental System Modeling Program. Editing assistance was provided by Stephen Butterworth and Maura Zimmerschied.

# 7.0 References

42 U.S.C. 6801 et seq. *Energy Conservation and Production Act of 1976*. Public Law 94-385. Accessed November 30, 2017, at http://uscode.house.gov/statutes/pl/94/385.pdf.

Burleyson CD, N Voisin, ZT Taylor, YL Xie, and I Kraucunas. 2017. "Simulated building energy demand biases resulting from the use of representative weather stations." *Applied Energy* 209:516-528. Available at https://doi.org/10.1016/j.apenergy.2017.08.244.

Dirks JA, WJ Gorrissen, JH Hathaway, DC Skorski, MJ Scott, TC Pulsipher, M Huang, Y Liu, and JS Rice. 2015. "Impacts of climate change on energy consumption and peak demand in buildings: A detailed regional approach." *Energy* 79:20-32. Available at https://doi.org/10.1016/j.energy.2014.08.081.

DOE – U.S. Department of Energy. 2013. *EnergyPlus Energy Simulation Software, Version 8.0*. Washington, D.C. Accessed November 30, 2017, at https://energyplus.net/.

EIA – Energy Information Administration. 2003*a*. *Commercial Buildings Energy Consumption Survey 2003*. U.S. Department of Energy, Washington, D.C. Accessed November 30, 2017, at https://www.eia.gov/consumption/commercial/data/2003/.

EIA – Energy Information Administration. 2003*b*. *2003 CBECS Building Characteristics and Consumption and Expenditures for All Buildings (Including Malls): Public Use Files* U.S. Department of Energy, Washington, D.C. Accessed July 3, 2018, at https://www.eia.gov/consumption/commercial/data/2003/pdf/layouts&formats.pdf.

EIA – Energy Information Administration. 2009*a*. *2009 RECS Variable and Response Codebook*. U.S. Department of Energy, Washington, D.C. Accessed July 3, 2018, at https://www.eia.gov/consumption/residential/data/2009/index.php?view=microdata\recs2009_public_codebook.xlsx.

EIA – Energy Information Administration. 2009*b*. *Residential Energy Consumption Survey 2009*. U.S. Department of Energy, Washington, D.C. Accessed November 30, 2017, at https://www.eia.gov/consumption/residential/data/2009/.

EIA – Energy Information Administration. 2018. *Monthly Energy Review.* U.S. Department of Energy, Washington, D.C.  Accessed June 5, 2018, at https://www.eia.gov/totalenergy/data/monthly/pdf/mer.pdf.

Hong T, W-K Chang, and H-W Lin. 2013. "A fresh look at weather impact on peak electricity demand and energy use of buildings using 30-year actual weather data." *Applied Energy* 11:333-350. Available at https://doi.org/10.1016/j.apenergy.2013.05.019.

Kraucunas I, L Clarke, J Dirks, et al. 2015. "Investigating the nexus of climate, energy, water, and land at decision-relevant scales: the Platform for Regional Integrated Modeling and Analysis (PRIMA)." *Climatic Change* 129(3-4):573-588. Available at https://doi.org/10.1007/s10584-014-1064-9.

NREL – National Renewable Energy Laboratory. 2011. *U.S. Department of Energy Commercial Reference Building Models of the National Building Stock*, February 2011. NREL/TP-5500-46861, Golden, CO. Available at https://www.nrel.gov/docs/fy11osti/46861.pdf

PNNL – Pacific Northwest National Laboratory. 2014. *Facility Energy Decision System Release 7.0 User's Guide*, December 2014. PNNL-SA-107079, Richland, WA. Accessed November 30, 2017, at http://www.pnl.gov/feds/pdfs/FEDS_7-0_user_guide.pdf.

PNNL – Pacific Northwest National Laboratory. 2015. *Methodology for Evaluating Cost-Effectiveness of Residential Energy Code Changes*, August 2015. PNNL-21294 Rev 1, Richland, WA. Available at https://www.energycodes.gov/sites/default/files/documents/residential_methodology_2015.pdf.

Voisin N, M Kintner-Meyer, R Skaggs, T Nguyen, D Wu, J Dirks, YL Xie, and M Hejazi. 2016. "Vulnerability of the US western electric grid to hydro-climatological conditions: How bad can it get?" *Energy* 115:1–12. Available at https://doi.org/10.1016/j.energy.2016.08.059.

Pacific Northwest
NATIONAL LABORATORY

U.S. DEPARTMENT OF
ENERGY

*Proudly Operated by* **Battelle** *Since 1965*

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99352
**1-888-375-PNNL** (7665)

www.pnnl.gov