Toward a Unified California Commercial Modeling Platform: Prototypes and Engine-Agnostic Scorecards

Mohammad Dabbagh, Fatemeh Yousefi, Nasim Mirian, Rahul Athalye, NORESCO
Andres Fergadiotti, Bach Tsan, Ryan McFadyen, Southern California Edison

ABSTRACT

A new commercial prototype building suite was developed to represent California's building stock for the purpose of improving and documenting model assumptions and harmonizing the assumptions between various policymaking agencies, including the California Energy Commission (CEC) and California Public Utilities Commission (CPUC). This set of prototype models are called the 'CalBEM Prototype Models' because they were developed under guidance from the California Building Energy Modeling (CalBEM) group[1]. Together with the models, each model input was documented in a document called the 'scorecard.' Each prototype scorecard summarizes building descriptions, thermal-zone internal loads, schedules, envelope and fenestration properties, HVAC and service hot water systems, controls, and other inputs; each field includes documented sources, units, ranges, and version tags, and can be filtered, compared, and batch-edited. A scorecard database was developed that draws from individual prototype scorecards and provides an engine-agnostic data repository for prototype inputs and metadata, which enables rapid input editing, traceable sources, and automated export to enable use with multiple simulation engines.

The selected CalBEM prototype models represents more than 80% of the constructed California commercial floor area captured in the Dodge construction-starts assessment for the building categories evaluated in this project. The estimate excludes building types that were not included in the initial prototype suite or that lacked sufficient stock-assessment basis for first-round prototype development. Five vintage bins were defined: four existing stock bins and one new-construction bin, with specifications derived from California-specific sources. Model energy consumption was validated using the 2022 Commercial End-Use Survey (CEUS). The models, scorecards, and the scorecard database were made available to the public through GitHub. Using the scorecards and the scorecard database, it is possible to develop the prototypes in different simulation engines. Moreover, the scorecard database allows the usage of an automated approach to developing models from revised inputs.

By decoupling model input data from any single engine, faster iteration, review, and sensitivity studies could be performed. Future versions of prototypes with improved inputs could be released in the same database format and could be used to automatically update prototype models. Applications include Title 24 evaluations, utility program design, custom-measure analysis, electrification, and grid-interactive assessments.

[1] https://calbem.ibpsa.us/calbem-prototypes/

Introduction

Modeling California's building stock

Whole-building energy modeling in California is used for Title 24 compliance, efficiency program design, custom measure analysis, and energy demand forecasting. Despite this breadth, California analyses have drawn on several established prototypes and modeling resources developed for different purposes—such as CBECC standard-design models for compliance, DEER-aligned prototypes for efficiency program analysis, and DOE/PNNL national prototypes for code development and benchmarking. Because these resources were created for different objectives and document their assumptions and inputs at different levels of detail, tracing prototype assumptions consistently across tools and applications is difficult. In particular, the documented source and basis for certain prototype inputs—such as loads, schedules, floor area, story count, and some rule-based standard-design assumptions—are not always easy for outside users to identify from a single, transparent source.

One important California application of whole-building energy modeling is building energy code compliance. The California Building Energy Efficiency Standards (Title 24, Part 6) were first adopted in 1976 and have been updated on a three-year cycle. The Standards include a prescriptive path and a performance path. Under the performance path for nonresidential buildings, the standard design model establishes the energy budget used for compliance comparison, varying by climate zone and building type. Compliance may be demonstrated by designing a building whose simulated annual energy use is below that budget. By law, the California Energy Commission (CEC) must provide approved software tools for this compliance approach. Today, the CEC provides public-domain compliance software (CBECC) for nonresidential buildings, which generates a "standard design" model and compares proposed design energy to the code budget. CBECC is fundamentally a compliance tool, not a prototype documentation platform. While the tool and Alternative Calculation Method (ACM) documentation are publicly available, outside users do not always find a single, consolidated record of the source and basis for some standard-design assumptions, which can make prototype review and cross-tool comparison difficult. This becomes especially important when compliance-oriented standard-design assumptions are referenced in broader policy, measure-development, or prototype-analysis workflows, where clearer documentation of their source and basis is needed.

Separately, the California Public Utilities Commission (CPUC) oversees ratepayer-funded energy efficiency portfolios. In that context, the Database of Energy Efficient Resources (DEER) provides standardized savings estimates and characterizations for common efficiency measures. DEER's data includes unit energy savings, useful life, and net-to-gross ratios. It is maintained by the CPUC Energy Division (with funding from ratepayers) and is used in planning and forecasting energy efficiency programs. In recent years, California has augmented DEER with prototype-based simulation tools: residential and commercial building prototypes in EnergyPlus format have been created to support analysis of efficiency measures. These prototypes cover all California climate zones (16 zones are used in Title 24), and they correspond to DEER's building-type taxonomy. Vintages are represented by code-year designations (e.g. 2008, 2013, 2016, etc.) or simply "existing vs. new" in different tools. This matters in California specifically because CEC, CPUC, utilities, consultants, and other stakeholders use prototype models in statewide policy, program, and compliance analyses; when the underlying prototype bases differ, predicted results can differ as well.

The California Commercial End-Use Survey (CEUS) is a recent statewide field survey of commercial buildings (2018–2022) conducted by the CEC. CEUS collected detailed data on building characteristics, equipment inventories, and hourly energy usage for over 24,000 buildings across California. Its sample was stratified by utility service territory, forecast zone, building type, and vintage (existing vs. new). The primary goal was to improve California's building energy demand forecasts by disaggregating demand to 20 fine-grained forecast zones. The CEUS findings (some publicly released, some proprietary) were used in this project to inform nonresidential prototype inputs and to support validation at the state and zone level. For example, CEUS provides floor-area-weighted end-use energy intensities and end-use shares by climate region. However, detailed consumption information (e.g. mean EUI by zone and end use) was not accessible.

Improved Representation of California's Building Stock

The CalBEM prototypes were developed to provide a more California-specific and more consistently documented representation of California's commercial building stock and energy consumption. A key feature of this work is documentation of inputs, their sources, and how the inputs were developed. This documentation serves as the foundation for the models and for future updates and changes. Further, the inputs from various prototypes were consolidated into a single scorecard database, which is an engine-agnostic repository that captures inputs (geometry, envelope, loads, schedules, HVAC/DHW intent), source and basis, units/ranges, and versioning. The approach aims to enable consistent implementation across tools, transparent review, and rapid iteration for agency workflows (CEC, CPUC), while maintaining traceability from outputs back to inputs.

Background and Related Work

Existing prototype ecosystems and limitations

Several prototype and stock-modeling resources were in use in California at the time this project was started. Most of these resources sought to represent the building stock of a given region. For example, the PNNL prototypes represent the United States building stock through a suite of prototype buildings. NREL's ResStock and ComStock are residential and commercial stock-modeling frameworks, respectively, that also represent the U.S. building stock, but they are stock-modeling frameworks rather than conventional prototype sets. The CBECC standard-design model set was originally designed to support testing and implementation of new features in CBECC. The needs assessment that preceded this project identified at least six model sets in use for policymaking analyses in California, each with its own assumptions. It also found several recurring limitations in existing model sets, including limited documentation of model inputs and parameters, difficulty accessing underlying models when they were implemented through wrapper software, and a lack of clear source or basis for some inputs. These findings helped motivate the development of a more unified and better-documented prototype framework for California policymaking. The following discussions describe the major prototype and stock-modeling resources used in California analyses in California and their specific use cases.

CEC/CBECC: For nonresidential Title 24 performance compliance, the CEC requires approved software that generates a standard design from the proposed design in accordance with ACM rules. The public-domain compliance software, CBECC, is used to model the proposed and standard designs for compliance comparison. For this paper, the key limitation is not that the software is inaccessible, but that the documented source and basis for some standard-design inputs are not always clear to outside users. In particular, this includes some inputs such as loads, schedules, floor area, number of stories, and certain other standard-design assumptions. This can make prototype review, reuse, and cross-tool comparison difficult, especially when compliance-oriented assumptions are referenced in broader policy, measure-development, or prototype-analysis workflows.
CPUC/DEER and prototype tools: DEER supports program planning with measure-level savings and has historically used DOE-2 and other program-analysis workflows. More recently, EnergyPlus prototypes aligned to DEER building types and California climate zones were published with Modelkit[2] templates for parametric generation (CPUC 2022; CPUC 2023). Public documentation on precise vintage coverage can be incomplete in accessible summaries, which limits cross-set comparisons without companion reports.
PNNL national prototypes: PNNL publishes prototypes for commercial buildings representing the United States building stock. These prototypes are based on ASHRAE climate zones. This prototype set has a national focus, i.e., it is not specific to California's climate zones, energy code, and building features such as number of stories and floor area. While this prototypes-set served as one of the sources of inputs in the development of the CalBEM prototypes, it could not be used directly for policymaking in California.
ResStock and ComStock: NREL's ResStock and ComStock are residential and commercial stock-modeling frameworks, respectively, representing the United States building stock. These models have a national focus and therefore are not specific to California's climate zones, energy code cycles, or building characteristics such as floor area, number of stories, and other California-specific prototype assumptions. While these resources served as useful reference resources during the development of the CalBEM prototypes, they could not be used directly as prototype sets for California policymaking.

This project was conceived because existing prototype sets in use in California did not have the same input assumptions or a unified set of assumptions representative of the building stock in California. Sources of model inputs and assumptions were not easily traceable, and it was difficult to understand how certain inputs were developed. Moreover, different policymaking agencies were using different sets of prototypes, with different inputs, resulting in different outcomes. The CalBEM prototypes were developed to provide a more California-specific and more consistently documented prototype framework, with rigorous documentation of inputs, a transparent development process, and models and documentation.

What "engine-agnostic" means

"Engine-agnostic" means that the core prototype assumptions are stored outside any single simulation engine. In the CalBEM scorecard database, site/climate identifiers, model geometry, envelope properties, internal loads and schedules, HVAC and service-water-heating system

[2] https://bigladdersoftware.com/projects/modelkit/

system-intent parameters, and reporting metrics are maintained as structured fields that can be mapped to multiple simulation engines without reauthoring the building description (CalBEM 2020; CPUC 2023). For example, a lighting power density, occupancy schedule, envelope U-value, outdoor-air rate, or HVAC system type is stored once in the scorecard database with its units, source, and prototype/vintage applicability; the EnergyPlus implementation maps those fields to EnergyPlus-specific objects and templates. The boundary between engine-agnostic and engine-specific content is drawn to keep core prototype data independent from the simulation engine, while translator templates create engine-ready models. In this project, the scorecard database and automation framework enable generation of engine-ready models from the same underlying specification. Other engines would require their own translator, but could use the same documented source assumptions. This differs from DOE/PNNL prototype scorecards, which are distributed as spreadsheets summarizing key model inputs alongside downloadable EnergyPlus models (DOE BECP 2017; PNNL 2021), and from state prototype systems such as CPUC's DEER/Modelkit framework, which supports parametric generation of EnergyPlus models but is still organized as an EnergyPlus-specific framework rather than a single engine-agnostic representation for multi-engine model generation (CPUC 2023).

Methodology

Building stock assessment

As prototypes are intended to represent the building stock, their geometry, envelope construction, HVAC systems, and associated input parameters should align as closely as possible with prevailing industry practices. To support this objective, a number of datasets were studied. Using the latest data from these sources, the main features of the building stock were determined. The primary datasets used in the study are described below.

California specific Construction Start Data by Dodge (Dodge, 2023).To characterize the distribution of floor space into various building types and geographical locations within California, Dodge construction data was examined. The dataset comprises building permit records issued since 1967 across multiple building categories throughout California. Statewide distributions of floor area and number of stories were derived based on categorized floor area and footprint bins. These distributions were subsequently used to refine the classification of various building types and to assess whether modifications to the existing prototype definitions were necessary in order to more accurately represent the characteristics of the current building stock.

Building types with the largest contributions to total floor area—including office, school, warehouse, retail, and hotel—were selected from the full building set for detailed geometric assessment.

California End Use Survey (CEUS) data (CEC, 2023)**.** The CEUS dataset was extensively utilized to determine HVAC system type for all prototypes during the building stock assessment phase. The dataset is California-specific and reflects relatively recent building stock conditions. In addition, it provides information disaggregated by predefined vintage bins, enabling age-dependent analysis. The CEUS data was also used to evaluate operating hours and update building schedules accordingly.

CBECS data (EIA, 2022)**.** This nationwide survey-based data, which contains fewer records for the Pacific region compared with CEUS, was used to estimate the distribution of HVAC system types, validate prototype geometries and roof shapes, characterize building envelope properties, and inform the definition of Office prototypes.

ComStock 2023 data. (NREL, 2023) This large dataset of simulated U.S. commercial buildings was analyzed to support the development of the Office prototype, verify prototype geometry and zoning, identify lighting technologies and HVAC systems.

Vintage bin analysis

As actual thermal properties of the California nonresidential building envelope were not available, a vintage-bin analysis was conducted to categorize the building stock into a manageable number of age groups. This process resulted in the definition of five distinct construction-period categories:**Error! Reference source not found.**

The **pre-1978** bin represents buildings constructed before California's first building energy standard was implemented in 1978, a period that was followed by high construction activity.

The **1978–1997** bin captures the early Title 24 era before major changes to window regulations, when older single-pane glazing systems were common.

The **1998–2007** bin reflects revised Title 24 window requirements that supported the use of multi-pane glazing from 2001, a period of high construction activity, and the approximate 25-year roof insulation replacement timeframe.

The **2008–2018** bin represents buildings constructed under stricter wall and roof insulation requirements introduced in 2008, with improved HVAC systems that are likely still in service. This period ends before the COVID 19 recession.

The **2019+** bin represents recent and new construction, reflecting current code requirements and construction practices

The categorization was informed by multiple considerations, including historical construction rates, macroeconomic cycles, revisions to energy regulations, temporal trends in window-to-wall ratio, and typical insulation replacement timelines. California-specific Construction Starts data3 (Dodge, 2023) were used to evaluate historical construction activity. Further details regarding this component of the analysis are provided by (Yousefi & Athalye, 2025) .

Error! Reference source not found.**Model inputs development**

Multiple sources were consulted to determine the input parameters for the prototype models. The primary sources are summarized in Table 1. In addition to these references, relevant online documentation, technical reports, and peer-reviewed literature covering various aspects of building performance were systematically reviewed. Feedback from the advisory group—obtained through weekly and quarterly meetings—was incorporated throughout the development process. Engineering judgment was applied to synthesize these inputs and to evaluate and finalize the selected parameter values.

Table 1: Main data sources

Datasets	Codes and Standards	Existing Prototypes	Other

[3] Provided to NORESCO by the CEC.

• Dodge 2023 • CEUS 2022 • ComStock 2023 • CBECS 2018	• California Title 24 (Part 6,4) • California Title 20 • ASHRAE 90.1 • ASHRAE 60.2 • CFR Title 10 • ASHRAE 170/OSHPD	• CEC prototypes (CBECC 2022.3.1) • PNNL Prototypes (ASHRAE 90.1-2022) • DEER prototypes (June 2024) • DOE 2004 Prototype • ORNL Office prototypes	• ACM • Title 24 appendices • CASE reports • PNNL prototypes' technical support documents • NREL's building stock documents • Design guides • State of industry reports • State forecast reports • ....

All required characteristics for the energy simulation of each prototype were documented in individual, human-readable Excel format scorecards. The scorecards were standardized and designed to comprehensively capture key modeling inputs, including space configuration and zoning, envelope characteristics, HVAC systems and associated efficiencies, thermostat setpoints, infiltration and ventilation rates, water usage, hourly operational schedules and internal loads including interior lighting, exterior lighting, plug loads, gas appliance loads, and elevator loads.

All scorecards share the following features:

- Standardized structure across all prototypes
- Human- and machine-readable table format
- Simulation engine–agnostic design
- Transparent documentation of references for every input parameter
- Comprehensive coverage of model inputs, with the exception of certain geometric design details and selected HVAC simulation parameters

These scorecards served as the primary input for the development of the prototype models and guided the overall modeling process.

Updates to Zoning Assumptions

While many prototype models retained zoning configurations consistent with the existing prototypes, several others were updated or were newly developed, including three Office prototypes, two Hotel prototypes, and the newly developed Grocery prototype model. The existing office prototypes typically employ a simplified core–perimeter zoning approach. However, this approach may not accurately represent internal heat gains, occupant behavior, or energy-use patterns. In particular, oversimplified zoning can lead to unrealistic ventilation and airflow turndown behavior, which can ultimately reduce the fidelity of simulation results. To address these concerns, the zoning configuration of the three office prototypes was revised. The new office set features detailed space layouts, refined thermal zoning beyond core-perimeter partitions (including high-load areas and AHU service boundaries), updated schedules and internal-load assumptions, and improved control approaches to better capture part-load operation and ventilation turndown. This configuration provides a more representative foundation for California offices across vintages and climates.

The Large Hotel prototype was another case where the original model included a simplified big-box basement that was a single zone. In the CalBEM version of the Large Hotel prototype,

the single zone was divided into conference rooms, offices, storage, and food preparation. In addition, the number of floors in both hotel prototypes was revised to align with the building stock assessment, resulting in corresponding changes to their overall zoning configurations. The Grocery prototype is entirely new and was developed based on the actual design characteristics of typical grocery stores.
As a result of the zoning updates, HVAC systems, internal loads, ventilation rates, peak water use, schedules, and other related inputs were comprehensively revised to remain consistent with the updated zoning and associated space-type allocations.

Model construction

The construction of the non-residential prototype models was carried out in four sequential steps:

1. Prepare base geometry models using SketchUp (Trimble Inc., 2023) or existing prototype models.
2. Integrate HVAC systems consistent with the scorecards into the geometry models using HVAC templates or by updating existing models in OpenStudio (NREL, 2025) or EnergyPlus (U.S. Department of Energy, 2025). At this stage, precise system efficiencies and capacities were generally not required, as these were later sized during the expansion step. For each prototype, three variations of the HVAC model were developed, corresponding to Vintage 1–3, Vintage 4, and New Construction.
3. Establish an automation framework with Processing scripts to automatically generate the Database, Load Input Data File (IDF), Template IDF, Envelope IDF, and Schedule IDF from the scorecards. IDF is the standard input file format for EnergyPlus. This step is described in detail in the following section.
4. Expand each prototype model to 16 climate zones using a custom-built parametric simulation infrastructure.

Figure 1 illustrates the process described above.

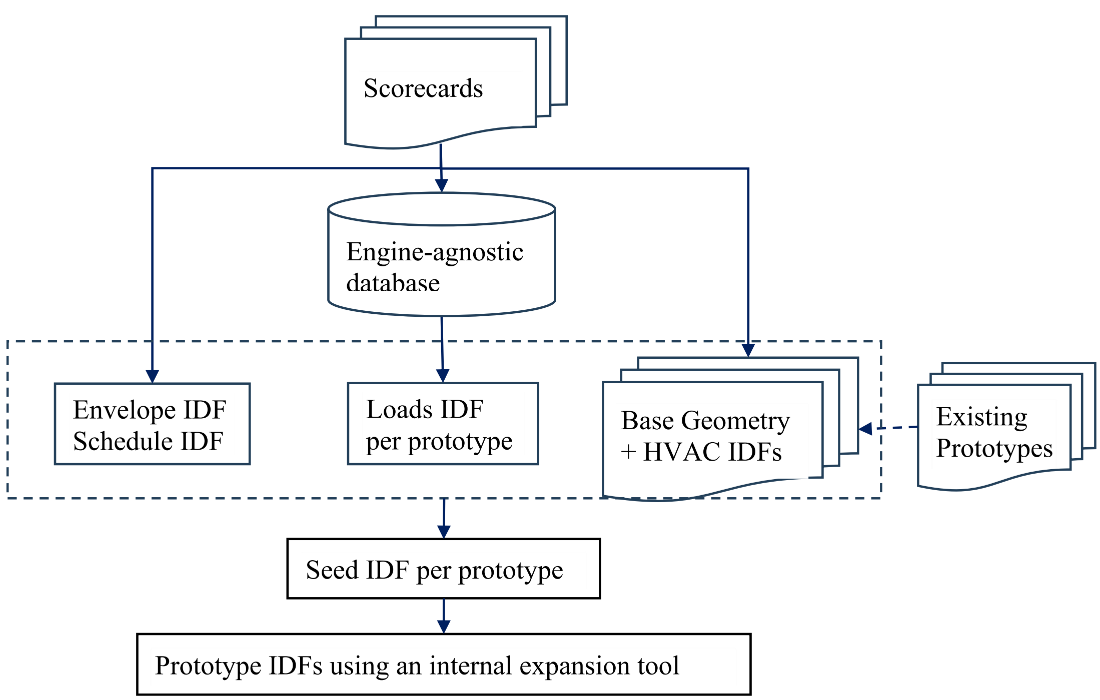

Figure 1: Prototype Model Generation flowchart

Automation pipeline

Developing and updating nonresidential prototypes to align with CEUS-2022 across 16 climate zones and multiple vintages quickly demonstrated that traditional, hand-built modeling workflows are too slow, difficult to review, and prone to transcription and unit-conversion errors. To address this, an automated, template-based pipeline was developed that converts standardized scorecard inputs into engine-ready models through a structured sequence of translation and validation steps. The workflow is designed to be engine-agnostic, repeatable, and transparent, with EnergyPlus v25.1 used as the reference implementation.

At the core of the pipeline is the Scorecard Database (Scorecard DB), which serves as the single source of truth for all prototype inputs. Each prototype includes hundreds of parameters spanning geometry, envelopes, internal loads, ventilation and infiltration, HVAC systems, service hot water, schedules, and setpoints. While the original scorecards are maintained in Excel for human readability and stakeholder review, the Scorecard DB stores the same information in a machine-readable structure that can be consumed by scripts and simulation frameworks. This separation allows model assumptions to be discussed and verified in plain language while ensuring that the data used to generate models is standardized and version-controlled.

The Scorecard DB is organized into three primary CSV layers that correspond to how simulation engines assemble building models.

- ENVELOPE.csv defines thermal properties such as U-values, SHGC, insulation levels, and infiltration by prototype, climate zone, and vintage.
- ZONE.csv contains zone-level loads and controls, including floor area, multipliers, lighting and plug loads, occupancy, ventilation rates, schedules, thermostat setpoints, and peak hot-water draws.
- WHOLE_BLDG.csv captures systems that are not zone-specific, such as water heating equipment, exterior lighting, and elevator loads.

These CSV layers are the machine-readable form of the scorecard inputs. Because they encode physical and semantic building parameters rather than engine-specific syntax, the schema functions as an engine-neutral intermediate representation. EnergyPlus support is implemented as one translator: the automation scripts read the normalized fields, apply unit conversions and naming conventions, and populate EnergyPlus template tokens to create complete IDF models. The same schema can drive additional engine-specific translators, for example, emitting SDD XML for CBECC or BDL for eQUEST, without altering the underlying inputs, since the single source of truth is preserved and only the engine backend is swapped. Portability to a given engine requires authoring that backend and is bounded by the engine's representational capabilities; translators reconcile capability and convention mismatches. Data validation checks enforce allowable ranges, data types, and unit consistency before any model is created.

.

The translation from scorecards to models is handled by a series of processing scripts written in Python. These scripts perform three main functions. First, they extract and normalize inputs from individual scorecards into the Scorecard DB, resolving naming differences and reusing common variables across prototypes. Second, they map database fields to template tokens used in the IDF generation framework. Third, they automate batch simulation and extract results for verification and calibration. By automating these steps, the pipeline eliminates repetitive manual entry, prevents mismatched units, and ensures that updates to the scorecards propagate consistently to all generated models.

The template-based IDF generation framework converts database values into engine-ready models. Each prototype has a master IDF template that contains placeholders for parameters that vary by climate zone, vintage, or construction type. These placeholders are populated programmatically using data from the Scorecard DB and a parametric run specification managed through an Excel-based "Run Manager." The Run Manager defines the simulation matrix (prototype × climate zone × vintage), generates the required IDFs, launches simulations, and aggregates outputs. Because the templates separate fixed geometry and topology from variable inputs, the same base model can be reused across hundreds of scenarios with minimal manual intervention.

Quality assurance is embedded throughout the pipeline. Automated pre-processing checks confirm that all template tokens are resolved that inputs fall within valid ranges, and that unit conversions are applied correctly. Post-simulation checks screen for runtime errors and warnings and evaluate unmet hours against ACM thresholds (no more than 150 hours per thermal zone). End-use results are benchmarked against CEUS 2022 data. Discrepancies beyond a defined threshold flag the prototype for documented review of its key drivers, internal loads, ventilation, infiltration, schedules, HVAC performance, and water-heating assumptions, rather than triggering automated tuning of inputs to the reference. Any input revisions arise from re-examining the underlying scorecard assumption against code and ACM defaults, not from minimizing the gap to CEUS.. This screening process ensures that large batches of models remain technically valid and that any departures from reference end-use intensities are surfaced and documented rather than silently absorbed.

The outputs of the pipeline include reusable IDF templates, fully populated climate- and vintage-specific models, structured simulation results, gap reports, and documented calibration decisions. Because the Scorecard DB is engine-agnostic, the same inputs can be rendered by different simulation tools (e.g., EnergyPlus, CBECC, or other wrappers) while maintaining consistent assumptions. This enables comparable results across projects, supports transparency for compliance and verification, and provides a scalable foundation for future code development, policy analysis, and resilience studies.

EnergyPlus implementation details

EnergyPlus v25.1 serves as the reference engine for this work, and the automation framework is tightly integrated with its macro processing, object hierarchy, and output structure. Prototype templates are written as modular IDF "parts" that correspond to envelope assemblies, internal loads, schedules, HVAC systems, and whole-building components. EnergyPlus macro language and Python scripts are used together: macros handle token replacement and conditional logic within templates, while Python manages database queries, unit conversions, batch generation, and file orchestration. This hybrid approach leverages EnergyPlus's native preprocessing capabilities while maintaining flexibility for large parametric runs.

Mapping between the Scorecard DB and EnergyPlus objects is handled through a structured variable dictionary. Each database field is linked to a specific IDF token and object parameter, ensuring that geometry, loads, and system specifications are populated consistently. For example, zone-level lighting power density values from ZONE.csv are mapped to Lights objects, ventilation rates to DesignSpecification:OutdoorAir, and envelope U-values to Construction definitions. HVAC system types are selected through template logic that assigns the appropriate air-loop, terminal, and coil configurations based on prototype and vintage. This mapping layer abstracts engine-specific syntax from the source data, allowing the same database to support multiple modeling platforms.

Naming and versioning are standardized to support traceability and reproducibility. Each generated model is tagged with identifiers for prototype, climate zone, vintage, and construction type, which are embedded in file names, output tables, and database records. Version control is applied to the Scorecard DB, templates, and processing scripts so that any change in assumptions can be tracked and reproduced. Because all IDFs are generated programmatically, it is possible to regenerate the full model set when inputs are updated, ensuring consistency across simulation cycles.

Automation extends to simulation management and result extraction. The Run Manager launches EnergyPlus runs in batch mode, monitors completion status, and parses output files to collect key metrics such as end-use energy, peak loads, and unmet hours. These results are written back to the scorecards for validation and to structured datasets for analysis. Automated checks flag models with excessive warnings, unresolved tokens, or unmet-hour violations, enabling rapid identification of issues within large simulation sets.

By combining standardized scorecards, a normalized database, template-based IDF generation, and automated validation, the EnergyPlus implementation transforms prototype modeling from a manual, case-by-case process into a scalable production workflow. This approach improves consistency, reduces errors, accelerates model development, and provides a transparent and reproducible foundation for California energy code analysis and future large-scale simulation efforts.

Model Validation

Two main data sources were used to compare the energy consumption reported by the prototype models with California's building stock data: CEUS and Portfolio Manager (Energy Star, 2025) . CEUS 2022 reports energy use data for twelve building categories, however, these categories do not cover all CalBEM prototypes. Therefore, Energy Star Portfolio Manager was also used to support the comparison of modeled energy consumption with existing building consumption in California.

Total energy use intensity (EUI), reported in kBtu/sf, was used as the primary metric for validation. For each prototype and climate zone, a vintage-weighted EUI was calculated using CEUS-derived vintage weights and compared with the target EUI from CEUS or Energy Star Portfolio Manager, using a ±25% acceptance range. When prototype-level EUI data was not available in CEUS, broader building category data were used instead. For example, the CEUS EUI for Lodging was used to evaluate both the Small Hotel and Large Hotel prototypes, which allowed for a little higher level of expected discrepancy.

For cases outside the acceptable range, model inputs were reviewed as much as possible to confirm that key assumptions, including internal equipment loads, occupancy levels, envelope U-values, schedules, and water use were reasonable. For example, gas and electric equipment loads and water use were revised for the Restaurant prototypes, while forklift loads and operating schedules were adjusted for the Warehouse prototype after they were found to be higher than expected. All revisions were checked to ensure that the updated inputs remained within the acceptable regulatory ranges.

Although the comparison was performed at building-type and climate-zone level, it was not always possible to bring every result within the acceptable range without making climate-zone-

level specific adjustment. Such adjustments were avoided when they would create inconsistent internal loads or schedule assumptions across climate zones. Therefore, the validation process focused on ensuring that prototype results were generally aligned with available benchmark data.
Figure 2 and Figure 3 show examples of how CEUS and national average data were used to validate the Small Hotel and Warehouse prototype models. The box-and-whisker charts show the acceptable EUI range, while the dots represent the CalBEM prototype results. For the Small Hotel prototype, most modeled results fall within the acceptable range, with only a few climate zones outside the range. New construction results are generally lower than existing building results due to more efficient envelope and HVAC assumptions.
For the Warehouse prototype, after the adjustments described above were applied, the modeled results generally remained within the national average range but fell outside the CEUS-based range in several climate zones. Since the underlying assumptions were reviewed and found to be reasonable, no further climate-zone-specific adjustments were made. A similar procedure was followed for each prototype to complete the validation process.

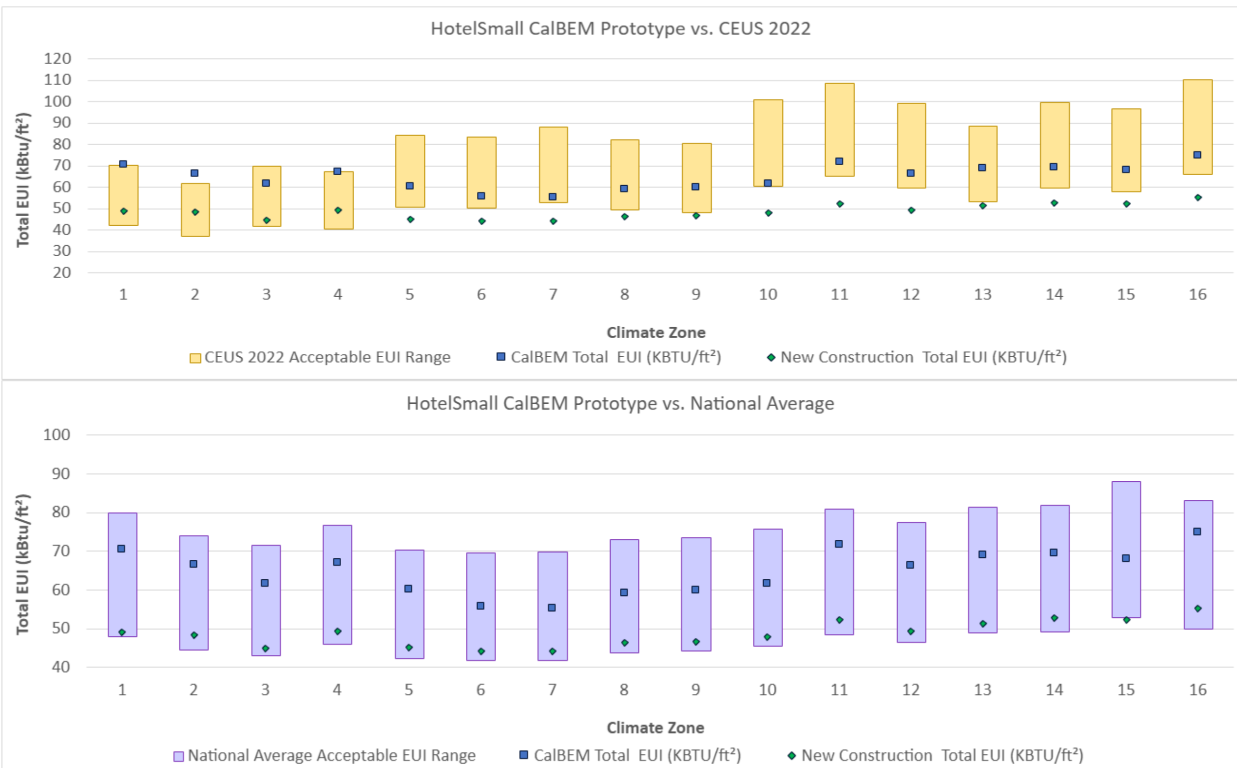

Figure 2: Comparison of Small Hotel Prototype EUI versus CEUS 2022 by Climate Zone

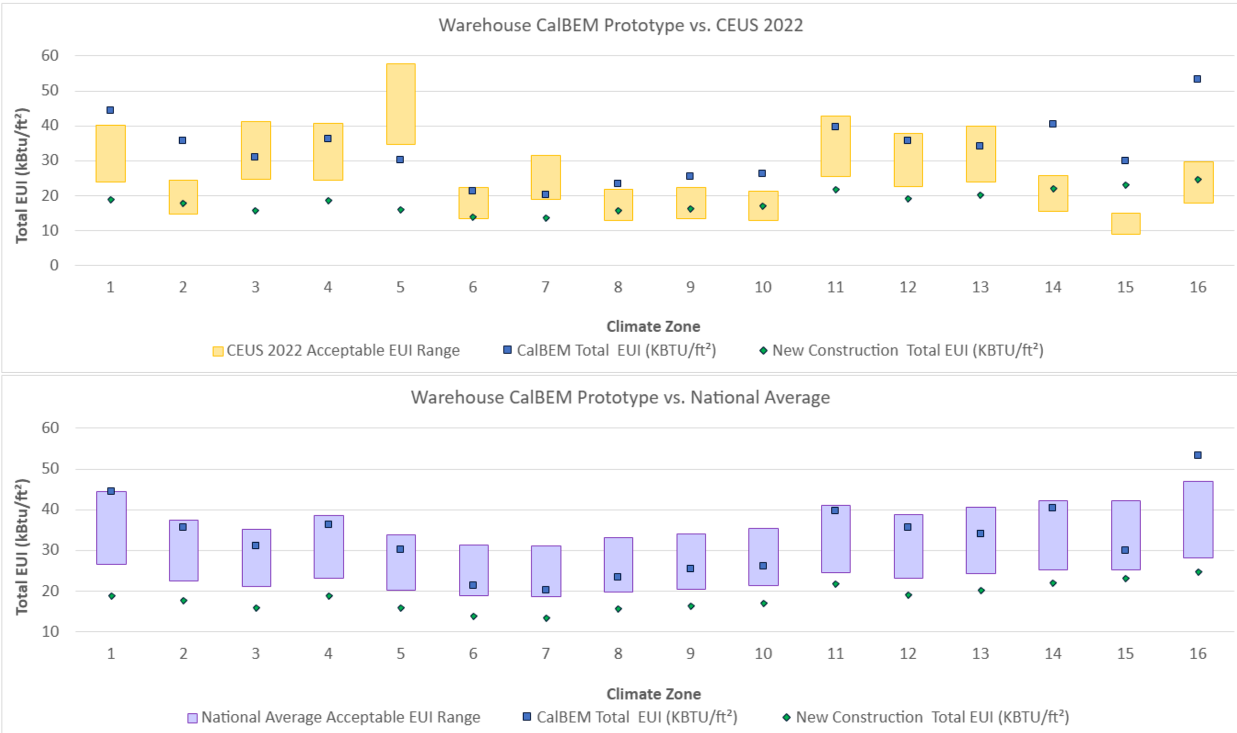

Figure 3: Comparison of Warehouse Prototype EUI versus CEUS 2022 by Climate Zone

Application

Online Hosting and Public Access

The CalBEM prototype suite and the associated scorecard specification are publicly accessible through the CalBEM website:

- Web portal (project overview and access):
  - CalBEM main site: https://calbem.ibpsa.us/
  - Prototype catalog: https://calbem.ibpsa.us/calbem-prototypes/

The website serves primarily as a public-facing entry point and provides links to the GitHub repository, which hosts the prototype models, specifications, and supporting documentation.

- Source repository (version control):
  - Root repository: https://github.com/NOR-Codes-Stds/CalBEM-Prototypes

The GitHub repository is the primary source for scorecards, translators, and engine-ready artifacts. Files are organized into three top-level directories: Models, Specifications, and Documentation. Version tags and commit history provide a record of changes and support reproducibility from published models back to input tables. The GitHub repository includes the following major directories:

- **Models**
  This directory contains CalBEM prototypes for nonresidential building types**.** Specifically, prototype models (IDF files) are generated with EnergyPlus 25.1, and are organized by building type, climate zone, and vintage bins.
- **Specifications**
  This directory houses the engine-agnostic scorecards and schema used to construct

EnergyPlus input files and to support translation to other engines. Each scorecard for nonresidential prototypes enumerates major model descriptions including geometry descriptors, ventilation and infiltration rates, envelope properties, internal loads, schedules, HVAC, and service hot water.

- **Documentation**
  This directory provides methodological and stock-composition references that underpin the nonresidential and multifamily prototypes. Key documents include a memorandum that outlines the prototype mix that best represents California's nonresidential stock. In addition, there is another technical document that describes the features evaluated to define vintage bins for California nonresidential buildings, provides the rationale for bin selection, and presents the distribution of non-residential floor area across the proposed bins.

Integrating CalBEM Prototypes into CEC and CPUC workflows

A key next step for this project is incorporating the newly developed CalBEM prototypes into the CEC and CPUC workflows. The fundamental issue identified at the conception of this project was that these agencies were using different underlying assumptions for the building stock. It is these assumptions that need to be harmonized across various use cases and analyses performed on California's building stock.

The scorecard database serves the purpose of harmonizing these assumptions across different use cases very well. The existing workflows for CEC and CPUC do not need to change in significant ways in order to incorporate the new modeling inputs. For CPUC, an effort is underway to create scorecards that parallel the CalBEM scorecards. Using the scorecard database, the DEER models could be updated using the assumptions developed under this project. Similarly, for the CEC, the scorecard database provides the inputs that can be incorporated into the CBECC prototypes in an automated fashion.

Summary

A new set of nonresidential prototype models was developed, called the "CalBEM" prototypes, which represent California's building stock. The inputs to models have been documented in a machine-readable format, so that the prototypes could be developed in different simulation engines. Using the machine-readable inputs, it is possible to integrate the same underlying assumptions about California's building stock into prototypes used by the CEC, CPUC, other agencies, practitioners, and stakeholders.

The prototypes, scorecards, and documentation have been shared on GitHub and is publicly accessible. The automated procedure of developing models from input specifications means changes to inputs can be traced to a defined record (scorecard field/value + provenance) and propagated deterministically to regenerated models, which reduces the likelihood of transcription-style inconsistencies that are typical of manual, case-by-case re-authoring. A unified, documented prototype suite with an engine-agnostic specification will allow analyses across California's policymaking agencies while preserving traceability from reported results back to inputs.

Future Work

Several updates have been planned for the CalBEM prototypes in 2026:

1. New prototypes: A few prototypes are being developed and will be added to the existing suite of prototypes: data center, refrigerated warehouse, controlled environment horticulture, and nursing homes. A mobile home prototype is also under consideration for future development.
2. Review and Enhancements: A detailed review of the published versions of the scorecards, inputs, and models is underway. After completion, the review findings will be incorporated and a second version of the prototype models will be published, planned for the end of 2026. Other possible enhancements to models, inputs, and the framework discovered during the first round of development are also planned to be included in the second version of the CalBEM prototype suite.
3. Expanded Automation Framework: The current automation framework provides the raw inputs and data but not the tools for generation models from inputs. Such a framework and the accompanying tools are being planned for the next iteration.

Acknowledgment

The authors would also like to express their gratitude to SCE Codes and Standards Program for funding this project. The authors wish to acknowledge the contribution of several Southern California Edison (SCE) staff.

References

CalBEM Prototypes. GitHub Repository, Models, Specifications, Documentation. https://github.com/NOR-Codes-Stds/CalBEM-Prototypes

CalBEM. CalBEM Website (IBPSA-US site). https://calbem.ibpsa.us/

CalBEM. CalBEM Prototypes Catalog. https://calbem.ibpsa.us/calbem-prototypes/

California Energy Commission. 2022 Building Energy Efficiency Standards (Title 24, Part 6). https://www.energy.ca.gov/programs-and-topics/programs/building-energy-efficiency-standards

California Energy Commission. CBECC-Com Nonresidential Compliance Software (EnergyPlus/OpenStudio + CSE). https://www.energy.ca.gov/programs-and-topics/programs/building-energy-code-compliance-software

California Energy Commission. California Commercial End-Use Survey (CEUS). https://www.energy.ca.gov/data-reports/surveys/california-commercial-end-use-survey-ceus

California Energy Commission. California Climate Zones Maps and Guidance. https://www.energy.ca.gov/maps

California Public Utilities Commission. Energy Efficiency Portfolios (overview and policy context

Database for Energy Efficient Resources). DEER Database (measures, UES, NTG, documentation). https://www.deeresources.com/

Dodge. 2023. Construction starts data. Retrieved from It is a paid date which was provided to NORESCO by CEC. https://www.construction.com

EIAU.S. Energy Information Administration. 2018 Commercial Buildings Energy Consumption Survey (CBECS).

NREL.2023. ComStock Reference Documentation: Version 1. https://docs.nlr.gov/docs/fy23osti/83819.pdf

Title 24 Stakeholder Controlled Environment Horticulture (CEH): 2028 Draft CASE Report, Feb. 2026. https://title24stakeholders.com/wp-content/uploads/2026/02/CEH_-2028_Draft_CASE_Report.pdf

U.S. Department of Energy Commercial Reference Building Models of the National Building Stock. https://docs.nlr.gov/docs/fy11osti/46861.pdf

U.S. DOE Building Energy Codes Program (BECP). Commercial Prototype Building Models (ASHRAE 90.1/IECC). https://www.energycodes.gov/prototype-building-models

Yousefi, F., and Athalye, R. 2025. Determining Vintage Bins for Prototype Development: An Analysis of California Commercial Buildings. 2025 ASHRAE Annual Conference.

AI-assisted Tools: Disclosures and Responsibility

AI-assisted tools used: CoPilot, ChatGPT
Used for: drafting text, editing, and summarizing.
Extent: [light editing / citation research].
Human review: The authors reviewed and verified all content and citations.