# DEnCity: An Open Multi-Purpose Building Energy Simulation Database

*Amir Roth, Department of Energy*
*Martha Brook, California Energy Commission*
*Elaine T. Hale, Brian L. Ball, Katherine Fleming & Nicholas Long,*
*National Renewable Energy Laboratory*

**ABSTRACT**

Building energy simulation is a multi-purpose tool for improving building energy efficiency. At a small scale, engineers and architects use simulation to design and certify energy-efficient buildings. At a large scale, policy makers analyze hundreds of thousands of simulation instances to investigate the effects of proposed changes on a building stock.

Simulation requires expertise and effort to set up and minutes or even hours to run, making it a poor fit for some small-scale use cases. In some of these, "just in time" simulation can be replaced with *pre-simulation*—large-scale simulation shrink-wrapped to answer small-scale questions. Of course, pre-simulation simulation comes with its own drawbacks, including inordinate simulation requirements and potentially poor coverage of candidate buildings.

In this paper, we articulate the need for a multi-purpose, open, and dynamic database of simulation inputs and results and present an implementation called **DEnCity—DOE's Energy City**. DEnCity enhances the value proposition of large-scale simulation in general and pre-simulation in particular. By virtue of being shared and general-purpose, DEnCity amortizes the cost and increases the resolution of a given analysis by leveraging relevant data points created for previous analyses and applications. We encourage both public institutions and private companies to contribute data sets to DEnCity and describe features that reduce the intellectual capital risk of doing so. We also describe an initial DEnCity application—a pre-simulation based prototype implementation of California's Building Energy Asset Rating System (BEARS).

## Motivation and Overview

Building energy simulation, bottom-up engineering calculation of building energy consumption from a detailed description, is a multi-purpose tool for improving the energy-efficiency of buildings. At the small (*i.e.*, individual building) scale, architects and engineers currently use simulation to design individual buildings and to certify them under voluntary programs like Leadership in Energy-Efficient Design (LEED). Emerging small-scale uses include design of building control systems, model-driven dynamic building control, continuous commissioning and diagnostics, and mandatory labeling. At the large (*i.e.*, many building) scale, policy developers analyze hundreds of thousands of simulations, typically parametric studies covering building prototypes representative of regional or national stock, to develop building energy-efficiency codes like ASHRAE's Standard 90.1 and design guides like its 30% beyond 90.1 Advanced Energy Design Guides (AEDGs). Simulation requires detailed information about a building's use and operation, *e.g.*, occupancy, lighting, and plug-load schedules as well as weather conditions. Consequently, most applications of simulation (*e.g.*, building design, asset labeling, stock analysis) use it to compare alternative designs under equivalent operating assumptions. Uses of simulation to predict actual energy consumption (*e.g.*, in dynamic building control) are less established.

Simulation models require considerable expertise, experience, and effort to set up. Similar expertise, experience, and effort is required to interpret simulation results. Simulations themselves take many minutes and even hours to run. These drawbacks make simulation a poor fit for a number of small-scale use cases. For a subset of these, simulation performed on-demand by the user at time-of-use can be replaced with *pre-simulation*—large scale simulation performed ahead-of-time by experts and re-packaged to answer a specific class of small-scale questions. ASHRAE's Standard 90.1 and AEDGs are canonical examples and success stories of this approach, leveraging the expertise of a few to create robust easy-to-follow prescriptive rules for many. More recently, pre-simulation has been successfully deployed in online interactive tools like DOE's 179D easy calculator and Energy Impact Illinois' EnCompass.

Of course, large-scale simulation comes with challenges as well. The primary one is getting good coverage over end-user buildings with a reasonable number of simulations. Individual large-scale analyses and applications struggle with this challenge because they typically start from scratch. Although numerous analyses and applications exist, each is isolated and exposes only the widgets, the prescriptive rules or regression equations, produced by analyzing the dataset to answer the application-specific questions. The raw underlying models and simulations themselves, which could potentially be re-analyzed for other purposes, are not shared. Often, they are discarded, having served their singular purpose.

This scenario, played again and again, illustrates the need for a public, open, and dynamic database of building energy simulation input models and results. Such a database would enhance the value proposition of large-scale simulation, by leveraging the raw data of previous analyses and applications to amplify the scope (breadth, depth, resolution, or all three) of new analyses and applications. This paper describes such a database—**DOE's Energy City (DEnCity)**. Pronounced 'density', DEnCity is meant to evoke the dense clouds of interesting regions in the universe of all possible building designs. The City spelling plays on the name SimCity™, a computer game featuring simulated buildings. We present an implementation of DEnCity based on DOE's open-source building energy simulation middleware, OpenStudio. We detail its initial population, which underpins a prototype pre-simulation based implementation of California's Building Energy Asset Rating System (BEARS). We list several potential future use cases—just as "big data" is revolutionizing analytics and services in other sectors, we envision it opening up new possibilities in the building space. Finally, we articulate the DEnCity value proposition and encourage both public and private institutions to contribute datasets to this common resource.

## Building Energy Simulation Datasets

Building energy simulation is nearly as old as computing itself. Simulation model and result datasets supporting design, research, and policy analysis are younger by a few hours. This section summarizes several notable instances including some larger, more recent ones.

### California Commercial End-Use Survey (CEUS)

In 2002 and 2003, the California Energy Commission (CEC), in cooperation with the four California investor-owned utilities, conducted on-site surveys of 2,790 commercial establishments throughout the state (Itron Inc., 2006). Data was collected on building characteristics, equipment operation, and energy consumption to support the design and planning of energy efficiency programs and the CEC's long-term energy demand forecasts. DOE2.2 models were constructed for each surveyed building site and calibrated to utility bills. These models were used to develop whole-building and end-use electric and gas energy intensities, sector-level hourly load profiles, and daily gas use profiles, and to evaluate the energy and demand impacts of energy-efficiency measures (EEMs).

### NREL 2007 Zero Net Energy (ZNE) Potential Study

In 2007, researchers at the National Renewable Energy Laboratory (NREL) completed a study of the potential of achieving zero net energy (ZNE) commercial buildings by 2025 (Griffith, Long, Torcellini, Judkoff, Crawley, & Ryan, 2007). The study examined a large number of prototype buildings representative of the 2003 Commercial Building Energy Consumption Survey population (CBECS) stock in 17 US climate zones. Each building was modeled at ASHRAE 90.1-2004 new construction standard-level to establish baseline performance, and then with combinations of EEMs projected to be commercially viable in 2025. EEMs considered included technologies like advanced lighting, HVAC, insulation, and fenestration (*e.g.*, dynamic windows) as well design measures like daylighting (*i.e.*, large north-south exposed perimeter zones, appropriate window to wall ratios, shading, and electric light dimming). The study used 115,680 EnergyPlus (US DOE) simulations and estimated that 47% of commercial floor space concentrated in mild, marine, and colder climates and in the education, retail, and warehousing sectors could be made ZNE.

### EPAct 2005/179D Easy Calculator

Section 179D of the 2005 Energy Policy Act establishes a tax deduction of up to $1.80 per square foot for qualifying commercial properties. To qualify, a building owner must demonstrate via simulation annual energy use 50% lower than an ASHRAE 90.1-2001 code-minimum baseline building. Partial deductions of $0.60 can be claimed for demonstrated savings in lighting, HVAC, and building envelope, individually.

The high cost of setting up a simulation relative to the size of the deduction discourages its use for small properties. To help owners of small buildings access the deduction as added incentive to upgrade, NREL created a web tool based on pre-simulation (Deru, Griffith, Leach, Bonnema & Hale, 2012). The building "universe" was divided into 192 major domains corresponding to 12 NREL Reference Building templates (Field, Deru, & Studer, 2010) at 16 climate zone representative locations. The baseline building for each domain is modeled at ASHRAE 90.1-2001 code level. Each domain is then populated with models in which this baseline is modified with a number of EEM packages for lighting (lower lighting power density), envelope (*e.g.*, lower U-values for walls, windows, and roofs), and for a small number of common HVAC system types (*e.g.*, higher furnace and coil efficiencies for packaged HVAC systems). Within each domain, the envelope upgrade parameter space was simulated in full—every combination of wall, roof, and window upgrade. The HVAC upgrade was similarly explored in full. However, envelope, HVAC, and lighting upgrades were not simulated together—as such, the tool supports only the partial deductions. Within each domain, the envelope and HVAC upgrade result sets were independently regressed to derive relative energy savings coefficients for each EEM. At time of use, the building type and location identify the domain and the corresponding set of coefficients, which are then used to calculate predicted energy savings and the deduction qualification.

### EnCompass

EnCompass (Energy Impact Illinois) is a pre-simulation based web tool for identifying retrofit opportunities in Chicago office buildings and for connecting building owners to retrofit financing, engineering, and contracting services. EnCompass uses data extracted from a large pre-simulated design space to quickly show building owners potential energy savings associated with various EEM packages. To build EnCompass, PositivEnergy Practice ran 278,000 EnergyPlus simulations. With a different focus than the 179D easy calculator, the EnCompass database covers a smaller region of the potential building universe in a more granular fashion. EnCompass includes only two building-location domains—office buildings representing downtown and suburban Chicago, both adapted from the NREL Reference Building templates. Within these, EnCompass explores not only envelope U-values and HVAC component efficiencies, but also square footage, vintage, window to wall ratio, and a greater variety of HVAC systems. It also evaluates these EEMs in all combinations. And whereas 179D uses regression to calculate energy savings over a single implicit baseline in each domain, EnCompass uses a "closest match" approach to look up EUIs for the building in both its current pre-retrofit and presumed post-retrofit states.

### Discussion

The primary advantage of pre-simulation is the expertise of those setting up the simulations and analyzing, interpreting, and shrink-wrapping the results. A typical end user of simulation is unlikely to have the same level of expertise, especially in cases of high-volume applications like mandatory labeling, which could involve many new and inexperienced end-users. Some of these benefits can be re-packaged and provided in a just-in-time fashion. For instance, a thick layer of defaults can simplify the model input creation process. Robust benchmarking and sanity checking on both inputs and outputs can be installed. However, the chance for outliers and special circumstances always exists with simulation, and experts are more adept at both spotting and diagnosing these cases. With pre-simulation, outliers and out-of-bounds situations are explicitly flagged, reducing confusion, misinterpretation of results, and misuse of simulation.

Another benefit of pre-simulation is its ability to produce value distributions as easily as point values. For many applications, distributions are more useful than point values because they quantify risk in addition to expected benefit. While none of the applications described above exploit value distributions, DEnCity's measured building counterpart, DOE's Building Performance Database (DBPD) (US DOE), uses it extensively to both provide anonymity for individual buildings and to support actuarial functions.

Another potential benefit is a reduction in the number or specificity of inputs to the building model creation process and a commensurate reduction in the burden on the auditor and modeler. Audit protocols are routinely optimized to prune unnecessary parameters. Analysis of a pre-simulated dataset performs this pruning automatically and systematically by identifying explicit input parameters (*i.e.*, parameters considered significant *a priori*) that actually have little influence on end use metrics in particular regions of the design space. Pre-simulation has a second input pruning mechanism that builds on its ability to provide value distributions. Namely, pre-simulation allows a user to obtain broader, less-specific distributions in response to fewer, less-specific inputs.

Calculation speed is another advantage, especially for interactive applications like EnCompass that engage the user in part by allowing her to select and deselect EEMs and watch the building's end-use energy consumption change instantaneously.

On the liability side of the ledger, any large-scale simulation solution must balance design universe resolution (which is needed to closely match the results of on-demand simulation) with design universe range (which is needed to cover a sufficient number of end user buildings). Providing uniformly high resolution over a wide range requires astronomical numbers of simulations. Analysis providers typically opt for resolution and drastically restrict the range (*i.e.*, scope) of their applications. Consider the 179D calculator. Supporting the partial lighting, envelope, and HVAC deductions required 225,000 simulations—over 800 envelope simulations plus nearly 300 HVAC simulations in each building-location domain. Supporting all combinations of envelope and HVAC EEMs together (a minimal requirement for the whole-building deduction) would require 800 *times* 300 simulations in each domain! That's 46,000,000 total runs! The 179D calculator also does not explicitly account for square footage, assuming that relative savings due to upgrades remain constant, and excludes buildings that differ substantially from NREL reference buildings. EnCompass considers only Chicago office buildings.

## DEnCity

Consider a scenario in which EnCompass is replicated for other building types in other cities. Collectively, the various EnCompass datasets could underpin a significantly more powerful 179D calculator and provide much of the data for an ASHRAE 90.1 update effort. Conversely, the data from an ASHRAE 90.1 update effort could feed EnCompass instances for various locations or an expanded calculator for 179D. Large datasets could also find new applications in benchmarking, calibration, early stage design guidance, and others tasks. This sharing, reuse, and multi-purposing of simulation data is the vision behind DEnCity. DEnCity is not limited to large parametric studies on models of hypothetical or reference buildings. Models of real buildings are valuable as well because they act as anchors for interesting and relevant regions of the building design space (regions which merit dense exploration) as well as for the (indirect) calibration of their design space neighborhood.

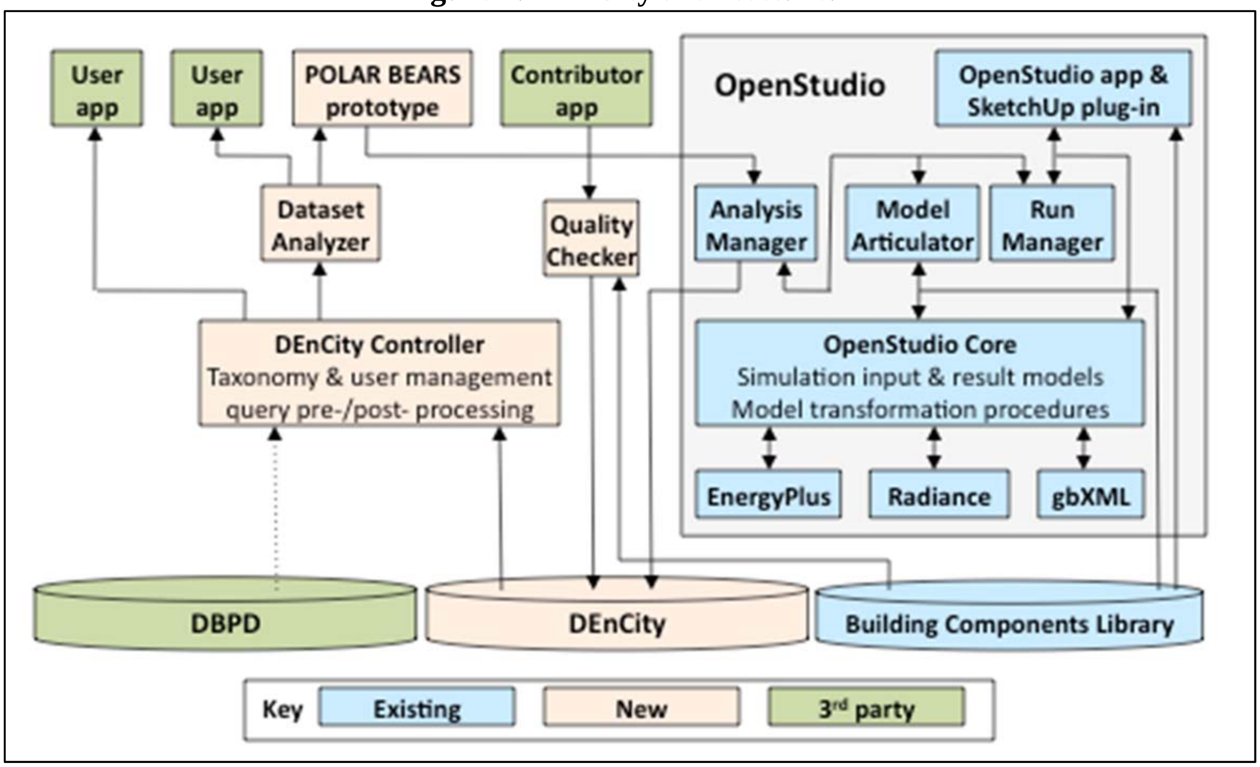
### Software Architecture

Figure 1 shows DEnCity's software architecture. DEnCity is an online database. In recent years, DOE has created several online databases for data that is relevant to building energy assessment. The Building Component Library (BCL) contains simulation objects like schedules, constructions, and equipment (Fleming, Long & Swindler, 2012). We expect that many of the models within DEnCity will contain BCL components. DBPD contains parameters and consumption data for actual buildings. We are aligning DEnCity's schema with DBPD's to facilitate applications that combine simulated and measured data.

DEnCity is implemented in mongoDB, an open-source NoSQL database that supports document storage. DEnCity stores full model input files; it does not permanently store full result files because these are huge and can always be recreated on-demand via simulation. DEnCity also extracts and indexes key input and result attributes (*e.g.*, square footage, weather file, HVAC system type, monthly energy consumption by fuel and end-use) for fast lookup and retrieval. DEnCity can incrementally grow the set of query-able model and result attributes as new applications and use cases appear. We are building a controller module for DEnCity that manages the schema and user permissions, and that performs query pre- and post-processing.

We use OpenStudio, NREL's open-source cross-platform simulation middleware, to create DEnCity's initial population. OpenStudio presents both building models and simulation results as "live" objects that can be manipulated programmatically (NREL). We use several OpenStudio modules specifically—AnalysisManager to define the population in terms of parameter value combinations, ModelArticulator to articulate full simulation models from attributes, and RunManager to simulate the resulting models using EnergyPlus on NREL's RedRock computing cluster. DEnCity is a not a simulation server. We imagine that dataset contributors will run simulations on their own resources. We will implement quality control for foreign model input files.

**Figure 1. DEnCity architecture.**




DEnCity is not an application server. We imagine that most client applications will query for summary result datasets that match certain attributes, and occasionally for input or input model/result file pairs, while performing their own analyses. Currently, we analyze and visualize data queried from DEnCity using a DatasetAnalyzer we are implementing in R (R Foundation). DatasetAnalyzer will be open source and made available to user applications.

### Value proposition

Like DBPD, DEnCity offers a clear value proposition for users of pre-simulation data—the potential for free, instantaneously available seed data for a given analysis. The more data in DEnCity, the more applicable and robust this proposition becomes. DOE and the California Energy Commission (CEC) are committed to supporting this value by providing the DEnCity infrastructure and datasets from their funded projects, and are working with other public organizations to do the same. And as with DBPD, DOE and CEC encourage private institutions to contribute data to DEnCity for the common good.

DEnCity appears to encourage "free riders"—users that only benefit from others' contributions and do not contribute their own data. However, direct benefits do accrue to those who contribute datasets to DEnCity. Contributors benefit from free storage and indexing of their datasets, and from the ability to use DEnCity's quality control, query, and analysis functions. They may also benefit because increasing the transparency of their raw data increases others' confidence in the analyses that build on that data. Contributors may also benefit from goodwill that comes with contributions. To enhance these admittedly intangible benefits, DEnCity includes features that minimize the cost of contribution in terms of exposure of intellectual capital. DEnCity contains only inert data, not analysis scripts, results, or summaries. Contributors of large parametric datasets can be confident that other users are not getting the insight and knowledge that they themselves derived from it, at least not without reinvesting in the analysis work. Beyond this, DEnCity tracks data ownership and will be able to allow owners to control the visibility of their data. For instance, owners can make only the summary attributes of their data, not the detailed simulation-ready model, visible to others. This kind of assurance may be especially important to contributors of real building models and is similar to, although slightly weaker than, the privacy assurance DBPD provides to data contributors.

### Are Building Models Reusable?

Whether people will choose to contribute to DEnCity is only a $32,000 question. The $64,000 question is whether simulations created in the service of one analysis is actually useful for another. Is a simulation model articulated from 40 high-level parameters really useful for design guidance? Can simulations models that embody different operating assumptions be used productively in the same analysis? Although the current DEnCity schema does not contain meta-data about the original purpose of the model, the templates used in its creation, we anticipate adding this information in order to both help users make better and more informed use of DEnCity data and claim the $64,000 prize.

## POLAR BEARS: a Pre-simulation-Based Asset Rating Methodology

The seed DEnCity population is a parametric dataset that underlies a pre-simulation based prototype implementation of California's Building Energy Asset Rating System (BEARS). A building asset rating attempts to capture operation-independent building performance, making a simulation based implementation both natural and *de rigueur*. Examples of such asset rating implementations include the USGBC's LEED Energy and Atmosphere Credit 1 performance path (USGBC), DOE's Home Energy Score (US DOE), and eventually the asset component of ASHRAE's Building Energy Quotient (ASHRAE).

Whereas the systems described above all use on-demand simulation, in this paper we propose an asset rating methodology based on pre-simulation and regression called Pre-simulated On Line Asset Rating (POLAR). Our proposed implementation of BEARS is naturally called POLAR BEARS. We aim to use pre-simulation to make the rating process more robust by (1) allowing modeling experts to map the rating space and flag outliers *a priori*, (2) providing ratings based on distributions of similar buildings rather than single building estimates, and (3) by simplifying the rating implementation software. The hypothesis is that pre-simulation can closely track ratings by traditional on-demand simulation and provide good coverage of the rating eligible building space with reasonable simulation requirements.

## The BEARS Program

BEARS will provide a rating metric that values the permanent energy-related components of a commercial building while fixing assumptions for building operations and occupant behavior. BEARS will be used in California within voluntary public goods efficiency programs (in concert with measured energy use ratings) to provide data to decision-makers on relative energy performance that can help target potential improvement opportunities. BEARS will also be used in future mandatory rating disclosure programs, where assessments of the relative energy performance of commercial buildings are needed to value energy efficiency within property transactions. BEARS aims to provide a consistent rating system that can be applied to all new, existing, or renovated commercial buildings in California.

The BEARS scale sets a benchmark of 100 at the whole building energy use of a 2010 Title-24 energy code compliant building. Zero on the BEARS scale indicates zero net energy, as defined by the CEC (California Energy Comission, 2012). Separate benchmarks for fifteen building types in each of sixteen California climate zones will be used to compute the BEARS metrics. Relative performance metrics will be computed for the whole building as well as key energy subsystems, such as lighting, cooling, heating, and water heating. BEARS will be piloted with commercial buildings in 2013.

## POLAR BEARS

POLAR BEARS is conceptually similar to the 179D easy calculator. The building universe is divided into domains by building type and California climate zone. Each building type-location domain is associated with a baseline Time Dependent Valuation Intensity (TDVI)—an energy use intensity metric that accounts for seasonal and time-of-use effects—codified in Title 24. The rating is calculated as TDVI(building-of-type-A) / $TDVI_{base}$(type-A). The use of intensities makes it easier to rate mixed-use buildings. For instance, the rating of a building that combines use types A and B is: $[\%(A)\times TDVI(A) + \%(B)\times TDVI(B)] / [\%(A)\times TDVI_{base}(A) + \%(B)\times TDVI_{base}(B)]$.

Because the BEARS scale is more granular than the 179D qualify-or-not binary scale, POLAR BEARS considers more parameters than the 179D calculator including building footprint shape, orientation, and dimensions, window to wall ratio, infiltration, and a greater variety of HVAC systems and lighting options. It also considers them in arbitrary combinations. To control the number of simulations, POLAR BEARS populates domains using Latin Hypercube Monte Carlo sampling (Giunta, Wojtkiewicz & Eldred, 2003). Even sparsely populated, POLAR BEARS domains may contain hundreds of thousands of building models, too many for regression over the entire domain to produce acceptable fit. Instead, given the characteristics of the building to be rated, POLAR BEARS executes a nearest neighbor query to extract the K most similar buildings. Influential parameters identified by regression over large portions of the space are weighed more heavily in the distance function. POLAR BEARS then regresses over this limited set to derive an equation for the TDVI.

## POLAR BEARS Experimental Prototype

For this paper, we evaluate the POLAR BEARS approach using a single building-location domain—small office buildings in the greater Fresno area. Even this single domain is not fully populated. As of this writing, it includes only buildings with rectangular footprints, gas furnace heating, and heat pump air conditioning. Table 1 lists the parameters, their value ranges, and their weights in the nearest neighbor function where higher coefficients correspond to more influential parameters. The fixed parameters are shaded in orange and the influential parameters are in green. Intuitively, these include floor area, insulation values, window to wall ratios, and cooling efficiency.

**Table 1. POLAR BEARS Characteristics for Small Office Buildings in Fresno**

| Characteristic | Value range | NNW | Characteristic | Value range | NNW |
|---|---|---|---|---|---|
| **Footprint shape** | **Rectangular** | | Window to wall ratio by façade | 0 — 1 | 10 |
| Floor area ($m^2$) | 10 — 1550 | 10 | Floor to sill height by façade (m) | 0 — 3 | 0 |
| Aspect ratio | 0.67—15 | 1 | Window U-factor by façade ($W/m^2K$) | 0 — 1 | 1 |
| Number of floors | 1—3 | 10 | Window SHGC by façade | 0 — 1 | 1 |
| Floor to floor height (m) | 2.5 — 10 | 1 | Window VT by façade | 0 — 1 | 0 |
| Orientation (degrees from north) | –45 — 45 | 1 | **Heating type** | **Gas furnace** | |
| Wall insulation R-value ($m^2K/W$) | 0.1 — 20 | 10 | **Cooling type** | **Heat pump** | |
| Attic insulation R-value ($m^2K/W$) | 0.1 — 20 | 10 | Heating efficiency | 0.5 — 1 | 1 |
| Interior lighting power density ($W/m^2$) | 0 — 30 | 10 | Cooling efficiency (COP) | 0.5 — 10 | 10 |
| Exterior lighting power (W) | 0 — 10,000 | 1 | Fan efficiency | 0.1 — 1 | 10 |

**Figure 2. Regression TDVI Over Model TDVI for 85 CEUS Office Buildings**

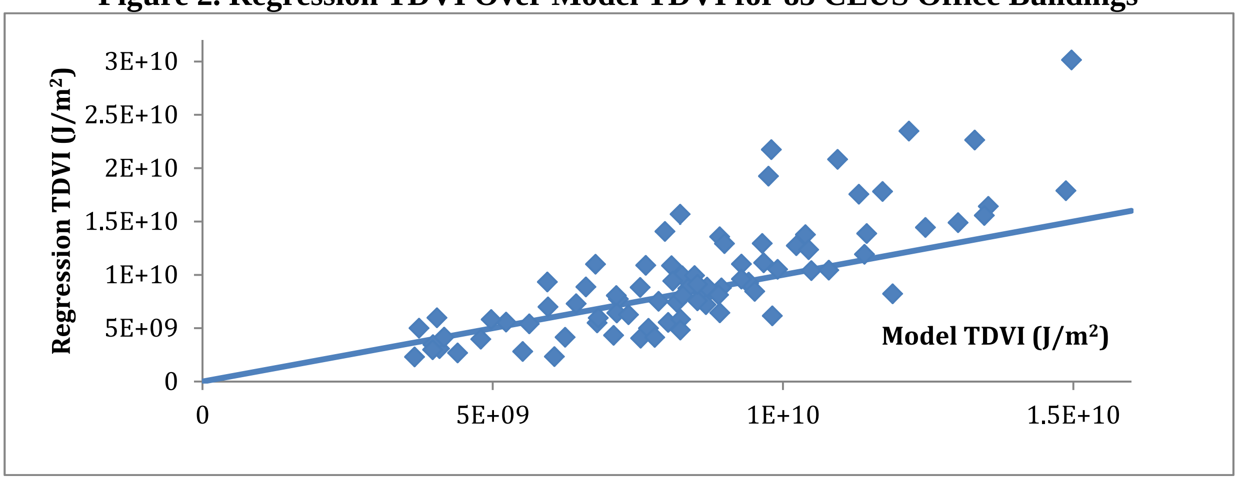
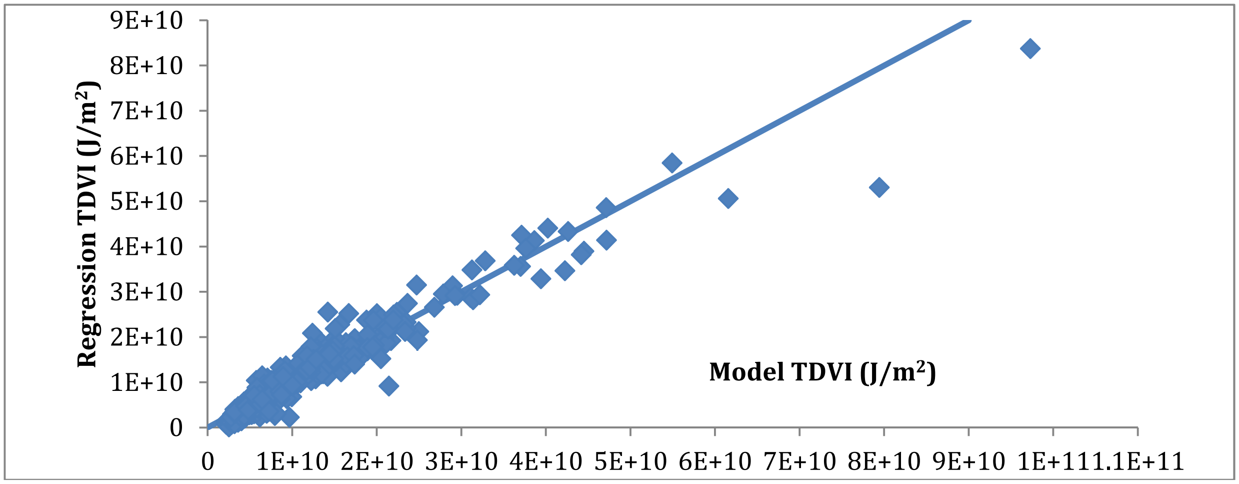


**Figure 3. Regression TDVI Over Model TDVI for 500 DEnCity Buildings**




Figures 2 and 3 show accuracy results (TDVI calculated from DEnCity Regression vs. TDVI calculated using a custom articulated model) for two datasets: 85 office buildings from CEUS, and 500 random buildings from the DEnCity database itself. For an asset rating with a 100-point scale, error of less than 1% is desirable. The CEUS experiment uses 100 neighbors and discards terms from the regression equation whose p-value is greater than 0.005. This usually results in an equation with 4 to 12 terms. Median error is 19% with a maximum of 121%. The median time to execute the nearest neighbor query and the regression is 9 seconds. The experiment on the random DEnCity sample uses 500 neighbors and discards regression terms with p-value of greater than 0.01, typically yielding 20 to 24 terms. Accuracy results here are better: 10% median and 93% maximum error. However, the time to execute the query and perform the regression is also longer: 127 seconds median.

This methodology is preliminary and we continue to refine it. Among the issues we have noted are: a number of outlier buildings with high energy consumption in the domain population, an inability to deal cleanly with building facades that have no windows (this problem is manifest in the CEUS experiment where many buildings have one or more window-less facades), and unrealistic over-representation of buildings with high aspect ratios. However, the biggest issue is the nearest neighbor function, which was found to return some poor neighbors using manual inspection of neighbor sets for several CEUS buildings. High-dimensionality nearest neighbor searches are notoriously inefficient. We are experimenting with a staged nearest-neighbor approach in which an initial pool of neighbors is selected using a low-dimension query on only the influential parameters. That pool is subsequently pruned using all parameters. Early returns are promising.

## Fallback Options

In some situations, POLAR BEARS cannot use regression and must fall back on direct simulation (of a full model articulated from the building's characteristics) to calculate the TDVI. This contingency is invoked when regression over the neighbor set results in a fit that fails to meet threshold requirements. It is similarly invoked when the rated building's characteristics do not yield a sufficient number of neighbor matches. In the latter case, POLAR BEARS will dynamically populate the neighborhood offline to avoid similar failures in the future. We expect these contingencies will be rare. Finally, the rated building may have some energy relevant characteristic that cannot be expressed using the limited list of rating parameters defined in the POLAR BEARS design space or even captured by the field audit protocol. In these cases, rating the building requires a custom, manually assembled simulation model. The expectation is that the 2013 pilot will account for a sufficient set of characteristics to handle a broad range of building features. If enough buildings are found to have a particular feature that falls outside the model universe, that feature could be added to the next version. The dynamic character of DEnCity supports this type of evolution.

### Discussion: POLAR BEARS vs. ESPM

ENERGY STAR Portfolio Manager (ESPM) (US EPA) is another building energy-efficiency rating system that calculates expected energy consumption by plugging values of building characteristics into regression equations. Despite its mechanical similarity to POLAR BEARS, the two systems are markedly different. ESPM calculates expected performance by regressing over the *actual* performance of similar buildings in the CBECS database; the EnergyStar score is the ratio of actual performance to expected performance. In this setup, the building "gets credit or blame" only for characteristics that the (regression) model excludes. For example, ESPM regresses on occupant density, subtracting (most of) the effect of this parameter from the rating. At the same time, it does not regress on thermostat set points and so these are conflated into the rating. This design makes ESPM applicable to a wide range of buildings, but also means that ENERGY STAR is not a pure asset rating—this is not an indictment, as this was never the intent. In contrast, POLAR BEARS calculates pure asset ratings. Predicted building energy use is the numerator which means the building gets credit or blame only for those characteristics that *are* explicit in the (regression) model. POLAR BEARS subtracts operating assumptions from the rating by fixing them within each building-type location domain. The cost of calculating pure asset ratings is the need to account for more characteristics—ESPM uses six parameters for office buildings, POLAR BEARS uses 35—and the commensurate need for a dataset much larger than CBECS.

## DEnCity Futures

A large repository of building energy models and their simulation results like DEnCity is a platform that enables new applications and uses. In closing, we explore several of these.

### Early-Stage Design Guidance

Many building characteristics that profoundly affect energy performance are determined early in the design process—siting, orientation, exterior geometry, and gross façade characteristics, to name several. At this stage, few of the hundreds or thousands of parameters needed for building simulation are known and simulation is too cumbersome to set up and too slow for the thousands of runs needed to assess the performance impacts of design decisions. A large database of simulation results can produce performance distributions with relatively few parameters, and do so quickly. With the proper analysis, the database could also underpin a module that actively guides the designer towards more energy-efficient regions in the design space.

### Stock Analyses on Actual Building Stock

Most technology potential studies and code and efficiency program impact analyses are currently performed using building prototypes that *represent* the building stock. However, as mandatory labeling programs proliferate and performance-based energy-efficiency codes are adopted, a growing number of actual buildings will have up-to-date energy simulation models. If these are publicly available in a standardized, query-able form, future analyses could be done on a growing subset of the *actual* building stock.

### Applications of DBPD and DEnCity

The DEnCity schema is designed to align with that of its real-building cousin DBPD with the idea of using the two in complementary, mutually reinforcing ways. The design guidance application described above could draw data from both databases, giving the designer a simulated performance map dotted with real world examples. Corresponding points in the two databases, an existing building's simulated and actual energy use, could be used to calibrate surrounding neighborhoods in DEnCity. In reciprocity, surrounding virtual neighborhoods in DEnCity could provide insight into performance and guidance on retrofit possibilities not yet measured and recorded in DPBD. Differences between corresponding points could identify simulation gaps and future research and development needs. Figure 1 shows the DEnCity controller module reading data from DBPD—the line is dotted because the link is, to date, hypothetical.

## Summary

DEnCity is an open database of building energy models and their simulation results. It is designed to align with DOE's Building Performance Database (DBPD), a database of real building consumption data that includes sufficient building characteristics to articulate full simulation models. DEnCity aims to improve the value proposition of large-scale building energy simulation by allowing simulations to be re-purposed and reused for future analyses. DOE and CEC plan to make the simulation data from their large scale analyses available in DEnCity and encourage others to do the same.

DEnCity is inspired by other projects that have used large-scale pre-simulation to overcome the drawbacks of just-in-time simulation for individual buildings. These limitations include potential lack of user expertise to create the simulation model and interpret its results, long simulation times, and the need to perform many simulations depending on the application. Previous applications include DOE's 179D tax credit calculator and EnCompass, an interactive retrofit exploration tool for Chicago office buildings. We are experimenting with a large-scale pre-simulation based implementation of California's asset rating.

## Bibliography


ASHRAE. *Building Energy Quotient (bEQ):* http://www.buildingeq.com/

California Energy Comission. (2012). *2011 Integrated Energy Policy Report.*

Deru, M., Griffith, B., Leach, M., Bonnema, E., and Hale, E. (2012). *179D Easy Calculator Technical Support Document*. Golden, CO, National Renewable Energy Laboratory.

Energy Impact Illinois. *EnCompass*. http://encompass.energyimpactillinois.org/

Field, K., Deru, M., & Studer, D. (2010). Using DOE Commercial Reference Buildings for Simulation Studies. *SimBuild* 2010.

Fleming, K., Long, N. & Swindler, A. (2012). Building Component Library: an Online Repository to Facilitate Building Model Creation. *ACEEE Summer Study on Energy Efficiency in Buildings*, 2012.

Giunta, A., Wojtkiewicz, S., and Eldred, M. (2003), Overview of modern design of experiments methods for computational simulations, *AIAA Journal*, 2003.

Griffith, B., Long, N., Torcellini, P., Judkoff, R., Crawley, D., & Ryan, J. (2007). *Assessment of the Technical Potential for Achieving Net Zero-Energy Buildings in the Commercial Sector.* Golden, CO: National Renewable Energy Laboratory.

Hale, E., Macumber, D., Weaver, E., and Shekhar, D. (2012). A Flexible Framework for Building Energy Analysis. *ACEEE Summer Study on Energy Efficiency in Buildings*, 2012.

Itron Inc. (2006). *California Commercial End Use Survey.* Consultant report.

NREL. *OpenStudio:* http://openstudio.nrel.gov/

R Foundation. *The R Project for Statistical Computing:* http://www.project-r.org/

US DOE. *DOE Buildings Performance Database:* http://www.buildingsperformance.net/

US DOE. *Buildings Technologies Program: EnergyPlus Energy Simulation Software:* http://www.energyplus.gov/

US DOE. *Building Technologies Program: Home Energy Score*. http://www1.eere.energy.gov/buildings/homeenergyscore/

US EPA. *ENERGY STAR: Portfolio Manager Overview*. http://www.energystar.gov/index.cfm?c=evaluate_performance.bus_portfoliomanager

USGBC. *LEED:* http://www.usgbc.org/LEED/