# Development of an Energy Score for Laboratories

Travis Walter and Josh Kace - Lawrence Berkeley National Laboratory Alison Farmer - International Institute for Sustainable Laboratories

August 18, 2023

# 1. Overview

This memo describes the development of a system for assigning a score to laboratory buildings that reflects its energy performance relative to its peers. We largely follow the methodology used by the U.S. Environmental Protection Agency (EPA) to develop the ENERGY STAR Score [1] but deviate when their method is not applicable, or when our data sources prevent us from doing so. Section 2 describes the datasets used for the analysis. Section 3 discusses the development of the linear regression model relating laboratory characteristics to energy performance. Section 4 details the use of the regression model's prediction for computing a score. Lastly, Section 5 discusses future development and refinement opportunities for the scoring system. This work was carried out by Lawrence Berkeley National Laboratory (LBNL) and the International Institute for Sustainable Laboratories (I2SL), but we also received valuable feedback from several stakeholders and members of the I2SL Labs2Zero Energy Score Technical Advisory Committee (TAC).

# 2. Data Sources

The primary source of data for this analysis was the dataset underlying the Laboratory Benchmarking Tool (LBT) [2], which is the largest known collection of energy-related information on laboratory buildings. We utilized only the subset of the dataset whose data has been lightly quality checked<sup>1</sup> by LBNL after being entered by users of the tool, leaving us with data for 990 laboratory buildings. For each building, the database contains over 100 data fields describing various characteristics relating to the laboratory's size, location, usage patterns, installed systems, and energy consumption. See the full list of LBT data fields and their descriptions [3] for more information. We carefully inspected the data and removed any data

<sup>&</sup>lt;sup>1</sup> Quality check procedures cannot identify all forms of user entry error within the LBT.

deemed to be unreliable (e.g., physically unrealistic) or otherwise not representative of laboratories in general (e.g., abnormally high or low values relative to other buildings).

We augmented the LBT dataset with weather data from Degree Days.net [4], which compiles temperature data from thousands of weather stations worldwide and calculates heating and cooling degree days (HDD and CDD) for a given location, time period, and base temperature. We downloaded HDD and CDD (both with 65F base temperature) data for each building in the LBT dataset using its location and the year in which energy consumption was measured. While the original LBT dataset includes data on building location (i.e., a proxy for climate) and the year of energy data measurement<sup>2</sup> (i.e., along with climate, a proxy for weather), we believed HDD and CDD to be more direct measurements of the meteorological conditions that would impact a building's energy consumption.

The combined dataset includes laboratory buildings with a variety of locations, use types, ages, operating characteristics, and levels of energy consumption. Many of the categorical fields in the dataset have unequal distribution of types. For example, Figure 1 shows that while the dataset contains buildings in 16 different climate zones, the large majority of them are in just 4 zones. While there are very few buildings with unknown climate zone, other categorical fields considered optional for setting up a building profile (fewer than 20 fields are required) have significant portions of the dataset with either unknown values, or with values of limited utility (e.g., other or combination). Fields that are not required tend to have significantly more missing data, including some fields that are now required but were previously optional.

<sup>&</sup>lt;sup>2</sup> Analysis assumes user-entered data for a given reporting year to be representative of a standard calendar year.



Figure 1: Histogram of climate zones in the dataset, including only buildings with data for the performance metric.

For numerical fields, many distributions are similarly unequal, and depending on the field, many buildings might not have data available. For example, Figure 2 shows a few buildings built before 1900 but 25% of buildings built after 2003, and that data for this field is available for only 650 of the 794 buildings with data for the performance metric.



Figure 2: Histogram of year built in the dataset, including only buildings with data for year built and the performance metric.

# 3. Fitting the Model

In order to develop a scoring system, we first constructed a model that estimates the typical energy consumption for a laboratory with a given set of characteristics. To fit this model, we assumed that the LBT dataset was a representative sample of the wider population of laboratory buildings for which the scoring system will be applied. It should be noted that without a comprehensive dataset containing information on all laboratory buildings in existence, we are unable to confirm whether this assumption of representativeness is true. However, since the LBT dataset is the largest known dataset of energy-related laboratory information, and since it appears to include buildings with a wide variety of types, locations, system characteristics, and usage patterns, we are confident this dataset can be used to derive a useful and trustworthy energy consumption model. The primary role of the model is to predict the performance metric upon which laboratories will be scored. We carefully considered several metrics, but narrowed the list down to two final candidates: site energy use intensity (EUI) and source EUI. Both candidate metrics are a measure of energy consumption that is normalized by the size (i.e., gross floor area) of the building, in order to avoid penalizing or rewarding buildings based on their size (i.e., an inherent property of the building that generally cannot be changed in order to improve performance). Site EUI measures the un-weighted sum of energy consumption from all fuels (e.g., electricity, natural gas, district steam). Source EUI also measures consumption from all fuels, but weights the consumption from each fuel according to the amount of primary energy used to generate and transport the energy to the building. While site EUI is more intuitive for building owners and operators (e.g., it can be read directly from utility bills) and can tend to encourage electrification of fossil-fuel systems, source EUI is more indicative of scope 1 and 2 greenhouse gas (GHG) emissions and is used for the ENERGY STAR Score (thereby providing consistency across building types for owners of portfolios of buildings). Thus, we selected source EUI as the performance metric upon which buildings will be scored. We removed the 5% of buildings without source EUI data from the dataset, and an additional 15% of buildings whose source EUI was estimated (not measured).

We next aimed to identify which data fields would be used as predictors in the model. With input from the Energy Score TAC, out of the more than 100 fields in the LBT dataset, we selected roughly 30 fields that were likely to have an impact on energy performance and that we considered to be a functional requirement for operating the laboratory (as opposed to the means by which those functional requirements are achieved). For example, we considered the number of occupants to be a functional requirement, but considered heating and cooling system types to be the means by which those occupants are served conditioned air (and thus not a functional requirement).

We further narrowed down the list of functional requirement fields by considering a combination of data quantity, data quality, and relationship to other fields. For example: We excluded vivarium area because despite nearly all buildings having data in this field, nearly all buildings had the same value (zero). We excluded total fume hood length from consideration because only 42% of buildings have data and because it is closely related to the number of ducted fume hoods (for which 73% of buildings have data). We excluded location-related fields (climate zone, latitude and longitude, etc.) and the year of energy measurement because despite high data availability, their effect on energy consumption is more directly explained by HDD and CDD. The resulting list of 14 candidate fields is as follows:

- Organization Type (e.g., academic, government, pharmaceutical)
- Predominant Lab Type (e.g., basic research, teaching, manufacturing)
- Predominant Lab Use (e.g., chemical, biological, physical)
- Year Built
- Total Occupant Density
- Occupied Hours / Week
- Ducted Fume Hood Density (i.e., number of hoods per sqft of lab area)

- Laboratory Occupied Minimum Air Change Rate
- Laboratory Area Ratio (i.e., proportion of gross floor area that is lab space)
- Biological Lab Area (i.e., proportion of lab space that is a biological lab)
- Chemistry Lab Area
- Physics/Engineering Lab Area
- HDD
- CDD

Next, we investigated correlations between each of the remaining candidate fields (individually) and the performance metric. Namely, for each field, we fit a linear regression model with that candidate field as the only predictor of source EUI and checked for statistical significance of the model coefficients (with a p-value threshold of 0.05). For numerical fields, we experimented with different transformations (e.g., logarithmic, piecewise linear), scaling and shifting, and with treating the numerical field as a categorical field by grouping the values into ranges (e.g., lab area < 0.2, 0.2-0.4, 0.4-0.6, etc.). For categorical fields, we experimented with several potential groupings of values when fitting the model. For example, Figure 3 shows a boxplot of source EUI for each of the lab types in the dataset. We experimented with combining different lab types into groups based on what makes physical sense and based on the similarity of their source EUI distributions. In general, categorical fields had to be grouped into relatively broad categories (e.g., roughly 3 or 4 values) before the coefficients of the model became statistically significant. These checks for individual significance narrowed the list of candidate fields down to the following 9 fields:

- Organization Type (regrouped from 13 values to 3: Academic, Government, and Other)
- Lab Type (regrouped from 9 values to 3: Manufacturing, Teaching, and Other)
- Lab Use (regrouped from 8 values to 2: Bio/Chem and Other)
- Occupied Hours / Week
- Occupied Minimum Air Change Rate
- Lab Area Ratio
- Biological lab area
- Physics/Engineering lab area
- CDD

When regrouping categorical fields, the "Other" value represents all values other than those listed, including buildings that are missing data, buildings with an unlisted value (i.e., buildings whose value is not among the list of options), and buildings whose value is a combination of values. For organization type, the "Other" category contains primarily corporate labs, but also some healthcare, unknown, unlisted, and combined types. For lab type, the "Other" category is predominantly research and development. For lab use, the "Other" category is mostly physics/engineering labs, but also some unlisted and combined types.



Figure 3: Boxplot of source EUI (kBtu/sqft) for each lab type, with the number of buildings of each type shown in parentheses.

With the short list of 9 candidate fields, each of which shows a statistically significant relationship to the performance metric, we next experimented with fitting linear regression models with various combinations of fields as predictors of source EUI. In some cases, multiple fields showed multicollinearity when included together in the same model (e.g., the primary lab use type is highly correlated with the proportion of lab area that is a biological lab), so we excluded that combination of fields from consideration as the model. We searched for the combination of fields that included as many fields as possible while avoiding multicollinearity, and arrived at a linear regression model that predicts source EUI (kBtu/sqft) with the following coefficients:

- Intercept: 290.5 kBtu/sqft
- Occupied Hours: 0.4473 (kBtu/sqft) / (hours/week)
- Lab Area Ratio: 2.979 (kBtu/sqft) / %
- CDD: 42.78 (kBtu/sqft) / (1000 degree-days)
- Lab Type = Manufacturing: +138.4 kBtu/sqft
- Lab Type = Teaching: -83.04 kBtu/sqft
- Lab Use = Bio/Chem: +74.50 kBtu/sqft

As an example calculation, consider a hypothetical laboratory building that is occupied 80 hours/week, has 40% of its gross floor used as a teaching lab that is not for biological/chemical research, and is operating in a location and during a year with 5000 CDD. For this building, the regression model predicts a source EUI of 290.5 +  $(0.4473 \times 80) + (2.979 \times 40) + (42.78 \times (5000/1000)) - 83.04 = 576.3 \text{ kBtu/sqft.}$ 

This model was fit to the 748 buildings in the dataset that had data available for all the numerical fields in the model. All model coefficients are statistically significant with p-value <= 0.028. The model explains only 15% of the variation in source EUI, but explains 87% of the variation in source energy.

Note that this regression model contains the combination of functional requirement fields that best predicts the source EUIs in the LBT dataset. While the LBT is the largest known collection of laboratory data, there is no guarantee it is representative of all laboratory buildings in existence. Similarly, it may be representative of labs of some types, or in some locations, etc., but might not contain data from enough labs of other types or in other locations. There are many reasons why fields that intuitively seem like they have an impact on source EUI might not be included in the model. For example, if many of the buildings in the dataset are missing data for a particular field, or if all buildings have data but almost all of them have the same value, that field might not show a statistically significant relationship to source EUI. This does not mean that field does not affect source EUI, it just means the model cannot discern its effect. Similarly, multiple fields can exhibit multicollinearity that makes the model unable to distinguish their effects. For example, hypothetically, if all the labs in one location were of one use type and all the labs in another location were of another use type (i.e., location and use type are highly correlated), then the model might be able to quantify the combined effect of both fields, but unable to separate the effect of location from the effect of use type. The LBT contains many fields that serve as proxies of other fields that more directly affect source EUI. For example, in reality, energy consumption is driven by the particular equipment used in a lab, but the lab's use type typically indicates which equipment is used, so the model might detect the effect of the proxy, but not the underlying fields. This does not imply those underlying fields have no effect.

# 4. Computing a Score

Using the regression model developed in Section 3, we followed the EPA's methodology [1] for computing a score. For each building in the dataset, we computed its EUI ratio as the measured source EUI divided by the source EUI predicted by the regression model using that building's fields as input to the model. The resulting ratios represent the proportion of model-predicted EUI that the building actually used (i.e., a ratio of 0.75 means the building used 75% as much energy as the model predicts for a building with the same occupancy hours, lab area ratio, CDD, etc.). We fit a gamma distribution to these ratios (see Figure 4), then used the fitted gamma distribution to generate a lookup table (see Table 1) that maps each range of EUI ratios to the

corresponding energy score. The score represents the percentage of buildings performing worse than a given building (i.e., a score of 100 indicates highest performance and a score of 1 indicates lowest performance). For example, consider the example building from Section 3 with a model-predicted source EUI of 576.3 kBtu/sqft, and assume that this building actually used 500 kBtu/sqft. The EUI ratio is computed as 500 / 576.3 = 0.8676. According to Table 1, this ratio corresponds to a score of 58.



Figure 4: Cumulative distribution function for EUI ratio. The blue circles represent the ratios computed from the dataset. The red line represents the gamma distribution fitted to the computed ratios.

Score	EUI Ratio Min	EUI Ratio Max
100	0.0000	0.2876
99	0.2876	0.3365
98	0.3365	0.3708
97	0.3708	0.3984
96	0.3984	0.4220
95	0.4220	0.4429
94	0.4429	0.4618
93	0.4618	0.4794

92	0.4794	0.4957
91	0.4957	0.5111
90	0.5111	0.5258
89	0.5258	0.5398
88	0.5398	0.5532
87	0.5532	0.5662
86	0.5662	0.5788
85	0.5788	0.5910
84	0.5910	0.6029
83	0.6029	0.6145
82	0.6145	0.6259
81	0.6259	0.6371
80	0.6371	0.6481
79	0.6481	0.6589
78	0.6589	0.6695
77	0.6695	0.6800
76	0.6800	0.6904
75	0.6904	0.7007
74	0.7007	0.7108
73	0.7108	0.7209
72	0.7209	0.7310
71	0.7310	0.7409
70	0.7409	0.7508
69	0.7508	0.7607
68	0.7607	0.7705
67	0.7705	0.7803
66	0.7803	0.7900
65	0.7900	0.7998
64	0.7998	0.8095
63	0.8095	0.8193
62	0.8193	0.8290
61	0.8290	0.8387
60	0.8387	0.8485
59	0.8485	0.8583
58	0.8583	0.8681
57	0.8681	0.8779
56	0.8779	0.8878
55	0.8878	0.8977
54	0.8977	0.9077
53	0.9077	0.9177
52	0.9177	0.9278
51	0.9278	0.9380

50	0.9380	0.9483
49	0.9483	0.9586
48	0.9586	0.9690
47	0.9690	0.9795
46	0.9795	0.9902
45	0.9902	1.0009
44	1.0009	1.0117
43	1.0117	1.0227
42	1.0227	1.0339
41	1.0339	1.0451
40	1.0451	1.0566
39	1.0566	1.0682
38	1.0682	1.0800
37	1.0800	1.0919
36	1.0919	1.1041
35	1.1041	1.1165
34	1.1165	1.1292
33	1.1292	1.1420
32	1.1420	1.1552
31	1.1552	1.1687
30	1.1687	1.1824
29	1.1824	1.1965
28	1.1965	1.2110
27	1.2110	1.2258
26	1.2258	1.2411
25	1.2411	1.2569
24	1.2569	1.2731
23	1.2731	1.2899
22	1.2899	1.3073
21	1.3073	1.3253
20	1.3253	1.3440
19	1.3440	1.3636
18	1.3636	1.3840
17	1.3840	1.4055
16	1.4055	1.4281
15	1.4281	1.4519
14	1.4519	1.4772
13	1.4772	1.5043
12	1.5043	1.5333
11	1.5333	1.5647
10	1.5647	1.5989
9	1.5989	1.6366

8	1.6366	1.6788
7	1.6788	1.7267
6	1.7267	1.7824
5	1.7824	1.8493
4	1.8493	1.9337
3	1.9337	2.0497
2	2.0497	2.2412
1	2.2412	inf

Table 1: Lookup table mapping each range of EUI ratios to the corresponding energy score.

#### 5. Next Steps

We are confident in the utility of this regression model and scoring system, but acknowledge that further scrutiny and refinement may be needed to achieve stakeholder buy-in and widespread adoption and use of the score. The next phase of our analysis will include additional review of the score computed for each of the buildings in the LBT dataset. We will check for indications that the score is treating any particular types of labs unfairly (e.g., whether labs of particular types or in particular locations tend to score abnormally higher or lower than their peers). If we identify any characteristics of labs that tend to result in inconsistent or non-intuitive scores, we will consider re-developing the model with those labs excluded from the dataset. For example, there may be cases where the dataset does not include sufficient coverage of certain lab types, uses, locations, etc., and we may conclude the score should not be used for those labs. If we identify areas of the database with especially low coverage, we may recommend further data collection that targets those areas.

We will also consider how our scoring system may be improved through the use of a physicsbased simulation model of a laboratory building (we have confirmed that an EnergyPlus model for laboratories is freely available from the Department of Energy). One potential use of a simulation model could be to use it to estimate the effect of particular fields that are included in the regression model (e.g., occupancy hours) on source EUI, then compare the magnitude of the resulting effect to the corresponding coefficient in our regression model (i.e., using the simulation model to independently derive a coefficient and comparing it to the coefficient learned from the LBT dataset). Another potential use could be to use the simulation model to derive coefficients for fields that were not included in the regression model but that intuitively should perhaps be included in the model (e.g., HDD), then augment the regression model with the coefficient derived from the simulation model.

Lastly, we will collect data from a handful of pilot laboratory sites that have additional data available than what is available in the LBT database. We will compute scores for those pilot labs then compare the score to the expected level of performance of that laboratory based on the

more-detailed data (i.e., we will check our computed scores for some pilot sites known to be poor- or well-performing and make sure the score is consistent).

# 6. References

[1] U.S. Environmental Protection Agency (EPA). "ENERGY STAR Score Technical Reference". April 2021.

https://portfoliomanager.energystar.gov/pdf/reference/ENERGY%20STAR%20Score.pdf.

[2] International Institute for Sustainable Laboratories. "Laboratory Benchmarking Tool". 2023. https://lbt.i2sl.org/.

[3] International Institute for Sustainable Laboratories. "Laboratory Benchmarking Tool: Data Fields". March 2023. https://lbt.i2sl.org/files/List%20of%20LBT%20Data%20Fields.pdf.

[4] BizEE Software. "Degree Days: Weather for Energy Saving". 2023. https://degreedays.net.

# 7. Appendix: Pilot Building Results

For initial testing of the score, two federal laboratory buildings were selected to evaluate their energy performance utilizing the above-described score methodology. These facilities were selected to provide insights on how well the score itself performed relative to facilities for which there was familiarity and a strong understanding of their internal systems and operations. Because of this consideration, it was determined that one building each from LBNL's Berkeley campus and NREL's Golden campus would serve as a strong starting point for piloting the score, considering that LBNL and NREL are both intimately involved with research around laboratory energy performance. Further criteria to select specific buildings on each campus were as follows:

- Reliable building-level meter data
- LBT required fields known
- Recent energy audit or energy projects (planned or completed)
- Confidence in lab space breakdown by type

First Test Building – LBNL's Building 84

Building 84 houses office and lab space for the Biosciences Area and the Earth and Environmental Sciences Area for Berkeley Lab. Genomic scientists work to better understand complex sequence motifs that control RNA transcription, DNA replication, and chromosome structure. On the Earth and Environmental Science side, climate scientists work on novel forms of climate modeling and pioneering work on carbon cycles. Built in 1997, the building still has original controls with limited capabilities, constant, 100% outside-air volume distribution and



reheat/cooling coils within spaces for temperature control. Lighting is primarily T8 with some LED retrofits throughout.



The EUI regression resulted in a predicted EUI of 472.2 kBtu/sqft. Percent lab area, Bio/Chem classification, and overall occupancy hours were the biggest contributors to the predicted EUI.



Since several years of data were readily available, the score was calculated over several years, including pre- and post-pandemic operations, as shown on the above chart. The score itself varied from 36 in 2017 to a peak of 66 during the pandemic.



Adjusting for occupancy hours for the pandemic period resulted in a slight leveling of the score, but not sufficient to fully normalize for extreme changes in occupancy present during the pandemic.

Due to available audit data, we were also able to model out the score assuming implementation of ECM and capital upgrades available in their most recent energy audit report. These hypothetical scores are shown in the graphic below:

	Measure	Description	Peak kW	kWh/yr	Therms/yr	Savings (\$/yt)	Installed Cost (\$)	Incentive Cost (5)	Net Measure Cost(\$)	Pay- Back (Yrs)
	2015- EEM-01	884 - Lighting - M - Upgrade Lighting Controls, Incl Emergency Lighting, Occupancy Sensing, and Zoning Controls	0.0	103,000	0	\$ 6,000	5 81,000	\$ 0	\$ 81,000	13.5
	2015- EEM-02	BS4 - Project - L - Integrate Plug Load with Advanced Lighting Controls	0.0	101,000	0	\$ 6,000	\$ 90,000	\$ 0	\$ 90,000	15.0
	2015- EEM-03	$\rm BB4$ - Project - L - Convert to Variable Air Volume Systems and Reduce Minimum Air Change Rates	0.0	150,000	19,000	<b>5</b> 21,000	5 1,100,000	\$ 19,000	\$ 1,081,000	51.5
If all ECMs implemented from	2015- EEM-04	884 - Project - L - Further Reduce Minimum ACR when Zone is Unoccupied	0.0	50,000	3,000	\$ 5,000	5 58,000	\$ 3,000	\$ 55,000	11.0
2018 Audit:	2015- EEM-05	B84 - OCx - M - Implement Automated Condenser Water Temperature Reset	0.0	5,000	0	\$ 400	\$ 4,000	\$ 0	\$ 4,000	10.0
	2019- EEM-01	994 - OCx - H - Decrease airflow to glass wash room 102 and disable conditioning	0.0	8,000	300	\$ 0	50	\$0	50	0.0
~36% Gas Reduction	2017- 12.01	894 - Main AHU and zone-level controls issues	0.0	U	0	\$0	50	\$ 0	\$ O	0.0
	2019- EEM-02	B84 - OCx - H - Decrease airflow to rooms 169 and 371	0.0	10,000	600	5.0	50	\$ 0	50	0.0
Score jumps from <b>40</b> to <b>62</b>	2019- EEM-03	BB4 - OCx - H - Rebalance all airflows and reduce fixed DSP sepoint	0.0	50,000	3,000	50	50	\$ 0	5.0	0.0
	2019 EEM-04	BS4 - Project - H - Convert existing system to DOAS and install zone FCUs or chilled beams	0.0	140,000	31,000	\$0	\$0	\$ 0	\$ 0	0.0
	2019- EEM-05	004 - Lighting - II - Replace existing compact florescent and linear fluorescent light foctures in the hallway with LEDs	2.5	22,110	0	50	50	\$ 0	50	0.0
	2019- EEM-06	904 - Lighting - M - Replace existing recessed linear fluorescent focures with LEDs	4.8	24,318	0	\$ 0	50	\$ 0	5.0	0.0
	2019- EEM-07	B84 - Lighting - $M$ - Replacing exitting fluorescent flutures in offices with LEDs	63	16,255	0	\$0	\$0	\$ 0	50	0.0
	2019- EEM-08	$\rm B94$ - OCx - H - Clean outside air intake louvers for BL-17 & 18	0.0	5,000	0	\$ 0	\$0	\$0	5.0	0.0
		Total	15.5	687,685	36,900	\$ 58,400	\$ 1,355,000	\$ 22,000	\$ 1,311,000	34.1

#### Second Test Building – NREL's Field Test Laboratory Building (FTLB)

The Field Test Laboratory Building is an amalgam of 40+ smaller laboratories mostly within the categories of chemistry and biology, with a primary focus on alternative and emerging fuel technologies. Built in 1982, the building systems are comprised of primary air handlers feeding fan coil units within spaces, with chilled water and hot water cooling and heating, respectively. Lighting is primarily T8 with some LED retrofits throughout.





The EUI regression resulted in a predicted EUI of 519.2 kBtu/sqft. Percent lab area, CDD, and Bio/Chem classification were the biggest contributors to the predicted EUI. This yielded a pilot energy score of 55 for FTLB.

Due to available audit data, we were also able to model out the score assuming implementation of ECM and capital upgrades available in their most recent energy audit report. These hypothetical scores are shown in the graphic below:



Overall, the scores derived for these pilot facilities were in line with the expectations of facility energy management staff. Caveats may be introduced for facilities with low occupancy hours (especially for the atypical occupancy of the pandemic period) as a result of the testing at Building 84.