Updating the Energy Score for Laboratories

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

May 1, 2025

1. Overview

This memo describes an update to a previously developed scoring system for laboratory buildings [1] using newly available data. Both the original and updated scoring systems were developed using a similar methodology to the one used by the U.S. Environmental Protection Agency (EPA) to develop the ENERGY STAR Score [2], and the primary data source was in both cases the dataset underlying the Institute for Sustainable Laboratories (I2SL) Laboratory Benchmarking Tool (LBT) [3]. Since the score was originally developed, significant efforts have been made to expand the LBT dataset through additional data collection. With the newly available data, we have re-evaluated the regression model underlying the scoring system, and have updated the score computations using the new model.

Section 2 of this memo describes the data used for the analysis. Section 3 discusses development of the updated regression model relating laboratory characteristics to energy performance. Section 4 details the use of the model's prediction for computing a score. Lastly, Section 5 discusses future development and refinement of the scoring system.

This work was primarily carried out by Lawrence Berkeley National Laboratory (LBNL) and the International Institute for Sustainable Laboratories (I2SL), but we received valuable feedback and advice from several stakeholders and members of the I2SL Energy Score Technical Advisory Committee (TAC).

2. Data Sources

The primary source of data for this analysis was the dataset underlying the International Institute for Sustainable Laboratories (I2SL) Laboratory Benchmarking Tool (LBT) [3], which is the largest known collection of energy-related information on laboratory buildings. For each building, the dataset contains over 100 data fields describing various characteristics related to the laboratory's size, location, usage patterns, installed systems, and energy consumption. See the full list of LBT data fields and their descriptions [4] for more information. We utilized a subset of the data that has been through cursory quality checking by LBNL after being entered by users of the tool. This check removes any data 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 most buildings).

As in the original analysis, we augmented the LBT dataset with weather data from Degree Days.net [5], 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. In cases where the year's end month was not available, it was assumed that the user-entered year corresponded to the calendar year.

Since the score was originally developed, I2SL has made significant data outreach efforts to expand the LBT dataset, with the goal of increasing the amount of recent data and of specifically targeting some types of lab buildings that were previously under-represented. Based on discussions with the I2SL Energy Score TAC, the type-specific outreach focused on laboratory buildings with large amount of vivarium space, on manufacturing facilities, and on facilities in cold-weather locations (ASHRAE Climate Zones 6 and higher). The outreach resulted in 323 buildings being added to the LBT dataset, including a significant increase in the targeted categories (6 more vivarium-dominated buildings, 3 more manufacturing facilities, and 14 more cold weather facilities).

In this update to the scoring system, we also revisited our criteria for the buildings to include in the score analysis. We implemented more stringent requirements than previously, with the goal of ensuring that the dataset more accurately reflects the general population of current laboratory buildings. We started with data for 1327 labs that had already passed cursory quality checking, and excluded additional data as follows:

- 511 records with data older than 2014
- 46 records with <15% or >95% of floor area reported to be laboratory space
- 22 records for campuses (rather than individual buildings)
- 77 records with estimated values for laboratory space area
- 121 records with estimated or missing energy data (including 59 embodied carbon-only records)
- 6 records based on other miscellaneous criteria

The resulting dataset consists of 544 laboratory buildings that we used for the remainder of the analysis.

3. Fitting the Model

As in the previous score development, we constructed a linear regression model that estimates the typical energy consumption for a laboratory with a given set of characteristics, assuming that the LBT dataset can serve as a representative sample of the wider population of laboratory buildings for which score will be computed. Without a comprehensive database of all laboratory buildings, we cannot confirm this assumption of representativeness, but we are confident that the LBT dataset can be used to derive a useful and trustworthy model of energy consumption. As with the previous scoring system, we selected source energy use intensity (EUI) as the performance metric upon which buildings will be scored.

We re-evaluated the use of each of the fields in the LBT dataset as potential predictor variables in the regression model (i.e., characteristics that will be used to estimate the expected source EUI for a given laboratory). With significant input from the Energy Score TAC, we identified 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). 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 fields if all (or nearly all) buildings have the same value of the field, and we excluded fields that are highly correlated with other fields. 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)
- Occupied Hours / Week
- Ducted Fume Hood Density (i.e., number of hoods per sqft of lab area)
- 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
- Vivarium Lab Area
- Dry Lab Area
- Other Lab Area
- Heating degree days (HDD)
- Cooling degree days (CDD)

We started by investigating correlations between each of the 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. 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. For categorical fields, we experimented with several potential groupings of values when fitting the model. After identifying fields that individually showed relationships to source EUI, we 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: 227 kBtu/sqft
- Ducted Fume Hood Density [number of hoods per unit area of lab space]: 35.0 (kBtu/sqft) / (number of hoods / 1000 sqft)
- Occupied Hours: 0.447 (kBtu/sqft) / (hours/week)
- Lab Area Ratio [fraction of gross area that is lab space]: 280 (kBtu/sqft) / 100%
- HDD: 12.0 (kBtu/sqft) / (1000 degree-days)
- CDD: 54.4 (kBtu/sqft) / (1000 degree-days)
- Lab Type = Manufacturing: +107 kBtu/sqft
- Lab Type = Teaching: -127 kBtu/sqft
- Vivarium Lab Area [fraction of lab space that is vivarium]: 165 (kBtu/sqft) / 100%

As an example calculation, consider a hypothetical teaching laboratory building with 0.3 ducted fume hoods per 1000 sqft of lab space, occupancy for 80 hours/week, net lab space amounting to 40% of its gross area, where 10% of the lab space is vivarium space, and weather conditions corresponding to 5000 HDD and 900 CDD for the year in question. For this building, the regression model predicts a source EUI of 227 + (35.0 x 0.3) + (0.447 x 80) + (280 x 40/100) + (12.0 x 5000/1000) + (54.4 x 900/1000) + (165 x 10/100) = 510.72 kBtu/sqft.

The above model was fit to the 544 buildings in the dataset. All model coefficients are statistically significant with p-value <= 0.069. The model explains only 11% of the variation in source EUI, but explains 80% 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. 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. The absence of a particular field from this model does not imply that the field has no effect on source EUI in a general laboratory building.

Compared to the previous version of the regression model, the new model has several similarities, but also some notable differences. The coefficients for the intercept, occupied hours, lab area ratio, CDD, and lab type are all in the same direction as before and have comparable magnitudes. There are two key differences between the previous and the new models:

- The previous model included a coefficient for biological and chemical laboratories, but the new model instead has a coefficient for vivarium area and ducted fume hood density.

- The previous model did not include a coefficient for HDD, but the new model does.

These key changes to the model represent relationships between predictors and source EUI that received significant consideration by the Energy Score TAC during previous score development. The absence of these predictors from the previous model was considered non-intuitive, and those concerns led to the concerted effort to collect more data. We believe the additional data collected (and the more stringent data quality standards) resulted in a more robust regression model, and therefore a more trustworthy scoring system.

4. Computing a Score

Using the regression model developed in Section 3, we followed the EPA's methodology [2] 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 characteristics). We fit a gamma distribution to these ratios, 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 510.72 kBtu/sqft, and assume that this building actually used 490 kBtu/sqft. The EUI ratio is computed as 490 / 510.72 = 0.9594. According to Table 1, this ratio corresponds to a score of 48.

Score	EUI Ratio Min	EUI Ratio Max
100	0.0000	0.3411
99	0.3411	0.3830
98	0.3830	0.4128
97	0.4128	0.4370
96	0.4370	0.4579
95	0.4579	0.4765
94	0.4765	0.4934
93	0.4934	0.5091
92	0.5091	0.5238
91	0.5238	0.5378
90	0.5378	0.5510
89	0.5510	0.5638
88	0.5638	0.5760

87	0.5760	0.5879
86	0.5879	0.5994
85	0.5994	0.6106
84	0.6106	0.6215
83	0.6215	0.6322
82	0.6322	0.6427
81	0.6427	0.6530
80	0.6530	0.6631
79	0.6631	0.6731
78	0.6731	0.6830
77	0.6830	0.6927
76	0.6927	0.7024
75	0.7024	0.7119
74	0.7119	0.7214
73	0.7214	0.7308
72	0.7308	0.7402
71	0.7402	0.7495
70	0.7495	0.7588
69	0.7588	0.7680
68	0.7680	0.7772
67	0.7772	0.7864
66	0.7864	0.7956
65	0.7956	0.8047
64	0.8047	0.8139
63	0.8139	0.8231
62	0.8231	0.8323
61	0.8323	0.8415
60	0.8415	0.8507
59	0.8507	0.8599
58	0.8599	0.8692
57	0.8692	0.8785
56	0.8785	0.8879
55	0.8879	0.8973
54	0.8973	0.9068
53	0.9068	0.9163
52	0.9163	0.9260
51	0.9260	0.9356
50	0.9356	0.9454
49	0.9454	0.9552
48	0.9552	0.9652
47	0.9652	0.9752
46	0.9752	0.9854

45	0.9854	0.9957
44	0.9957	1.0060
43	1.0060	1.0166
42	1.0166	1.0272
41	1.0272	1.0380
40	1.0380	1.0490
39	1.0490	1.0602
38	1.0602	1.0715
37	1.0715	1.0830
36	1.0830	1.0947
35	1.0947	1.1067
34	1.1067	1.1189
33	1.1189	1.1313
32	1.1313	1.1440
31	1.1440	1.1570
30	1.1570	1.1703
29	1.1703	1.1840
28	1.1840	1.1980
27	1.1980	1.2124
26	1.2124	1.2272
25	1.2272	1.2424
24	1.2424	1.2582
23	1.2582	1.2745
22	1.2745	1.2914
21	1.2914	1.3090
20	1.3090	1.3272
19	1.3272	1.3463
18	1.3463	1.3663
17	1.3663	1.3872
16	1.3872	1.4093
15	1.4093	1.4326
14	1.4326	1.4574
13	1.4574	1.4839
12	1.4839	1.5124
11	1.5124	1.5432
10	1.5432	1.5769
9	1.5769	1.6140
8	1.6140	1.6556
7	1.6556	1.7028
6	1.7028	1.7579
5	1.7579	1.8241
4	1.8241	1.9079

3	1.9079	2.0232
2	2.0232	2.2143
1	2.2143	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, including differences between the scores produced by the original and new scoring systems. 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.

6. References

[1] T. Walter, J. Kace, and A. Farmer. "Development of an Energy Score for Laboratories". August 2023. https://lbt.i2sl.org/files/Development%20of%20an%20Energy%20Score%20for%20Laboratories.docx.pdf.

[2] U.S. Environmental Protection Agency (EPA). "ENERGY STAR Score Technical Reference". April 2021. https://portfoliomanager.energystar.gov/pdf/reference/ENERGY%20STAR%20Score.pdf.

[3] International Institute for Sustainable Laboratories. "Laboratory Benchmarking Tool". 2025. https://lbt.i2sl.org/.

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

[5] BizEE Software. "Degree Days: Weather for Energy Saving". 2025. https://degreedays.net.