ENERGY STAR® Performance Ratings
Technical Methodology for Data Center

This document presents specific details on the EPA's analytical result and rating methodology for Data Center. For background on the technical approach to development of the energy performance ratings, refer to *Energy Performance Ratings – Technical Methodology* (http://www.energystar.gov/ia/business/evaluate_performance/General_Overview_tech_methodology.pdf).

Model Release Date

June 2010

Portfolio Manager Definition

Data Center applies to spaces specifically designed and equipped to meet the needs of high density computing equipment such as server racks, used for data storage and processing. Typically these facilities require dedicated uninterruptible power supplies and cooling systems. Data Center functions may include traditional enterprise services, on-demand enterprise services, high performance computing, internet facilities, and/or hosting facilities. Often Data Centers are free standing, mission critical computing centers. When a data center is located within a larger building, it will usually have its own power and cooling systems. The Data Center space is intended for sophisticated computing and server functions; it should not be used to represent a server closet or computer training area.

Reference Data

The Data Center regression model is based on survey data collected by EPA. EPA relies on publicly available external data sets to develop rating models where feasible, but a sufficiently robust set of data center energy consumption information was not available. In its effort to collect survey data, EPA coordinated with major industry associations, including Uptime Institute, Green Grid, 7x24 Exchange, and AFCOM, to inform their members and encourage participation.

EPA collected data from stand alone data center facilities as well as those enclosed within larger buildings. In addition to collecting energy consumption data, EPA consulted with industry associations and partners to determine the specific operating parameters that were likely to influence energy consumption, and developed a list of operating characteristics that were requested from survey participants.

Data Filters

Four types of filters are applied to define the peer group for comparison and to overcome any technical limitations in the data: Building Type Filters, EPA Program Filters, Data Limitation Filters, and Analytical Filters. A complete description of each of these categories is provided in Section V of the general technical description document: *Energy Performance Ratings –*

Technical Methodology. **Table 1** presents a summary of each filter applied in the development of the Data Center model, the rationale behind the filter, and the resulting number of observations after the filter is applied. After all filters are applied, the remaining data set has 61 observations.

Table 1
Summary of Data Center Model Filters

Condition for Including an Observation in the Analysis	Rationale	Number Remaining
Must have complete data for energy use and operating characteristics	EPA Program Filter - Complete data is necessary for analysis.	120
Must provide IT Energy measured at the output of the UPS Meter	EPA Program Filter – In order to develop an equitable comparison between facilities, all IT Energy consumption must be measured at the same location.	108
Must be a Stand Alone data center	Analytical Filter – Data for Stand Alone data centers was more robust and resulted in higher significance for regression models. Using stand alone facilities is also more consistent with the process used by EPA for other space types.	61

Dependent Variable

The dependent variable in the Data Center analysis is Power Usage Effectiveness, or PUE. This variable is unique to the Data Center space type. Source energy use intensity (source EUI) is used for the majority of EPA's rating models, but was not considered to be the best choice for a Data Center model. Source EUI can vary widely for data center facilities, and does not take into consideration the varying densities of IT equipment that can be present in these facilities.

EPA consulted with industry associations and data center operators to identify an appropriate metric to evaluate energy use in data center facilities. The dependent variable of PUE is defined as:

$$\text{PUE} = \text{Total Energy} / \text{IT Energy}$$

where both Total Energy and IT Energy are expressed in Source kBtu.

Total Energy includes the annual energy consumption for all fuels at the Data Center. In many cases, the only energy consumption at Data Centers is electricity. However, it is important to capture any other fuel use (e.g. chilled water, natural gas), in order to evaluate the total energy performance of the facility. This practice is consistent with all EPA rating models.

IT Energy is defined as the total amount of energy required by the server racks, storage silos, and other IT equipment in the Data Center. For the purposes of ENERGY STAR, this should be measured at the output of the Uninterruptible Power Supply (UPS).

EPA considered alternate locations for measuring IT Energy consumption, and requested data from both the UPS and PDU meters from survey participants. Measurements at the PDU meter or closer to the racks can provide a more accurate representation of IT Energy. However, these measurements are still not commonplace in the industry, and were not provided by a large number of survey participants. EPA prefers a common metric that can be used by the majority of Data Center operators.

By setting PUE as the dependent variable, the regressions analyze the key drivers of PUE – those factors that explain the variation in Power Usage Effectiveness in Data Centers.

Independent Variables

The EPA survey contained numerous building operation questions that were identified as potentially important for Data Centers. Based on a review of the available variables in the data, in accordance with the EPA criteria for inclusion[1], EPA analyzed the following variables:

- Building Square Footage
- Data Center Square Footage
- Tier Level (four levels denoting increasing equipment redundant capacity)
- Number of racks
- UPS Utilization
- Annual IT Energy
- Building Type (Stand alone data center vs. Enclosed in another building)
- Data Center type (options included Hosting, Hybrid, Internet, Traditional, and Telecom)
- HDD
- CDD

EPA performed extensive review on all of these operational characteristics. In addition to reviewing each characteristic individually, characteristics were reviewed in combination with each other (e.g., IT Energy / Square Foot). Based on analytical results and residual plots, variables were examined using different transformations (such as the natural logarithm). The analysis consisted of multiple regression formulations. These analyses were structured to find the combination of statistically significant operating characteristics that explained the greatest amount of variance in the dependent variable: PUE.

Based on the Data Center regression analysis, the following characteristic was identified as the key explanatory variable that can be used to estimate the expected average PUE in Data Centers:

- Annual IT Energy

[1] For a complete explanation of these criteria, refer to *Energy Performance Ratings – Technical Methodology* (http://www.energystar.gov/ia/business/evaluate_performance/General_Overview_tech_methodology.pdf).

IT Energy Analysis
The Annual IT Energy variable warrants additional discussion, because it is unlike most operating characteristics included in EPA's rating models. Most variables can be entered as a single value that remains relatively constant throughout the year (i.e. Square Foot, Hours of Operation). IT Energy, on the other hand, must be metered on a regular basis. Users can enter monthly entries in Portfolio Manager. The total use over a period of one year is calculated and used in the regression equation.

The regression analysis shows that facilities with higher IT Energy loads have lower PUE values on average. This phenomenon can be understood as an economy of scale: the larger, more intensive data centers can have more opportunities for efficiencies than their smaller counterparts. Similar behavior has been observed in other types of commercial buildings, where larger buildings may use less energy per square foot. The relationship between PUE and Annual IT Energy was only observed up to a certain Annual IT Energy value, after which the average PUE values level off (i.e. there are no longer economies of scale beyond a certain size). Therefore, the adjustment of IT Energy within the model is applied over that range, and capped at a maximum adjustment at the value of 0.4 Source TBtu. That is, Data Centers with IT energy values of greater than 0.4 TBtu will receive the same regression adjustment as Data Centers with IT energy equal to 0.4 TBtu.

Model Testing and Variables not Correlated with PUE
The Data Center rating model includes fewer operating characteristics than most EPA rating models. It was determined that the Annual IT Energy is the primary factor influencing PUE, and that other operating characteristics do not show any statistically significant correlation with PUE.

Climate is one characteristic that was examined closely. EPA found no statistically significant relationship between heating and cooling degree days and PUE, which was initially considered to be surprising. However, upon further review, it was determined that the energy required for cooling a data center is dominated by the high internal loads generated by the IT equipment, and that climate has a relatively low contribution to the building cooling load.

EPA also examined Tier level, a measure of redundancy of equipment capacity, and Data Center type, with options that included Hosting, Hybrid, Internet, Traditional, and Telecom. These characteristics were both excluded from the final Data Center model. The dependence of PUE on both variables was not shown to be statistically significant. Additionally, the variables were determined to be hard to define. Operators reported that there could be multiple Tiers or Data Center types within one facility. For Tier, it was also determined that data centers may have unnecessarily high Tier levels, and normalization for Tier could provide a disincentive for efficient design.

Regression Modeling Results

The final regression is a weighted ordinary least squares regression across the data set of 61 observations. The dependent variable is PUE. The independent variable is centered relative to the mean value, presented in **Table 2**. The final model is presented in **Table 3**. The model variable is significant at the 95% confidence level or better, as shown by the significance level (a p-level of less than 0.05 indicates 95% confidence).

The model has an adjusted R^2 value of 0.0988, indicating that this model explains 9.88% of the variance in PUE for Data Center facilities. Because the final model is structured with PUE as the dependent variable, the explanatory power of IT Energy is not included in the R^2 value, thus this value appears artificially low. Re-computing the R^2 value in units of source energy[2], demonstrates that the model actually explains 92.2% of the variation of source energy of Data Centers. This is an excellent result for a statistically based energy model.

Detailed information on the ordinary least squares regression approach, the methodology for performing weather adjustments, and the independent variable centering technique is available in the technical document: *Energy Performance Ratings – Technical Methodology*.

Table 2
Descriptive Statistics for Variables in Final Regression Model

Variable	Full Name	Mean	Minimum	Maximum
PUE	Power Usage Effectiveness	1.924	1.362	3.598
Annual IT Energy	Annual IT Energy (in Source TBtu)	0.2091	0.0129	0.7204

Table 3
Final Regression Modeling Results

Dependent Variable		Power Usage Effectiveness (PUE)		
Number of Observations in Analysis		61		
Model Adjusted R^2 value		0.0988		
Model F Statistic		7.579		
Model Significance (p-level)		0.0078		
	Unstandardized Coefficients	**Standard Error**	**T value**	**Significance (p-level)**
(Constant)	1.924	0.0523	36.81	0.0000
C_Annual IT Energy	-0.9506	0.3453	-2.753	0.0078

Note:
- *The prefix C_ on Annual IT Energy indicates that the value is centered. The centered variable is equal to difference between the actual value and the observed mean. The observed mean Annual IT Energy is 0.2091 Source TBtu.*
- *Annual IT Energy is computed in Source Energy, and entered in Tera Btu (TBtu):*
  Annual IT Energy kWh(3.412kBtu/kWh)*(3.34 Source/Site Electric)*(1 TBtu/10^9kBtu)*
- *The Annual IT Energy adjustment is capped at 0.4 TBtu*

[2] The R^2 value in Source Energy is calculated as: 1 – (Residual Variation of Y) / (Total Variation of Y). The residual variation is sum of (Actual Source Energy_i – Predicted Source $\text{Energy}_i)^2$ across all observations. The Total variation of Y is the sum of (Actual Source Energy_i – Mean Source $\text{Energy})^2$ across all observations.

Data Center Lookup Table

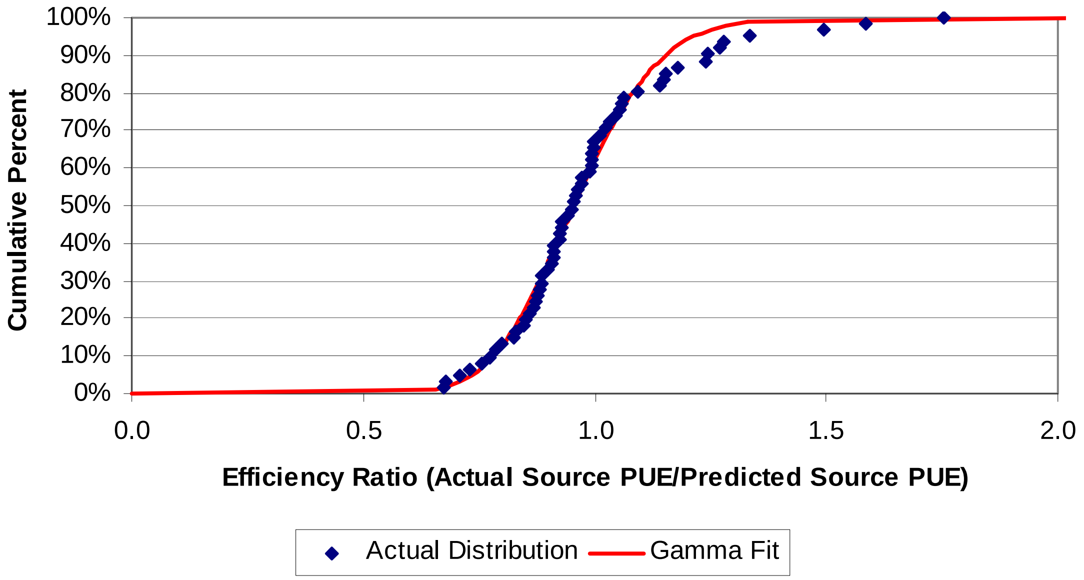
The final regression model (presented in **Table 3**) yields a prediction of PUE based on a building's Annual IT Energy. Some buildings in the data sample use more energy than predicted by the regression equation, while others use less. The *actual* PUE of each building observation is divided by its *predicted* PUE to calculate an energy efficiency ratio:

Energy Efficiency Ratio = Actual PUE/ Predicted PUE

A lower efficiency ratio indicates that a building uses less energy than predicted, and consequently is more efficient. A higher efficiency ratio indicates the opposite.

The efficiency ratios are sorted from smallest to largest and the cumulative percent of the population at each ratio is computed using the individual observation weights from the dataset. Figure 1 presents a plot of this cumulative distribution. A smooth curve (shown in red) is fitted to the data using a two parameter gamma distribution. The fit is performed in order to minimize the sum of squared differences between each building's actual percent rank in the population and each building's percent rank with the gamma solution. The final fit for the gamma curve yielded a shape parameter (alpha) of 43.93 and a scale parameter (beta) of 0.0219. For this fit, the sum of the squared error is 0.0555.

Figure 1:
Data Center Distribution



The final gamma shape and scale parameters are then used to calculate the efficiency ratio at each percentile (1 to 100) along the curve. For example, the ratio on the gamma curve at 1% corresponds to a rating of 99; only 1% of the population has a ratio this small or smaller. The ratio on the gamma curve at the value of 25% will correspond to the ratio for a rating of 75; only 25% of the population has ratios this small or smaller. The complete lookup table is presented at the end of the document. In order to read this lookup table, note that if the ratio is less than 0.6569 the rating for that building should be 100. If the ratio is greater than or equal to 0.6569 and less than 0.6879 the rating for the building should be 99, etc.

Example Calculation

As detailed in the document *Energy Performance Ratings – Technical Methodology*, there are five steps to compute a rating. The following is a specific example with the Data Center model:

Step 1 – User enters building data into Portfolio Manager

For the purposes of this example, sample data is provided.

- Energy data
  - Total annual electricity = 15,000,000 kWh
  - Total annual natural gas = 20,000 therms
  - Note that this data is actually entered in monthly meter entries
- Operational data
  - Annual IT Energy = 8,500,000 kWh

Step 2 – Portfolio Manager computes the Actual Power Usage Effectiveness

In order to compute the actual PUE, Portfolio Manager must convert each fuel from the specified units (e.g. kWh) into Site kBtu, and must convert from Site kBtu to Source kBtu.

- Convert the meter data entries into site kBtu
  - Electricity: (15,000,000 kWh)*(3.412kBtu/kWh) = 51,180,000 kBtu Site
  - Natural gas: (20,000 therms)*(100kBtu/therm) = 2,000,000 kBtu Site
- Apply the source-site ratios to compute the source energy
  - Electricity:
    51,180,000 SitekBtu*(3.34 Source kBtu/Site kBtu) = 170,941,200 kBtu Source
  - Natural Gas:
    2,000,000 Site kBtu*(1.047 Source kBtu/Site kBtu) = 2,094,000 kBtu Source
- Combine source kBtu across all fuels
  - 170,941,200 kBtu + 2,094,000 kBtu = 173,035,200 kBtu Source

Portfolio Manager must also convert Annual IT energy (kWh) into Site kBtu, and then from Site kBtu to Source kBtu.

- Convert the Annual IT Energy into site kBtu
  - (8,500,000 kWh)*(3.412kBtu/kWh) = 29,002,000 kBtu Site
- Apply the source-site ratios to compute the source energy
  - 29,002,000 Site kBtu *(3.34 Source kBtu/Site kBtu) = 96,866,680 kBtu Source

Then, Portfolio Manager calculates the actual PUE.

- Divide total source energy by Annual IT Energy to get the Actual PUE
  - PUE = 173,035,200 kBtu / 96,866,680 kBtu = 1.786

Step 3 – Portfolio Manager computes the Predicted Power Usage Effectiveness

Portfolio Manager uses the building data entered under Step 1 to compute centered values for each operating parameter. The centered value is entered into the Data Center regression equation to obtain a predicted PUE.

- Calculate centered variables
  - Subtract the reference centering IT Energy value from the measured value
    - Annual IT Energy = 96,866,680 kBtu /(10^9 TBtu/kBtu) = 0.0969 TBtu
    - Annual IT Energy - 0.2091 = 0.0969 - 0.2091 = -0.1122
- Compute predicted PUE using the regression equation:
  - PUE = 1.924 – 0.9506 * (Centered Source TBtu)
  - PUE = 1.924 – 0.9506 *(-0.112) = 2.031

Step 4 – Portfolio Manager computes the energy efficiency ratio

The energy efficiency ratio is equal to: Actual PUE/ Predicted PUE

- Ratio = 1.786/2.031 = 0.8795

Step 5 – Portfolio Manager looks up the efficiency ratio in the lookup table

Starting at 100 and working down, Portfolio Manager searches the lookup table for the first ratio value that is larger than the computed ratio for the building.

- A ratio of 0.8795 is greater than 0.8770 (requirement for 71) but less than 0.8810 (requirement for 70)
- ***The rating is 71***

Attachment

Table 4 lists the energy efficiency ratio cut-off point for each rating, from 1 to 100.

Table 4
Lookup Table for Data Center

Rating	Cumulative Percent	Energy Efficiency Ratio		Rating	Cumulative Percent	Energy Efficiency Ratio	
		> =	<			> =	<
100	0%	0.0000	0.6569	50	50%	0.9548	0.9584
99	1%	0.6569	0.6879	49	51%	0.9584	0.9620
98	2%	0.6879	0.7082	48	52%	0.9620	0.9657
97	3%	0.7082	0.7236	47	53%	0.9657	0.9694
96	4%	0.7236	0.7364	46	54%	0.9694	0.9731
95	5%	0.7364	0.7474	45	55%	0.9731	0.9768
94	6%	0.7474	0.7571	44	56%	0.9768	0.9805
93	7%	0.7571	0.7659	43	57%	0.9805	0.9843
92	8%	0.7659	0.7739	42	58%	0.9843	0.9880
91	9%	0.7739	0.7814	41	59%	0.9880	0.9919
90	10%	0.7814	0.7883	40	60%	0.9919	0.9957
89	11%	0.7883	0.7949	39	61%	0.9957	0.9996
88	12%	0.7949	0.8011	38	62%	0.9996	1.0035
87	13%	0.8011	0.8071	37	63%	1.0035	1.0075
86	14%	0.8071	0.8127	36	64%	1.0075	1.0115
85	15%	0.8127	0.8182	35	65%	1.0115	1.0156
84	16%	0.8182	0.8235	34	66%	1.0156	1.0198
83	17%	0.8235	0.8285	33	67%	1.0198	1.0240
82	18%	0.8285	0.8335	32	68%	1.0240	1.0282
81	19%	0.8335	0.8383	31	69%	1.0282	1.0326
80	20%	0.8383	0.8429	30	70%	1.0326	1.0370
79	21%	0.8429	0.8475	29	71%	1.0370	1.0415
78	22%	0.8475	0.8520	28	72%	1.0415	1.0461
77	23%	0.8520	0.8563	27	73%	1.0461	1.0508
76	24%	0.8563	0.8606	26	74%	1.0508	1.0556
75	25%	0.8606	0.8648	25	75%	1.0556	1.0605
74	26%	0.8648	0.8689	24	76%	1.0605	1.0656
73	27%	0.8689	0.8730	23	77%	1.0656	1.0708
72	28%	0.8730	0.8770	22	78%	1.0708	1.0761
71	29%	0.8770	0.8810	21	79%	1.0761	1.0816
70	30%	0.8810	0.8849	20	80%	1.0816	1.0873
69	31%	0.8849	0.8888	19	81%	1.0873	1.0932
68	32%	0.8888	0.8926	18	82%	1.0932	1.0994
67	33%	0.8926	0.8964	17	83%	1.0994	1.1058
66	34%	0.8964	0.9002	16	84%	1.1058	1.1125
65	35%	0.9002	0.9039	15	85%	1.1125	1.1195
64	36%	0.9039	0.9076	14	86%	1.1195	1.1269
63	37%	0.9076	0.9113	13	87%	1.1269	1.1348
62	38%	0.9113	0.9150	12	88%	1.1348	1.1432
61	39%	0.9150	0.9186	11	89%	1.1432	1.1521
60	40%	0.9186	0.9223	10	90%	1.1521	1.1619
59	41%	0.9223	0.9259	9	91%	1.1619	1.1725
58	42%	0.9259	0.9295	8	92%	1.1725	1.1842
57	43%	0.9295	0.9331	7	93%	1.1842	1.1974
56	44%	0.9331	0.9367	6	94%	1.1974	1.2126
55	45%	0.9367	0.9403	5	95%	1.2126	1.2306
54	46%	0.9403	0.9439	4	96%	1.2306	1.2530
53	47%	0.9439	0.9475	3	97%	1.2530	1.2831
52	48%	0.9475	0.9512	2	98%	1.2831	1.3315
51	49%	0.9512	0.9548	1	99%	1.3315	>1.3315