ENERGY STAR® Portfolio Manager®

Technical Reference

ENERGY STAR Score for Wastewater Treatment Plants in the United States

OVERVIEW

The ENERGY STAR Score for Wastewater Treatment Plants applies to primary, secondary, and advanced treatment facilities with or without nutrient removal capacity. The objective of the ENERGY STAR score is to provide a fair assessment of the energy performance of a property relative to its peers, taking into account the climate, weather, and business activities at the property. To identify the aspects of building activity that are significant drivers of energy use and then normalize for those factors, a statistical analysis of the peer building population is performed. The result of this analysis is an equation that will predict the energy use of a property, based on its experienced business activities. The energy use prediction for a building is compared to its actual energy use to yield a 1 to 100 percentile ranking of performance, relative to the national population.

- **Property Types.** The ENERGY STAR score for wastewater treatment plants applies to primary, secondary, and advanced treatment facilities, but does not apply to drinking water treatment or distribution utilities.
- **Reference Data.** The analysis for wastewater treatment facilities is based on survey data collected by the American Waterworks Association Research Foundation (AwwaRF).
- **Adjustments for Weather and Business Activity.** The analysis includes adjustments for:
    - Influent Flow
    - Influent Biological Oxygen Demand
    - Effluent Biological Oxygen Demand
    - Plant Load Factor
    - Whether or not the facility has trickle filtration
    - Whether or not the facility has nutrient removal
    - Weather and Climate (using Heating and Cooling Degree Days, retrieved based on Zip code)
- **Release Date.** The ENERGY STAR score for wastewater treatment plants was released in October 2007.

This document presents details on the development of the 1 - 100 ENERGY STAR score for wastewater treatment plants. More information on the overall approach to develop ENERGY STAR scores is covered in our Technical Reference for the ENERGY STAR Score, available at www.energystar.gov/ENERGYSTARScore. The subsequent sections of this document offer specific details on the development of the ENERGY STAR score for wastewater treatment plants:

REFERENCE DATA & FILTERS

For the ENERGY STAR score for wastewater treatment plants, the reference data used to establish the peer building population in the United States is based on survey data collected by the American Waterworks Association Research Foundation (AwwaRF) under a project agreement with CDH Energy. AwwaRF referenced EPA's Office of Water database of wastewater treatment plants in order to draw a statistically representative sample population. The EPA Permit Compliance System provided contact information and flow data for the plants. With the use of this sample population, AwwaRF, in coordination with the California Energy Commission (CEC) and the New York State Energy Research and Development Authority (NYSERDA) funded the survey and research effort to analyze energy use of these plants.

To analyze the facility energy and operating characteristics in this survey data, 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, Program Filters, Data Limitation Filters, and Analytical Filters. A complete description of each of these categories is provided in our Technical Reference for the ENERGY STAR Score, at www.energystar.gov/ENERGYSTARScore. The wastewater treatment plant approach differs slightly from the other property types because the reference data is the AwwaRF survey. As such, the "Building Filter" is referred to as the "Plant Filter". Because the survey is conducted on wastewater treatment plants only, this filter serves to remove observations with incomplete survey responses. ***Figure 1*** presents a summary of each filter applied in the development of the ENERGY STAR score for wastewater treatment plants, the rationale behind the filter, and the resulting number of properties in the data set after the filter is applied. After all filters are applied, the remaining data set has 257 plants.

Figure 1 – Summary of Filters for the ENERGY STAR Score for Wastewater Treatment Plants

Condition for Including an Observation in the Analysis	Rationale	Number Remaining
Wastewater Plant Filter	Plant Filter – Applied to remove any plants with missing data under any of the elements. This is the starting set for the analysis.	289
Estimated natural gas use < 10% of total energy use[1]	Data Limitation Filter – ENERGY STAR scores must be based on actual billed energy consumption. 10% limitation is applied to reduce potential error on estimated consumption.	278
Average daily wastewater flows > than 0.6 million gallons per day (MGD)	Analytical Limitations Filter – values determined by AwwaRF to be statistical outliers.	272
Average influent BOD (biological oxygen demand) level > 30; < 1,000	Analytical Limitations Filter – values determined by AwwaRF to be statistical outliers.	265
Treatment plant electricity use greater than 100,000 kWh	Analytical Limitations Filter – values determined by AwwaRF to be statistical outliers.	260
Average effluent BOD level > 0	Analytical Limitations Filter – values determined by AwwaRF to be statistical outliers.	257

[1] For some observations in the AwwaRF survey, values were reported for total natural gas expenditures but not for total natural gas consumption. In these cases AwwaRF estimated natural gas consumption using a flat national rate for natural gas ($0.874/therm). To reduce error from this estimation, EPA only retained those observations where the estimated natural gas consumption was less than 10% of the total energy consumption.

Of the filters applied to the reference data, some result in constraints on calculating a score in Portfolio Manager and others do not. Data Limitation Filters account for limitations in the data availability, but do not apply in Portfolio Manager. Analytical Filters are used to eliminate outlier data points or different subsets of data, and may or may not affect eligibility. In some cases, a subset of the data will have different behavior from the rest of the properties, in which case an Analytical Filter will be used to determine eligibility in Portfolio Manager (e.g., average daily wastewater flow, average influent BOD level, and average effluent BOD level in wastewater treatment plants). In other cases, Analytical Filters exclude a small number of outliers with extreme values that skew the analysis, but do not affect eligibility requirements. A full description of the criteria you must meet to get a score in Portfolio Manager is available at www.energystar.gov/EligibilityCriteria.

VARIABLES ANALYZED

To normalize for differences in business activity, we perform a statistical analysis to understand what aspects of building activity are significant with respect to energy use. The filtered reference data set described in the previous section is analyzed using a weighted ordinary least squares regression, which evaluates energy use relative to plant activity (e.g., influent flow, load factor, and climate). This linear regression yields an equation that is used to compute energy use (also called the dependent variable) based on a series of characteristics that describe the business activities (also called independent variables). This section details the variables used in the statistical analysis for wastewater treatment plants.

Dependent Variable

The dependent variable is what we try to predict with the regression equation. For the wastewater treatment plant analysis, the dependent variable is energy consumption expressed in source energy use intensity (source EUI). For wastewater treatment plants, this is defined as the total source energy use of the property divided by the average influent flow (in gallons per day). The regressions analyze the key drivers of source EUI – those factors that explain the variation in source energy use per unit flow through the treatment plant.

Independent Variables

The survey data included numerous questions on plant operation including influent and effluent water quality, nutrient removal processes, trickle filtration, UV disinfection, sludge processes, digester gas recovery and general treatment level parameters. AwwaRF reviewed the responses for completeness and performed basic statistical analysis to understand responses for key parameters. Having assessed the data, AwwaRF developed regression equations using step-wise regression. Parameters with high significance, as judged through a t-test, were included in the equation. Different transformations (such as natural logarithm, abbreviated as Ln) were examined for each of the variables. Based on the regression analysis, AwwaRF and EPA identified the following eight key explanatory variables that can be used to estimate the expected average source EUI (kBtu/gpd) of wastewater treatment plants.[2]

- Natural log of average influent flow
- Natural log of average influent biological demand (BOD_5) concentration
- Natural log of average effluent biological demand (BOD_5) concentration
- Natural log of influent load factor
- Fixed film trickle filtration process (yes/no)

[2] For a complete explanation of these criteria, refer to our Technical Reference for the ENERGY STAR Score, at www.energystar.gov/ENERGYSTARScore.

- Nutrient removal (yes/no)
- Natural log of heating degree days
- Natural log of cooling degree days

Although AwwaRF led the development of the data collection and regression analysis, EPA worked closely with AwwaRF to help develop the ENERGY STAR score for wastewater treatment plants. In this process, EPA reviewed the analysis conducted by AwwaRF and proposed slight modifications to the proposed methodology, in order to align the results with EPA's standard methodology. In the review process, EPA assessed the overall project approach to make sure the reference population was a statistically representative sample of the U.S. population of wastewater treatment plants and that the general formulation of the regression equation was consistent with EPA's approach to developing ENERGY STAR scores. The analysis conducted by AwwaRF was thoroughly reviewed and replicated to confirm that a comprehensive and statistically rigorous investigation of variables that influence energy consumption was conducted.

Testing

In addition to thoroughly reviewing the survey data and analysis conducted by AwwaRF and CDH, subsequent testing of the final regression equation was performed by both EPA and AwwaRF to assess the utility and accuracy of the equation. AwwaRF, in coordination with the NYSERDA, examined 16 plants of varying sizes. In addition, historical data from the Sheboygan, Wisconsin wastewater treatment plant was reviewed in order to assess the metric at a single facility over time. The results of the testing phase support the final equation as a useful methodology for assessing energy performance.

EPA also tested the validity of the final ENERGY STAR score using supplemental data supplied by EPA's Region 1 Office. EPA Region 1 engaged five regional wastewater utilities to participate in the pilot project, facilities with a variety of sizes, locations, and probable energy consumption patterns were chosen. The results of the pilot project also supported the methodology for assessing energy performance of wastewater treatment plants.

It is important to reiterate that the final regression equation is based on the nationally representative reference data, not data previously entered into Portfolio Manager.

REGRESSION EQUATION RESULTS

The final regression is a weighted ordinary least squares regression across the filtered data set of 257 observations. The dependent variable is source EUI (kBtu/gpd). Each independent variable is centered relative to the mean value, presented in ***Figure 2***. The final equation is presented in ***Figure 3***. All variables in the regression equation are significant at the 90% confidence level or better, as shown by the significance levels (a p-level of less than 0.10 indicates 90% confidence).

The regression equation has a coefficient of determination (R^2) value of 0.3876, indicating that this equation explains 38.76% of the variance in source EUI for wastewater treatment plants. Because the final equation is structured with energy per flow as the dependent variable, the explanatory power of flow 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[3], demonstrates that the equation

[3] 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 Energy$_i$)^2 across all observations. The Total variation of Y is the sum of (Actual Source Energy$_i$ – Mean Source Energy)^2 across all observations.

actually explains 80.8% of the variation of source energy of wastewater treatment plants. This is an excellent result for a statistically-based energy model.

Detailed information on the ordinary least squares regression approach is available in our Technical Reference for the ENERGY STAR Score, at www.energystar.gov/ENERGYSTARScore.

Figure 2 - Descriptive Statistics for Variables in Final Regression Equation

Variable	Mean	Minimum	Maximum
Source Energy Use (kBtu/gallons per day)	10.13	0.8344	49.70
Ln (Average Influent Flow)	1.862	-0.4308	5.784
Ln (Influent Biological Oxygen Demand)	5.204	3.800	6.585
Ln (Effluent Biological Oxygen Demand)	1.660	-1.204	4.736
Ln (Plant Load Factor)	4.171	2.855	4.690
Presence of trickle filtration (0 if no, 1 if yes)	0.1790	0.0000	1.000
Presence of nutrient removal (0 if no, 1 if yes)	0.4591	0.0000	1.000
Ln (Heating Degree Days)	8.724	6.775	9.324
Ln (Cooling Degree Days)	6.500	4.554	8.089

Figure 3 - Final Regression Results

Summary				
Dependent Variable	Source Energy Intensity (kBtu/gallons per day)			
Number of Observations in Analysis	257			
R^2 value	0.3876			
Adjusted R^2 value	0.3679			
F Statistic	19.62			
Significance (p-level)	0.0000			
	Unstandardized Coefficients	**Standard Error**	**T value**	**Significance (p-level)**
Constant	10.13	0.3101	32.66	0.0000
C_Ln (Average Influent Flow)	-0.9421	0.2449	-3.846	0.0002
C_Ln (Influent Biological Oxygen Demand)	4.876	0.7759	6.284	0.0000
C_Ln (Effluent Biological Oxygen Demand)	-2.082	0.4195	-4.963	0.0000
C_Ln (Plant Load Factor)	-4.668	1.236	-3.778	0.0002
C_Presence of trickle filtration	-2.577	0.8255	-3.122	0.0020
C_Presence of nutrient removal	1.235	0.6634	1.861	0.0639
C_Ln (Heating Degree Days)	2.355	1.214	1.939	0.0536
C_Ln (Cooling Degree Days)	1.243	0.7434	1.672	0.0959

Note: The prefix C_ on each variable indicates that it is centered. The centered variable is equal to difference between the actual value and the observed mean. The observed mean values are presented in ***Figure 2****.*

ENERGY STAR SCORE LOOKUP TABLE

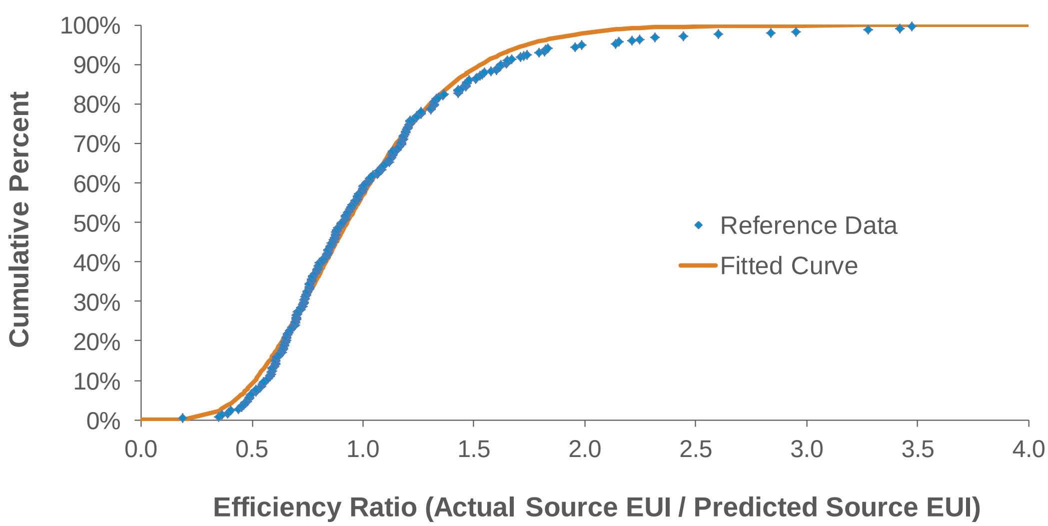
The final regression equation (presented in ***Figure 3***) yields a prediction of source EUI based on a building's operating characteristics. Some buildings in the reference data sample use more energy than predicted by the regression equation, while others use less. The *actual* source EUI of each reference data observation is divided by its *predicted* source EUI to calculate an energy efficiency ratio:

$$Energy\ Efficiency\ Ratio = \frac{Actual\ Source\ EUI}{Predicted\ Source\ EUI}$$

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 reference data set. ***Figure 4*** presents a plot of this cumulative distribution. A smooth curve (shown in orange) 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 5.830 and a scale parameter (beta) of 0.1688. For this fit, the sum of the squared error is 0.1131.

Figure 4 – Distribution for Wastewater Treatment Plants



The final gamma shape and scale parameters are 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 score 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 score of 75; only 25% of the population has ratios this small or smaller. The complete score lookup table is presented in ***Figure 5***.

Figure 5 – ENERGY STAR Score Lookup Table for Wastewater Treatment Plants

ENERGY STAR Score	Cumulative Percent	Energy Efficiency Ratio	
		> =	<
100	0%	0.0000	0.2699
99	1%	0.2699	0.3168
98	2%	0.3168	0.3496
97	3%	0.3496	0.3757
96	4%	0.3757	0.3980
95	5%	0.3980	0.4177
94	6%	0.4177	0.4356
93	7%	0.4356	0.4520
92	8%	0.4520	0.4674
91	9%	0.4674	0.4818
90	10%	0.4818	0.4955
89	11%	0.4955	0.5085
88	12%	0.5085	0.5211
87	13%	0.5211	0.5332
86	14%	0.5332	0.5449
85	15%	0.5449	0.5562
84	16%	0.5562	0.5673
83	17%	0.5673	0.5780
82	18%	0.5780	0.5886
81	19%	0.5886	0.5989
80	20%	0.5989	0.6091
79	21%	0.6091	0.6191
78	22%	0.6191	0.6289
77	23%	0.6289	0.6386
76	24%	0.6386	0.6482
75	25%	0.6482	0.6577
74	26%	0.6577	0.6671
73	27%	0.6671	0.6764
72	28%	0.6764	0.6856
71	29%	0.6856	0.6948
70	30%	0.6948	0.7039
69	31%	0.7039	0.7129
68	32%	0.7129	0.7219
67	33%	0.7219	0.7309
66	34%	0.7309	0.7399
65	35%	0.7399	0.7488
64	36%	0.7488	0.7577
63	37%	0.7577	0.7667
62	38%	0.7667	0.7756
61	39%	0.7756	0.7845
60	40%	0.7845	0.7934
59	41%	0.7934	0.8024
58	42%	0.8024	0.8114
57	43%	0.8114	0.8204
56	44%	0.8204	0.8294
55	45%	0.8294	0.8385
54	46%	0.8385	0.8476
53	47%	0.8476	0.8567
52	48%	0.8567	0.8659
51	49%	0.8659	0.8752

ENERGY STAR Score	Cumulative Percent	Energy Efficiency Ratio	
		>=	<
50	50%	0.8752	0.8846
49	51%	0.8846	0.8940
48	52%	0.8940	0.9035
47	53%	0.9035	0.9130
46	54%	0.9130	0.9227
45	55%	0.9227	0.9325
44	56%	0.9325	0.9424
43	57%	0.9424	0.9523
42	58%	0.9523	0.9625
41	59%	0.9625	0.9727
40	60%	0.9727	0.9831
39	61%	0.9831	0.9936
38	62%	0.9936	1.0043
37	63%	1.0043	1.0152
36	64%	1.0152	1.0262
35	65%	1.0262	1.0375
34	66%	1.0375	1.0489
33	67%	1.0489	1.0606
32	68%	1.0606	1.0725
31	69%	1.0725	1.0847
30	70%	1.0847	1.0972
29	71%	1.0972	1.1099
28	72%	1.1099	1.1230
27	73%	1.1230	1.1364
26	74%	1.1364	1.1502
25	75%	1.1502	1.1644
24	76%	1.1644	1.1791
23	77%	1.1791	1.1942
22	78%	1.1942	1.2099
21	79%	1.2099	1.2262
20	80%	1.2262	1.2431
19	81%	1.2431	1.2607
18	82%	1.2607	1.2791
17	83%	1.2791	1.2984
16	84%	1.2984	1.3187
15	85%	1.3187	1.3402
14	86%	1.3402	1.3630
13	87%	1.3630	1.3872
12	88%	1.3872	1.4133
11	89%	1.4133	1.4415
10	90%	1.4415	1.4722
9	91%	1.4722	1.5060
8	92%	1.5060	1.5438
7	93%	1.5438	1.5866
6	94%	1.5866	1.6365
5	95%	1.6365	1.6963
4	96%	1.6963	1.7717
3	97%	1.7717	1.8751
2	98%	1.8751	2.0457
1	99%	2.0457	>2.0457

EXAMPLE CALCULATION

As detailed in our Technical Reference for the ENERGY STAR Score, at www.energystar.gov/ENERGYSTARScore, there are five steps to compute a score. The following is a specific example for the score for wastewater treatment plants:

1 User enters building data into Portfolio Manager

- 12 months of energy use information for all energy types (annual values, entered in monthly meter entries)
- Physical building information (size, location, etc.) and use details describing building activity (average flow, etc.)

Energy Data	Value
Electricity	1,307,400 kWh
Fuel Oil	17,578 gallons

Property Use Details	Value
Average Influent Flow (MGD)	2.968
Average Influent BOD (mg/l)	160.4
Average Effluent BOD (mg/l)	10.17
Plant Design Flow Rate (MGD)	5
Fixed Film Trickle Filtration Process	0 (No)
Nutrient Removal	0 (No)
HDD (provided by Portfolio Manager, based on Zip code)	4941
CDD (provided by Portfolio Manager, based on Zip code)	756

2 Portfolio Manager computes the actual source EUI

- Total energy consumption for each fuel is converted from billing units into site energy and source energy
- Source energy values are added across all fuel types
- Source energy is divided by average influent flow to determine actual source EUI

Computing Actual Source EUI

Fuel	Billing Units	Site kBtu Multiplier	Site kBtu	Source kBtu Multiplier	Source kBtu
Electricity	1,307,400 kWh	3.412	4,460,849	3.14	14,007,065
Fuel Oil	17,578 gallons	138.874	2,441,127	1.01	2,465,538
				Total Source Energy (kBtu)	16,472,604
				Actual Source EUI (kBtu/gpd)	**5.550**

3 Portfolio Manager computes the predicted source EUI

- Using the property use details from Step 1, Portfolio Manager computes each building variable value in the regression equation (determining the natural log or density as necessary).
- The centering values are subtracted to compute the centered variable for each operating parameter.
- The centered variables are multiplied by the coefficients from the regression equation to obtain a predicted source EUI.

Computing Predicted Source EUI

Variable	Actual Plant Value	Reference Centering Value	Plant Centered Variable	Coefficient	Coefficient * Centered Variable
Constant	--	--	--	10.13	10.13
Ln (Average Influent Flow)	1.088	1.863	-0.7751	-0.9421	0.7302
Ln (Influent Biological Oxygen Demand)	5.078	5.204	-0.1263	4.876	-0.6160
Ln (Effluent Biological Oxygen Demand)	2.319	1.656	0.6634	-2.082	-1.381
Ln (Plant Load Factor)	4.084	4.171	-0.0874	-4.668	0.4079
Trickle filtration (0 if no, 1 if yes)	0.0000	0.1790	-0.1790	-2.577	0.4613
Nutrient removal (0 if no, 1 if yes)	0.0000	0.4591	-0.4591	1.235	-0.5670
Ln (Heating Degree Days)	8.505	8.724	-0.2187	2.355	-0.5150
Ln (Cooling Degree Days)	6.628	6.500	0.1280	1.243	0.1592
			Predicted Source EUI (kBtu/gpd)		**8.809**

4 Portfolio Manager computes the energy efficiency ratio

- The ratio equals the actual source EUI (Step 2) divided by predicted source EUI (Step 3)
- Ratio = 5.550 / 8.809 = 0.6300

5 Portfolio Manager uses the efficiency ratio to assign a score via a lookup table

- The ratio from Step 4 is used to identify the score from the lookup table
- A ratio of 0.6300 is greater than or equal to 0.6289 and less than 0.6386
- ***The ENERGY STAR score is 77***