ENERGY STAR® Performance Ratings
Technical Methodology for Wastewater Treatment Plant

This document presents specific details on the EPA's analytical result and rating methodology for Wastewater Treatment Plant. 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
October 2007

Portfolio Manager Wastewater Treatment Plant Definition

A Wastewater Treatment Plant is a facility that is designed to treat municipal wastewater. The level of treatment at a plant will vary based on the biological oxygen demand (BOD) limits and the specific processes involved. This space type in Portfolio Manager is appropriate for primary, secondary, and advanced treatment facilities with or without nutrient removal. Treatment processes may include biological, chemical, and physical treatment. This space type is best applied to wastewater treatment facilities of 150 million gallons per day (MGD) or smaller. This space type does not apply to *drinking water treatment and distribution utilities*.

Reference Data

The Wastewater Treatment Plant regression model 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.

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*.

The Wastewater Treatment Plant approach differs slightly from the other models 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. **Table 1** presents a summary of each filter applied in the development of the Wastewater Treatment Plant Model. After all filters are applied, the remaining data set has 257 observations.

Table 1 Summary of Wastewater Model Filters		
Condition for Including Observation in the Analysis	**Rationale**	**Remaining Number of Plants**
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 less than 10% of total energy use[1]	Data Limitation Filter – Energy performance models must be based on actual billed energy consumption. 10% limitation is applied to reduce potential error on estimated consumption.	278
Average daily wastewater flows greater 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 greater than 30 and less than 1000	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 greater than 0	Analytical Limitations Filter – values determined by AwwaRF to be statistical outliers.	257

Dependent Variable

The dependent variable in the Wastewater Treatment Plant analysis is source energy use intensity (source EUI). This is equal to the total source energy use of the facility (kBtu) divided by the average influent flow (in gallons per day). By setting source EUI as the dependent variable, the regressions will analyze the key drivers of source EUI – those factors that influence the variation in source energy per unit flow through the treatment plant.

[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.

Independent Variables

General Overview:
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. Upon collecting the survey data, AwwaRF reviewed the responses for completeness and performed basic statistical analysis to understand responses for key parameters. Having assessed the data, AwwaRF developed a model using step-wise regression. Parameters with high significance, as judged through a t-test, were included in the model. Different transformations (such as natural logarithm) 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.

- 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)
- 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 national energy performance rating for Wastewater Treatment Plants. In this process, EPA reviewed the analysis conducted by AwwaRF and proposed slight modifications to the proposed rating 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 US population of Wastewater Treatment Plants and that the general formulation of the regression models was consistent with EPA's approach to model development. 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.

Model Testing:
In addition to thoroughly reviewing the survey data and analysis conducted by AwwaRF and CDH, subsequent testing of the final model was performed by both EPA and AwwaRF to assess the utility and accuracy of the model. 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 model as a useful methodology for assessing energy performance.

EPA also tested the validity of the final rating model using supplemental data supplied by EPA's Region 1 Office. EPA Region 1 engaged five regional wastewater utilities to participate in the pilot project, including 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.

Regression Modeling Results

The final regression is an ordinary least squares regression across the filtered data set of 257 observations. The dependent variable is source EUI (source energy use per gallon of treatment per day). Each independent variable is centered relative to the mean value, presented in **Table 2** below. The final model is presented in **Table 3**. All model variables are significant at the 90% confidence level or better, as shown by the significance levels (a p-value of less than 0.10 indicates 90% confidence). The model has an R^2 value of 0.388, indicating that this model explains 38.8% of the variability in source energy per flow for Wastewater Treatment Plants. Because the final model 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[2], demonstrates that the model actually explains 81% of the variation in source energy at Wastewater Treatment Plants. This is an excellent result for statistically based energy models.

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 document: *Energy Performance Ratings – Technical Methodology*.

Table 2
Descriptive Statistics Variables in Final Regression Model

Variable	Full Name	Mean	Minimum	Maximum
SrcEUI	Source energy use per flow (kBtu/gallons per day)	10.13	0.8344	49.70
Ln(inf_average)	Natural log of the average influent flow (LN(million gallons per day))	1.862	-0.4308	5.784
Ln(inf_BOD)	Natural log of the influent biological oxygen demand (LN(mg/l))	5.204	3.800	6.585
Ln(eff_BOD)	Natural log of the effluent biological oxygen demand (LN(mg/l))	1.660	-1.204	4.736
Ln(inf_lf)	Natural log of the plant load factor (ln(100*average influent flow/plant design flow rate))	4.171	2.855	4.690
Process_tf	Presence of trickle filtration (0 if no, 1 if yes)	0.1790	0.0000	1.000
Treat_nr	Presence of nutrient removal (0 if no, 1 if yes)	0.4591	0.0000	1.000
Ln(HDD)	Natural log of the value for heating degree days	8.724	6.775	9.324
Ln(CDD)	Natural log of the value for cooling degree days	6.500	4.554	8.089

Note:
- *Statistics are computed over the filtered data set (n=257 observations)*
- *The mean values are used to center variables for the regression*

[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 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.

Table 3 Final Regression Modeling Results				
Dependent Variable	Source Energy Intensity (kBtu/gallons per day)			
Number of Observations in Analysis	257			
Model R^2 value	0.3876			
Model F Statistic	19.62			
Model Significance (p-level)	0.0000			
	Unstandardized Coefficients	**Standard Error**	**T value**	**Significance (p-level)**
Intercept	10.13	0.3101	32.66	0.0000
C_ln_inf_average	-0.9421	0.2449	-3.846	0.0002
C_ln(Inf_BOD)	4.876	0.7759	6.284	0.0000
C_ln(eff_BOD)	-2.082	0.4195	-4.963	0.0000
C_ln(inf_lf)	-4.668	1.236	-3.778	0.0002
C_process_tf	-2.577	0.8255	-3.122	0.0020
C_treat_nr	1.235	0.6634	1.861	0.0639
C_LN_HDD	2.355	1.214	1.939	0.0536
C_LN_CDD	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 Table 2*
- *Full variable names and definitions are presented in Table 2*

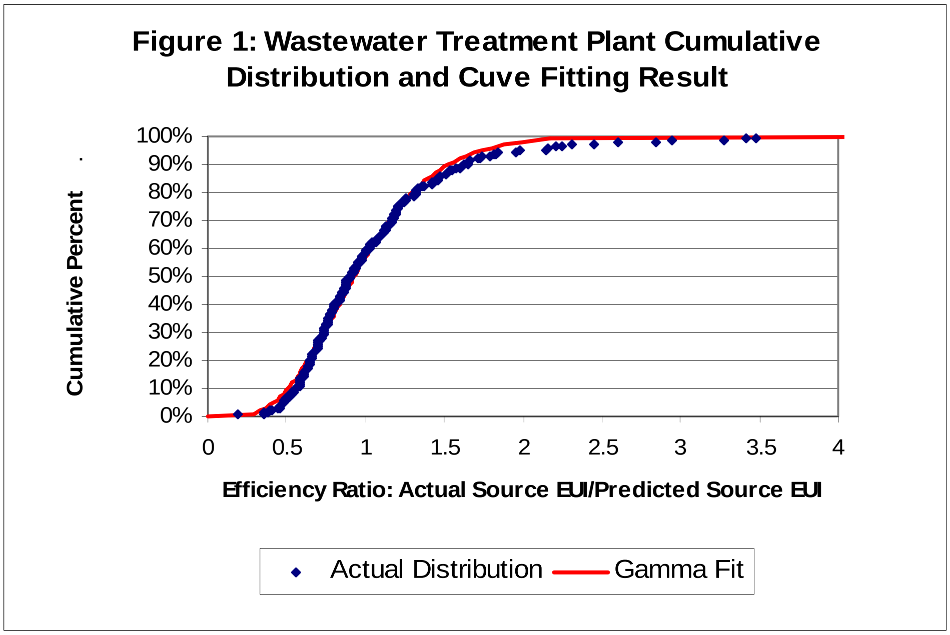
Wastewater Treatment Plant Lookup Table

The final regression model (presented in **Table 3**) yields a prediction of source EUI based on the plant's operating constraints. Some Wastewater Treatment Plants in the sample population use more energy than predicted by the regression equation, while others use less. The *actual* source EUI of each observation in the sample population is divided by its *predicted* source EUI to calculate an energy efficiency ratio:

Energy Efficiency Ratio = Actual Source EUI / Predicted Source EUI

A lower efficiency ratio indicates that a plant 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. Each observation in this population was weighted equally. **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 plant's actual percent rank in the population and each plant's percent rank with the gamma solution. The final fit for the gamma curve yielded a shape parameter (alpha) of 5.829 and scale parameter (beta) of 0.1687. For this fit, the sum of the squared error is 0.1144.



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 0.286218 the rating for that plant should be 100. If the ratio is greater than or equal to 0.286218 and less than 0.336001 the rating for the plant 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 Wastewater Treatment Plant model:

Step 1 – User enters plant data into Portfolio Manager
For the purposes of this example, sample data is provided.

- Energy data
  - Total annual electricity = 1,307,400 kWh
  - Total annual fuel oil = 17,578 gallons
  - Note that this data is actually entered in monthly meter entries
- Operational data
  - Average influent flow = 2.968 (MGD)
  - Average influent BOD = 160.4 (mg/l)
  - Average effluent BOD = 10.17 (mg/l)
  - Plant design flow rate = 5 (MGD)

- Fixed film trickle filtration process = No (0)
- Nutrient removal = No (0)
- HDD (provided by Portfolio Manager, based on zip code) = 4941
- CDD (provided by Portfolio Manager, based on zip code) = 756

Step 2 – Portfolio Manager computes the Actual Source Energy Use Intensity

In order to compute actual source EUI, 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: (1,307,400kWh)*(3.412 kBtu/kWh) = 4,460,849 kBtu Site
  - Fuel Oil: (17,578 gallons)*(138.874 kBtu/gallon) = 2441201 kBtu Site
- Apply the source-site ratios to compute the source energy
  - Electricity:
    4,460,849 Site kBtu*(3.34 Source kBtu/Site kBtu) = 14,899,235 kBtu Source
  - Fuel Oil:
    2,441,201 Site kBtu*(1.01 Source kBtu/Site kBtu) = 2,465,613 kBtu Source
- Combine Source kBtu across all fuels
  - 14,899,235 kBtu + 2,465,613 kBtu = 17,364,848 kBtu
- Divide total source energy by average influent flow
  - Source EUI = (17,364,848 kBtu / 2.968 MGD)/ 1000000
    = 5.851 kBtu/gallons per day

Step 3 – Portfolio Manager computes the Predicted Source Energy Intensity

Portfolio Manager uses the plant data entered under Step 1 to compute centered values for each operating parameter. These centered values are entered into the Wastewater Treatment Plant regression equation to obtain a predicted rating.

- Calculate centered variables
  - Use the operating characteristic values to compute each variable in the model.
    (e.g. LN(Average Influent Flow) = LN(2.968) = 1.088)
  - Subtract the reference centering value form the calculated variable.
    (e.g. LN(Average Influent Flow) - 1.863 = 1.088 - 1.863 = -0.7750)
  - These calculations are summarized in **Table 4**
- Compute predicted source energy use intensity
  - Multiply each centered variable by the corresponding coefficient in the model
    (e.g. Coefficient*CenteredLN(AvgInfluentFlow) = -0.9421*-0.7750= 0.7301)
  - Take the sum of these products (i.e. coefficient*CenteredVariable) and add to the intercept (this yields a predicted source EUI of 8.810 kBtu/gallons per day)
  - This calculation is summarized in **Table 5**.

Step 4 – Portfolio Manager computes the energy efficiency ratio

The energy efficiency ratio is equal to: Actual Source EUI / Predicted Source EUI

Ratio = 5.851 kBtu/flow / 8.810 kBtu/flow = 0.6641

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 that is larger than the computed ratio for the Plant.

- A ratio of 0.6641 is less than 0.666926 (requirement for a 78) but greater than 0.656495 (requirement for a 79)
- ***The rating is a 78***

Table 4 **Example Calculation – Computing Plant Centered Variables**				
Operating Characteristic	Formula to Compute Variable	Plant Variable Value	Reference Centering Value	Plant Centered Variable (Variable Value - Center Value)
Ln (inf_average)	LN(Influent Average)	1.088	1.863	-0.7750
Ln (inf_BOD)	LN(Influent BOD)	5.078	5.204	-0.1260
Ln (eff_BOD)	LN(Effluent BOD)	2.319	1.656	0.6630
Ln (inf_lf)	LN(Influent Load Factor)	4.084	4.171	-0.0870
Process_tf	Yes/No (1/0) Trickle Filtration	0.0000	0.1790	-0.1790
Treat_nr	Yes/No (1/0) Nutrient Removal	0.0000	0.4591	-0.4591
Ln(HDD)	LN(# heating degree days)	8.505	8.724	-0.2190
Ln(CDD)	LN(# cooling degree days)	6.628	6.500	0.1280
Note: *The load factor is computed as: 100*(Average Influent Flow)/(Plant Design Flow Rate)*				

Table 5 **Example Calculation – Computing Predicted Source EUI**			
Operating Characteristic	Centered Variable	Coefficient	Coefficient * Centered Variable
Constant (intercept)	NA	10.13	10.13
Ln (inf_average)	-0.7750	-0.9421	0.7301
Ln (inf_BOD)	-0.1260	4.876	-0.6144
Ln (eff_BOD)	0.6630	-2.082	-1.380
Ln (inf_lf)	-0.0870	-4.668	0.4061
Process_tf	-0.1790	-2.577	0.4613
Treat_nr	-0.4591	1.235	-0.5670
Ln(HDD)	-0.2190	2.355	-0.5157
Ln(CDD)	0.1280	1.243	0.1591
Predicted Source EUI (kBtu/gallon per day)			***8.810***

Attachment

Table 6 lists the energy efficiency ratio cut-off point for each rating, from 1 to 100.

Table 6
Lookup Table for EPA Wastewater Treatment Plant Rating

Rating	Cumulative Percent	Energy Efficiency Ratio		Rating	Cumulative Percent	Energy Efficiency Ratio	
		> =	<			>=	<
100	0%	0	0.286218	50	50%	0.928121	0.938026
99	1%	0.286218	0.336001	49	51%	0.938026	0.948008
98	2%	0.336001	0.370690	48	52%	0.948008	0.958074
97	3%	0.370690	0.398440	47	53%	0.958074	0.9682300
96	4%	0.398440	0.422088	46	54%	0.9682300	0.978484
95	5%	0.422088	0.442994	45	55%	0.978484	0.988843
94	6%	0.442994	0.461923	44	56%	0.988843	0.999315
93	7%	0.461923	0.479353	43	57%	0.999315	1.009909
92	8%	0.479353	0.495604	42	58%	1.009909	1.020632
91	9%	0.495604	0.510904	41	59%	1.020632	1.031494
90	10%	0.510904	0.525420	40	60%	1.031494	1.042505
89	11%	0.525420	0.539277	39	61%	1.042505	1.053675
88	12%	0.539277	0.552575	38	62%	1.053675	1.065016
87	13%	0.552575	0.565393	37	63%	1.065016	1.076538
86	14%	0.565393	0.577795	36	64%	1.076538	1.088254
85	15%	0.577795	0.589833	35	65%	1.088254	1.100178
84	16%	0.589833	0.601552	34	66%	1.100178	1.112325
83	17%	0.601552	0.612989	33	67%	1.112325	1.124710
82	18%	0.612989	0.624176	32	68%	1.124710	1.13735
81	19%	0.624176	0.635140	31	69%	1.13735	1.150265
80	20%	0.635140	0.645906	30	70%	1.150265	1.163474
79	21%	0.645906	0.656495	29	71%	1.163474	1.177001
78	22%	0.656495	0.666926	28	72%	1.177001	1.190870
77	23%	0.666926	0.677216	27	73%	1.190870	1.205108
76	24%	0.677216	0.687380	26	74%	1.205108	1.219745
75	25%	0.687380	0.697431	25	75%	1.219745	1.234817
74	26%	0.697431	0.707382	24	76%	1.234817	1.250361
73	27%	0.707382	0.717244	23	77%	1.250361	1.266420
72	28%	0.717244	0.727029	22	78%	1.266420	1.283043
71	29%	0.727029	0.736745	21	79%	1.283043	1.300285
70	30%	0.736745	0.746402	20	80%	1.300285	1.318212
69	31%	0.746402	0.756009	19	81%	1.318212	1.336896
68	32%	0.756009	0.765573	18	82%	1.336896	1.356425
67	33%	0.765573	0.775103	17	83%	1.356425	1.37690
66	34%	0.775103	0.784604	16	84%	1.37690	1.398442
65	35%	0.784604	0.794085	15	85%	1.398442	1.421196
64	36%	0.794085	0.803552	14	86%	1.421196	1.44534
63	37%	0.803552	0.813011	13	87%	1.44534	1.471092
62	38%	0.813011	0.822469	12	88%	1.471092	1.498726
61	39%	0.822469	0.831932	11	89%	1.498726	1.528593
60	40%	0.831932	0.841405	10	90%	1.528593	1.561154
59	41%	0.841405	0.850896	9	91%	1.561154	1.597026
58	42%	0.850896	0.860409	8	92%	1.597026	1.637074
57	43%	0.860409	0.869950	7	93%	1.637074	1.682554
56	44%	0.869950	0.879525	6	94%	1.682554	1.735396
55	45%	0.879525	0.889140	5	95%	1.735396	1.798812
54	46%	0.889140	0.898802	4	96%	1.798812	1.878762
53	47%	0.898802	0.908515	3	97%	1.878762	1.988462
52	48%	0.908515	0.918286	2	98%	1.988462	2.169378
51	49%	0.918286	0.928121	1	99%	2.169378	>=2.16378