ENERGY STAR® Performance Ratings
Technical Methodology for Retail Store

This document presents specific details on the EPA's analytical result and rating methodology for Retail Store. 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 Retail Store Definition

Retail Store applies to stores of at least 5,000 square feet in gross floor area used to conduct the retail sale of consumer products goods, with the exception of Electronics Stores, Convenience Stores, and Supermarkets. The total gross floor area should include all supporting functions such as kitchens and break rooms used by staff, storage areas, administrative areas, elevators, stairwells, atria, etc. Stores must be free standing or located in strip centers. Stores located in enclosed malls are not eligible, with the exception of mall anchors. Retail segments typically eligible for benchmarking include: Department Stores, Discount Stores, Supercenters, Warehouse Clubs, Drug Stores, Dollar Stores, Home Center/Hardware Stores, and Apparel/Hard Line Specialty Stores (i.e. books, clothing, office products, toys, home goods). Electronics stores and convenience stores are not eligible at this time. Note that supermarkets are rated according to a different model; please refer to the Technical Methodology for Supermarket.

Reference Data

The Retail Store regression model is based on data from the Department of Energy, Energy Information Administration's 2003 Commercial Building Energy Consumption Survey (CBECS). Detailed information on this survey, including complete data files, is publicly available at: http://www.eia.doe.gov/emeu/cbecs/contents.html.

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 Limitation 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 Retail Store 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 182 observations.

The reasons for applying filters on the use and quantity of propane are worthy of additional discussion. In CBECS, major fuel use is reported in exact quantities of consumption. However, if a building uses propane, the amount of propane is reported according to the variable PRAMT8, which uses ranges rather than exact quantities (e.g. less than 100 gallons, 100 to 500 gallons, etc.) Therefore, the quantity must be estimated within the range. To limit error associated with this estimation, EPA applies two limits to the propane quantity.

1. The quantity of propane expressed by PRAMT8 must be 1000 gallons or smaller.
2. The value of propane cannot account for more than 10% of the total source energy use. Because the exact quantity of propane is not reported, this cap ensures that the quantity of propane entered will not introduce undue error into the calculation of total energy consumption. In order to apply this 10% limitation, the value at the high end of the propane category is employed (e.g. for the category of less than 100, a value of 99 is used). If the 10% cap is not exceeded, then EPA will use the value at the middle of the range to calculate total energy use (e.g. for the category of less than 100, a value of 50 is used).

Table 1
Summary of Retail Model Filters

Condition for Including an Observation in the Analysis	Rationale	Number Remaining
PBAPLUS8=42	Building Filter – CBECS defines building types according to the variable "PBAPLUS8." Retail Stores are coded as PBAPLUS8=42.	291
Must operate for at least 30 hours per week	EPA Program Filter – Baseline condition for being a full time Retail Store.	282
Must operate for at least 10 months per year	EPA Program Filter – Baseline condition for being a full time Retail Store.	267
Retail activity must characterize more than 50% of the floor space[1]	EPA Program Filter – In order to be considered part of the Retail peer group, more than 50% of the building must be defined by retail activity.	259
Must have square foot <=1,000,000	Data Limitation Filter – CBECS masks actual values above 1,000,000, using regional averages.	259
If propane is used, the amount category (PRAMTC8) must equal 1, 2, or 3	Data Limitation Filter – Cannot estimate propane use if the quantity is "greater than 1000" or unknown.	250
If propane is used, the maximum estimated propane amount must be 10% or less of the total source energy	Data Limitation Filter – Because propane values are estimated from a range, propane is restricted to 10% of the total source energy.	243
Must not use chilled water	Data Limitation Filter – CBECS does not collect quantities of chilled water.	241
Must have square foot >= 5,000	Analytical Limitation – Analysis could not model behavior for buildings smaller than 5,000ft^2.	182

[1] If the variable ONEACT8=1, this indicates that one activity occupies 75% or more of the building. If the variable ONEACT8=2, then the building can specify up to 3 activities (ACT18, ACT28, ACT38). One of these activities must be retail (PBAX8=15), and must account for more than 50% of the floor area.

Dependent Variable

The dependent variable in the Retail analysis is source energy use intensity (source EUI). This is equal to the total source energy use of the facility divided by the gross floor area. By setting source EUI as the dependent variable, the regressions analyze the key drivers of source EUI – those factors that explain the variation in source energy per square foot in a Retail Store.

Independent Variables

General Overview:

The CBECS data contain numerous building operation questions that EPA identified as potentially important for Retail Stores. Based on a review of the available variables in the CBECS data, in accordance with the EPA criteria for inclusion[2], EPA analyzed the following variables[3]:

- SQFT8 – Square footage
- RGSTRN8 – Number of cash registers
- WKHRS8 – Weekly hours of operation
- NWKER8 – Number of employees during the main shift
- PCNUM8 – Number of personal computers
- SRVNUM8 – Number of servers
- PRNTRN8 – Number of printers
- RFGWIN8 – Number of walk-in refrigeration units
- RFGOPN8 – Number of open refrigerated cases
- RFGRSN8 – Number of residential refrigerators
- RFGCLN8 – Number of closed refrigerated cases
- RFGVNN8 – Number of refrigerated vending machines
- FDRM8 – Commercial food preparation area
- SNACK8 – Snack bar
- FASTFD8 – Fast food or small restaurant
- CAF8 – Cafeteria or large restaurant
- NFLOOR8 – Number of floors
- HDD658 – Heating degree days
- CDD658 – Cooling degree days
- HEATP8 – Percent heated
- COOLP8 – Percent cooled

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., Heating Degree Days * Percent Heated). As part of the analysis, some variables were reformatted to reflect the physical relationships of building components. For example, the number of workers on the main shift is typically evaluated in a density format. The number of

[2] 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).

[3] Note that the 8 at the end of all variables indicates that the 2003 CBECS survey is the eighth survey conducted by the Energy Information Administration.

workers *per square foot* (not the gross number of workers) is expected to be correlated with the energy use per square foot. In addition, 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: source EUI.

Based on the Retail Store regression analysis, the following nine characteristics were identified as key explanatory variables that can be used to estimate the expected average source EUI ($kBtu/ft^2$) in a Retail Store:

- Natural log of gross square foot
- Weekly operating hours
- Number of workers per 1,000 square feet
- Number of personal computers (PCs) per 1,000 square feet
- Number of cash registers per 1,000 square feet
- Number of walk in refrigeration units per 1,000 square feet
- Number of open and closed refrigeration cases per 1,000 square feet
- Heating degree days times Percent of the building that is heated
- Cooling degree days times Percent of the building that is cooled

In addition to the variables listed above, EPA requested and funded the collection of an additional variable "STORE8", which places each observation into one of five categories: Discount Store, Drugstore, Home Center/Hardware Store, Department Store, and Other type of Store. EPA performed extensive analysis of these categories, including using these categories to create interactive regression terms with the other variables in the analysis. Based on these analyses, the category of Store was *not* determined to be statistically significant. The CBECS data lists supermarkets and convenience stores separately from Retail Stores. Therefore, neither of these types of buildings is eligible to rate as a Retail Store according to the model discussed herein. A separate model and technical description are available for Supermarket.

Model Testing:
In addition to the analysis of CBECS data, EPA performed subsequent testing on supplemental data for approximately 600 stores shared with EPA by 10 retail organizations. The results of testing and analysis of this dataset showed that the performance distribution of the test stores was similar to that of the CBECS 2003 observations. This analysis also confirmed that the CBECS categories under "Store8" are not significant. This supplemental data helped EPA verify that the Retail Store regression model provides a valid assessment of energy performance across a variety of Retail Stores. The rating model can be applied to most retail stores including: Department Stores, Discount Stores, Supercenters, Warehouse clubs, Drug Stores, Dollar Stores, Home Centers/Hardware Stores, and Apparel/Hard Line Specialty Stores. However, the analysis showed that the Retail Store model cannot be used to evaluate the energy performance of Electronics Stores. The plug load requirement of these facilities makes it impossible to perform a peer comparison with other retailers.

Finally, the supplemental data included a variety of stand alone retail stores, retail stores in strip malls, and anchor establishments at enclosed malls. Several of the organizations who shared data with EPA had facilities in more than one of these categories. Analysis across all three types of stores did not identify a bias, and therefore confirmed that the Retail Store model is appropriate for rating free standing retail stores, retail stores located within strip mall facilities, and anchor establishments located at enclosed malls.

It is important to reiterate that the final regression model is based on the nationally representative CBECS data, not the supplemental data collected by EPA. The supplemental data served to verify that the CBECS-based regression model provides a valid assessment of energy performance in Retail Stores.

Regression Modeling Results

The final regression is a weighted ordinary least squares regression across the filtered data set of 182 observations. The dependent variable is source EUI. Each independent variable is centered relative to the mean value, presented in **Table 2**. 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-level of less than 0.10 indicates 90% confidence). The model has an R^2 value of 0.71, indicating that this model explains 71% of the variance in source EUI for Retail Store buildings. Because the final model is structured with energy per square foot as the dependent variable, the explanatory power of square foot 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[4], demonstrates that the model actually explains 94.4% of the variation of source energy of Retail Stores. 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*.

[4] 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 2
Descriptive Statistics for Variables in Final Regression Model

Variable	Full Name	Mean	Minimum	Maximum
SrcEUI	Source Energy per Square Foot	153.1	6.660	1009
LNSqFt	Natural Log of Square Foot	9.371	8.517	13.02
WkHrs	Weekly Operating Hours	63.74	30.00	168.0
WkrDen	Number of Workers per 1000 ft^2	0.6279	0.2500	4.000
PCDen	Number of Computers per 1000 ft^2	0.3149	0.0000	2.000
RgstrDen	Number of Cash Registers per 1000 ft^2	0.1905	0.0000	1.400
WalkinDen	Number of Walk-in Refrigerators per 1000 ft^2	0.0038	0.0000	0.1110
RfgCommDen	Number of Open and Closed Refrigerators per 1000 ft^2	0.0450	0.0000	1.000
HDDxPH	Heating Degree Days x Percent Heated	3811	0.0000	9625
CDDxPC	Cooling Degree Days x Percent Cooled	972.1	0.0000	5206

Note:
- *Statistics are computed over the filtered data set (n=182 observations)*
- *Values are weighted by the CBECS variable ADJWT8*
- *The mean values are used to center variables for the regression*

Table 3
Final Regression Modeling Results

Dependent Variable	Source Energy Intensity (kBtu/ft^2)			
Number of Observations in Analysis	182			
Model R^2 value	0.710			
Model F Statistic	46.74			
Model Significance (p-level)	0.0000			
	Unstandardized Coefficients	**Standard Error**	**T value**	**Significance (p-level)**
(Constant)	153.1	5.685	26.93	0.0000
C_LNSqFt	20.19	9.315	2.167	0.0316
C_Wkhrs	1.373	0.4209	3.263	0.0013
C_WkrDen	61.76	15.54	3.975	0.0001
C_PCDen	70.60	20.80	3.394	0.0009
C_RgstrDen	249.1	33.79	7.372	0.0000
C_WalkinDen	720.2	379.6	1.897	0.0595
C_RfgCommDen	81.90	44.34	1.847	0.0665
C_HDDxPH	0.0113	0.0036	4.274	0.0000
C_CDDxPC	0.0125	0.0073	1.725	0.0863

Note:
- *The regression is a weighted ordinary least squares regression, weighted by the CBECS variable "ADJWT8".*
- *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.***

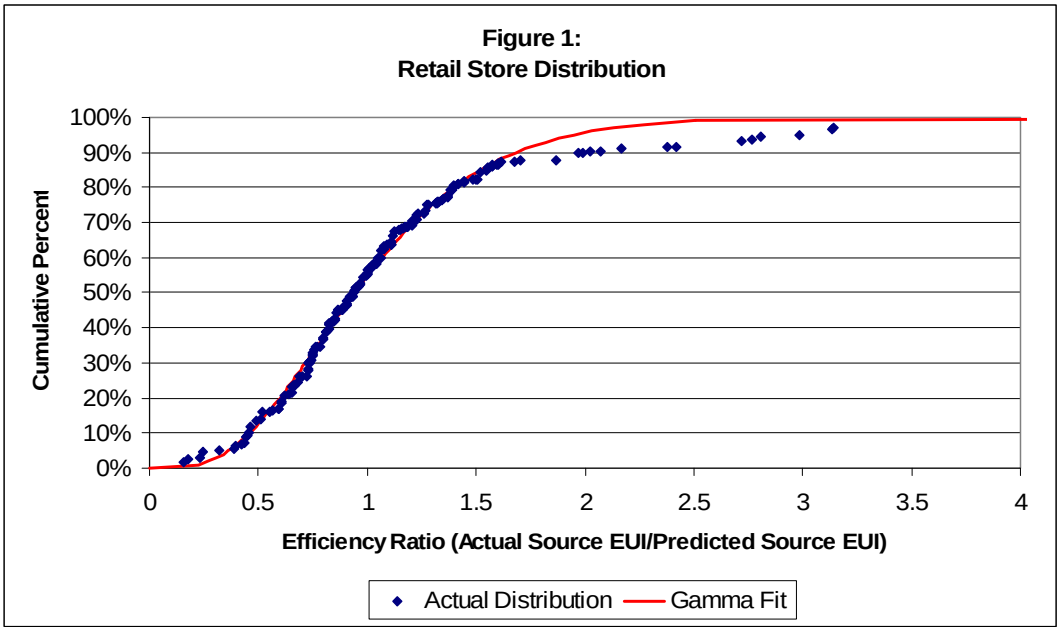
Retail Store Lookup Table

The final regression model (presented in **Table 3**) yields a prediction of source EUI based on a building's operating constraints. Some buildings in the CBECS data sample use more energy than predicted by the regression equation, while others use less. The *actual* source EUI of each CBECS observation 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 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 CBECS 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 4.2595 and a scale parameter (beta) of 0.2397. For this fit, the sum of squared error is 0.074.



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% is 0.6707. 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 0.2243 the rating for the building

should be 100. If the ratio is greater than or equal to 0.2243 and less than 2.5066 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 Retail Store 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 = 400,000 kWh
  - Total annual natural gas = 180 therms
  - Note that this data is actually entered in monthly meter entries
- Operational data
  - Gross floor area (ft^2) = 50,000
  - Weekly operating hours = 70
  - Workers on main shift[5] = 8
  - Number of personal computers = 3
  - Percent heated = 100%
  - Percent cooled = 100%
  - Number of cash registers = 6
  - Number of walk-in refrigeration/freezer units = 0
  - Number of open and closed refrigeration/freezer cases = 7
  - HDD (provided by Portfolio Manager, based on zip code) = 3850
  - CDD (provided by Portfolio Manager, based on zip code) = 2300

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: (400,000kWh)*(3.412kBtu/kWh) = 1,364,800 kBtu Site
  - Natural gas: (180 therms)*(100kBtu/therm) = 18,000 kBtu Site
- Apply the source-site ratios to compute the source energy
  - Electricity:
    1,364,800 Site kBtu*(3.34 Source kBtu/Site kBtu) = 4,558,432 kBtu Source
  - Natural Gas:
    18,000 Site kBtu *(1.047 Source kBtu/Site kBtu) = 18,846 kBtu Source
- Combine source kBtu across all fuels
  - 4,558,432 kBtu + 18,846 kBtu = 4,577,278 kBtu
- Divide total source energy by gross floor area
  - Source EUI = 4,577,278 kBtu/50,000ft^2 = 91.5 kBtu/ft^2

[5] This represents typical peak staffing level during the main shift. For example, in a retail store if there are two daily 8 hour shifts of 15 workers each, the Workers on Main Shift value is 15.

Step 3 – Portfolio Manager computes the Predicted Source Energy Intensity

Portfolio Manager uses the building data entered under Step 1 to compute centered values for each operating parameter. These centered values are entered into the Retail Store regression equation to obtain a predicted source EUI.

- Calculate centered variables
  - Use the operating characteristic values to compute each variable in the model. (e.g. LN(Square Foot) = LN(50,000) = 10.82) .
  - Subtract the reference centering value from calculated variable (e.g. LN(Square Foot) - 9.371 = 10.82 - 9.371 = 1.449).
  - 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(Square Foot) = 20.19*1.449=29.26)
  - Take the sum of these products (i.e. coefficient*CenteredVariable) and add to the constant (this yields a predicted source EUI of 148.6 kBtu/ft^2)
  - 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 = 91.5/148.6 = 0.6157

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 0.6157 is less than 0.6237 (requirement for 79) but greater than 0.6116 (requirement for 80)
- ***The rating is 79***

Table 4 Example Calculation – Computing Building Centered Variables				
Operating Characteristic	Formula to Compute Variable	Building Variable Value	Reference Centering Value	Building Centered Variable (Variable Value - Center Value)
LN(SqFt)	LN(Square Foot)	10.82	9.371	1.449
WkHrs	Weekly Operating Hours	70.00	63.74	6.260
WkrDen	(#Workers/ft^2*1000)	0.1600	0.6279	-0.4679
PCDen	(#Computers/ft^2*1000)	0.0600	0.3149	-0.2549
RgstrDen	(#Registers/ft^2*1000)	0.1200	0.1905	-0.0705
WalkinDen	(Walk-in/ft^2*1000)	0.0000	0.0038	-0.0038
RfgCommDen	(#Open&Closed/ft^2*1000)	0.1400	0.0450	0.0950
HDDxPH	(HDD*Percent Heated)	3850	3811	39.00
CDDxPC	(CDD*Percent Cooled)	2300	972.1	1328

Note
- *Densities are always expressed as the number per 1,000 square feet*
- *The center reference values are the weighted mean values from the CBECS population, show in* ***Table 2***

Table 5 Example Calculation – Computing predicted Source EUI			
Operating Characteristic	Centered Variable	Coefficient	Coefficient * Centered Variable
Constant (intercept)	NA	153.1	153.1
LN(SqFt)	1.449	20.19	29.26
WkHrs	6.260	1.373	8.595
WkrDen	-0.4679	61.76	-28.90
PCDen	-0.2549	70.60	-18.00
RgstrDen	-0.0705	249.1	-17.56
WalkinDen	-0.0038	720.2	-2.737
RfgCommDen	0.0950	81.90	7.781
HDDxPH	39.00	0.0113	0.4407
CDDxPC	1328	0.0125	16.60
Predicted Source EUI (kBtu/ft^2)			***148.6***

Attachment

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

Table 6
Lookup Table for Retail Rating

Rating	Cumulative Percent	Ratio	Rating	Cumulative Percent	Ratio
100	0%	<0.224337	50	50%	0.954258
99	1%	0.274272	49	51%	0.966245
98	2%	0.309965	48	52%	0.978348
97	3%	0.338985	47	53%	0.990575
96	4%	0.364015	46	54%	1.002935
95	5%	0.386355	45	55%	1.015437
94	6%	0.406745	44	56%	1.028092
93	7%	0.425649	43	57%	1.040909
92	8%	0.443381	42	58%	1.053899
91	9%	0.460165	41	59%	1.067074
90	10%	0.476165	40	60%	1.080446
89	11%	0.491508	39	61%	1.094028
88	12%	0.506293	38	62%	1.107833
87	13%	0.520598	37	63%	1.121876
86	14%	0.534487	36	64%	1.136173
85	15%	0.548013	35	65%	1.150742
84	16%	0.561222	34	66%	1.165601
83	17%	0.574152	33	67%	1.180770
82	18%	0.586834	32	68%	1.196271
81	19%	0.599298	31	69%	1.212127
80	20%	0.611567	30	70%	1.228365
79	21%	0.623665	29	71%	1.245014
78	22%	0.635610	28	72%	1.262105
77	23%	0.647420	27	73%	1.279672
76	24%	0.659111	26	74%	1.297756
75	25%	0.670697	25	75%	1.316400
74	26%	0.682192	24	76%	1.335653
73	27%	0.693607	23	77%	1.355569
72	28%	0.704953	22	78%	1.376212
71	29%	0.716241	21	79%	1.397654
70	30%	0.727481	20	80%	1.419975
69	31%	0.738683	19	81%	1.443273
68	32%	0.749854	18	82%	1.467659
67	33%	0.761003	17	83%	1.493262
66	34%	0.772139	16	84%	1.520240
65	35%	0.783268	15	85%	1.548780
64	36%	0.794398	14	86%	1.579110
63	37%	0.805537	13	87%	1.611512
62	38%	0.816692	12	88%	1.646341
61	39%	0.827869	11	89%	1.684051
60	40%	0.839075	10	90%	1.725238
59	41%	0.850317	9	91%	1.770704
58	42%	0.861603	8	92%	1.821568
57	43%	0.872938	7	93%	1.879461
56	44%	0.884329	6	94%	1.946894
55	45%	0.895784	5	95%	2.028047
54	46%	0.907309	4	96%	2.130689
53	47%	0.918912	3	97%	2.272085
52	48%	0.930600	2	98%	2.506554
51	49%	0.942379	1	99%	>2.506554