DEVELOPMENT OF ENERGY STAR® ENERGY PERFORMANCE INDICATOR FOR COMMERCIAL BREAD AND ROLL BAKERIES

GALE A. BOYD, MATT DOOLIN, AND RAGHAV SABOO

DUKE UNIVERSITY, SOCIAL SCIENCE RESEARCH INSTITUTE

BOX 90989, DURHAM, NC 27708

JUNE 15, 2016

SPONSORED BY THE U.S. ENVIRONMENTAL PROTECTION AGENCY AS PART OF THE ENERGY STAR PROGRAM.

ACKNOWLEDGMENTS

This work was sponsored by the U.S. EPA Climate Protection Partnerships Division's ENERGY STAR Industrial program under the direction of Elizabeth Dutrow and Walt Tunnessen with support from Joshua Smith at ICF International. The research has benefited from comments by company representatives participating in the *ENERGY STAR Commercial Baking Industry Focus, the American Bakers Association,* and the *McDonald's Bakery and Baked Goods Sustainability Team*. The results presented in this paper have been screened to ensure that no confidential data are revealed.

CONTENTS

DEVELOPMENT OF ENERGY STAR® ENERGY PERFORMANCE INDICATOR FOR COMMERCIAL BREAD AND ROLL BAKERIES

GALE A. BOYD, MATT DOOLIN, AND RAGHAV SABOO

ABSTRACT

Organizations that implement strategic energy management programs undertake a set of activities that, if carried out properly, have the potential to deliver sustained energy savings. Energy performance benchmarking is a key activity of strategic energy management and one way to enable companies to set energy efficiency targets for manufacturing facilities. The opportunity to assess plant energy performance through a comparison with similar plants in its industry is a highly desirable and strategic method of benchmarking for industrial energy managers. However, access to energy performance data for conducting industry benchmarking is usually unavailable to most industrial energy managers. The U.S. Environmental Protection Agency (EPA), through its ENERGY STAR program, seeks to overcome this barrier through the development of manufacturing sector-based plant energy performance indicators (EPIs) that encourage U.S. industries to use energy more efficiently. This report describes work with the Commercial Baking industry to provide a plant-level indicator of energy efficiency for facilities that produce various types of bread, rolls, and related products in the United States. Consideration is given to the role that performance-based indicators play in motivating change; the steps necessary for indicator development, from interacting with an industry in securing adequate data for the indicator; and actual application and use of an indicator when complete. How indicators are employed in EPA's efforts to encourage industries to voluntarily improve their use of energy is discussed as well. The report describes the data and statistical methods used to construct the EPI for plants within the industry, specifically large commercial bakeries producing bread, rolls, frozen dough and related products. The individual equations are presented, as are the instructions for using those equations as implemented in an associated Microsoft Excel-based spreadsheet tool.

1 INTRODUCTION

ENERGY STAR was introduced by EPA in 1992 as a voluntary, market-based partnership to reduce air pollution and greenhouse gas emissions associated with energy use through increased energy efficiency (U.S. Environmental Protection Agency 2015). This government program enables industrial and commercial businesses as well as consumers to make informed decisions that save energy, reduce costs, and protect the environment. For businesses, a key step in improving energy efficiency is to institutionalize a strategic approach to energy management. Drawing from management standards for quality and environmental performance, EPA developed the *ENERGY STAR Guidelines for Energy Management* that identifies the components of successful energy management practices (U.S. Environmental Protection Agency 2003).

These include:

- Commitment from a senior corporate executive to manage energy across all businesses and facilities operated by the company;
- Appointment of a corporate energy director to coordinate and direct the energy program and multi-disciplinary energy team;
- Establishment and promotion of an energy policy;
- Development of a system for assessing performance of the energy management efforts, including tracking energy use as well as benchmarking energy in facilities, operations, and subunits therein;
- Conduct of assessments to determine areas for improvement;
- Setting of goals at the corporate, facility, and subunit levels;
- Establishment of an action plan across all operations and facilities, as well as monitoring successful implementation and promoting the value to all employees; and
- Pursue recognition and rewards for the success of the program.

Of the major steps in energy management program development, benchmarking energy performance by comparing current energy performance to a baseline or a similar entity is critical. In manufacturing, it may take the form of detailed comparisons of specific production lines or pieces of equipment, or it may be performed at a broader system level by gauging the performance of a single manufacturing plant with respect to its industry. Regardless of the application, benchmarking enables companies to determine whether better energy performance could be expected. It empowers them to set goals and evaluate their reasonableness.

Boyd, Dutrow, and Tunnessen (2008) describe the evolution of a statistically based plant energy performance indicator for the purpose of benchmarking manufacturing energy use for ENERGY STAR.

Boyd and Lee (2016) describe the basic approach used in developing such an indicator, including the concept of normalization and how variables are chosen to be included in the analysis. To date, ENERGY STAR has developed statistical indicators for a wide range of industries (U.S. Environmental Protection Agency 2015). This report describes the basic concept of benchmarking and the statistical approach employed in developing performance-based energy indicators for the commercial baking industry, the evolution of the analysis done for this industry, the final results of this analysis, and ongoing efforts by EPA to improve the energy efficiency of this industry and others.

2 BENCHMARKING THE ENERGY EFFICIENCY OF INDUSTRIAL PLANTS

Among U.S. manufacturers, few industries participate in industry-wide plant benchmarking. The petroleum and petrochemical industries each support plant-wide surveys conducted by a private company and are provided with benchmarks that address energy use and other operational parameters related to their facilities. Otherwise, most industries have not benchmarked energy use across their plants. As a result, some energy managers find it difficult to determine how well their plants might perform.

In 2000, EPA began developing a method for developing benchmarks of energy performance for plant-level energy use within a manufacturing industry. Discussions yielded a plan to use a source of data that would nationally represent manufacturing plants within a particular industry, create a statistical model of energy performance for the industry's plants based on these data along with other available sources for the industry, and establish the benchmark for the comparison of those best practices, or best-performing plants, to the industry. The primary data sources would be the Census of Manufacturing, Annual Survey of Manufacturing, and Manufacturing Energy Consumption Survey collected by the Census Bureau, or data provided by trade associations and individual companies when warranted by the specific industry circumstances and participation.

3 EVOLUTION OF THE COMMERICAL BAKING EPI

In early 2012, EPA decided to expand its existing work with food processors by starting an industrial focus in commercial bakeries. This was done with early involvement and interest from the industry trade association, the American Bakers Association. Shortly after, EPA produced an energy guide entitled *Energy Efficiency Improvement and Cost Saving Opportunities for the Baking Industry*, which highlighted cost saving energy efficiency opportunities throughout bakeries at the process level.

3.1 Using Census Data

The first draft of the Commercial Bread and Roll Bakery EPI (hereafter, referred to as the commercial baking EPI for brevity) was based on data reported to the U.S. Census Bureau under the six-digit NAICS code 311812, specifically covering the inputs and outputs of bread and bread-type rolls. After the initial analysis, the following variables were included in the model to account for energy usage:

- Flour (Input/Capacity)

- Production worker hours (Utilization)
- Total value of production (Production)
- White hearth bread value (Product mix)
- Dark wheat bread value (Product mix)
- Other breads value (Product mix)
- Rolls, hamburger, and hot dog buns value (Product mix)
- Other rolls value (Product mix)
- Other production value (Product mix)

Early results indicated that *other production* was the largest product mix contributor to energy use, hinting that the Census data could be missing key information or capturing more facilities that would fall outside the scope of traditional commercial bakeries. The effects of other variables were of the expected direction and magnitude. For example, total value of production showed a positive, significant impact on energy, indicating that more expensive products would use more energy through some combination of increased baking times, longer mixing times, or some other process-related component.

Industry participants were critical of the Census data-based model for a variety of reasons. Production data was based on sales rather than physical quantity values. Sales data could be impacted by different product markups by brand. In addition, data on the amount or sales of frozen products were not included. Frozen products have different baking times and require large freezers, both of which would impact the energy usage of a plant that was not accounted for in the Census-based model. Therefore, industry participants suggested that they could voluntarily provide data for a new analysis, with support from the industry trade association.

3.2 Using Industry-Supplied Data

To move forward with industry-supplied data, discussions centered on what variables are suspected to contribute the most to energy use and what data all facilities are currently tracking. The focus participants indicated that most plants track pounds of raw dough and baked weight of final products; the raw dough values would be greater than the sum of all final baked product weights. Whole wheat breads were not suspected to use any more energy in production than white breads, so no distinction was made between these two types. Major product types included hearth breads, pan breads, rolls, English muffins, bagels, frozen dough, and other products. Information on the weight of scrapped product was also collected, but found to be very insignificant in number and data accuracy was questionable. Instead, the ratios between raw dough and final baked product weights capture wasted product as the amount of dough needed to make a certain amount/type of bread is fairly standard across the industry. More product types and industry-supplied data helped ensure that data used in the analysis would be within the commercial bakery scope, focusing on breads and excluding pastries, pies, cakes, etc. Data on amounts of frozen products, the presence of freezers in the facility, and whether the production area was conditioned were collected to pinpoint operational differences between facilities that could result in different energy profiles for plants with similar product totals. Finally, electricity and fuel data were supplied directly from facilities and different types of fuels were identified to make the appropriate source energy conversions. All data provided for EPI research & development were covered

under company-specific non-disclosure agreements with Duke University. No data were supplied to EPA.

The earliest model using industry-supplied data consisted of 317 plant-years of data from 2010-2012, covering 135 unique facilities across 17 companies. Commercial bakeries primarily use electricity for mixing, machining, freezing, and air conditioning, and use fuels for baking and heating. These distinct energy sources for these processes allowed electricity and fuels to be split, and the efficiency of each energy source was evaluated in the model before being aggregated into one overall efficiency score (discussed in greater detail in the following section). Weather variables, including heating degree days (HDD) and cooling degree days (CDD), were added to account for the effect different climates would have on electricity and fuel consumption via heating/cooling requirements. Additionally, weather conditions affect chiller efficiencies and oven efficiency due to the amount of outside air used. Early results were in line with expectations and are detailed below:

- Plant Size
    - There is evidence of economies of scale in electricity and fuel use, meaning that a plant producing twice as much product would not use twice as much energy. Amounts of raw dough processed were used as a proxy for plant capacity.
- Electricity
    - The percentage of frozen products played a large role in energy consumption, capturing the effects of operating freezers and freezer capacities.
    - Air conditioning the facility did not have a significant impact on energy usage.
    - There were small differences in electricity usage per product, with pan breads and bagels using the least and hearth bread and English muffins using the most.
    - Cooling degree days (CDD) have a small but significant impact.
    - Electricity intensity (electricity per lb. of raw dough) varies more widely than fuel intensity, but this is due to different product requirements rather than inefficiencies.
- Fuels
    - Frozen dough was the least fuel intensive, as there is no baking involved. Other frozen products have lower fuel use, but not to the same extent, implying partial baking occurred.
    - Hearth breads and English muffins use the most fuels, with a larger difference than electricity.
    - Heating degree days (HDD) have a small but significant impact.

Through industry testing and feedback, some changes were made to the model to correct some identified problems, including the strength of the air conditioning variable, plants that reported no lost weight from raw dough to final product, and how to more accurately account for products that are partially baked and then frozen. The first issue dealt with the air conditioning result since it is counterintuitive that operating air conditioning would not have a statistical impact on electricity consumption. Many facilities reported air conditioning, but they only conditioned the office and employee areas and not areas of production. The definition was altered to only include if the manufacturing areas of the facility were air conditioned. This change indicated that conditioning the production space increases expected electricity consumption by 33%, although there was a relatively small number of facilities that condition these areas.

Next, a limited number of companies did not track final baked product weight and instead used the amount of raw dough that went in to each product. Reporting these raw dough product weights would result in a summed total matching the raw dough inputs; this would imply that all products were being frozen without any baking, and should be classified as frozen dough even though that is not the case. Since the final baked product weight data were not being tracked at the company level, the average ratio of raw dough to final product for each product in the industry sample was used to estimate total baked weights. The ratio was approximately equal to 0.90, indicating that 10% of raw dough weight was lost as evaporated moisture during the baking process. Focus participants agreed that this ratio was accurate in most cases, and impacted companies did not report producing any par-baked products that would affect this assumption. Updating these numbers more accurately reflected the energy usage for each product type.

Finally, the effect of partial baking and amount of frozen products were linked, and required multiple steps to resolve in the model. Several data entry errors for frozen product percentages (e.g., a 1.00 indicating 100% frozen product was interpreted as 1% frozen product) were corrected, which strengthened the effect of frozen products on energy usage and clarified the extent of partial baking. Additionally, many facilities that were reporting high percentages of frozen product were only producing products labeled as "other." Many of these plants were identified as producing frozen pastries, pies, cakes, and other sweets, and were therefore dropped from the analysis. Following data adjustments, two additional variables were added that captured partial baking. The first was a ratio of total product weight to processed raw dough. If this variable was equal to one, then no baking occurred and all products were frozen dough. This variable would be larger for partially baked products, as less moisture would be baked out due to a shorter baking time and the products would be heavier than similar fully baked products. Intuitively, partial baking would require less energy, and this is confirmed by the negative coefficient on this variable for both the fuel and electricity models. The second variable was the percentage of frozen product less the percentage that is frozen dough. Since frozen dough is already being accounted for in the product mix, also including it in the frozen product variable overstated the effect of frozen dough and understated the impact of products that were partially baked and then frozen. By capturing the percentage of partially baked products, the impact is more accurately captured in the model without requiring additional data collection. Summary statistics for the final variables utilized in the model are detailed in Table 1 below:

Table 1 Summary Statistics

Variable	Mean	Standard Deviation
Electricity (MWh)	7,109	4,421
Total Fuels (MMBtu)	49,410	29,433
Raw Dough Processed (million lbs.)	66	43
Final Baked Weight to Raw Dough	86%	5%
Hearth Breads (million lbs.)	1.2	6.5
Pan Breads (million lbs.)	30	28
Rolls (million lbs.)	17	17
English Muffins (million lbs.)	2.4	8.7
Bagels (million lbs.)	1.2	5.4
Frozen Dough (million lbs.)	3.3	25

Other Products (million lbs.)	1.4	4.9
Percentage of Frozen Product	6%	20%

3.3 Modeling electricity and fuel use separately

There are instances in which modeling total energy use is the most appropriate approach, particularly when there are substantial opportunities to meet production energy requirements by using fuels instead of electricity, or when there is onsite electricity generation from combined heat and power (CHP) where typically more fuel is used and less electricity purchased. This would result in a plant appearing very fuel inefficient and very electric efficient; examples of the converse are possible. However, when certain products are inherently more (or less) electric or fuel intensive, then it may be appropriate to represent the electricity and fuel use separately since those production differences can more readily be accounted for in the analysis. Separating the energy forms may also improve the ability to measure weather effects, since higher cooling degree days (CDD) will be associated with higher cooling loads and electricity use; conversely, heating degree days (HDD) will be associated with heating loads and fuel use.[1]

The value of separately modeling electricity and fuel was quite high in the large commercial bakery sector: since some products are frozen, the process would be more electric intensive. There is also a sub-category of frozen products that are not baked, so would be both more electric intensive and less fuel intensive. To capture these product mix effects, a separate analysis of the two energy forms was needed. In addition, there are few, if any, applications of CHP in this industry, so this particular reason to model total energy is not a concern. Industry-provided data allowed for separate electric and thermal analyses in the third version of the draft model.

The analysis of each energy form follows the same general approach as would be taken for a total energy analysis, resulting in an individual measure of energy efficiency performance. Since the percentile rankings of these individual measures of efficiency are based on a probability distribution, each with its own variance, the Energy Performance Score (EPS) of the total energy use would also be derived from these separate variances. If electric efficiency was independent of (unrelated to) fuel efficiency, then the relevant variance for the sum of electricity use and fuel use would be the sum of the underlying variances, but this is unlikely to be the case. The energy management of the firm (plant) might make a particular location more (or less) efficient in both cases, making the efficiencies correlated. Either the joint distribution would need to be explicitly modeled, or another way to obtain the efficiency distribution for total energy would be necessary. In the case of the commercial baking EPI, the latter was chosen.

The general form of the underlying EPI equation is

$$E = f(Y, X;\ \theta) + \varepsilon \qquad (1)$$

or in this case

[1] There are exceptions to this pattern, e.g., electricity used in the heating system, fuels driving adsorption chillers, etc.

$$\ln(E) = a + \sum_{j=1}^{n} b_j \, ln(y_j) + \sum_{j=1}^{m} c_j \, x_j + \varepsilon \quad (2)$$

Where Y includes measure of activities, X is plant characteristics like weather and product mix, and the vector of parameters to be estimated is $\theta = (a, b_*, c_*, \sigma^2)$.

We compute the estimate of efficiency as $\hat{\epsilon}_{i,t}$, for every plant, i, and year, t, from the parameter estimates, which are denoted by a $\frown$, from

$$\ln(E) - \hat{a} + \sum_{j=1}^{n} \widehat{b_j} \ln(y_j) + \sum_{j=1}^{m} \widehat{c_j} \, x_j = \hat{\epsilon}_{i,t} \quad (3)$$

For models using ordinary least squares (OLS) estimates, such as is the case for this industry, we have estimated the variance of the error term of equation (1), and we can compute the probability that the difference between actual energy use and predicted average energy use is no greater than this computed difference under the assumption that the efficiency, ε, is normally distributed with zero mean and variance σ^2, i.e., $\varepsilon \sim N(0, \sigma^2)$

$$\text{EPS} = (1 - \Pr(\varepsilon \leq \hat{\epsilon}_{i,t})) \cdot 100 \quad (4)$$

One minus this probability, multiplied by 100, is the *Energy Performance Score* (EPS), and is the *percentile ranking of the energy efficiency of the plant.*[2]

However, the EPI has two types of energy, so it is necessary to have $\hat{\epsilon}_{i,t,e}$ and $\hat{\epsilon}_{i,t,f}$, where e and f represent electricity and fuels. The sum of two normally distributed variables is not necessarily normal, unless they are uncorrelated. It would be preferable to compute the analog of $\hat{\epsilon}_{i,t,e}$ and $\hat{\epsilon}_{i,t,f}$, but for the sum of electricity and fuels. To do this, we need to account for the fact that the equations are estimated in log form, convert the predicted values for the energy use into levels, convert them to common units so they can be added together, and have a method to compute the probability in equation (4) that is the basis for the EPS.

While it is true that the predicted value of the natural log of energy use, $\widehat{\ln(E)}$, is

$$\widehat{\ln(E)} = \hat{a} + \sum_{j=1}^{n} \widehat{b_j} \ln(y_j) + \sum_{j=1}^{m} \widehat{c_j} \, x_j \quad (5)$$

we need the predicted *level of energy use* in order to add electricity and fuel together. For an OLS, the estimate of the predicted level of energy use – i.e., the expected value of E – is not the exponential of $\widehat{\ln(E)}$, but is

$$\hat{E} = e^{\left(\widehat{\ln(E)} + \frac{\sigma^2}{2}\right)} \quad (6)$$

σ^2 is the OLS error variance estimated from (2).

For notational simplicity, we denote $E_{i,t}$, $\widehat{E_{i,t}}$, $F_{i,t}$, and $\widehat{F_{i,t}}$ to be the actual and predicted pairs for electricity use and fuel use, respectively, for each plant and year. We can compute the estimates of

[2] By ENERGY STAR convention, the EPS is 100 for the lowest value of energy intensity, representing efficiency. In statistics, the lowest (left-most value of the density and distribution) is zero and the largest (right-most value) is 100%. To create the EPS we use the simple transformation.

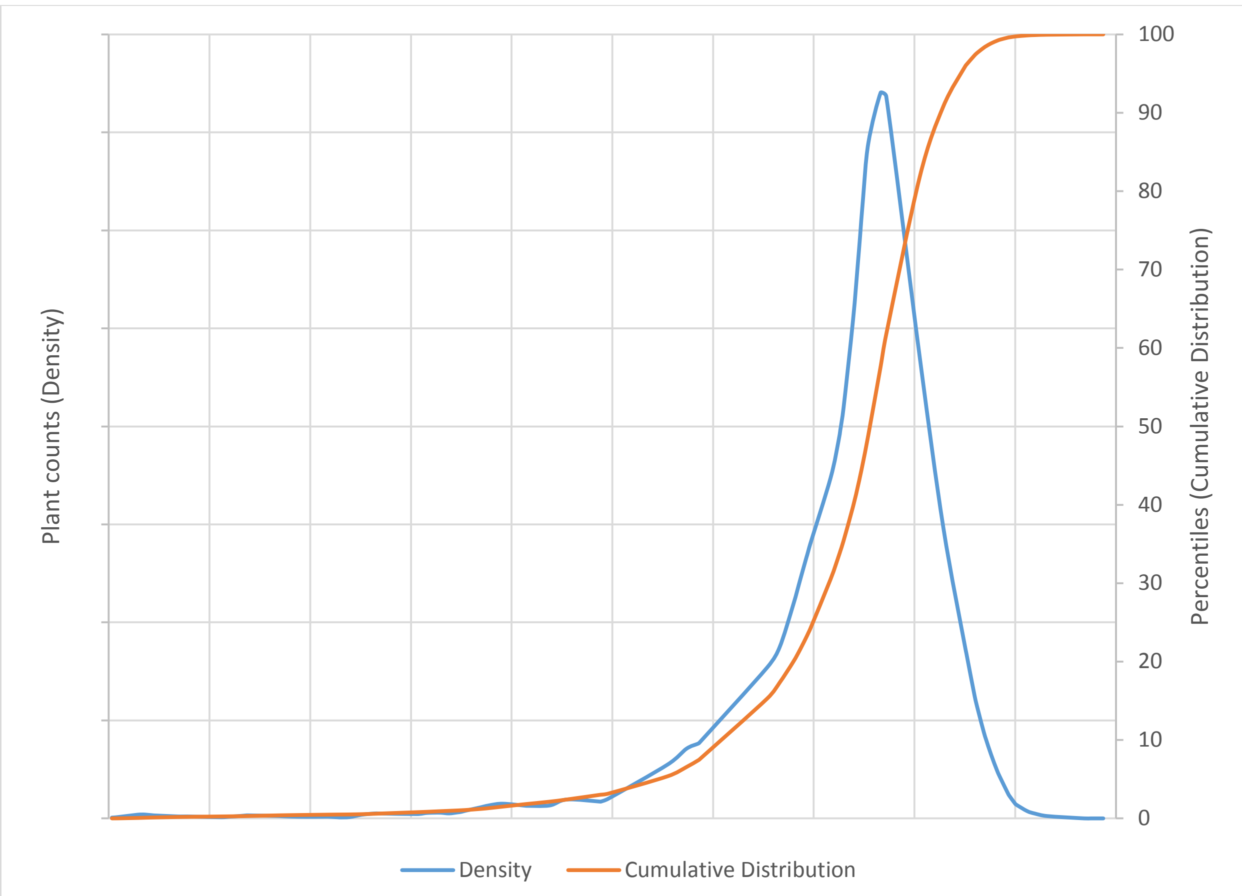
plant level total energy efficiency by adding actual electricity and fuel use and subtracting the predicted levels from equation (6) above.

$$\left(E_{i,t} \cdot C + F_{i,t}\right) - \left(\widehat{E_{i,t}} \cdot C + \widehat{F_{i,t}}\right) = \widehat{\epsilon_{i,t}} \qquad (7)$$

C is the unit conversion of electricity to source MMBtu, since the model estimate for electric efficiency is in units other than source MMBtu.

To allow for the possibility that the distribution of $\epsilon_{i,t}$ from equation (7) is not a simple normal distribution, we estimate the distribution non-parametrically via a kernel density. Kernel density estimation is a flexible approach to computing the density function, similar in concept to a "smoothed histogram." The support points for the non-parametric estimate for the density of the plant level efficiencies, $\epsilon_{i,t}$, are then used to compute the cumulative distribution function via numerical integration over the support points. An example is shown in Figure 1 (Boyd and Lee 2016). The kernel density (blue) and associated distribution (orange) of a set of actual efficiency estimates is obviously not a normal distribution. The cumulative distribution can be converted to a lookup table for the percentile corresponding to any value of $\epsilon_{i,t}$.

Figure 1 Example of a Kernel Density and Associated Cumulative Distribution Estimate for Energy Efficiency in Metal Based Durables (source Boyd and Lee 2016)



4 FINAL MODEL ESTIMATES

This section presents the final equations used for the EPI, based on the methods and evolution described above. Stylized results that provide additional interpretation are also given.

4.1 Statistical Estimates and Elasticities

The final equations for electricity and fuels are shown in equations 8 and 9 and final results are shown in Table 2 below. All product mix ratios are compared to the electricity or fuels needed for pan breads, as this was omitted to avoid perfect multicollinearity. Elasticities are reported along with coefficients to more easily understand the impact each variable has on energy usage. Using elasticity values, a 1% increase in each variable would lead to an electricity/fuels increase equal to the given elasticity. For example, a 10% increase in raw dough would increase predicted electricity by 7.4% and fuels by 7.2%. For the ratio variables, the same applies, but any increases are in comparison to a similar increase to pan breads, the omitted product type. For example, a 10% increase in rolls would increase predicted electricity 0.43% more than a 10% increase in pan breads. For the air conditioned dummy variable, moving from 0 (no AC) to 1 (AC) would increase predicted electricity by 33%. The coefficient for amount of raw dough indicates economies of scale, as processing twice as much raw dough would only increase energy usage by approximately 70%, as opposed to 100% if there were not any benefits to larger plant sizes. These results are consistent with industry expectations regarding energy needed for different product types. Table 2 provides a summary of the variables and their impacts on electricity and fuel EPI models.

$$\begin{aligned}\ln(electricity) = {} & \alpha + \beta_1 \ln(rawdough) + \beta_2 \left({}^{total\ product\ wt}/_{rawdough}\right) + \\ & \beta_3 \left({}^{hearth\ breads}/_{total\ product\ wt}\right) + \beta_4 \left({}^{rolls}/_{total\ product\ wt}\right) + \\ & \beta_5 \left({}^{bagels}/_{total\ product\ wt}\right) + \beta_6 \left({}^{frozen\ dough}/_{total\ product\ wt}\right) + \beta_7 CDD + \\ & \beta_8 AirConditioned + \beta_9 \%\ of\ frozen\ product\ (exclude\ frozen\ dough) + \varepsilon \end{aligned} \tag{8}$$

$$\begin{aligned}\ln(fuels) = {} & \alpha + \beta_1 \ln(rawdough) + \beta_2 \left({}^{total\ product\ wt}/_{rawdough}\right) + \\ & \beta_3 \left({}^{hearth\ breads}/_{total\ product\ wt}\right) + \beta_4 \left({}^{english\ muffins}/_{total\ product\ wt}\right) + \\ & \beta_5 \left({}^{bagels}/_{total\ product\ wt}\right) + \beta_6 \left({}^{frozen\ dough}/_{total\ product\ wt}\right) + \beta_7 HDD + \varepsilon \end{aligned} \tag{9}$$

Table 2 Model Results and Variable Impacts

	Electricity Model			Fuels Model		
Variables	Coefficient	Standard Error	Elasticity	Coefficient	Standard Error	Elasticity
Log Raw Dough	0.744**	0.028	0.744	0.716**	0.0268	0.716
Baked Weight Ratio	-0.737*	0.365	-0.635	-0.485*	0.351	-0.418
Hearth Bread Ratio	0.257*	0.107	0.009	0.493**	0.099	0.0174

Rolls Ratio	0.143*	0.065	0.043	-	-	-
English Muffin Ratio	-	-	-	0.172*	0.096	0.00818
Bagel Ratios	-0.090	0.090	-0.0045	0.118*	0.083	0.006
frozdoughRatio2	0.572**	0.119	0.0190	-1.947**	0.105	-0.0648
Annual HDD	-	-	-	0.000069**	0.000	0.288
Annual CDD	0.000029**	0.000	0.0478	-	-	-
Air Conditioning Dummy	0.336**	0.062	0.336	-	-	-
Percentage of Frozen Product Excluding Frozen Dough	0.297**	0.093	0.0191	-	-	-
R-Squared	0.809			0.820		

** Significant at the 99% level; * Significant at the 90% level

4.2 Stylized Results

When only examining the raw data on energy intensity, we see that the range of performance is quite wide for both electricity and fuel usage. The EPI analysis shows that this observation taken by itself is actually misleading; after accounting for additional factors, the range of performance is much narrower. The red lines in Figures 2-4 take the raw energy intensity data and transform it into the kernel density distribution of plants that lie above or below the average energy electricity, fuel, and total intensity of 1.33 MMBtu/1000 pounds of raw dough, 0.91 MMBtu/1000 pounds of raw dough, and 2.25 MMBtu/1000 pounds of raw dough represented as a percent difference. The full range of intensity differences exceed 100% on either side for both energy types. The blue lines representing the kernel density from the EPI analysis tells a different story. Most of those differences come from differences that can be accounted for in the analysis, more or less of different product types, different climates, more frozen products, etc. The range of actual efficiency, after these differences are accounted for, is narrower. This is consistent with the results of a meta-analysis of EPI studies for other industries (Boyd 2016). In the commercial baking industry, the difference in total energy consumption between an "average" plant (score of 50) and an "efficient" plant (score of 75) is roughly 17%. This matches closely with other food processing industries in which energy costs are smaller than other production costs.

Figure 2 Comparing the Distribution of Electricity Intensity to Efficiency

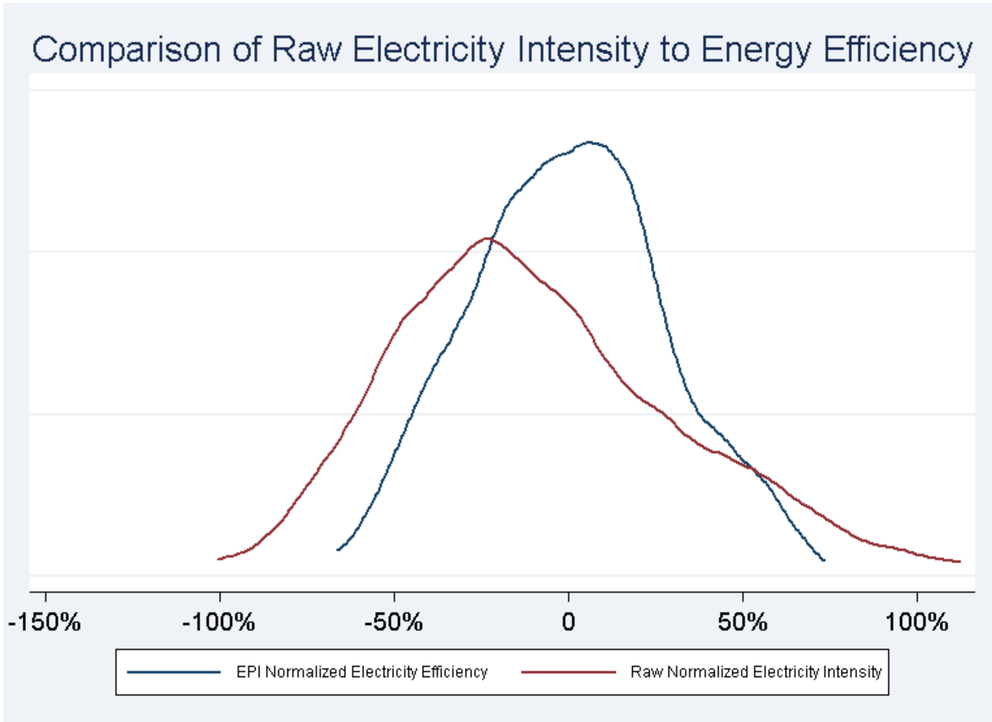


Figure 3 Comparing the Distribution of Fuels Intensity to Efficiency

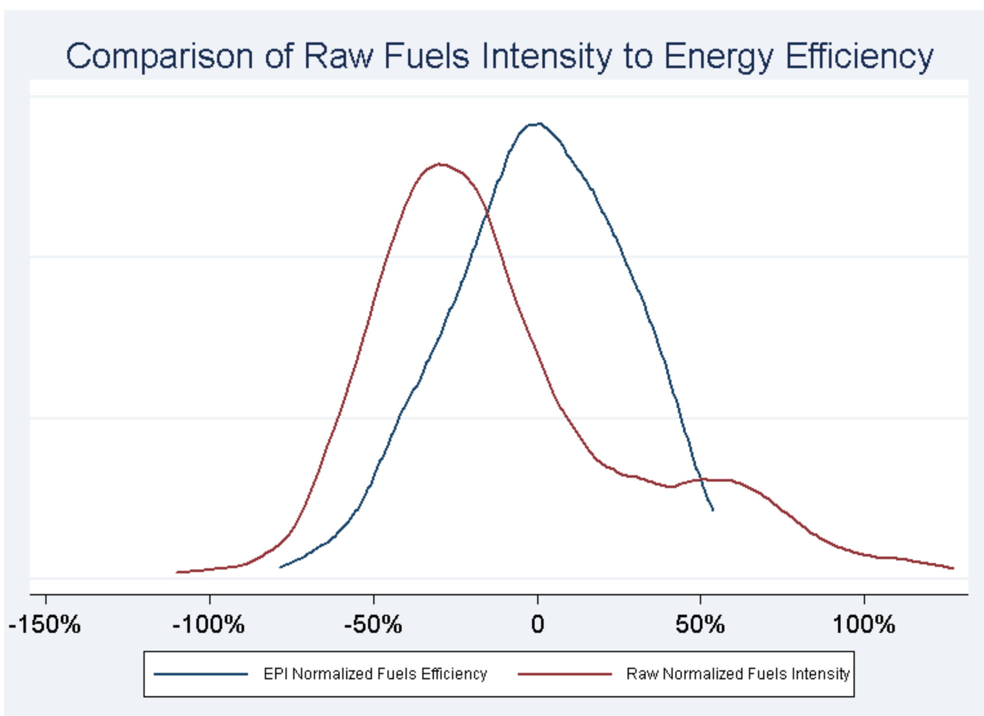
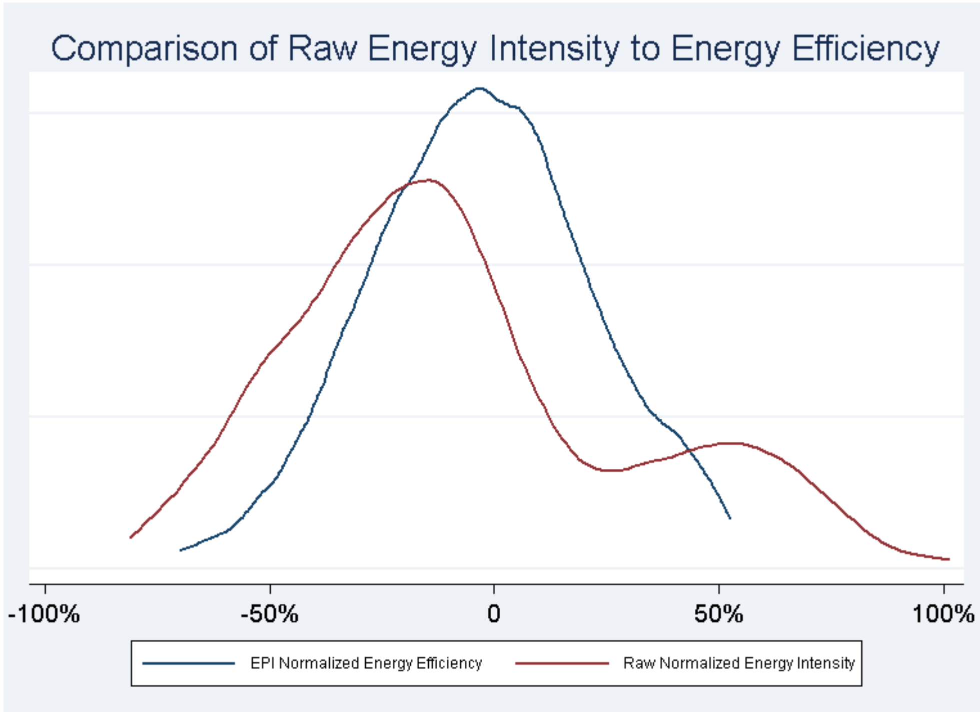


Figure 4 Comparing the Distribution of Total Energy Intensity to Efficiency



5 ASSESSING POTENTIAL COVERAGE AND BIAS IN THE EPI BAKING INDUSTRY DATABASE

Since this EPI is based on a set of data that was voluntarily supplied by the industry participants, for the EPI to be useful to those that may not have directly participated in the Focus, the ***coverage*** and possible ***bias*** of this data set should be determined. Additionally, for EPA to use the EPI to determine eligibility for ENERGY STAR certification, the EPI's underlying data set must be evaluated for industry coverage and bias. ***Coverage*** addresses whether the data represents a large enough portion of the industry to be meaningful. Anecdotally, the American Bakers Association and industry participants felt that most of the large companies and many of the smaller ones were involved. Our analysis provides supporting evidence that this appears to be the case. ***Bias*** addresses the concern that, even if coverage is considered adequate, the fact that coverage will not be 100% means it is important to determine if some part of the industry is over- or under-represented. Preliminary analysis for coverage and bias was conducted early in the study but yielded uncertain results and was delayed until the final model was completed and approved by industry. This final analysis addresses two possible sources of biases: product mix and plant size. The analysis below indicates that product mix bias is not a concern, at least at the level at which we are able to measure. However, the data are clearly skewed to the larger plants and so a size cutoff is necessary for ensuring accurate benchmarking and for ENERGY STAR certification. The EPI is therefore considered applicable to Large Commercial Bakeries (defined below) and ENERGY STAR certification would be offered only to plants that meet the size minimum. While the analysis shows that this cutoff may eliminate many very small establishments, the EPI is still applicable to the

majority of industry production. The following describes the data and methods used to assess the coverage and bias issues.

The basis for the coverage and bias analysis is the published data from the 2012 Economic Census[3] (NAICS Code 311812). This analysis is complicated by the fact that Census reports plant data on employment and sales, and the EPI data set includes production in physical units, specifically raw dough processed. In order to make a comparison, an estimate of raw dough is derived from the sales data.

To determine coverage for our sample of commercial bakeries compared to the larger collection of Census data, plant sizes and product volume were considered. Using Census published data on value of shipments (i.e. sales) from different sized commercial bakeries (by employment counts) and a product cost estimate (i.e. average cost per pound), an average product weight amount per plant in each size category was estimated for each size commercial bakery designated by total employees. An example of a Census size category would be all plants with between 20 to 49 employees, 50 to 99 employees, 100 and 249 employees, etc. These estimates of physical production, by size category, helped determine where our sample fits in relation to the Census facilities. The following described the step-by-step process:

1. For each standard Census size category, based on employment counts, the total sales value was divided by the number of establishments in that category to yield the average sales per plant.
2. For each size category, the average sales value per plant was divided by the average price per lb. of white bread[4] ($2.30) and then divided by the loss factor of 0.9 to yield the average baked weight/average raw dough weight. This results in an estimate of average raw dough processed per plant in each category.
    a. One might argue that this price is too high because commercial level sales would be at a lower "wholesale price."
    b. However, some products are likely more expensive on a per lb. basis than simple white bread, so it is also possible that the opposite is true, and the price is too low.
    c. Given that both a. and b. could be true, the choice of white bread as the benchmark price was used and is intended to keep the analysis simple and straightforward, but still be seen a reasonable "average price."
3. The average raw dough per plant for the different sizes is compared to the smallest and largest plants in the EPI data set. We found that:
    a. The smallest plants in the EPI data were consistent with the Census size category of 50-99 employees.
    b. The rest of the EPI sample was consistent with the larger Census plant categories.
4. Taking the raw dough per plant in each size category and multiplying by the number of plants in that category yields an estimate of raw dough production in each category and for the entire industry.

[3] Information from the industry trade "Red Book" publication was also examined but the way those data were compiled was not conducive to a clear comparison. The Red Book is also self-reported data that covers other aspects of the industry beyond those product types the EPI is focused on.
[4] Source: http://www.numbeo.com/cost-of-living/country_result.jsp?country=United+States accessed on 2/27/2016

Since the EPI data appear to reflect Census establishments with 50+ employees, we then looked at the Census sales number for plants with more (and less) than 50 employees. Although the majority of the plants in the Census data are below 50 employees, plants with greater than 50 employees account for 87% of the total value of shipments in the industry. This is typical of many industries; while there may be many very small establishments, particularly in an industry like baking, large scale production dominates total industry output. Comparing the raw dough estimate from the Census data for plants with more than 50 employees with the total EPI production data in the same year, the EPI sample accounts for 78% of the raw dough produced at facilities with greater than 50 employees nationwide. This implies that the EPI sample covers 68% of the industry as a whole (78% * 87%).

As expected, the EPI sample is weighted towards the large-scale commercial baking facilities. To address this bias in the EPI data and for ENERGY STAR certification requirements, a minimum production value of 8 million pounds of raw dough is established. This value is based on the average of the bottom 5% of the EPI dataset.[5] For comparison purposes, the average raw dough production in the Census category of 50-99 employees was 8.4 million lbs.

To assess possible bias in product mix, the analysis was simpler. When broadly looking at product mixes, the sample of industry supplied data had more categories than the Census information. Census product data were split into three types: bread, rolls, and other categories including pies and cakes that were not part of the EPI focus. When just looking at the Census data on total value from bread and rolls, the breakdown equates to roughly 60% bread (wheat, white, etc. and frozen) and 40% rolls (muffins, bagels, croissants, and frozen). To replicate this number within the EPI sample, pan and hearth breads were combined and rolls, muffins, and bagels were combined. Frozen dough was split between the two groups. In this case, the product mix from the sample nearly mirrors the available Census product mix (see Table 3 below). This indicates that the EPI sample is representative from a product mix standpoint and is not biased towards plants that produce specific products.

Table 3 EPI Sample and Census Product Mixes

Product type	EPI Sample	Census
Pan Breads	54%	-
Hearth Breads	2%	-
Rolls	29%	-
English Muffins	4%	-
Bagels	2%	-
Frozen Dough	6%	-
Other	2%	-
Bread (white, wheat, rye, including frozen)	59%	59%
Rolls (muffins, bagels, croissants) including frozen	38%	41%

6 SCORING COMMERCIAL BAKING PLANT EFFICIENCY

[5] Computed this way to avoid possible disclosure of confidential values for a single plant.

This section describes the spreadsheet tool that is created based on the above analysis. Suggestions for how to use the tool and interpret the results are also shown below.

6.1 How the EPI Works

The commercial baking EPI scores the energy efficiency of commercial bakeries that process more than 8 million pounds of raw dough annually. To use the tool, the following information must be available for a plant.

- Total energy use
    - Electricity (converted to source MMBtus by the spreadsheet tool)
    - Fuel use for all fuel types in physical units or MMBtu (converted to source MMBtus by the spreadsheet tool)
- Final Product Weights in lbs.
    - Raw dough processed
    - Pan breads produced
    - Hearth breads produced
    - Rolls produced
    - English muffins produced
    - Bagels produced
    - Frozen dough produced
    - Other products
- Weather
    - Heating degree days (HDD)
    - Cooling degree days (CDD)
- Plant Operations
    - Is the production area air conditioned?
    - What percent of total products are frozen?

Based on these data inputs, the EPI will report an Energy Performance Score (EPS) for the plant in the current time period that reflects the relative energy efficiency of the plant compared to that of the industry. The EPS is a percentile score on a scale of 1–100. An EPS of 75 means a particular plant is performing better than 75% of the plants in the industry, on a normalized basis. ENERGY STAR defines the 75th percentile as the benchmark for efficiency, so plants that score 75 or better are classified as efficient. The model also estimates what the energy use would be for an "average" plant (defined as the 50th percentile) with the same production characteristics. This overall score is complemented with similar efficiency scores for electricity and fuels consumption. While the underlying model was developed from industry-supplied data, it does not contain or reveal any confidential information.

6.2 Spreadsheet Tool

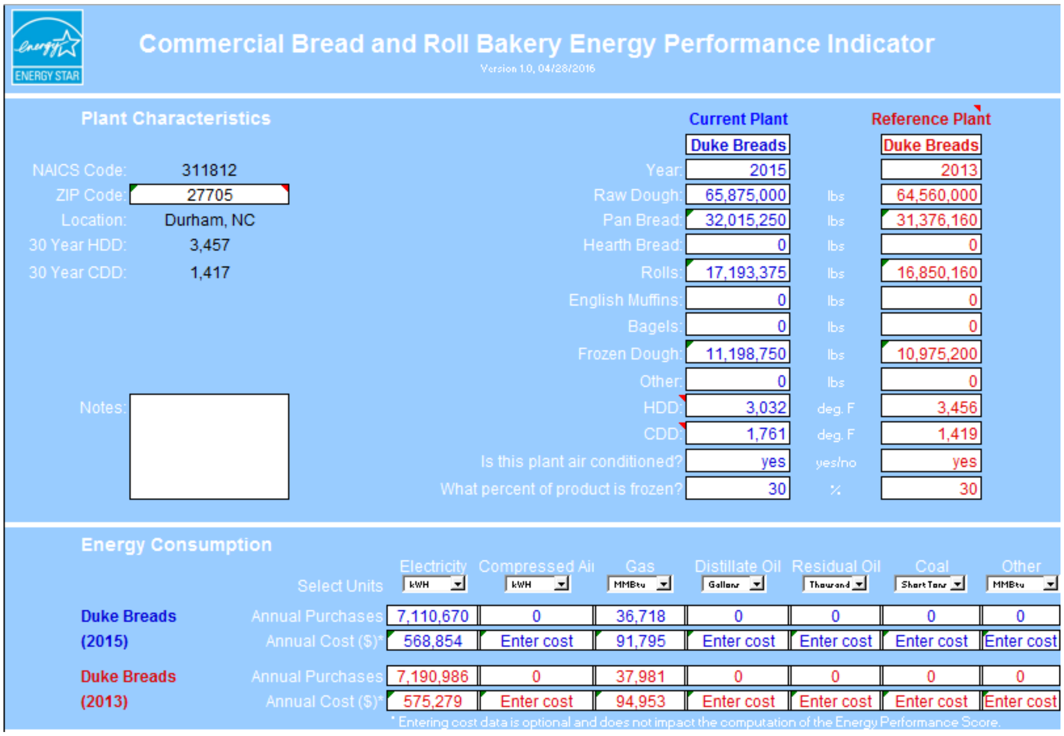
To facilitate the review and use by industry energy managers, a spreadsheet was constructed to display the results of the EPI for an arbitrary[6] set of plant-level inputs. Energy managers were

[6] In other words, for plant data that may not originally have been in the data set used to estimate the model equations.

encouraged to input data for their own plants and then provide comments. A version of this spreadsheet corresponding to the results described in this report is available from the EPA ENERGY STAR web site.[7] Example inputs and outputs of the spreadsheet tool are shown in Figures 5-8.

[7] http://www.energystar.gov/epis

Figure 5 Input Section of the Commercial Baking EPI Spreadsheet Tool



ENERGY STAR

Commercial Bread and Roll Bakery Energy Performance Indicator

Version 1.0, 04/28/2016

Plant Characteristics

NAICS Code:	311812
ZIP Code:	27705
Location:	Durham, NC
30 Year HDD:	3,457
30 Year CDD:	1,417
Notes:	

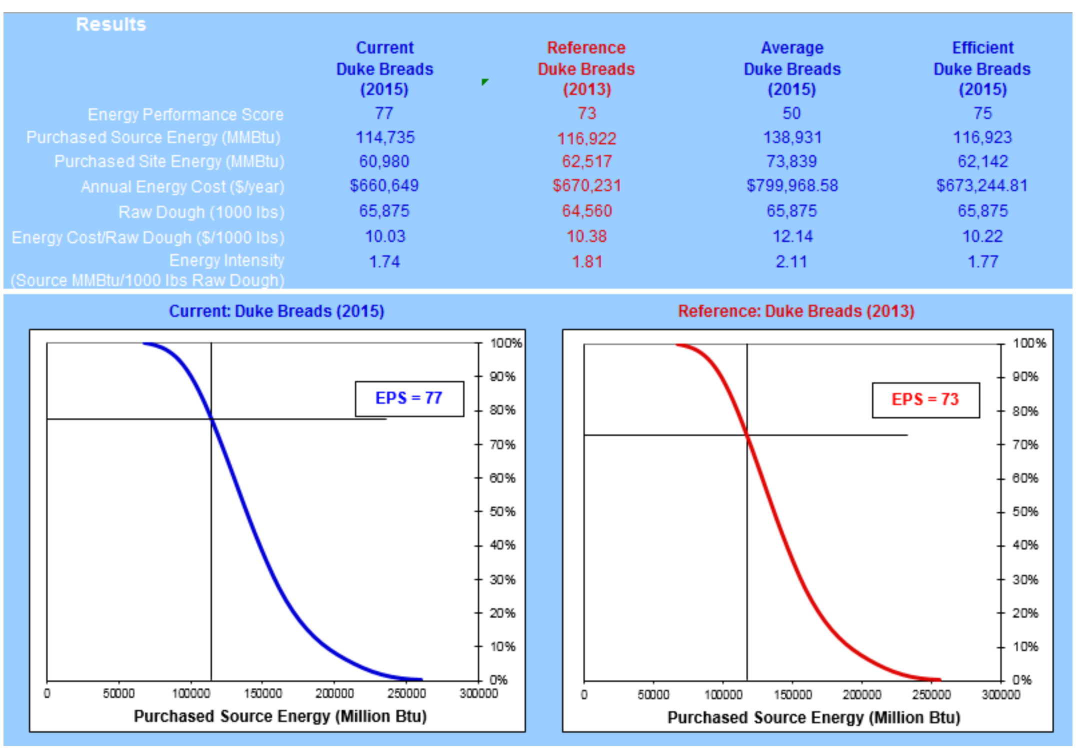
	Current Plant		Reference Plant
	Duke Breads		Duke Breads
Year	2015		2013
Raw Dough	65,875,000	lbs	64,560,000
Pan Bread	32,015,250	lbs	31,376,160
Hearth Bread	0	lbs	0
Rolls	17,193,375	lbs	16,850,160
English Muffins	0	lbs	0
Bagels	0	lbs	0
Frozen Dough	11,198,750	lbs	10,975,200
Other	0	lbs	0
HDD	3,032	deg. F	3,456
CDD	1,761	deg. F	1,419
Is this plant air conditioned?	yes	yes/no	yes
What percent of product is frozen?	30	%	30

Energy Consumption

		Electricity	Compressed Air	Gas	Distillate Oil	Residual Oil	Coal	Other
	Select Units	kWH	kWH	MMBtu	Gallons	Thousand	Short Tons	MMBtu
Duke Breads (2015)	Annual Purchases	7,110,670	0	36,718	0	0	0	0
	Annual Cost ($)*	568,854	Enter cost	91,795	Enter cost	Enter cost	Enter cost	Enter cost
Duke Breads (2013)	Annual Purchases	7,190,986	0	37,981	0	0	0	0
	Annual Cost ($)*	575,279	Enter cost	94,953	Enter cost	Enter cost	Enter cost	Enter cost

* Entering cost data is optional and does not impact the computation of the Energy Performance Score.

Figure 6 Output Section of the Commercial Baking EPI Spreadsheet Tool



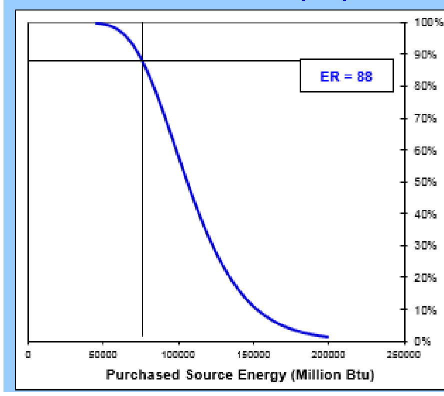
Results

	Current Duke Breads (2015)	Reference Duke Breads (2013)	Average Duke Breads (2015)	Efficient Duke Breads (2015)
Energy Performance Score	77	73	50	75
Purchased Source Energy (MMBtu)	114,735	116,922	138,931	116,923
Purchased Site Energy (MMBtu)	60,980	62,517	73,839	62,142
Annual Energy Cost ($/year)	$660,649	$670,231	$799,968.58	$673,244.81
Raw Dough (1000 lbs)	65,875	64,560	65,875	65,875
Energy Cost/Raw Dough ($/1000 lbs)	10.03	10.38	12.14	10.22
Energy Intensity (Source MMBtu/1000 lbs Raw Dough)	1.74	1.81	2.11	1.77

Figure 7 Electricity Output Section of the Commercial Baking EPI Spreadsheet Tool

Electricity Results	Duke Breads (2015)	Duke Breads (2013)	Average Duke Breads (2015)	Efficient Duke Breads (2015)
Electric Rating	88	85	50	75
Annual Electricity Cost ($/year)	$568,854	$575,279	$789,333	$654,440
Raw Dough (1000 lbs)	65,875	64,560	65,875	65,875
Electricity Cost/Raw Dough ($/1000 lbs)	8.64	8.91	11.98	9.93
rchased Electricity (Source MMBtu):	76,181	77,042	105,708	87,643
Purchased Electricity (Site MMBtu):	24,262	24,536	33,665	27,912
Purchased Electricity (kWh):	7,110,670	7,190,986	9,866,658	8,180,505
Purchased Electric Energy Intensity (Source MMBtu/1000 lbs raw dough)	1.16	1.19	1.60	1.33
Purchased Electric Energy Intensity (kWh/1000 lbs raw dough)	107.94	111.38	149.78	124.18

Current: Duke Breads (2015)



Reference: Duke Breads (2013)

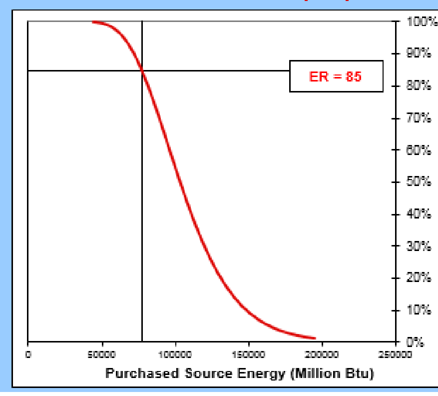
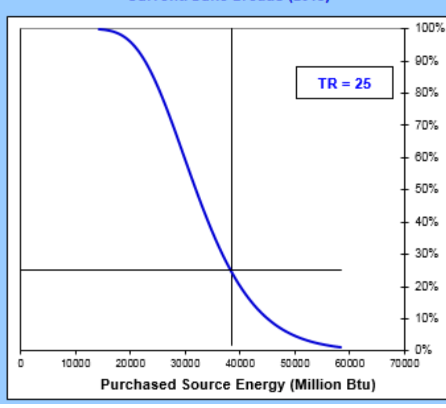


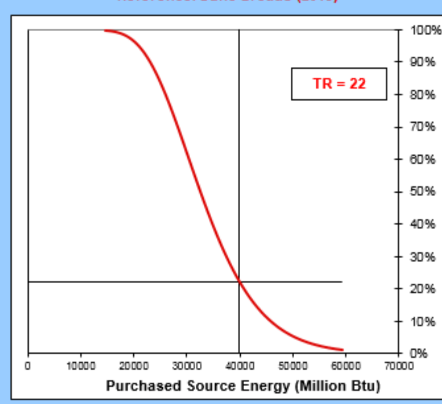
Figure 8 Thermal Output Section of the Commercial Baking EPI Spreadsheet Tool

Thermal (Fuels) Results	Duke Breads (2015)	Duke Breads (2013)	Average Duke Breads (2015)	Efficient Duke Breads (2015)
Thermal Rating	25	22	50	75
Annual Thermal Energy Cost ($/year)	$91,795	$94,953	$76,209	$63,482
Raw Dough (1000 lbs)	65,875	64,560	65,875	65,875
Thermal Energy Cost/Raw Dough ($/1000 lbs)	1.39	1.47	1.16	0.96
Purchased Thermal (Source MMBtu)	38,554	39,880	32,008	26,663
Purchased Thermal (Site MMBtu)	36,718	37,981	30,484	25,393
Source Thermal Energy Intensity (Source MMBtu/1000 lbs raw dough)	0.59	0.62	0.49	0.40

Current: Duke Breads (2015)



Reference: Duke Breads (2013)



6.3 Use of the ENERGY STAR Commercial Bread and Roll Bakery EPI

EPIs are developed to provide industry with a unique metric for evaluating energy performance that will lead plants to take new steps to improve their energy performance. To promote the use of EPIs, EPA works closely with the manufacturers through an ENERGY STAR Industrial Focus on energy efficiency in manufacturing to promote strategic energy management among the companies in this industry. The EPI is an important tool that enables companies to determine how efficiently each of the plants in the industry is using energy and whether better energy performance could be expected. The EPI and the Energy Performance Score also serve as the basis for ENERGY STAR recognition. Plants that score a 75 or higher become eligible for ENERGY STAR certification.

EPA recommends that companies use the EPIs on a regular basis. At a minimum, it is suggested that corporate energy managers benchmark each plant on an annual basis. A more proactive plan would provide for quarterly use (rolling annual basis) for every plant in a company. EPA suggests that the EPI score be used to set energy efficiency improvement goals at both the plant and corporate levels. The EPIs also can be used to inform new plant designs by establishing energy intensity targets.

The models described in this report are based on the performance of the industry for a specific period of time. One may expect that energy efficiency overall will change as technology and business practices change, so the models will need to be updated. EPA plans to update these models every few years, contingent on newer data being made available and industry use and support of the EPI tools.

6.4 Steps to Compute a Score

All of the technical information described herein is built into spreadsheets available from EPA (http://www.energystar.gov/epis). Anyone can download, open the EPI spreadsheets, and enter, update, and manage data as they choose. The following details each step involved in computing an EPS for a plant.

1. *User enters plant data into the EPI spreadsheet*

- Complete energy information includes all energy purchases (or transfers) at the plant for a continuous 12-month period. The data do not need to correspond to a single calendar year.
- The user must enter specific operational characteristic data. These characteristics are those included as independent variables in the analysis described above.

2. *EPI computes the Total Source Energy (TSE) Use*

- TSE is computed from the metered energy data.
- The total site energy consumption for each energy type entered by the user is converted into source energy using the site-to-source conversion factors.
- TSE is the sum of source energy across all energy types in the plant.
- TSE per relevant unit of production is also computed.

3. EPI computes the Predicted "Best Practice"[8] TSE

- Predicted "Best Practice" TSE is computed using the methods above for the specific plant.
- The terms in the regression equation are summed to yield a predicted TSE.
- The prediction reflects the expected minimum energy use for the plant, given its specific operational characteristics.

4. EPI compares Actual TSE to Predicted "Best Practice" TSE

- A lookup table maps all possible values of TSE that are lower than the Predicted "Best Practice" TSE to a cumulative percent in the population.
- The table identifies how far the energy use for a plant is from best practice.
- The lookup table returns a score on a scale of 1-to-100.
- The Predicted TSE for a median and 75th percentile plant is computed based on the plant-specific characteristics.
- A score of 75 indicates that the building performs better than 75% of its peers.
- Plants that earn a 75 or higher may be eligible to earn the ENERGY STAR.

[8] The model computes the "best practice" for frontier models and "average practice" for ordinary least squares. Steps 3 and 4 are similar for the OLS models, except that the prediction is for the average energy use and the percentiles are relative to the average (i.e., 50th percentile).

7 REFERENCES

Boyd, G., E. Dutrow and W. Tunnessen (2008). "The Evolution of the Energy Star Industrial Energy Performance Indicator for Benchmarking Plant Level Manufacturing Energy Use." Journal of Cleaner Production, Volume 16 (Issue 6): 709-715.

Boyd, G. and J. Lee (2016). Measuring Plant Level Energy Efficiency and Technical Change in the U.S. Metal-Based Durable Manufacturing Sector Using Stochastic Frontier Analysis. 6th International Symposium on Energy Challenges & Mechanics (ISECM). Inverness, Scotland, United Kingdom.

Boyd, G. A. (2016). "Comparing the statistical distributions of energy efficiency in manufacturing: meta-analysis of 24 Case studies to develop industry-specific energy performance indicators (EPI)." Energy Efficiency: 1-22.

U.S. Environmental Protection Agency (2003). Guidelines for Energy Management. Washington, DC, U.S. Environmental Protection Agency. Available at: www.energystar.gov/guidelines.

U.S. Environmental Protection Agency (2015). Reducing Greenhouse Gas Emissions by Advancing Industrial Energy Efficiency 2000—2015. EPA Document Number 430-R-15-006.