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

DEVELOPMENT OF ENERGY STAR® ENERGY PERFORMANCE INDICATOR FOR FLUID MILK PROCESSING

GALE A. BOYD, MATT DOOLIN, YIFANG GAO, AND SU ZHANG

DUKE UNIVERSITY, SOCIAL SCIENCE RESEARCH INSTITUTE

BOX 90989, DURHAM, NC 27708

AUGUST 30, 2018

ACKNOWLEDGMENTS

This work was sponsored by the United States Environmental Protection Agency (U.S. EPA), Office of Atmospheric Programs, Climate Protection Partnerships Division. The research benefited from the comments of many company representatives participating in the *ENERGY STAR Focus on Energy Efficiency in Dairy Processing,* as well as input from the International Dairy Food Association (IDFA). We would also like to thank Walt Tunnessen (U.S. EPA) and Joshua Smith (ICF) for their helpful comments.

The results presented in this report were prepared while the authors were Special Sworn Status researchers at the Triangle Research Data Center, a member of the Federal Statistical Research Data Center Network. All results presented here were reviewed and do not reveal any confidential information. The results do not reflect the views of the Census Bureau.

CONTENTS

DEVELOPMENT OF ENERGY STAR® ENERGY PERFORMANCE INDICATOR FOR FLUID MILK PROCESSORS

GALE A. BOYD, MATT DOOLIN, YIFANG GAO, AND SU ZHANG

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 fluid milk processing industry to provide a plant-level indicator of energy efficiency for facilities that produce various types of fluid milk 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, including 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 an EPI for plants within the fluid milk processing industry, specifically those that produce fluid milk, cottage cheese, yogurts, creams, etc. 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[1]

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 identify the components of a successful energy management program (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;
- Appoint a corporate energy director to coordinate and direct the energy program and multi-disciplinary energy team;
- Establish and promote an energy policy;
- Develop 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;
- Assess performance and set goals at the corporate, facility, and subunit levels;
- Create action plans across all operations and facilities, as well as monitor successful implementation and promote 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

[1] The introductory, background material presented in this report follows substantially from prior documentation for other studies of industry energy efficiency. For early examples, see Boyd, G. A. (2005). Development of a Performance-based Industrial Energy Efficiency Indicator for Automobile Assembly Plants. Argonne IL, Argonne National Laboratory**:** May 2005, Boyd, G. A. (2006). Development of a Performance-based Industrial Energy Efficiency Indicator for Cement Manufacturing Plants. Argonne IL, Argonne National Laboratory.

companies to determine whether better energy performance could be expected. It empowers managers to set more informed goals and evaluate their reasonableness.

(Boyd, Dutrow et al. 2008) describe the evolution of a statistically based plant energy performance indicator (EPI) for the purpose of benchmarking manufacturing energy use for ENERGY STAR. Boyd (2016) describes 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 a performance-based energy indicator for the dairy 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 energy 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. A handful of industry associations, such as the Portland Cement Association, provide energy use comparisons to their members. 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 are performing.

In 2000, EPA began developing a method for producing 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 carefully defined industry, create a statistical model of energy performance for the industry's plants based on these data along with other available data sources for the industry, and establish an energy performance benchmark for the industry. The primary data sources would be the Census of Manufacturing, Annual Survey of Manufacturers, 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. Since then, EPA's ENERGY STAR program has coordinated the development of multiple EPIs across a wide variety of industrial sectors.

3 EVOLUTION OF THE FLUID MILK PROCESSING EPI

The EPA was approached by the International Dairy Foods Association (IDFA), which wanted to partner with the ENERGY STAR program to launch an "Industrial Focus" for the dairy sector. In 2010, EPA launched the ENERGY STAR Focus on Energy Efficiency in Dairy Processing (Dairy Focus) to advance energy management and energy-efficiency within the broader dairy industry.[2] During the first phase of the Dairy Focus, EPA and IDFA encouraged the industry to participate in the ENERGY STAR Challenge for Industry and contribute to the development of an "energy guide."

In 2011, EPA released an energy guide entitled *Energy Efficiency and Cost Saving Opportunities for the Dairy Processing Industry (Brush, Masanet et al. 2011)* that highlights cost saving energy efficiency opportunities throughout dairy facilities at the process level. The Dairy Focus then turned its attention to developing an Energy Performance Indicator (EPI).

3.1 Initial Scope and Sector Breakdown

The highly specialized nature of dairy facilities quickly led to the conclusion that one EPI could not serve the entire industry. Therefore, EPIs were planned for fluid milk processing facilities and ice cream producers. After examining data from the 2002 Census of Manufacturers, product types that could be used in the fluid milk processing model were identified. Starting with the broad Census Bureau definition of dairy product manufacturing under the NAICS code 3115, the large number of identified products were reduced in the first draft of the EPI to only include fluid milk products, which include semi soft cultured dairy products like yogurt and cottage cheese, but not frozen or hard products (cheese). Since many participating dairy processors in the Dairy Focus also manufactured juice at fluid milk facilities, this product was added to the model as well. After early reviews, a weather variable was added to account for the geographical differences of the plants. The first sample included 258 plants. The following variables were included in the model to statistically account for differences in energy usage:

- Total Employees
- Production Worker Hours
- Total Milk Inputs
- Value Share of Total Production of All Other Products in the Plant
- Share of Juice Production
- Share of Bulk Milk Production
- Share of Packaged Milk Production
- Share of Cottage Cheese Production
- Share of Yogurt Production
- Share of Ultra High Temperature (UHT) Milk Production

[2] See: www.energystar.gov/dairyprocessingfocus for more information.

- Cooling Degree Days

Given the large number of products, the initial EPI model used a new approach that incorporated a binary variable as well as a share variable for each product type. These binary variables indicated whether the plant made that product at all, regardless of the amount. The idea was that different products have different requirements that would affect fixed energy consumption due to the equipment or steps needed to make the product. For example, some products may need different refrigeration levels or require more process steps. Since data were not available for the exact energy usage for specific product processes, the binary variables were intended to address the energy differences. The share variables then would account for differences in quantities of the product types. Early results of the model indicated that cottage cheese was the most energy intensive product and that bulk milk and juice were the least energy intensive. The variables "total employment" and "production worker hours" were used as proxies for plant utilization and captured the differences in downstream product processing. Since greater processing requires more time, and more energy, in the plant, more worker hours per product would be expected. This labor impact on energy consumption was debated during the model testing process. Some companies either felt the impact of these variables was too large or did not agree with the statistical results showing the significant relationship between labor values and energy. Model results with production labor variables removed showed a clear case of omitted variable bias as the error term increased. However, the goal was to benchmark energy use, and it was decided that including labor productivity in the energy benchmark could be misleading. To correct for this and adjust the impact of labor on predicted energy values, production worker hours and total employees were set to their respective sample mean intensities (i.e., average labor productivity) to simulate the distribution of energy efficiency based on the average labor intensity using the OLS regression model that included labor. The impact from these variables is now set to the total production values multiplied by the mean intensities and the difference from the industry mean values. Due to this change in the construction of the efficiency distribution, kernel density points had to be created to obtain a predicted energy distribution and is detailed more in the section 3.4 below.

3.2 Energy from Packaging

Following the completion of the first draft EPI, the industry was concerned that the EPI did not account for energy differences between different packing and container types. If everyone used the same type of packaging, the model would not need an adjustment as the amount of energy used for packaging would simply be based off their total production, which was included in the original analysis. However, many different container types are used, including paper cartons, glass bottles, plastic jugs (manufactured both on and off site), and no packaging at all for bulk milk products that leave the facility in tankers. Differences in container mix could cause different energy use and resulting scores for similar facilities and would need to be accounted for in the final model.

Although the Census Bureau did not collect data on the quantity of containers shipped or purchased, data on the material cost for each different packaging type was available. Material costs for purchased paperboard, glass, metal, plastic jugs, and plastic resins (presumably used to blow mold plastic jugs in house) were added into the model as value shares of total packaging costs. Ideally, the cost shares are consistent with the physical quantity shares. For example, a facility that has a 75% cost share for plastic jugs likely ships 75% of their final products in plastic jugs. Industry feedback determined this assumption was reasonable although not exact, and packing material cost data were used to account for energy differences in product packaging.

The first model utilizing packaging cost shares produced results that some industry reviewers indicated were counterintuitive. Namely paperboard packaging was more energy intensive than plastic blow molding. Even though the difference was small, the group began to explore why the statistical model would yield these results. The results are in the appendix. When the model was re-estimated using data from 2012, blow molding was found to be more energy intensive, with paper and purchased jugs having similar energy impacts. It is important to note that these differences were small, but the 2012 results presented below, are consistent with industry expectations.

3.3 Data Update and Final Model

During the lengthy testing and discussion phase of the model, more recent Census data became available and it was decided to update the model to 2012 data. A simpler approach was taken for the updated draft by removing the binary variables for each product type, only leaving the product share variables. Additionally, the newer data combined multiple products into a single variable without quantity values. Quantities were imputed using inflation-adjusted price estimations based off previous Census data. Finally, new data on milk substitute products, including soy, almond, and rice milks that represented a small amount of the industry but are very energy intensive, were included in the analysis.

The same packaging variables were included in the new model but the results were different. Blow molding was more energy intensive than paperboard containers, agreeing with industry expectations, although the difference between the two was not statistically significant.[3] Changes to the packaging coefficients addressed the main concerns that the industry had but an additional problem arose: bulk milk products do not use packaging and bulk milk specialty plants were receiving inflated scores due to a higher predicted energy value that included energy attributed to packaging. To correct this, the packaging value share variables were interacted with the percentage of products that were not listed as bulk product. For example, a bulk milk specialty plant would have a value of zero for the share of non-bulk products – i.e., 100% would be bulk milk and 0% would be non-bulk – that would be interacted with the packaging coefficients, meaning there would not be any expected energy from packaging. A facility that produced 10% bulk milk would interact the applicable packaging coefficients with 0.90, meaning only 90% of their product is packaged and less energy would be expected. After these adjustments, the final model was ready with a sample comprised of 160 observations. Generally,

[3] This means that the estimates of the energy in blow molding and paperboard were not sufficiently precise to say that the energy use was different with a high degree of confidence, not that packaging didn't impact the energy use.

the final sample consists of the larger processors in the country and covers approximately 59% of fluid milk processors in the United States.[4] More details on the statistical methods used are detailed below.

3.4 Statistical Approach

The methodology underlying this analysis presumes that there is some reduced form of relationship between plant-level energy use and the various plant input and output characteristics examined above. We assume that this relationship can be approximated by a functional form that is amenable to statistical estimation using data from a cross section or panel of plants within some "reasonably defined" industry group, in this case fluid milk processors. Depending on the form of the statistical model, discussed in more detail below, the actual plant energy use can then be compared to the predicted average, given the plant's characteristics. How far the actual energy use is above or below the predicted average is the plant's measure of efficiency. In statistical terms, the difference between actual and predicted energy use is equal to the residual of the statistical model for plants that are in the sample; alternatively, this difference is an out-of-sample prediction when the statistical model is applied to other data. It is in this out-of-sample context that we expect the model to be most often used, i.e., to compute energy efficiency using data for plant-level operations that were not in the statistical analysis, possibly from a different year. If that is the case, then the model is measuring current performance against a prior "benchmark year." If we further assume that the estimated distribution of efficiency from the statistical model is static, then the out-of-sample prediction of efficiency can be converted to a percentile (ranking) of efficiency based on the estimated distribution. The approach applied here is similar to guidance from ISO 50001 regarding the creation of EnPI,[5] although the ENERGY STAR EPI approach predates the release of ISO 50001 (Boyd, Dutrow, & Tunnessen 2008).

The concept of the analysis that supports the EPI can be easily described in terms of the standard linear regression model, which is reviewed in this section. Consider the example of a production process that has a fixed energy component and a variable energy component. A simple equation for this can be written as

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

where E is the measure of total source energy (total Btu of fuel use, plus electricity use converted to Btu based on average U.S. thermal plant efficiency including line losses), Y is either production or a vector of production-related activities, X is a vector of plant characteristics, β is a parameter vector (the normalization factors), and ε is the measure of relative plant efficiency.

Given data on energy use and production, the parameters can be fit via a linear regression model. Since the actual data may not be perfectly measured, and this simple relationship between energy and production is only an approximation of the "true" relationship, linear regression estimates of the parameters rely on the proposition that any departures in the plant data from Equation 1, which cannot be directly observed, are randomly distributed within the population and uncorrelated with the plant

[4] It should be noted that the data source used for the analysis includes all plants in the industry. A smaller sub-set of the data was used for the analysis since some of the data elements that were used for the model were not reported by every plant.
[5] Both ENERGY STAR and ISO 50001 use the term Energy Performance Indictor. Since ENERGY STAR began publically using the term first, ISO adopted the acronym "EnPI" to limit confusion.

production and characteristics. This strong assumption implies that the actual relationship includes a random error term ε that follows a normal (bell-shaped) distribution. For simplicity, we assume that the function f ( ) is linear in the parameters, but allow for non-linear transformations of the variables. In this case, production activity enters the equation in log form, as does the energy variable.

$$ln(E) = f(ln(Y), X; \beta) + \varepsilon \qquad (2)$$

$$ln(E) = a + \sum_{i=1}^{n} b_i \, ln(y_i) + \sum_{i=1}^{m} c_i \, X_i + \varepsilon \qquad (3)$$

This means that ε can be interpreted as percentage deviations in energy, rather than absolute. This has implications for the model results since we now think of the distributional assumptions in terms of percent, rather than absolute level. In either case of a linear or log-linear functional form, standard measures of statistical significance provide a test for whether or not to include a particular characteristic. In other words, one can test if two different plant characteristics have different energy implications, in a statistically identifiable way.

Energy Performance Score (EPS)

Assuming we are using a model that has been estimated in one of the case studies in the out-of-sample context described above, and we have data for a plant in a year different from the study data year, we can compute the difference between the actual energy use and the predicted average energy use from equation (3).

$$ln(E) - (a + \sum_{i=1}^{n} b_i \ln(y_i) + \sum_{i=1}^{m} c_i \, X_i) = \hat{e} \qquad (4)$$

For the models using ordinary least square (OLS), we have also 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 ε is normally distributed with zero mean and variance, σ^2, which is estimated via OLS.

$$\Pr(\varepsilon \geq \hat{e}) \qquad (5)$$

We take probability in (5) and subtract it from the value of one[6] and multiply by 100. This is the Energy Performance Score (EPS), and is the percentile ranking of the energy efficiency of the plant. Since this ranking is based on the distribution of inefficiency for the entire industry, but normalized to the specific systematic factors of the given plant, this statistical model allows the user to answer the hypothetical but very practical question, "How does my plant compare to everyone else's in my industry, if all other plants were similar to mine?"

The final equation for the fluid milk processing EPI is shown below in equation 6.

[6] We subtract the probability in (5) from 1 to reflect the fact that a low value of $\hat{e}$ is "good" and we want that to result in a higher EPS.

$$\ln(energy) = \alpha + \beta_1 \ln(total\ production) + \beta_2 \ln(production\ worker\ hours) + \beta_3 \ln(total\ employees) + \beta_4 CDD + \beta_5 \left(\frac{cottage\ cheese}{total\ production}\right) + \beta_6 \left(\frac{packaged\ and\ bulk\ milk}{total\ production}\right) + \beta_7 \left(juice\ product\ value/_{tvs}\right) + \beta_8 \left(\frac{dairy\ substitute\ value}{tvs}\right) + \left[1 - \left(\frac{bulk\ milk}{total\ production}\right)\right] * \left[\beta_9 \left(\frac{paper\ packaging\ value}{packaging\ expenses}\right) + \beta_{10} \left(\frac{resin\ and\ plastics\ value}{packaging\ expenses}\right) + \beta_{11} \left(\frac{purchased\ jug\ value}{packaging\ expenses}\right) + \beta_{12} \left(\frac{glass\ and\ metal\ packaging\ value}{packaging\ expenses}\right)\right] + \varepsilon \tag{6}$$

While the analysis clearly found a relationship between the labor inputs and energy use, the EPI uses the average labor productivity as the basis for the benchmark rather than excluding it from the regression model, to avoid omitted variable bias. The result is shown in equation 7, where $\overline{\left(\frac{production\ worker\ hours}{total\ production}\right)}$ and $\overline{\left(\frac{total\ employees}{total\ production}\right)}$ are the sample mean labor productivities.

$$\mathrm{Ln}(energy) = \alpha + \beta_1 \ln(total\ production) + \beta_2 \ln\left(\overline{\left(\frac{production\ worker\ hours}{total\ production}\right)} \cdot total\ production\right) + \beta_3 \ln\left(\overline{\left(\frac{total\ employees}{total\ production}\right)} \cdot total\ production\right) + \beta_4 CDD + \beta_5 \left(\frac{cottage\ cheese}{total\ production}\right) + \beta_6 \left(\frac{packaged\ and\ bulk\ milk}{total\ production}\right) + \beta_7 \left(juice\ product\ value/_{tvs}\right) + \beta_8 \left(\frac{dairy\ substitute\ value}{tvs}\right) + \left[1 - \left(\frac{bulk\ milk}{total\ production}\right)\right] * \left[\beta_9 \left(\frac{paper\ packaging\ value}{packaging\ expenses}\right) + \beta_{10} \left(\frac{resin\ and\ plastics\ value}{packaging\ expenses}\right) + \beta_{11} \left(\frac{purchased\ jug\ value}{packaging\ expenses}\right) + \beta_{12} \left(\frac{glass\ and\ metal\ packaging\ value}{packaging\ expenses}\right)\right] + \varepsilon \tag{7}$$

Since the ε in equation 7 is no longer normally distributed, we cannot use the estimated variance from the regression analysis to compute the probability in equation 5 to get the EPS. Instead, we use equation 7 to compute $\hat{e}$ for each observation in the sample, as in equation 4. A kernel density is then fit to these simulated plant efficiencies. The support points of the non-parametric kernel are numerically integrated to generate the cumulative function needed to compute the EPS.

4 FINAL MODEL ESTIMATES

This section presents the final results 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 Variable Impacts

Table 1 shows the average plant production, sales, energy, employment, and energy intensity. The standard deviations are fairly large, reflecting the diversity of size and energy performance in this sector.

Table 1 Plant average sample statistics for 2012

	Production (thousands lbs)	Value of Shipments (thousand $)	Total Energy (MMBtu)	Employment	MMBtu/lb
Plant Average	217,800	128,900	1,597	182	842
St. Dev.	159,100	105,000	1,535	112	685

Final parameters estimated are shown in Table 2. For interpretation, the variables can be grouped into logged variables and product/value ratios as the effect on energy usage for these are interpreted differently. Logged variables, including total production, production worker hours, and total employees, can be interpreted the same as elasticities. Using the coefficients, a 1% increase in each variable would lead to a total energy increase equal to the appropriate coefficient. For example, a 10% increase in total fluid milk production would increase the predicted energy use by 5.3% given the coefficient value of 0.525. Product mix variables are measured in unit increases as opposed to percentages, and any increases are in comparison to a similar increase in yogurts and “other” dairy products (creams, flavored milks, etc.). Energy coefficients for yogurts and other products were observed to be not significantly different from each other and were bundled together as the omitted product share to avoid perfect multicollinearity. For example, a 10% increase in cottage cheese production would increase predicted energy by 15.83 units more than a 10% increase to yogurt and other production, holding all other factors constant. Conversely, a 10% increase in packaged and bulk milk production would decrease predicted energy by 4.95 units. This negative coefficient implies that fluid packaged and bulk milk products are the least energy-intensive products in our sample, with cottage cheese being the most energy intensive. The same principle applies to the value shares for packaging. These results are consistent with industry expectations regarding energy needed for different product and container types. Table 3 provides a summary of the variable coefficients and their impacts on the EPI model.

Table 2 Fluid Milk Model Results

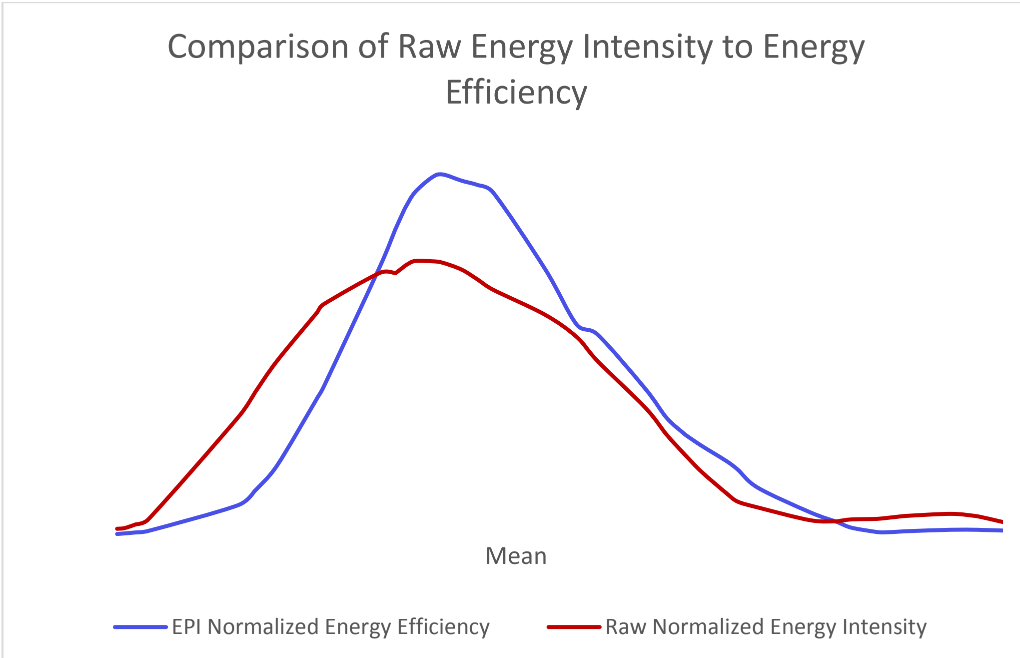
Variables	Coefficient	Standard Error
Log Total Quantity	0.525**	0.069
Cottage Cheese Product Share	1.583**	0.478
Packaged and Bulk Milk Product Share	-0.495**	0.117
Juice Products Value Share	1.093*	0.591
Dairy Substitutes Value Share	8.097*	3.381
Log Production Worker Hours	0.244**	0.0666
Log Total Employees	0.227**	0.078
Annual Cooling Degree Days (CDD)	0.0000201	0.00003
Paper Packaging Value Share	-0.203	0.165
Resin and Plastic Packaging Value Share	-0.182	0.154
Purchased Jug Value Share	-0.277*	0.147
Glass and Metal Packaging Value Share	1.158**	0.394
Employment / lb	0.0011	0.0007
Labor hours / lb	1.3	0.9
R-Squared	0.853	

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

4.2 Stylized Results

When examining the raw data on energy intensity (energy/total product) for fluid milk processing, the range of performance is quite wide. Table one shows that the energy intensity coefficient of variation, i.e., the standard deviation divided by the mean, is 0.8. In other words, the standard deviation is almost as large as the sample mean. The EPI analysis shows that this observation taken by itself is misleading; some differences in intensity are due to external factors and, after accounting for additional factors, the range of performance is narrower. Most of those differences come from differences that can be accounted for in the analysis, more or less of different product and packaging types, different climates, 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 fluid milk industry, the difference in total energy consumption between an "average" plant (score of 50) and an "efficient" plant (score of 75) is roughly 24%. These results show a lower spread than other food processing sectors, but still fall in line with many other previously studied industries and illustrate the opportunity for energy efficiency improvements. The red line in Figure 1 below takes the raw energy intensity data and transforms it into the kernel density distribution of plants that lie above or below the average total energy intensity, and are represented on a log scale. The blue line represents the kernel density from the EPI analysis, graphically showing the narrower range discussed above after additional differences in plant characteristics are accounted for.

Figure 1 Comparing the Distribution of Energy Intensity to Efficiency



5 SCORING FLUID MILK PLANT EFFICIENCY

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

5.1 How the Fluid Milk Processing EPI Works

The fluid milk processing plant EPI scores the energy efficiency of plants that produce fluid milk products including bulk or packaged milk, cottage cheese, and yogurts. To use the tool, the following information must be available for a plant:

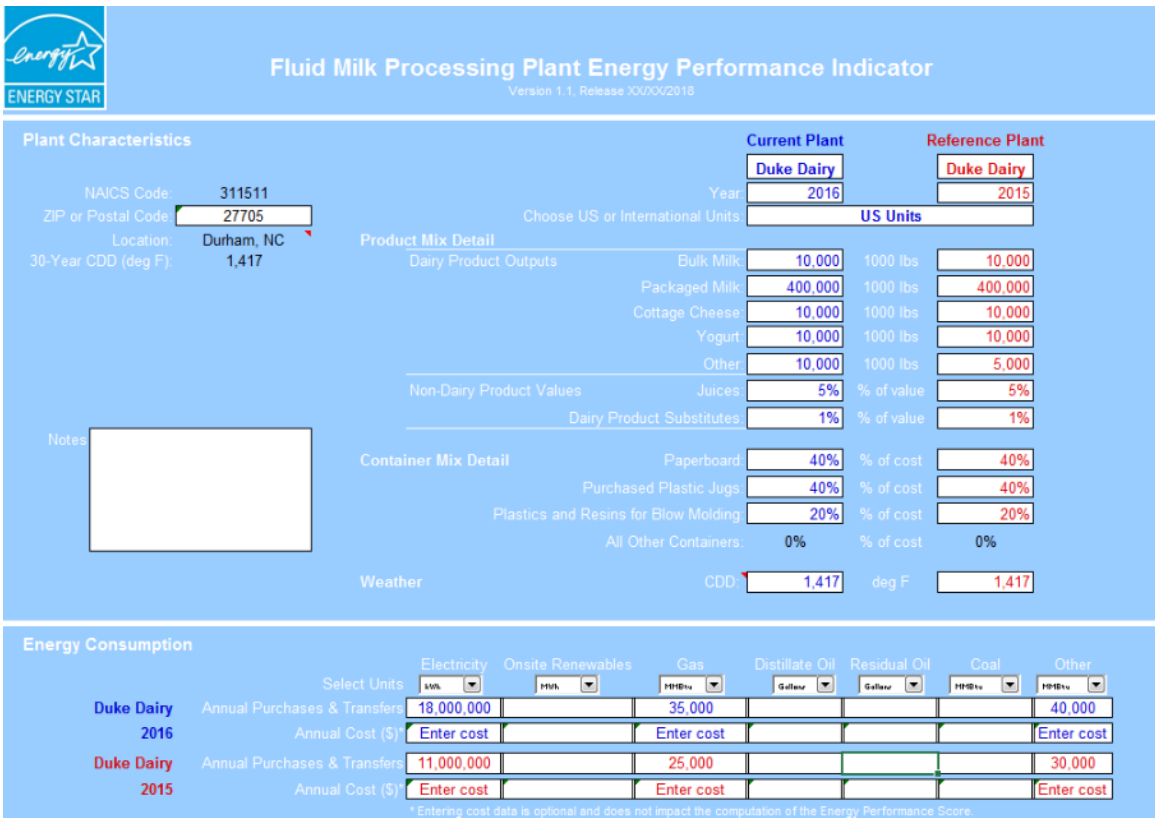
- Total Energy Use
  - Electricity (converted to source MMBtus by the spreadsheet tool)
  - Fuel use for all fuel types in physical units or MMBtu (converted to source MMBtus by the spreadsheet tool)
- Weather
  - Cooling degree days (CDD)
- Product Mix Variables
  - Total production of bulk dairy products
  - Total production of packaged dairy products
  - Total production of cottage cheese
  - Total production of yogurt
  - Total production of other fluid milk products (includes creams, eggnog, sour creams, whipped topping, half and half, eggnog, and buttermilk)
  - Juice product value share
  - Dairy substitute product value shares (soy milks, rice milks, etc.)
- Packaging Mix Variables
  - Paperboard value share of total packaging expenses
  - Purchased plastic jug value share of total packaging expenses
  - Plastics and resins purchased for blow molding value share of total packaging expenses
  - Other container type value share of total packaging expenses

Based on these data inputs, the EPIs 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.

5.2 Spreadsheet Tool

To facilitate the review and use by industry energy managers, a spreadsheet-based tool was constructed to display the results of the EPIs for an arbitrary[7] set of plant-level inputs. Energy managers in the dairy industry were encouraged to test the EPIs by inputting data for their own plants and then provide comments on the results to the developers. After testing, a final version of this spreadsheet-based tool corresponding to the results described in this report was placed on the EPA ENERGY STAR web site for industry use.[8] Example inputs and outputs of the spreadsheet-based tool are shown in Figures 2-3.

Figure 2 Input Section of the Fluid Milk Processing EPI Spreadsheet Tool



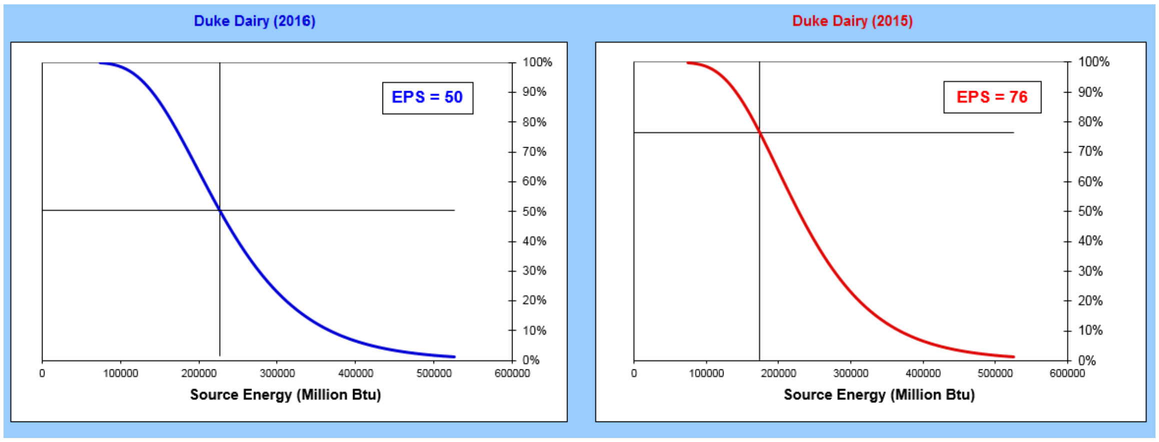
[7] In other words, for plant data that may not originally have been in the data set used to estimate the model equations.

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

Figure 3 Output Section of the Fluid Milk Processing EPI Spreadsheet Tool

Results	Your Current Plant Duke Dairy 2016	Your Reference Plant Duke Dairy 2015	Average Plant Duke Dairy 2016	Efficient Plant Duke Dairy 2016
Display energy results in: US Units				
Energy Performance Score (EPS)	50	76	50	75
Source Energy (MMBtu)	226,955	174,100	227,443	176,369
Site Energy (MMBtu)	116,180	92,532	116,430	90,285
Annual Energy Cost ($/year)	$0	$0	$0	$0
Total Production (1000 lbs)*	440,000	440,000	440,000	440,000
Energy Cost/Total Production ($/1000 lbs)*	$0.00	$0.00	$0.00	$0.00
Energy Intensity (Source MMBtu/1000 lbs)*	0.52	0.40	0.52	0.40

*Production units in the Results section match those selected in the above Plant Characteristics section.



5.3 Use of the ENERGY STAR Fluid Milk Processing 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 within an industry 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 improve these models every few years, contingent on newer data being made available and industry use and support of the EPI tools.

5.4 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 steps detail how to compute 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 characteristics 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 "Average Practice"[9] TSE

- Predicted "Average 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 "typical" energy use for the plant, given its specific operational characteristics.

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

- A lookup table maps all possible values of TSE that are lower than the Predicted "Average Practice" TSE to a cumulative percent in the population.
- The table identifies how far above or below the energy use for a plant is from predicted level.

---

[9] The model computes the "best practice" for frontier models and "average practice" for ordinary least squares. Steps 3 and 4 are similar for the frontier models, except that the prediction is for the minimum energy use and the percentiles are relative to the best (i.e., 99^{th} percentile).

- 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.

6 REFERENCES

Boyd, G., E. Dutrow and W. Tunnesen (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. A. (2005). Development of a Performance-based Industrial Energy Efficiency Indicator for Automobile Assembly Plants. Argonne IL, Argonne National Laboratory**:** May 2005.

Boyd, G. A. (2006). Development of a Performance-based Industrial Energy Efficiency Indicator for Cement Manufacturing Plants. Argonne IL, Argonne National Laboratory.

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.

Brush, A., E. Masanet and E. Worrell (2011). "Energy Efficiency Improvement and Cost Saving Opportunities for the Dairy Processing Industry."

U.S. Environmental Protection Agency (2003). Guidelines for Energy Management. Washington, DC, U.S. Environmental Protection Agency.

U.S. Environmental Protection Agency (2015). Reducing Greenhouse Gas Emissions by Advancing Industrial Energy Efficiency 2000—2015. O. o. A. P. C. P. P. Division. Washington DC.

APPENDIX

One explanation for why paperboard might be the most energy intensive form of packaging was that paperboard was more common for small quantities, thus resulting in more processing to package a similar volume. For example, filling eight cardboard pints of milk takes more energy than making and filling one plastic gallon of milk. The filling repetition of smaller paperboard containers could use more energy per volume, thus making smaller packing more energy intensive. Further analysis was conducted analyzing engineering estimates provided by industry and reported energy usage for carton fillers and blow molding from a USDA report on a process engineering simulation model for fluid milk plants (Tomasula, et al 2013). Table A below shows energy estimates gathered from the listed sources:

Table A Energy Estimates for Packaging

Process step	Energy use	Units	Source:
Blow Mold	0.045	kWh per jug	Erba et al 1997
Unit Fill	0.01	kWh per unit	NRCAN 2001

Using these estimates, we compared plants that use blow molded gallon jugs with plants that use paperboard, but with smaller container sizes. Specifically, these estimates allowed us to reconcile the analysis estimates with the implied "*fill rate*," which we will define as the *average number of units that must be filled to package one gallon of milk*. To reconcile the EPI and engineering estimates we:

1) Computed the average energy per gallon from the NRCAN estimates, including blow molding.
2) Applied the percent difference between blow molding (BM) and paperboard carton (PBC), based on the analysis, to the total energy per gallon in step 1.
3) Subtracted BM from PBC to get the difference in packaging energy use implied by the analysis.
4) Added the BM energy use estimate to the analysis implied difference in packaging energy from step 3. This is done since we want to know how much fill energy is needed to account for *both* the (removal of the) BM operations and the EPI implied difference.
5) Computed the number of units that would need to be filled to account for the energy in step 4.
6) Divided the number of units in step 5 by the number of ounces in a gallon to get the average volume (oz.) per unit filled.

The results are shown in Table B.

Table B Reconciling EPI and USDA/NRCAN Packaging Estimates with Implied Unit Fill Rates

Process step	Analysis Δ (%)	USDA/NRCAN reference (kWh/Gal)	Δ from BM (kWh/Gal)	Δ from BM + BM energy (kWh/Gal)	Fill rate (units per gallon)	Average oz. per unit filled
Blow mold	0%	0.464				
Paperboard	14%	0.535	0.071	0.116	12.6	10.2

Since the unit filling energy is estimated to be 0.01 kWh per unit, for the analysis and engineering estimates to be consistent, plants with paperboard packaging would need to fill 11.6 *more* units per gallon, for a total of 12.6, to account for the predicted higher energy use. This implies an average volume per unit filled of 10.2 oz.; somewhere between pint and half-pint container size.

The other theories for the differences in packaging energy use were examined, including the price volatility of resins for blow molding, and the type of energy intensive products exclusively shipped in paperboard containers. However, no additional evidence was found to support these theories. Discussions on this issue continued throughout the testing of the model.