Development of Performance-based Industrial Energy Efficiency Indicators for Food Processing Plants

Gale A. Boyd, PhD
Sr. Research Scholar
Department of Economics, Duke University
Box 90097, Durham, NC 27708

June 2011

Sponsored by the U.S. Environmental Protection Agency as part of the ENERGY STAR program.

ACKNOWLEDGMENTS

The work described in this report was sponsored by the U.S. Environmental Protection Agency, Office of Atmospheric Programs, Climate Protection Partnership Division as part of the *ENERGY STAR Focus on Energy Efficiency in Food Processing*. The research discussed in this report has benefited from the active review of food processing companies participating in the *ENERGY STAR Focus.*

The research in this paper was conducted while the author was a Special Sworn Status researcher of the U.S. Census Bureau at the Triangle Census Research Data Center. Research results and conclusions expressed are those of the authors and do not necessarily reflect the views of the Census Bureau or the sponsoring agency. This paper has been screened to insure that no confidential data are revealed. Support for this research at the Triangle RDC from NSF (awards no. SES-0004335 and ITR-0427889) is also gratefully acknowledged.

Development of Performance-based Industrial Energy Efficiency Indicators for Food Processing Plants

Gale A. Boyd

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. One key management opportunity is determining an appropriate level of energy performance for a plant through comparison with similar plants in its industry. Performance-based indicators are one way to enable companies to set energy efficiency targets for manufacturing facilities. The U.S. Environmental Protection Agency (EPA), through its ENERGY STAR program, is developing plant energy performance indicators (EPIs) to encourage a variety of U.S. industries to use energy more efficiently. This report describes work with the food processing industry to provide a plant-level indicator of energy efficiency for facilities that produce various types of food 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 selected segments of the food processing industry: juice, frozen fried potatoes, tomato products, and cookie and cracker bakeries. The individual equations are presented, as well as instructions for using those equations as implemented in an associated Microsoft Excel workbook.

1 Introduction

ENERGY STAR was introduced by EPA in 1992 as a voluntary, market-based partnership to reduce air pollution through increased energy efficiency. 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. A key step in improving corporate energy efficiency is to institutionalize strategic energy management. Based on continuous improvement principles and practices of leading corporate energy programs, EPA developed the ENERGY STAR Guidelines for Energy Management to identify the components of successful energy management (EPA 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 audits 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
- Provision of rewards for the success of the program.

These principles subsequently have been incorporated in the International Standards Organization's standard for energy management, known as ISO 50001.

Of the major steps in energy management program development, benchmarking energy use by comparing current energy performance to that of 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 higher organizational level by gauging the performance of a single manufacturing plant 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 Tunnessen (2007) 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. This report describes the basic concept of benchmarking and the statistical approach employed in developing performance-based energy indicators for several segments of the food processing industry, the evolution of the analysis done for these segments of 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 circumstance and participation.

2.1 Scope of an Indicator — Experience with the Food Processors

EPA initiated discussions about developing a plant-level benchmark with food processors. Companies with facilities located within the United States were invited to participate in discussions. At the outset, the term "plant benchmark" was used. Industry engineers routinely develop benchmarks at many levels of plant operation, but they expressed concern that using the word "benchmark" would be confusing and could imply a particular process or tool. For this reason, it was decided that a simple descriptive term would be clearer; thus, ENERGY STAR plant energy performance indicator (EPI) was adopted. The scope for the EPI is a plant-level indicator, not process-specific, and it relates plant inputs in terms of all types of energy use to plant outputs as expressed in a unit of production and/or material processed. Discussion with industry representatives helped to define the energy focus of the model.

The model was designed to account for major, measurable impacts that affect a plant's energy use. The starting point for EPI development was Census data for industrial plants. For the food industry, the basic inputs included information on energy use, total production (physical), amount of material input in the form of preprocessed inputs, the total value of shipments, the shares of product types, and production labor person hours. The actual data used for each of the industry segments depended on the information available from Census and on the results of the statistical analysis.

Ideally the approach to developing an EPI identifies those factors that most directly influence energy use and applies them to normalize the energy use. The most basic normalization is for production level, i.e. energy use per unit of product. Other factors may influence the level of energy use per unit of product, including specific product types, and quality and choice of materials used in production (e.g., amount of raw vs. preprocessed inputs). Including these other factors in the statistical model allows one to construct alternative "benchmarks" of the basic concept of energy use per unit of product. This ideal situation may be limited due to the availability of data or simple limits of the capacity of the methodology to incorporate all of the possible options. The options and data under consideration for the analysis of food industry energy use are as follows.

Production: The industry can be grouped into a wide range of product segments. The initial focus was on frozen and canned fruits and vegetables. This was identified as too broad by the industry representatives at the first focus meeting. Since there are very few plants that produce multiple products in several different segments, we constructed separate EPIs for three distinct product segments: frozen fried potatoes, non-frozen juice, and tomato-based products (paste, sauce, catsup, etc). While separating plants into the three groups effectively controls for the broad differences in product type, there are still issues regarding the measurement of production and differences in product type within each market segment. The Census data provide total value and quantity of product shipped for each plant in each segment; physical measures of production were preferred.

The different product types may have different energy requirements. The role of product types is explored for each segment listed above.

Materials: Data on the use of raw and preprocessed materials can also be included in the analysis, to the extent that they have direct correlation with energy. However, the level of raw material use may not reflect what types of downstream processing different products may require. Since some plants produce products directly from produce, this is likely to have a different energy impact.

Capacity: A source of industry-wide data on plant capacities was not available. If trade associations or other industry sources have this type of information, it could be incorporated in a future analysis. The book value of capital is available from the Census, but would be difficult to apply in this setting.

Utilization: Without direct measurement of plant capacity and physical product, a simple measure of utilization is not possible. However, labor hours may provide a proxy of plant utilization. Labor data may also capture differences in downstream product processing, i.e. differences in the raw production and a fabricated final product. These data are available from Census and can be tested during model development.

The primary focus of this analysis is plants that produce foodstuffs from raw and preprocessed materials in order to manufacture intermediate or final products of a consumer or commercial nature. The U.S. Bureau of Census defines food processing in several segments, and we draw the analysis from several different categories. The first category, *Frozen Fruit, Juice, and Vegetable Manufacturing (NAICS 311411),* comprises establishments primarily engaged in manufacturing frozen fruits, frozen vegetables, and frozen fruit juices, ades, drinks, cocktail mixes and concentrates. Within this category, our focus is on 3114114 Frozen french-fried potatoes and other frozen potato products (patties, puffs, etc.):

- 31141143B1 Frozen french-fried potatoes
- 31141144C1 Other frozen potato products (patties, puffs, etc.).

The second category is *Fruit and Vegetable Canning (NAICS 311421).* This U.S. industry segment comprises establishments primarily engaged in manufacturing canned, pickled, and brined fruits and vegetables. Examples of products made in these establishments are canned juices; canned jams and jellies; canned tomato-based sauces, such as catsup, salsa, chili, spaghetti, barbeque, and tomato paste; pickles, relishes, and sauerkraut. We focus on juices and tomato-based sauces.

311421 Canned vegetable juices, fruit juices, nectars, and concentrates include the following products.

- 311421J111 Canned orange juice, single strength
- 311421J221 Canned apple juice, single strength
- 311421J231 Canned grapefruit juice, single strength
- 311421J241 Canned prune juice, single strength
- 311421J251 Other canned whole fruit juices and mixtures of whole fruit juices

- 311421J261 Canned nectars, single strength
- 311421JYWV Canned fruit juices, nectars, and concentrates, nsk
- 311421A111 Canned tomato juice (including combinations containing 70 percent or more tomato juice)
- 311421A121 Other canned vegetable juices
- 311421AYWV Canned vegetable juices, nsk
- 311421M111 Fresh orange juices and nectars, single strength
- 311421M121 Other fresh juices and nectars, single strength
- 311421MYWV Fresh fruit juices and nectars, single strength, nsk
- 311421M131 Concentrated fruit juice (except for fountain use)
- 311421J271 Fruit juices, concentrated, hot pack
- 312111A111 Fruit drinks, cocktails, and ades, containing some real juice (with added sugar, citric acid, etc.)

311421D Canned tomato products include the following products.

- 311421D111 Canned spaghetti, pizza, and marinara sauces, with or without other added ingredients, except salsa, including those with less than 20 percent meat
- 311421D221 Canned tomato sauce, except pulp, puree, and paste, 7.1 oz to 10 oz (8 oz tall, etc.)
- 311421D231 Canned tomato sauce, except pulp, puree, and paste, other sizes
- 311421D261 Canned chili sauce
- 311421D271 Canned barbecue sauce
- 311421D291 Canned tomato pulp and puree
- 311421D3A1 Canned salsa, 16 oz
- 311421D3B1 Canned salsa, 7 oz to 12 oz
- 311421D3C1 Canned salsa, other sizes
- 311421DYWV Catsup and other tomato sauces, pastes, etc., nsk
- 311421D241 Canned catsup, 14 oz to 32 oz
- 311421D251 Canned catsup, all other sizes (including individual serving sizes)
- 311421D281 Canned tomato paste

As mentioned above, the initial focus on developing a food processing EPI was on all canned and frozen fruits and vegetables, but the industry participants found this approach to be much too broad. A series of separate studies for 8 different categories was initiated. Of those, 3 received industry testing and feedback. A wide range of food products are not covered here, but may be the subject of future analysis.

The second focus on developing an EPI was in baked products. Based on a positive response from the Biscuit and Cracker Manufacturers Association (B&CMA), we chose *Crackers, pretzels, biscuits, and related products* NAICS 311821. This sector produces the following products.

- 311821 1331 Graham crackers
- 311821 1111 Saltines

- 311821 1341 Cracker meal and crumbs
- 311821 1221 Cracker sandwiches
- 311821 1391 Other crackers and related products
- 311821 4111 Sandwich cookies
- 311821 4331 Marshmallow cookies
- 311821 4341 Creme-filled cookies
- 311821 4221 Chocolate chip cookies
- 311821 4351 Oatmeal cookies
- 311821 4361 All other cookies and wafers

The model is based on total source energy, defined as the total Btus of purchased/transferred fuels, steam, and hot water, plus the total amount of purchased/transferred electricity converted from kWh to Btu at roughly the average rate of conversion efficiency for the entire U.S. electric grid.[1] Source energy is used to more closely align our energy measure with the underlying goals of the EPA ENERGY STAR program, pollution reduction at the source. For this reason a kWh of electricity is treated as the equivalent energy at the production source.

2.2 Data Sources

This analysis uses confidential plant-level data from two sources: the Census of Manufacturers (CM) and the Manufacturing Energy Consumption Survey (MECS) maintained by the Center for Economic Studies (CES), U.S. Bureau of the Census (Census). The CM includes the non-public, plant-level data that are the basis of government-published statistics on manufacturing. The CM includes economic activity — for example, labor, energy, plant and equipment, materials costs, and total shipment value of output —for all plants during the years of the Economic Census. The MECS is also used. MECS is a detailed survey of energy use for a sample of plants in the ASM and CM.

Under Title 13 of the U.S. Code, these data are confidential; however, CES allows academic and government researchers with Special Sworn Status to access these confidential micro-data under its research associate program at one of nine designated Census Research Data Centers. The confidentiality restrictions prevent the disclosure of any information that would allow for the identification of a specific plant's or firm's activities. Aggregate figures or statistical coefficients that do not reveal the identity of individual establishments or firms can be released publicly. The ability to use plant-level data, rather than aggregate data, significantly enhances the information that can be obtained about economic performance, particularly when examining the issue of energy efficiency.

<u>Variable Specific Data Sources and Transformations</u>

---

[1] EPA ENERGY STAR updates the source energy conversion periodically. As a result, the analyses underlying juice, frozen fried potatoes, and tomato products use a source energy conversion of 10,236 Btu/kWh, and the analysis for cookie and cracker bakeries uses 11,396 Btu/kWh.

- Data for total value of shipments and labor (person hours) were taken from the CM for 2002.
- Production of different product types (using 10-digit NAICS product codes) was taken from the 2002 CM product trailer files.
- Material input (using 7-digit NAICS material codes) was taken from the 2002 CM material trailer files.
- Electricity use was taken from the 2002 ASM, which was available for every plant in the dataset.
- Fuel use was taken from the 2002 MECS for those plants included in the MECS sample by converting the physical units for every fuel type into Btu content and summing. For all other plants, fuel use was imputed from the cost of fuels as reported in the 2002 ASM using the price of fuels based on the total cost and total Btu for each plant and averaging over all plants in the specific sample.

3 Statistical Approach

The goal of this study was to develop an estimate of the distribution of energy efficiency across the industry. Efficiency is the difference between the actual energy use and "best practice," i.e., the lowest energy use achievable. What is achievable is influenced by operating conditions that vary between plants, so the measure of best practice must take these conditions into account. Statistical models are well-suited for accounting for these types of observable conditions but typically are focused on average practice, not best practice. However, stochastic frontier regression analysis is a tool that can be used to identify "best practice." This section provides the background on the stochastic frontier, a discussion on the review process and evolution of the model's equations, and the final model estimates.

3.1 Stochastic Frontier

The concept of the stochastic frontier analysis that supports the EPI can be easily described in terms of the standard linear regression model, which is reviewed in this section. A more detailed discussion on the evolution of the statistical approaches for estimating efficiency can be found in Greene (1993). Consider at first the simple example of a production process that has a fixed energy component and a variable energy component. A simple linear equation for this can be written as

$$E_i = \alpha + \beta y_i \tag{1}$$

where
E = energy use of plant i and
y = production of plant i.

Given data on energy use and production, the parameters α and β 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 may only be an approximation of the "true" relationship, linear regression estimates of the parameters rely on the proposition that any departures in the plant data from Eq. 1 are "random." This implies that the actual relationship, represented by Eq. 2, includes a random error term ε that follows a normal (bell-shaped) distribution with a mean of 0 and variance of σ^2. In other words, about half of the actual values of energy use are less than what Eq. 1 would predict and half are greater.

$$E_i = \alpha + \beta\, y_i + \varepsilon_i$$
$$\varepsilon \sim N\,(0, \sigma^2) \tag{2}$$

The linear regression gives the average relationship between production and energy use. If the departures from the average, particularly those that are above the average, are due to energy inefficiency, we would be interested in a version of Eq. 1 that gives the "best" (lowest) observed energy use. For example, consider that capacity utilization can influence the energy use per unit of production due to the fixed and variable components of plant energy use (see Figure 1). A regression model can find the line that best explains the average response of energy use per unit of production to a change in utilization rates. The relationship between the lowest energy consumption per unit of production relative to changes in utilization can be obtained by shifting the line downward so that all the actual data points are on or above the line. This "corrected" ordinary least squares (COLS) regression is one way to represent the frontier.

While the COLS method has its appeal in terms of simplicity, a more realistic view is that not all the differences between the actual data and the frontier are due to efficiency. Since we recognize that there may still be errors in data collection/reporting, effects that are unaccounted for in the analysis, and that a linear equation is an approximation of the complex factors that determine manufacturing energy use, we still wish to include the statistical noise, or "random error," term v_i in the analysis – but also add a second random component u_i to reflect energy inefficiency.[2] Unlike the statistical noise term, which may be positive or negative, this second error term will follow a one-sided distribution. If we expand the simple example of energy use and production to include a range of potential effects, we can write a version of the stochastic frontier model as energy use per unit of production as a general function of systematic economic decision variables and external factors,

$$E_i = h(Y_i, X_i, Z_i;\, \beta) + \varepsilon_i \tag{3}$$
$$\varepsilon_i = u_\mathrm{i} - v_\mathrm{i} \qquad\qquad v \sim N\,[0, \sigma_\mathrm{v}^2]\,,$$

where
E = TSE, total source energy (or other measure of total fuel and electricity);
Y = production, measured by dollar shipments or physical production;
X = systematic economic decision variables (i.e., labor-hours worked, materials processed, plant capacity, or utilization rates);
Z = systematic external factors (e.g., heating and cooling loads); and

[2] By random we mean that this effect is not directly measurable by the analyst, but that it can be represented by a probability distribution.

β = all the parameters to be estimated.

We assume that energy (in)efficiency u is distributed according to one of several possible one-sided statistical distributions,[3] for example exponential, half normal, or truncated normal. It is then possible to estimate the parameters of Eq. 3, along with the distribution parameters of u.

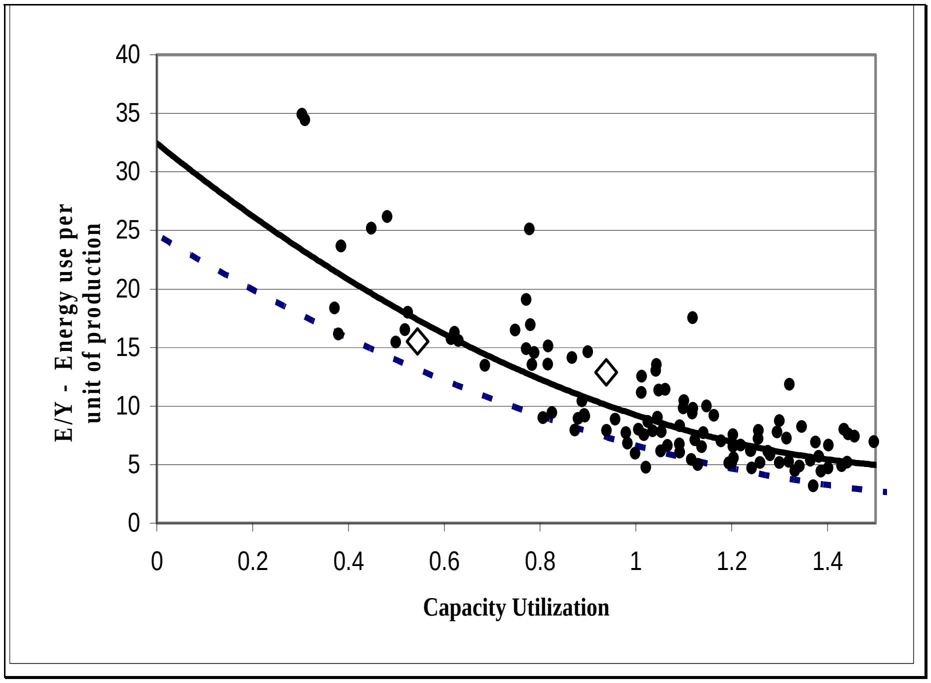


Figure 1 COLS and Frontier Regression of Energy Use per Unit of Production against Capacity Utilization

One advantage of the approach is that the parameters used to normalize for systematic effects and describe the distribution of efficiency are jointly estimated. The standard regression model captures the behavior of the average (see solid line in Figure 1), but the frontier regression (the dotted line in Figure 1) captures the behavior of the best performers. For example, if the best performing plants were less sensitive to capacity utilization because they use better shutdown procedures, then the estimated slope of the frontier capacity utilization curve would not be as steep as the slope for the average plants.

Another advantage of this method is that we can test if the differences in energy use, represented by the terms u and v, are statistically significant. If the estimated variance of u is small, we can conclude that the simpler statistical model in Eq. 2 is valid, and base our measurements on those results. Therefore the frontier yields a more general analysis that allows for either a one-sided (skewed) distribution representing efficiency or a more "normal" (bell-shaped) distribution. If the former is the case, then we interpret that as meaning the many plants are close to one another in terms of energy use, with a smaller number being "further" from the group of good performers. In the latter case, that of the bell-shaped, normal efficiency distribution, we have a few "good performers,"

[3] We also assume that the two types of errors are uncorrelated, $\sigma_{u,v} = 0$.

a large number of "average" plants, and a few "poor performers." In either case we have a statistical approach to assign a ranking for the plants.

For simplicity, we assume that the function *h( )* is linear in the parameters, but allow for non-linear transformations of the variables. In particular, production, materials, and labor enter the equation in log form, as does the energy variable. This means that the terms *u* and *v* can easily 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. When there is wide variation in plant scale, this seems appropriate and may avoid possible heteroscedasticity in either or both error terms.

Given data for any plant, we can rearrange Eq. 3 into Eq. 4 to compute the difference between the actual energy use and the predicted frontier energy use:

$$E_i - \left[h(Y_i, X_i, Z_i\,;\,\beta)\right] = u_{\mathrm{i}} - v_{\mathrm{i}} \tag{4}$$

In the case where the frontier model is appropriate, we have estimated the probability distribution of *u*. Eq. 5 represents the probability that the plant inefficiency is greater than this computed difference:

$$\begin{gathered}\text{Probability}\left[\textit{energy inefficiency} \geq E_i - \left(h(Y_i, X_i, Z_i\,;\,\beta)\right)\right] = \\ 1 - F(E_i - h(Y_i, X_i, Z_i\,;\,\beta))\end{gathered} \tag{5}$$

F( ) is the cumulative probability density function of the appropriate one-sided density function, i.e., gamma, exponential, truncated normal, etc. The value 1 - *F( )* in Eq. 5 defines the EPI score and may be interpreted as a *percentile ranking of the energy efficiency* of the plant. In practice, we only can measure $E_i - h(Y_i, X_i, Z_i\,;\,\beta) = u_{\mathrm{i}} - v_{\mathrm{i}}$, so this implies that the EPI score $1 - F\left(E_i - h(Y_i, X_i, Z_i\,;\,\beta)\right) = 1 - F\left(u_{\mathrm{i}} - v_{\mathrm{i}}\right)$ is affected by the random component of v_i; that is, the score will reflect the random influences that are not accounted for by the function *h(*)*.

In the case where the frontier model is not appropriate, there is no *u* term and corresponding estimate, only *v*.

$$E_i - \left[h(Y_i, X_i, Z_i\,;\,\beta)\right] = v_{\mathrm{i}} \tag{6}$$

We can drop the minus sign for *v* since the normal distribution is two sided. The estimate of the variance $v \sim N\,[0, \sigma_v^2]$ can be used in Eq. 5 where *F( )* is now the cumulative probability density function of a standard normal distribution.

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 plants in my industry, *if all other plants* were similar to mine?"

3.2 Evolution of the Model

The model evolved over a period of time, based on comments from industry reviewers and subsequent analyses. The initial model was based on data from 2002. Industry participants were given an opportunity to test and comment on each version of the model via the annual focus meetings and quarterly conference calls, and personal communications. Companies were asked to input actual data for all of their plants and then to determine whether the results were consistent with any energy efficiency assessments that may have been made for these plants. The resulting comments improved the EPI. This section summarizes this review process and the actions taken vis-à-vis the EPI analysis. Table 1 summarizes the inputs for the final models for the four sub-sectors of the food processing industry.

Table 1: Food Processing Plant Characteristics and Products Included in the EPI

- Frozen Fried Potato
  - Total production
  - Use of on-site frozen product warehouses
- Juice
  - Total production
  - Share of citrus production (citrus vs. non-citrus)
  - Share of concentrate
  - Share of fresh product (NAICS 311421M111, 311421M121, or 311421MYWV)
  - Share of canned[4] product
  - Share of drinks and juice-ades
  - Share of "other non-juice" products
  - Share of material input as concentrate (NAICS 311421M131 or 311421J271
  - Share of material input as frozen fruit
- Tomato-based sauces, etc.
  - Total production
  - Total production worker hours
  - Share of production as tomato paste
  - Share of inputs as fresh product
- Cookies and Crackers
  - Cookies (total production)
  - Crackers (total production)
  - Marshmallow cookies (% of plant total)
  - Other (% of plant total)

The main areas of discussion included product mix and use of preprocessed inputs. Additional industry segment-specific issues were raised as well. For juice this included discussion of citrus products, and for frozen vegetables the issue of product warehouses. The first version of the model included a wide range of product segments in a single model. Industry participants felt that was far too broad and that segment-specific tools

[4] All shelf stable products regardless of form of packaging, including bottled, canned, and juice box.

were needed. Several segments were included in preliminary analysis, but only these three received testing and comments. Early on it was identified that frozen fried potato products could not be included with other frozen vegetables because the additional process step of frying was an energy intensive one. Analysis confirmed this and a separate model was developed. For juice processing, initially all fruit (and vegetable) juice was treated the same. Comments from a citrus processor identified the need for those plants to use additional processing because of the volume of citrus peels and associated energy use in processing. Analysis confirmed that citrus-based products are statistically more energy intensive, but does not reveal if the citrus peel is the reason for the difference. The model for juice includes an adjustment for citrus and non-citrus production. For juice and tomato products, both the output and the input (but not both) may be in the form of a concentrate or paste, respectively. The use / production of concentrates or pastes have implications for the plant energy use. Intuitively, producing either of these requires more water to be driven from the product, increasing the energy use. Conversely, using a preprocessed input would lower the energy requirements. The statistical results confirm this.

In all cases, plants with extreme values of the energy-output ratio, value-output ratio (price), or labor output ratio were dropped from the analysis. Only plants with at least 50% of their production for that particular product segment are included in the analysis. A variable that reflects the fraction of total value in the primary product relative to the total value of shipments is included in the model if it was statistically significant.

The apparent presence of outliers and the small sample for fried potatoes led to some additional screening and data adjustments. Comparison of the ratio of inputs to final product (yield) showed some plants with a ratio greater than one. These plants also tended to have lower energy/output ratios. Any plant with a yield exceeding unity had the production values adjusted to reflect an average yield based on the other data points. There were a few additional outliers where this procedure was not effective. These plants were excluded from the model by using fixed effects and the corresponding coefficient was suppressed for disclosure purposes.

During the development process for the cookie and cracker EPI, several member companies from B&CMA commented that the model should address differences in production processes in addition to end products. This led to testing the impact of several different types of products that reflect different production methods, such as “sandwich” cookies and crackers and “wire cut” cookies (represented in the data by chocolate chip and oatmeal cookies). It was determined that none of these products had significantly different energy use despite different production processes, except marshmallow, which was added as a category to the final model.

3.3 Model Estimates

This section presents the current model estimates for each of the four industry segments: frozen fried potatoes, juice processing, tomato products, and cookies and crackers. Several alternatives for specification of *h( )* and for the distribution of the error term u were tried. Only the “preferred” model estimates are presented.

Frozen Fried Potatoes:

The final version of the fried potatoes equation for TSE is

$$\ln(energy) = A + \beta_1 \ln(production) + \beta_2 Dummy(no\ warehouse) + \varepsilon \qquad (7)$$

where
Energy = total source energy
Production = production of fried potatoes (1000 lbs)
Dummy = dummy variable =1 when there are no frozen product warehouse and 0 otherwise
and
β = vector of parameters to be estimated.

The variable ε is distributed as N(0, σ^2).

A dummy variable was included in the model for some outliers rather than dropping them from the sample. This was done in order to maintain the sample from previous model results cleared by Census (for testing purposes). This dummy variable was suppressed for clearance purposes. The estimated parameters of the model are shown in Table 2. Sample size is 27 plants. The production parameter and intercept shown are significant at the 99% level. The dummy for warehouse is only significant at the 90% level in a one tailed test. Estimates of the frontier resulted in extremely small variance estimates of *u*, so the simpler OLS model is used in this segment.

Table 2 Frozen Fried Potato Energy Model Estimates

Variable	Estimate	Standard Error	t-ratio
Log Production*	0.9100907	0.045367	20.06
Outlier	suppressed		
Dummy for no warehouses	-0.1473023	0.109256	-1.35
Constant	2.546295	0.529202	4.81

Error Distribution Parameters	
σ^2	. 0659
R – square	.9673
F(3, 25)	226.87

Juice (non-frozen):

The final version of the juice processing energy equation for TSE is

$$
\begin{aligned}
\ln(energy) = {} & A + \beta_1 \ln(production) + \beta_2 Share\,of\ Juice\ production\,to\,total\ value \\
& + \beta_3 Share\,of\ \ juice\ from\,citrus + \beta_4 Share\,of\ canned\ product \\
& + \beta_5 Share\,of\ concentrate \\
& + \beta_6 Share\,of\ \ fresh\ product + \beta_7 Share\,of\ \ juice\,and\ drink\,"ades" \\
& + \beta_8 Share\,of\ material\ as\ concentrate \\
& + \beta_9 Share\,of\ material\ as\ frozen\,and\ preprocessed + \varepsilon
\end{aligned} \tag{8}
$$

where

Energy	=	total source energy (MMBTU);
Production	=	total juice production (1000 gallons)
Share of juice	=	ratio of total juice production to total value ($)
Share of citrus	=	ratio of citrus juice production to total juice (1000 gallons)
Share of canned	=	ratio of canned/bottled (shelf-stable) production to total juice production (1000 gallons)
Share of concentrate	=	ratio of concentrate production to total juice production (1000 gallons)
Share of fresh	=	ratio of fresh production to total juice production (1000 gallons)
Share of drinks	=	ratio of juice drinks and “ades” production to total juice production (1000 gallons)
Share of concentrate	=	ratio of concentrate use to total materials ($)
Share of frozen	=	ratio of frozen and preprocessed fruit use to total materials ($)

and

β = vector of parameters to be estimated.

The variable ε is distributed as $N(0, \sigma^2)$.

The estimated parameters of the model are shown in Table 3. Sample size is 44 plants. All parameters shown with an asterisk are significant at the 99% level. All variables are jointly significant from zero. Estimates of the frontier resulted in extremely small variance estimates of *u*, so the simpler OLS model is used in this segment.

Table 3 Juice Processing Energy Model Estimates

Variable	Estimate	Standard Error	t-ratio
Log Production*	0.844615	0.095432	8.85
Share of juice*	-1.94352	1.366	-1.42
Share of citrus*	1.021825	0.435901	2.34
Share of canned	1.047428	0.871968	1.2
Share of concentrate*	2.1622	0.863798	2.5
Share of fresh	0.431051	0.910018	0.47
Share of drinks*	2.218101	1.320205	1.68

Share of concentrate	-0.10256	0.279434	-0.37
Share of frozen	-0.61882	1.116042	-0.55
Constant*	3.782562	1.716484	2.2

Error Distribution Parameters	
σ^2	0.6334
R – square	0.8003
F(9, 34)	15.14

<u>Tomato Products:</u>

The final version of the tomato products equation for TSE is

$$\ln(energy) = A + \beta_1 \ln(production) + \beta_2 \ln(production\ person\ hours) \\ + \beta_3 Share\ of\ fresh\ inputs\ to\ production + \beta_4 Share\ of\ product\ as\ paste + \varepsilon \tag{9}$$

where

Energy	=	total source energy (MMBTU);
Production	=	total production (1000 lbs)
Person hours	=	total production worker person hour (1000 hours)
Share of fresh	=	ratio of fresh inputs to total production (1000 lbs)
Share of paste	=	ratio of paste production to total (1000 lbs)

and

β = vector of parameters to be estimated.

The variable ε is distributed as N(0, σ^2).

The estimated parameters of the model are shown in Table 4. Sample size is 40 plants. All parameters shown with an asterisk are significant at the 99% level in a two-tailed test. Estimates of the frontier resulted in extremely small variance estimates of *u*, so the simpler OLS model is used in this segment.

Table 4 Tomato Processing Energy Model Estimates

Variable	Estimate	Standard Error	t-ratio
Log Production*	0.355233	0.122204	2.91
Log Person hours*	0.756751	0.17298	4.37
Ratio of fresh inputs to total production*	0.116953	0.056078	2.09
Share of production as paste*	1.563001	0.421234	3.71
Constant*	3.336597	0.844661	3.95

Error Distribution Parameters	
σ^2	0.6335
R – square	0.8008
F(4, 35)	35.17

Cookies and Crackers:

The final version of the cookie and cracker equation for TSE is

$$\ln(energy) = A + \beta_1 \ln(production) + \beta_2(\%\, value\, other) + \beta_3 Share\ of\ cookies + \beta_4 Share\, of\ marshmallow + \varepsilon \tag{10}$$

where

Energy	=	total source energy (MMBTU);
Production	=	total production (1000 lbs)
% value other	=	ratio of value of products other than cookies and crackers to total value of all production (dollars)
Share of cookies	=	ratio of production of cookies to total production of cookies and crackers
Share of marshmallow	=	ratio of production of marshmallow cookies to total production of cookies and crackers

and

β = vector of parameters to be estimated.

The variable ε is distributed as N(0, σ^2).

The estimated parameters of the model are shown in Table 5. Sample size is 64 plants. All parameters shown with an asterisk are significant at the 99% level in a two-tailed test. Estimates of the frontier resulted in extremely small variance estimates of *u*, so the simpler OLS model is used in this segment.

Table 5 Cookie and Cracker Baking Energy Model Estimates

Variable	Estimate	Standard Error	t-ratio
Log Production*	0.714887	0.040518	17.64
% value other *	-1.02694	0.435916	-2.36
Share of cookies *	-0.52269	0.154447	-3.38
Share of marshmallow *	5.077321	2.86239	1.77
Constant*	5.285898	0.580692	9.10
Error Distribution Parameters			
σ^2	0.29569		
R – square	0.8735		
F(5, 58)	80.09		

4 Judging Food Processing Plant Energy Efficiency

4.1 How the EPI Works

The food processing plant EPIs rate the energy efficiency of four segments – frozen fried potatoes, juice processing, and tomato products processing plants, and cookie and cracker bakeries – based in the United States. To use the tool, the following information must be available for a plant.

- Total energy use
    - Electricity in kWh (converted to Btus by the spreadsheet)
    - Fuel use for all fuel types in physical units or Btu
- Frozen fried potatoes
    - frozen fried potatoes production (1000 lbs)
    - frozen product warehouse
- Juice processing
    - juice production (1000 gallons)
    - ratio of total juice production value to total production value ($)
    - ratio of citrus juice production to total juice production (gallons)
    - ratio of canned/bottled (shelf-stable) production to total juice production (gallons)
    - ratio of concentrate production to total juice production (gallons)
    - ratio of fresh production to total juice production (gallons)
    - ratio of juice drinks and "ades" production to total juice production (gallons)
    - ratio of concentrate use to total materials ($)
    - ratio of frozen and preprocessed fruit use to total materials ($)
- Tomato products
    - total tomato products production (1000 lbs)
    - total production worker person hour (1000 hours)
    - ratio of fresh inputs to total production (lbs)
    - ratio of paste production to total production (lbs)
- Cookies and crackers
    - total cookie production (1000 lbs)
    - total cracker production (1000 lbs)
    - ratio of production of marshmallow cookies to total cookie and cracker production (% on weight basis)
    - ratio of production value of all "other" products to total value of production (% of plant total production value)

Based on these data inputs, these food processing EPIs will report a score for the plant in the current time period that reflects the relative energy efficiency of the plant compared to that of the industry. It is a percentile score on a scale of 0–100. Plants that rate 75 or better are classified as efficient (ENERGY STAR defines the 75th percentile as efficient.) An Energy Performance Score of 75 means a particular plant is performing better than 75% of the plants in the industry. The model also reports on the average plant in the industry (defined as the 50th percentile). While the underlying model was developed from data for U.S.-based plants, it does not contain or reveal any confidential information.

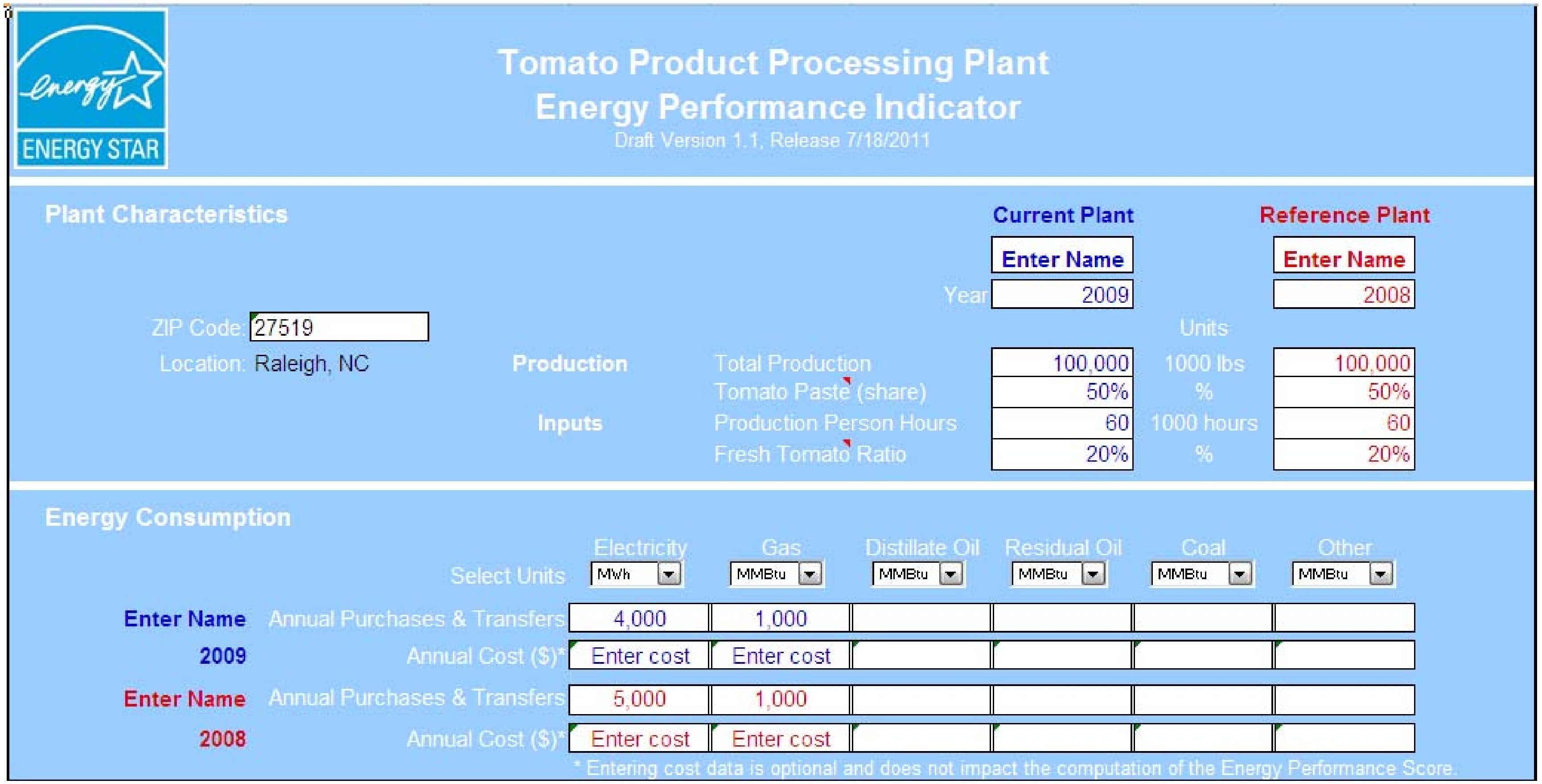
4.2 Spreadsheet Tool

To facilitate the review and use by industry energy managers, a spreadsheet was constructed to display the results of the EPI for an arbitrary[5] set of plant-level inputs. The spreadsheet accepts the raw plant-level inputs described above, computes the values for *h( ),* and then displays the results from the appropriate distribution functions for the models presented in Eqs. 7, 8, 9, and 10. The energy managers were encouraged to input data for their own plants and then provide comments. A version of these spreadsheets is available from the EPA ENERGY STAR web site.[6] Examples of the input section of each spreadsheet are shown in Figures 2-5. The results section examples are shown in Figure 6-9.

---

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

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



ENERGY STAR

Tomato Product Processing Plant
Energy Performance Indicator
Draft Version 1.1, Release 7/18/2011

Plant Characteristics

ZIP Code: 27519
Location: Raleigh, NC

		Current Plant	Units	Reference Plant
		Enter Name		Enter Name
	Year	2009		2008
Production	Total Production	100,000	1000 lbs	100,000
	Tomato Paste (share)	50%	%	50%
Inputs	Production Person Hours	60	1000 hours	60
	Fresh Tomato Ratio	20%	%	20%

Energy Consumption

		Electricity	Gas	Distillate Oil	Residual Oil	Coal	Other
	Select Units	MWh	MMBtu	MMBtu	MMBtu	MMBtu	MMBtu
Enter Name	Annual Purchases & Transfers	4,000	1,000				
2009	Annual Cost ($)*	Enter cost	Enter cost				
Enter Name	Annual Purchases & Transfers	5,000	1,000				
2008	Annual Cost ($)*	Enter cost	Enter cost				

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

Figure 2 Input Section of the Tomato Product EPI Spreadsheet Tool

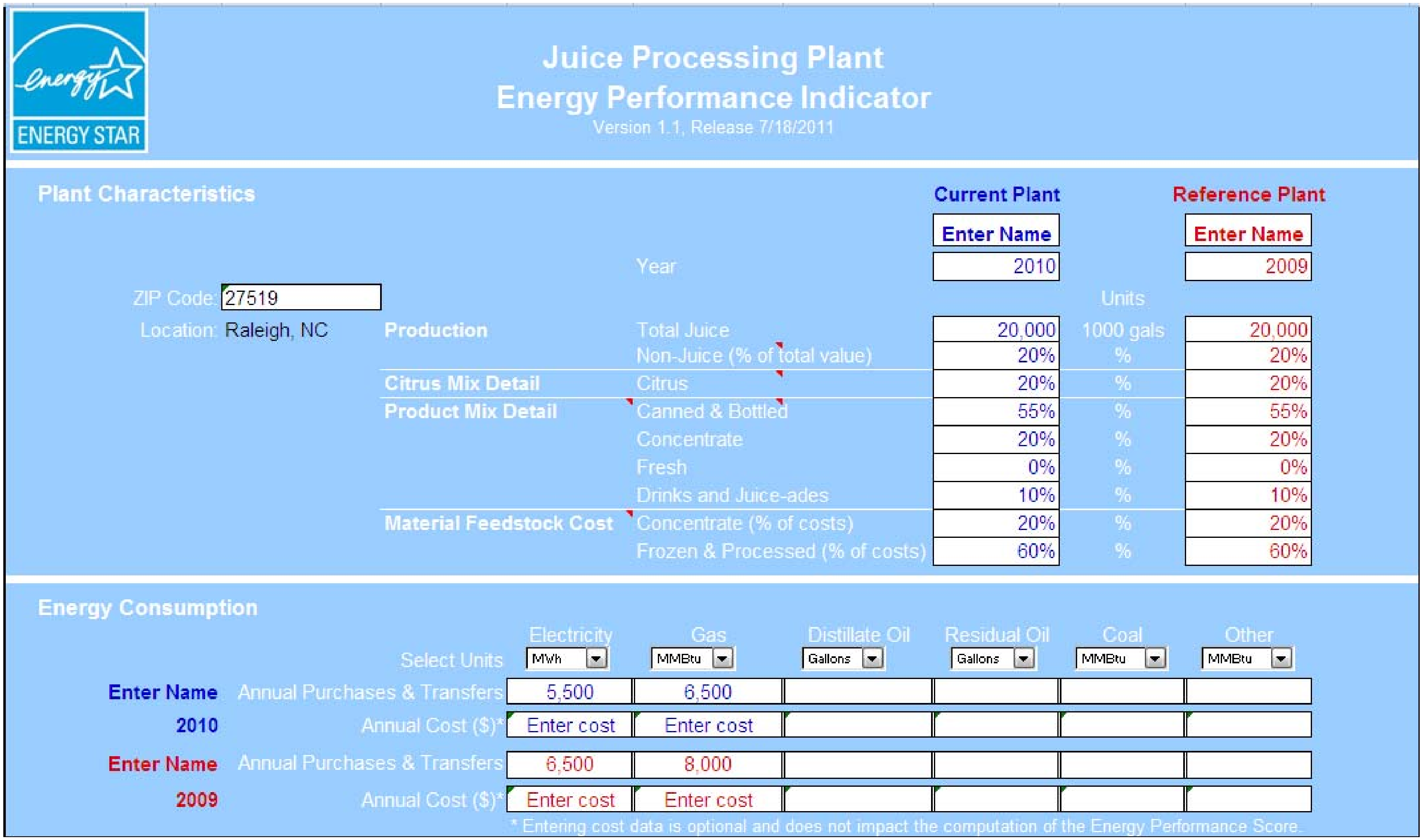
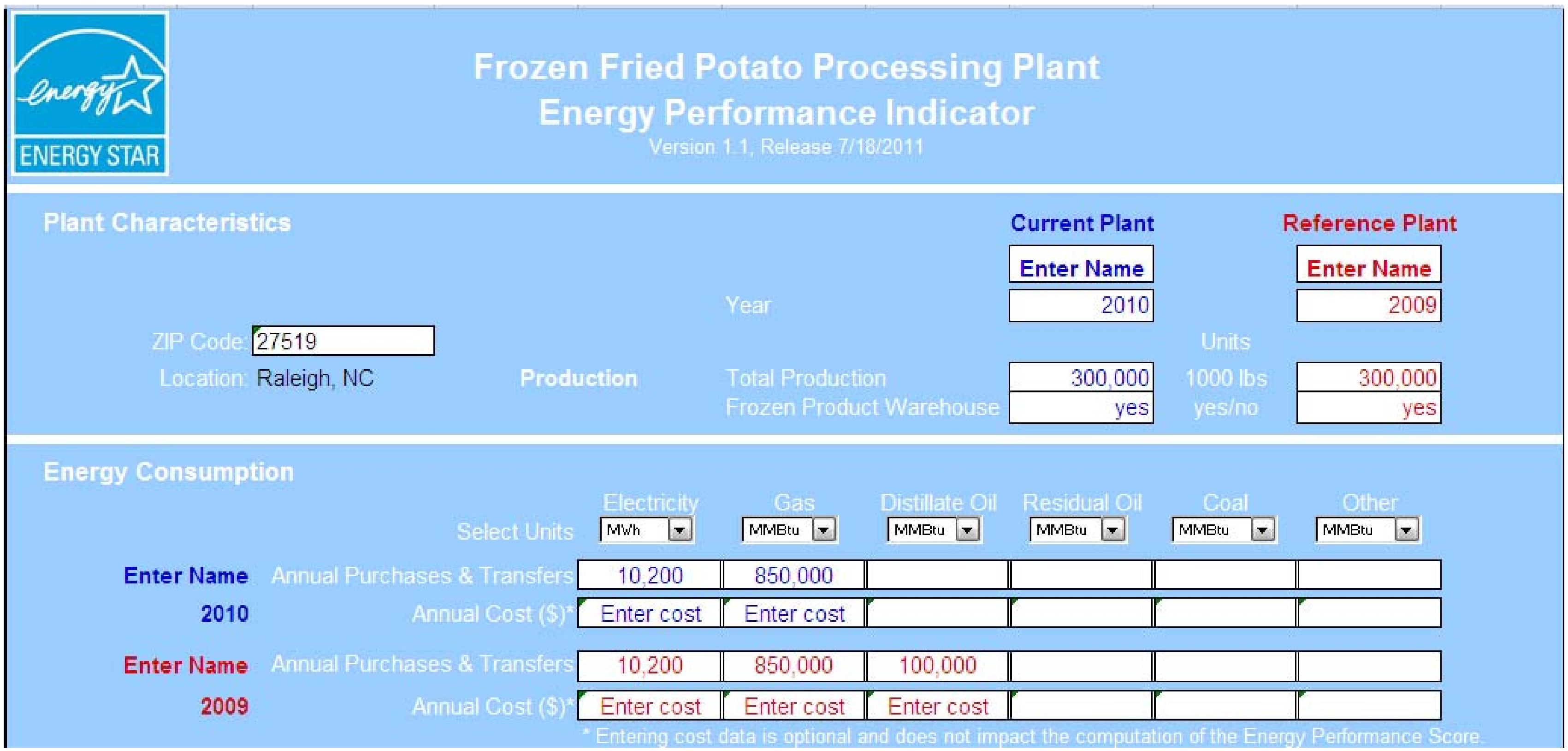


Figure 3 Input Section of the Juice Processing EPI Spreadsheet Tool


Frozen Fried Potato Processing Plant
Energy Performance Indicator
Version 1.1, Release 7/18/2011

ENERGY STAR

Plant Characteristics

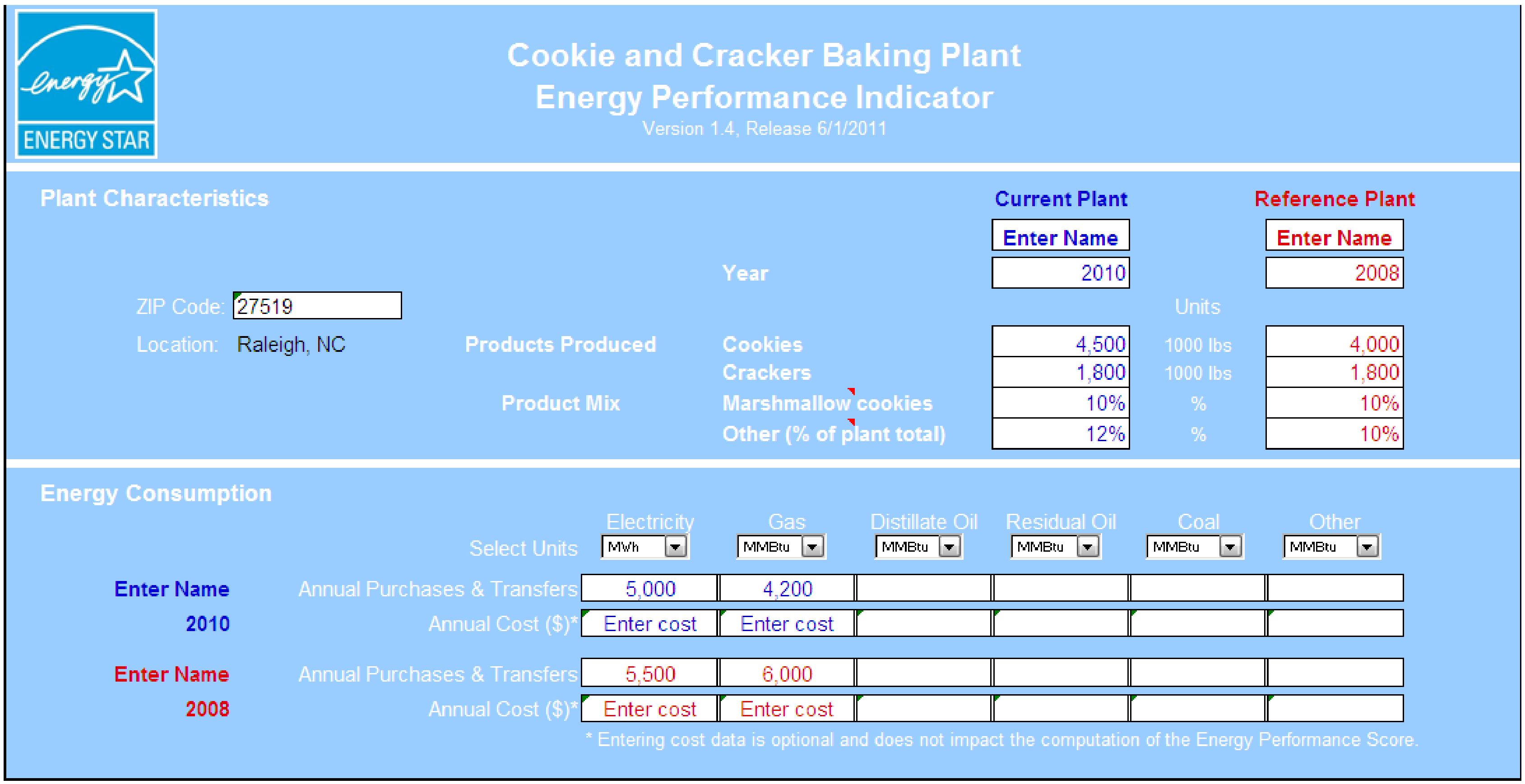
		Current Plant	Units	Reference Plant
		Enter Name		Enter Name
	Year	2010		2009
ZIP Code: 27519				
Location: Raleigh, NC	Production: Total Production	300,000	1000 lbs	300,000
	Frozen Product Warehouse	yes	yes/no	yes

Energy Consumption

		Electricity	Gas	Distillate Oil	Residual Oil	Coal	Other
	Select Units	MWh	MMBtu	MMBtu	MMBtu	MMBtu	MMBtu
Enter Name	Annual Purchases & Transfers	10,200	850,000				
2010	Annual Cost ($)*	Enter cost	Enter cost				
Enter Name	Annual Purchases & Transfers	10,200	850,000	100,000			
2009	Annual Cost ($)*	Enter cost	Enter cost	Enter cost			

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

Figure 4 Input Section of the Frozen Fried Potato Processing EPI Spreadsheet Tool



Cookie and Cracker Baking Plant
Energy Performance Indicator
Version 1.4, Release 6/1/2011

ENERGY STAR

Plant Characteristics

ZIP Code: 27519
Location: Raleigh, NC

		Current Plant	Units	Reference Plant
		Enter Name		Enter Name
	Year	2010		2008
Products Produced	Cookies	4,500	1000 lbs	4,000
	Crackers	1,800	1000 lbs	1,800
Product Mix	Marshmallow cookies	10%	%	10%
	Other (% of plant total)	12%	%	10%

Energy Consumption

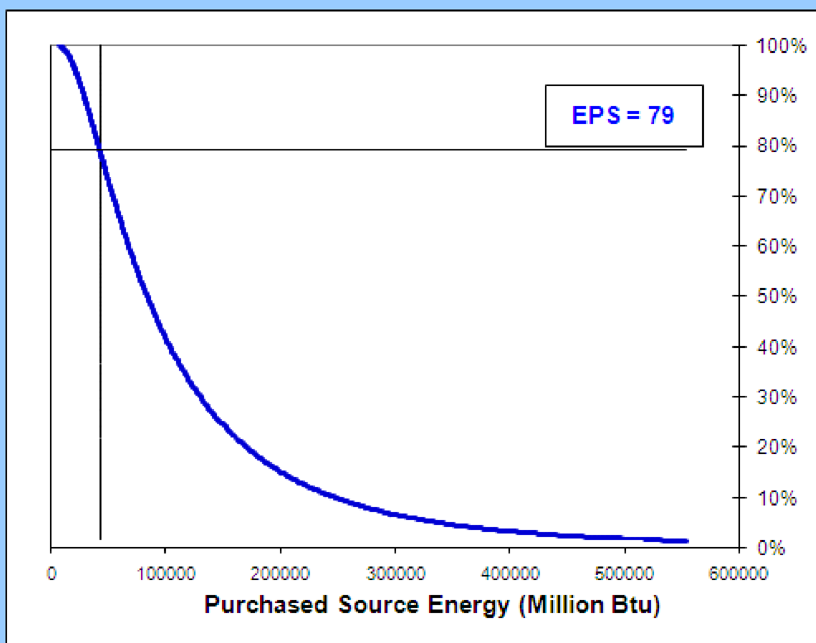
		Electricity	Gas	Distillate Oil	Residual Oil	Coal	Other
	Select Units	MWh	MMBtu	MMBtu	MMBtu	MMBtu	MMBtu
Enter Name	Annual Purchases & Transfers	5,000	4,200				
2010	Annual Cost ($)*	Enter cost	Enter cost				
Enter Name	Annual Purchases & Transfers	5,500	6,000				
2008	Annual Cost ($)*	Enter cost	Enter cost				

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

Figure 5 Input Section of the Cookies and Cracker Bakery EPI Spreadsheet Tool

Results	Your Current Plant Enter Name 2009	Your Reference Plant Enter Name 2008	Average Plant Enter Name 2009	Efficient Plant Enter Name 2009
Energy Performance Score (EPS)	79	71	50	75
Annual Energy Cost ($/year)	$0	$0	$0	$0
Energy Cost/Total Production ($/1000 lbs)	$0.00	$0.00	$0.00	$0.00
Purchased Source Energy (MMBtu)	42,075	52,344	83,254	46,845
Purchased Site Energy (MMBtu)	14,648	18,060	28,984	16,308
Energy Intensity (Source MMBtu/1000 lbs)	0.42	0.52	0.83	0.47

Enter Name (2009)



Enter Name (2008)

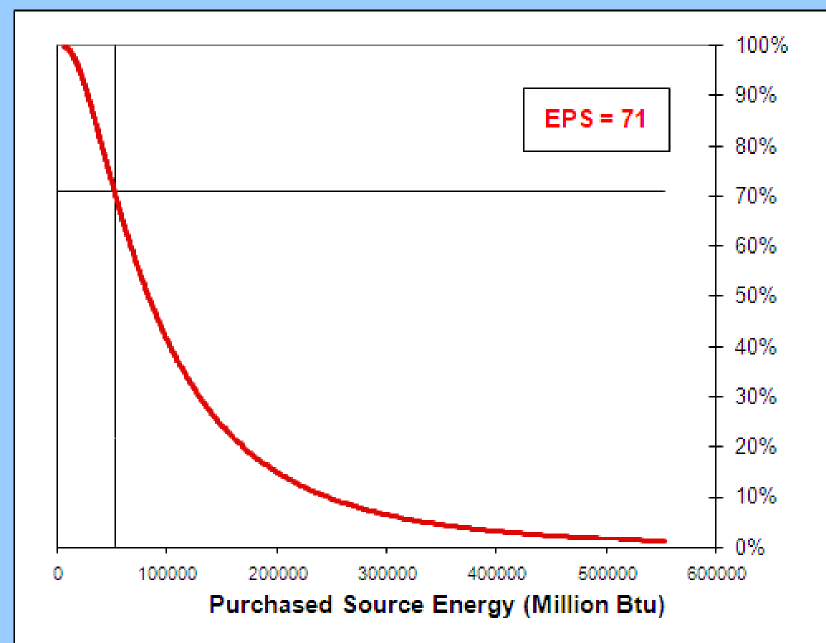


Figure 6 Output Section of the Tomato Products EPI Spreadsheet Tool

Results	Your Current Plant Enter Name 2010	Your Reference Plant Enter Name 2009	Average Plant Enter Name 2010	Efficient Plant Enter Name 2010
Energy Performance Score (EPS)	77	70	50	75
Annual Energy Cost ($/year)	$0	$0	$0	$0
Energy Cost/Total Production ($/1000 gal)	$0.00	$0.00	$0.00	$0.00
Purchased Source Energy (MMBtu)	62,978	74,747	113,014	65,779
Purchased Site Energy (MMBtu)	25,266	30,178	45,340	26,390
Energy Intensity (Source MMBtu/1000 gal)	3.15	3.74	5.65	3.29

Enter Name (2010) | Enter Name (2009)

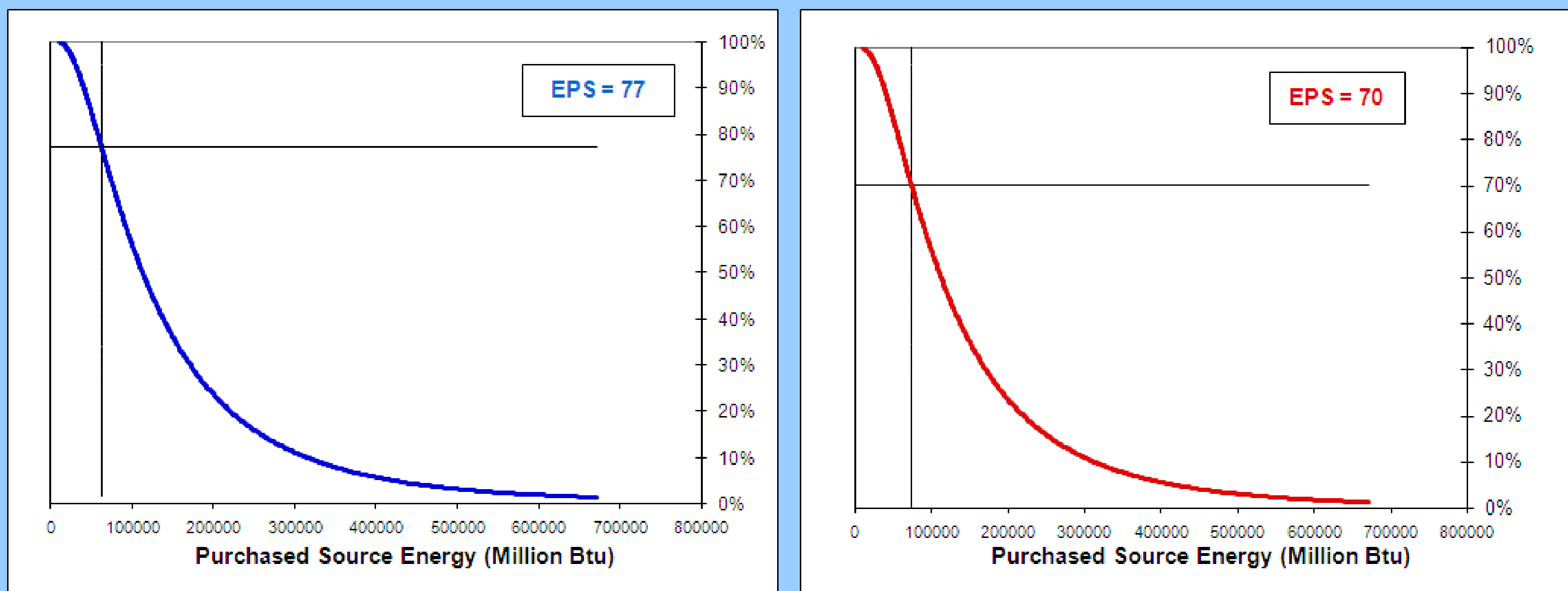
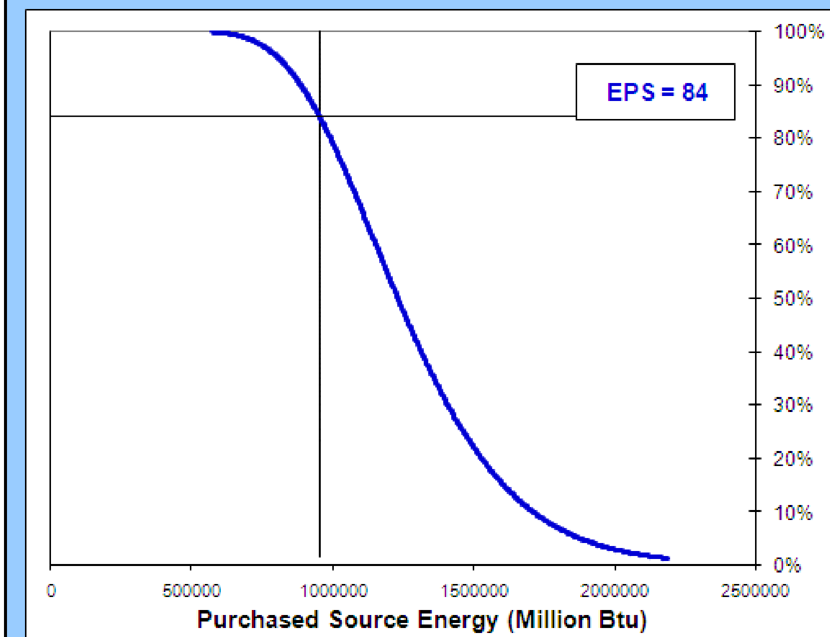


Figure 7 Output Section of the Juice Processing EPI Spreadsheet Tool

Results	Your Current Plant Enter Name 2010	Your Reference Plant Enter Name 2009	Average Plant Enter Name 2010	Efficient Plant Enter Name 2010
Energy Performance Score (EPS)	84	73	50	75
Annual Energy Cost ($/year)	$0	$0	$0	$0
Energy Cost/Total Production ($/1000 lbs)	$0.00	$0.00	$0.00	$0.00
Purchased Source Energy (MMBtu)	954,742	1,054,742	1,231,740	1,034,500
Purchased Site Energy (MMBtu)	884,802	984,802	1,141,509	958,718
Energy Intensity (Source MMBtu/1000 lbs)	3.18	3.52	4.11	3.45

Enter Name (2010)



Enter Name (2009)

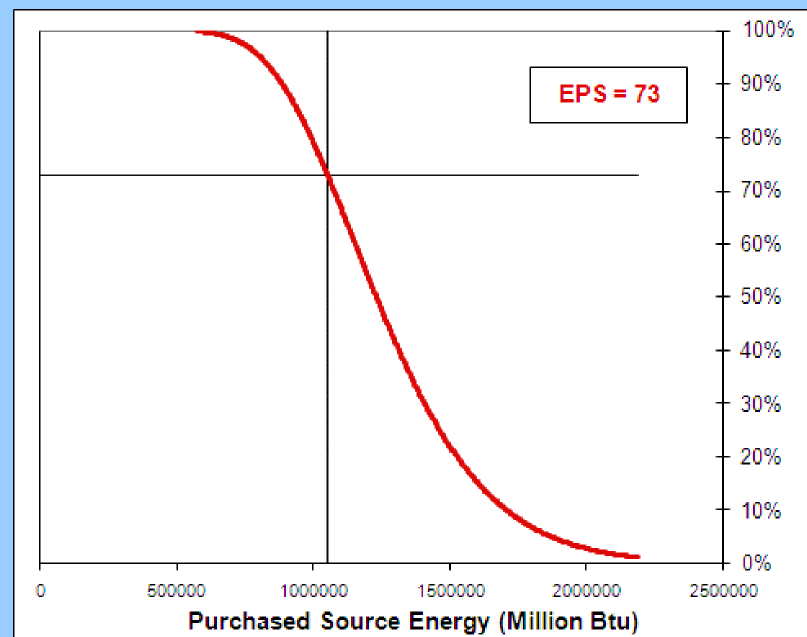
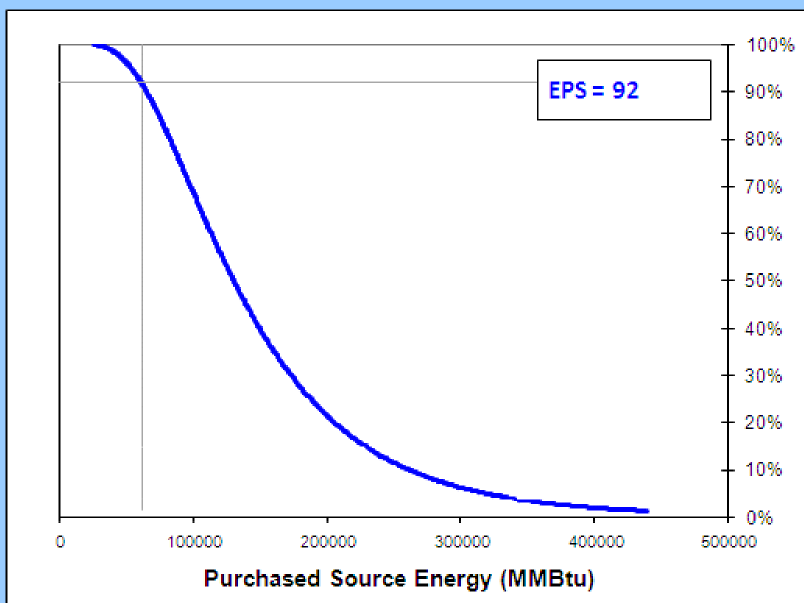


Figure 8 Output Section of the Frozen Fried Potato Processing EPI Spreadsheet Tool

Results

	Your Current Plant Enter Name 2010	Your Reference Plant Enter Name 2008	Average Plant Enter Name 2010	Efficient Plant Enter Name 2010
Energy Performance Score (EPS)	92	86	50	75
Annual Energy Cost ($/year)	$0	$0	$0	$0
Energy Cost/Total Production ($/1000 lbs)	$0	$0	$0	$0
Purchased Source Energy (MMBtu)	61,180	68,678	129,976	89,800
Purchased Site Energy (MMBtu)	21,260	24,766	45,166	31,205
Energy Intensity (Source MMBtu/1000 lbs)	9.71	11.84	20.63	14.25

Enter Name (2010)



Enter Name (2008)

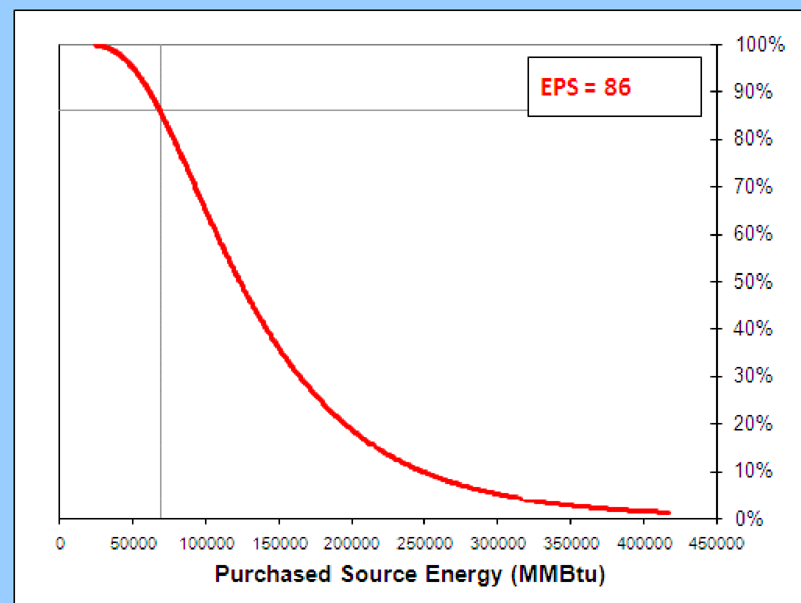


Figure 9 Output Section of the Cookies and Cracker Bakery EPI Spreadsheet Tool

4.3 Use of the ENERGY STAR Food Processing EPIs

After several years of work with the food processing companies, the ENERGY STAR food processing EPIs for these industry segments are now complete, as are spreadsheet tools for calculating EPI scores. EPA intends to use the EPIs to motivate improvement in energy use in U.S.-based manufacturing. 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.

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 and Energy Performance Score be used to set energy efficiency improvement goals at both the plant and corporate levels.

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.

4.4 Steps to Compute an Energy Performance 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 EPI score 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 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 source to site 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"[7] 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 building, given its specific operational constraints.

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 an Energy Performance 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.

5 References

Boyd, G., E. Dutrow, et al. (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. and W. Tunnessen (2007). "Motivating Industrial Energy Efficiency through Performance-based Indicators." *ACEEE Summer Study On Energy Efficiency In Industry Improving Industrial Competitiveness: Adapting To Volatile Energy Markets, Globalization, And Environmental Constraints*, White Plains, New York, ACEEE.

EPA, 2003, *Guidelines for Energy Management*, U.S. Environmental Protection Agency, Washington, DC; available online at http://www.energystar.gov/index.cfm?c=guidelines. guidelines_index.

Greene, W.H., 1993, "The Econometric Approach to Efficiency Analysis," pp. 68–119 in *The Measurement of Productive Efficiency: Techniques and Applications*, H. Fried, et al., (editors), Oxford University Press, NY.

[7] 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).