DEVELOPMENT OF AN ENERGY STAR® ENERGY PERFORMANCE INDICATOR FOR DISTILLED SPIRITS

GALE A. BOYD AND MATT DOOLIN

DUKE UNIVERSITY, SOCIAL SCIENCE RESEARCH INSTITUTE

BOX 90989, DURHAM, NC 27708

JOSHUA S. SMITH

ICF, INC.

1902 RESTON METRO PLAZA, RESTON, VA 20190

JUNE 12, 2023

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

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 Distilled Spirits Production*. We would also like to thank Daniel Macri (U.S. EPA) for comments and review.

CONTENTS

DEVELOPMENT OF AN ENERGY STAR® ENERGY PERFORMANCE INDICATOR FOR DISTILLED SPIRITS

GALE A. BOYD, MATT DOOLIN, AND JOSHUA SMITH

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 distilled spirits industry to provide a plant-level indicator of energy efficiency for distilleries in the United States and Canada. 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 facilities operating within the distilled spirits industry. 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 is 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, and 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 distilled spirits 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 Manufactures, Annual Survey of Manufactures, 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.

A multi-variable benchmark like that developed by ENERGY STAR provides advantages over simple energy intensity metrics by creating an even basis for all plants in an industry, facilitating an apples-to-apples comparison. The benchmark normalizes for the non-controllable and market differences between plants, such as product mix, size, geographic location, etc., to get an accurate measure of energy performance based on the operating behaviors of each site. Comparing sites based on energy intensity alone requires that plants already be similar, whereas the multi-variable benchmark can compare dissimilar plants to one another and focus on the operating practices that differentiate their performance.

3 EVOLUTION OF THE DISTILLED SPIRITS EPI

In the fall of 2019, EPA organized a focus for distilled spirits manufacturing within its ENERGY STAR program, as a subset of the broader food processing focus. ENERGY STAR was contacted by state energy and environmental offices expressing a desire for this industry to have an EPI so distilleries could be eligible for ENERGY STAR plant certification and be engaged in improving their energy efficiency programs. Between 2020 and 2023, ENERGY STAR program staff and industry collaborated on developing the EPI detailed below and an industry-specific energy guide entitled *Energy Efficiency and Cost Saving Opportunities for Distilleries* (Stuckrath and Worrell 2022).

3.1 Using Census Data

Draft Version 1.0 of the Distilled Spirits EPI was based on data reported to the U.S. Census Bureau under the six-digit NAICS code 312140, specifically covering the inputs and outputs of distilleries. The sample was comprised of a panel dataset, utilizing data from the 2002, 2007, and 2012 Census of Manufactures (CMF) that included all distilled spirit types with the exception of brandy. Due to the differences in input materials and product composition, the Census Bureau classified brandy as a wine product rather than a distilled liquor. Distilleries generally use electricity and fuels for distinct processes within the plant, allowing for electricity and fuels to be split into two separate models. Draft Version 1.0 of the EPI developed separate statistical models for each energy source before aggregating them into one overall efficiency score.

After the initial analysis of the Census data and external sources, as described below, the following variables were included in the models to account for electricity and fuel usage in the industry, respectively.

Electricity Model Variables:

- Total Distilled Production Volume
- Onsite Wastewater Treatment
- Cooling Degree Days (CDD)

Fuel Model Variables:

- Total Distilled Production Volume
- Share of Bottled Volume to Production Volume
- Share of Purchased Neutral Spirits to Production
- Heating Degree Days (HDD)

Some distilleries remove moisture from grains and other byproducts before they are removed from the site, which is an energy intensive process. Although a variable for distillers' dried grains and solubles was included in the Census of Manufactures form, the quality of the responses was deemed questionable based on company responses to other variables. Given the large energy impact that byproduct handling can have at a distillery, an effort was undertaken to supplement the quality of the

Census data with external data collected from participating focus companies regarding the byproduct dewatering and disposal process. However, this additional data did not represent enough of the sample or provide additional clues as to how the missing information could be determined based on other plant characteristics. Ultimately, a variable capturing how distilleries handled their byproducts was unable to be included in this first version of the EPI.

Variables for degree days and onsite wastewater treatment were collected from external sources, not included in the Census data. Both cooling degree days and heating degree days were sourced from the ENERGY STAR Portfolio Manager Degree Days Calculator and matched with plants within the Census data based on ZIP code. The presence of an onsite wastewater treatment plant was determined by using the EPA's Enforcement and Compliance History Online (ECHO) database. Sites were filtered by the six-digit NAICS code for distilleries (312140) and matched with the existing sample of Census data. A wastewater treatment plant was identified based on the presence of specific permits related to the Clean Water Act that would indicate onsite treatment. The match rate between sites identified in the ECHO database and the Census data was not perfect.

Industry participants were critical of the results of the Census-based model for a variety of reasons, the largest being the exclusion of a variable to account for the method of byproduct handling mentioned above. Various focus companies estimated that dryhouse operations (the process that removes water from spent grain) could contribute as much as 40% of the total energy used at the distillery, meaning that without a variable to account for different handling methods, the scores from this model would be heavily biased against companies with onsite drying.

Other comments centered around the large impact that the weather variables were having that was not consistent with the expectations of the energy managers. Additionally, the data from 2012 pre-dates the large increase in craft distillers in the United States, leaving an underrepresentation of smaller plants in the underlying data. Without an option to address these concerns with the existing dataset, industry focus participants volunteered to provide data for a more robust analysis.

3.2 Using Industry-Supplied Data

Moving forward with industry-supplied data, discussions within the distilled spirits focus centered on what variables are suspected to contribute the most to energy use and what data all facilities are currently tracking. Variables were classified into the six categories listed below; the types of data collected in each category are briefly described. A voluntary data collection form was developed with the industry and distributed via the Focus on Energy Efficiency in Distilled Spirits Production and the Distilled Spirits Council of the United States (DISCUS). All data provided for EPI research and development was covered under company-specific non-disclosure agreements with Duke University. No data were provided to EPA.

- Energy
    - Electricity and fuels of all types
- Process Inputs
    - Raw materials and purchased spirits
- Distilling

    - Total production volumes and distillation processes
- Support Processes
    - Byproduct processing methods and environmental controls
- Bottling
    - Total volume and bottling method
- Other Onsite Uses
    - Climate control and presence of additional onsite operations

With a dataset more tailored to the needs of the EPI, Draft Version 1.1 was developed using many variables in an attempt to capture as much variation across the industry as possible. The distinct differences between electricity and fuels seen in Draft Version 1.0 were no longer present statistically in Draft Version 1.1, so a total energy model was used instead. Variables included total distilled production, the percentage of moisture removed from the byproducts, adjustments for specific raw material inputs, cooling degree days (CDD), and a binary variable for the presence of air pollution control equipment. Also, a squared term was included to capture the non-linear relationship between the total distilled production variable and energy consumption. Total inputs were included as a way to differentiate between spirit types, with an emphasis on capturing the differences in the physical weights and additional handling required for heavier input types of spirits such as agave when producing tequila. In addition to specific share variables for certain input types, a yield variable was included for total input weight to total output. Specific input share variables included in the model were the amount of barley, corn, rye, and purchased spirits. Variables that were tested but did not have a statistically significant impact included on-site wastewater treatment, bottling operations, and other space uses such as tasting rooms, offices, etc. Although these processes do increase energy usage, the impact could not be determined due to small sample sizes within the data or potentially being in the statistical noise when compared to the larger energy drivers for the industry.

Industry feedback on the Draft Version 1.1 model results were varied, with the effects of some variables aligning with expectations and others having a larger impact than expected. Increasing production was found to have a larger energy benefit for smaller companies. Similar to other industries, there are general efficiency improvements that can be made as scale increases (economies of scale), but this might come with diminishing returns. This effect was seen by some reviewers to be biased against the largest plants in the distribution, setting unrealistic benchmarks. Although the squared term was intended to capture the non-linear relationship between production and energy, it was not adequate given the extreme production variations within the industry. Also, the reviewers felt predicted energy impacts of air pollution control equipment and the share of rye in the total input mix were too large and not reflective of industry knowledge. For example, although the presence of different types of air pollution control measures can use additional energy, the model estimated that air pollution equipment could account for as much as 40% of total energy at the distillery. Focus participants also were not aware of any process reasons that would explain the drastic increase in predicted energy usage by adding rye to the mash bill.

Estimates from focus participants regarding the energy needed to operate a dryhouse and remove all of the moisture from their byproducts were 30%-40% of total energy at the plant. The estimated coefficient fell within this range at 36%, implying Draft Version 1.1 of the model was able to account for one of the largest shortcomings of the first draft based on Census data. Another effect seen in the second draft model was that distilleries that had a higher percentage of their total distilled production originating from redistilled purchased spirits received lower predicted energy results. Because a goal of the focus was to produce one benchmarking tool that could account for the variety of distillery configurations, having a method that could successfully account for distilleries that have limited or no on-site fermentation, mashing, etc. was a desirable result. Overall, the model captured some of the main energy drivers but also had some inexplicable results that could have resulted from spurious correlations in the dataset.

Based on the detailed industry feedback and, potentially, unrealistic energy impacts in Draft Version 1.1, a simplified approach was taken when developing Draft Version 1.2. Focusing solely on the main energy drivers within the industry that had robust energy coefficients from the previous model helped remove some of the more "suspect" energy estimates. The estimated coefficients in the model provide information on the relationship between the independent variables (in this case, production quantities, moisture removal, etc.) on the dependent variable (in this case total energy). Specifically, the sign and magnitude of the coefficients determine the impact that the independent variables have on the predicted value of the dependent variable. Only three variables were tested in the updated analysis: total distilled production volumes, purchased and redistilled spirits, and the percentage of moisture removed from byproducts. In order to account for the challenge of the large variations in production volumes within the distilled spirits industry, a linear spline was included in the model, allowing for different economies of scale based on the production volume of the distillery. This approach accounts for the large variation in production quantities present within the industry between the smaller craft distillers and the larger macro distilleries better than the previous squared term approach. More information on the linear spline is detailed in section 3.4 below. Other coefficients for the moisture removal and purchased spirits variables were robust across the different models and within the expectations of the focus participants. Previously tested variables for wastewater treatment, bottling quantities, air pollution control equipment, and input weight ratios were reintroduced into the model one at a time but did not have consistent or statistically significant results.

Overall feedback for this model was positive and progressed closer to the initial expectations that focus participants had based on their knowledge of their facilities. One final issue arose based on the additional energy requirements for plants in colder climates. Since colder weather by itself does not increase the required energy for the production processes, weather would only impact the energy consumption if some spaces required climate control due to colder temperatures that would not be controlled for in warmer climates. To adjust for this, the heating degree days (HDD) variable for the plant location was interacted with a binary variable that indicated whether or not the plant had climate control in these spaces.

The final sample consisted of 59 observations with data from 11 companies and 26 individual distilleries. Although the underlying sample is small compared to the total *number* of distilleries in the country, it comprises a substantial amount of the total volume, especially whiskey production. Total whiskey production in the sample is roughly 75% of the total volume reported to the Alcohol and Tobacco Tax and Trade Bureau (TTB) for the year 2020 (Alcohol and Tobacco Tax and Trade Bureau, 2023). More details on the statistical methods used in the analysis are detailed below.

3.3 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, distilleries. 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, although the ENERGY STAR EPI approach predates the release of ISO 50001 (Boyd, Dutrow, & Tunnessen 2008)[2].

The concept of the analysis that supports the EPI can be easily described in terms of the ordinary least squares (OLS) linear regression model, which is reviewed in this section. 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.[3]

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 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

[2] Both ENERGY STAR and ISO 50001 use the term Energy Performance Indicator. Since ENERGY STAR began publicly using the term first, ISO adopted the acronym "EnPI" to limit confusion.

[3] Interpreting ε as the measure of relative plant efficiency is a strong assumption. Stochastic frontier analysis (SFA) has been employed to decompose the error term into random noise and efficiency. The analysis presented below did test the SFA approach and the maximum likelihood estimates were unable to decompose the error term, so the simpler OLS model was used.

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 with parameters estimated from the underlying data (denoted with " ^ ") 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 ($\hat{e}$) between the actual energy use and the predicted average energy use from equation (3).

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

For the models using ordinary least squares (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 from equation (4) under the assumption that ε is normally distributed with zero mean and variance, $\widehat{\sigma^2}$, which is estimated via OLS.

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

We take probability in equation (5), subtract it from the value of one,[4] and multiply by 100. This is the Energy Performance Score (EPS), which 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 plants in my industry, if all other plants were similar to mine?"

The final equation for the regression model underlying the distilled spirits EPI is shown in equation (6). The impact of production on energy was measured using a linear spline, where the five knots represent different size "bins" in the linear relationship, indicated as k_n. Linear spline is described in more detail in section 3.4 below.

[4] 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.

$$\ln(energy) = \alpha + \beta_1 \ln(total\ production - k_1) + \beta_2 \ln(total\ production - k_2) + \beta_3 \ln(total\ production - k_3) + \beta_4 \ln(total\ production - k_4) + \beta_5 \ln(total\ production - k_5) + \beta_6(moisture\ removal\ pct) + \beta_7\left(\frac{redistilled\ purchased\ spirits}{total\ production}\right) + \beta_8(climate\ control\ dummy * \frac{HDD}{HDD\ adjustment}) + \varepsilon \quad (6)$$

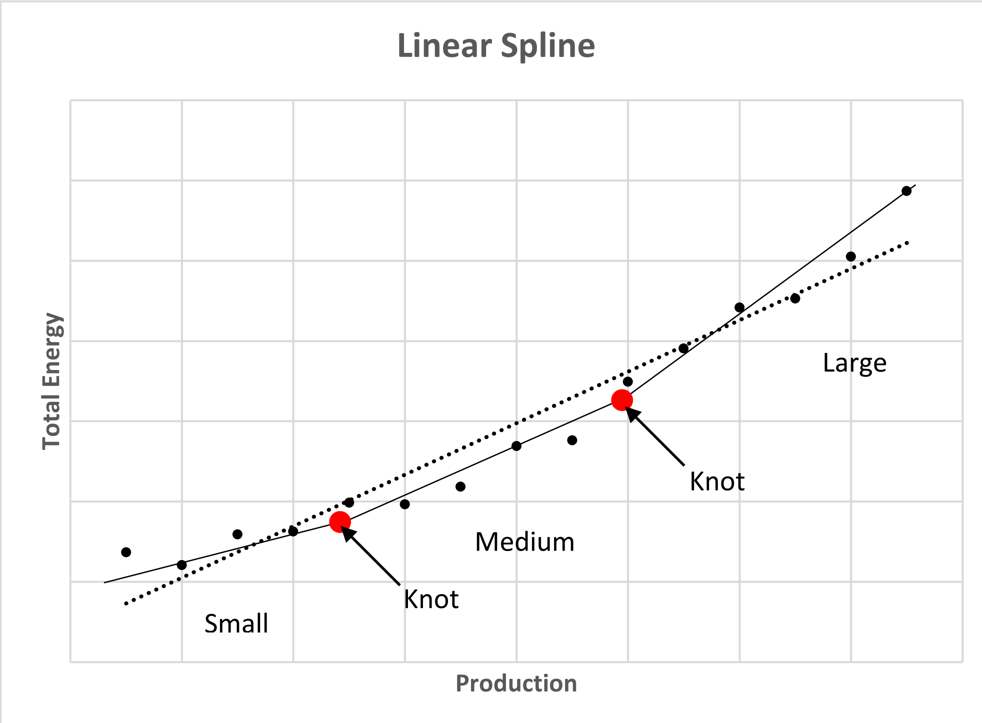
After the final analysis was completed, it was discovered that some plants outside of the United States and Canada, the geographic scope of the ENERGY STAR program, were skewing the model distribution, resulting in fewer than 25% of the plants within the United States and Canada being in the top quartile. Since ENERGY STAR certification is not offered outside of these two countries, and it was not clear how similar the operating conditions of these plants were to U.S. and Canadian plants, they were removed from the final Energy Performance Score (EPS) distribution, but remained in the underlying regression model. Removing these observations after running the model means a kernel density is used to create a non-parametric distribution of plant efficiencies for the remaining observations. The kernel density is fit to the simulated plant efficiencies and the support points of the non-parametric kernel are numerically integrated to generate the cumulative function needed to compute the EPS.

3.4 Linear Spline

Plants that produce higher volumes of spirits may experience economies of scale in the sense that doubling production may not double energy use. Essentially, the energy intensity for plants with lower production may be higher than for those with higher volumes. Initial analysis of the data strongly suggested that higher production volume does result in lower average energy intensity, but industry review comments also indicated that this “advantage” may differ across the wide range of distillery sizes. The energy advantage of increased production for a very small plant may be substantial, but the energy advantage of increased production for a very large plant may have diminishing returns. There is a size at which the advantages of higher production have reached their maximum, and higher production volumes no longer result in lower average energy intensities.

To capture this empirical and anecdotal evidence, the EPI analysis employs a very flexible representation of the production/energy relationship called a linear spline regression. This statistical model allows for the relationship between energy and production to differ by creating different “bins” of data. To illustrate, a hypothetical dataset of production and energy is split into 3 bins of equal sizes based on production: small, medium, and large. The demarcation between the bins are called “knots;” the number of knots is always one less than the number of bins. If a standard linear regression model is fit to the entire data set, the model underestimates energy for the largest and smallest plants, while overestimating energy use for the medium plants. This would be very undesirable for an energy benchmark. In the linear spline, each bin has its own energy-to-production relationship. Small plants’ energy use grows more slowly with production; medium plants’ energy use grows more rapidly; and the energy use in the largest plants grows most rapidly. The linear spline meets at each knot (i.e., no discontinuity), but has a different slope on either side of the knot. In this example, the two-knot, three-bin linear spline does a much better job of fitting the entire dataset.

Figure 1 Linear Spline Illustration



In the example above, there are three bins with two knots that are equally spaced in terms of size of production; i.e., bins have the same size, and also number of data points in each bin. This example was purely hypothetical and an equal number of data points in a bin with equal widths is highly unlikely in the real world. The number of bins in the example was chosen to illustrate the hypothesis that small, medium, and large plants have diminishing returns in the energy production relationships. Empirically we do not know what constitutes "small" vs "large," nor do we know if the resulting model estimates will follow our hypothesis regarding diminishing returns.

To estimate a linear spline regression, the number and position of the knots for a linear spline must be determined by the analyst. This essentially requires choosing the number of bins and their width. The width need not be the same for each bin. While there are various methods for choosing the "optimal" number and width, the number and position using knots for the distilled spirits analysis were placed at the 20th, 40th, 60th, and 80th percentiles of the log of production.[5] This results in the same number of data points in each bin, but varying widths. Five bins were selected to increase the flexibility in the analysis, but not "slice the data too finely."

4 FINAL MODEL ESTIMATES

[5] Other placement was tried and the results were compared by plotting the resulting regression residuals, which are directly related to the 1-100 Energy Performance Score. This placement appeared to provide a good overall fit in terms of flexibly capturing the energy production relationship, given the relatively small dataset.

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, energy, and energy intensity for the full sample, while Tables 2 and 3 supply the same averages but split distilleries in the dataset based on whether or not they remove moisture from their byproducts. The average intensity for plants that remove moisture from byproducts is lower than for those that do not remove moisture due to correlations between moisture removal and distillery size where there are large efficiency differences. In other words, larger plants are more likely to remove moisture from the byproducts and have a lower intensity due to their larger volumes, even though moisture removal demands more energy. These average intensities are similar to other published benchmarking numbers for the industry. The standard deviations are large for all three tables, reflecting the vast diversity of size and energy performance in this industry.

Table 1 Plant Average Sample Statistics (Full Sample)

	Production (proof gallons)	Total Energy (MMBtu)	Energy Intensity (MMBtu/proof gallon)
Plant Average	14,431,680	603,348	0.0590
St. Dev.	19,710,930	706,444	0.0391

Table 2 Plant Average Sample Statistics (0% Moisture Removal)

	Production (proof gallons)	Total Energy (MMBtu)	Energy Intensity (MMBtu/proof gallon)
Plant Average	3,123,107	126,997	0.0735
St. Dev.	5,531,670	171,215	0.0572

Table 3 Plant Average Sample Statistics (Greater than 0% Moisture Removal)

	Production (proof gallons)	Total Energy (MMBtu)	Energy Intensity (MMBtu/proof gallon)
Plant Average	21,656,600	907,683	0.0497

St. Dev.	22,065,170	751,141	0.0159

Final parameter estimates are shown in Table 4. When we examine the coefficients for the five production-variable bins we see that the smallest plant bin has a coefficient of 0.59, while the next bin has a coefficient of 0.85. This means that a 10% increase in production for the smallest plants increases energy use by ~6%, but by ~8.5% for the next largest group. As production volumes increase, the predicted energy intensity reduction decreases. The model shows there are energy advantages to increasing production that are mainly concentrated around the smallest plants in the industry. When an already large plant increases production volume, the efficiency advantage is much smaller or nearly non-existent. The largest three bins have coefficients of 0.94, 0.55, and 0.98. While the coefficient for the fourth bin is lower than the adjacent two, this coefficient has a rather high standard error and reflects higher uncertainty of the precise value. In fact, all three of these coefficients are not statistically different than 1.0. However, the EPI does not impose a one-to-one relationship for these largest groups of plants, but rather lets the actual data determine the benchmark. It should be noted that, while the slope of the fourth bin is lower than for the third and fifth, it still reflects increased energy use as production changes. It remains an open question whether plants of this size do have some unique benefits or if these estimates are an artifact of the particular data sample.

Moisture removal percentage can be similarly interpreted. The coefficient of 0.34 means that with a 10% increase in the byproduct moisture removed, energy usage would increase by 3.4%. A distillery that removes 100% moisture would use 34% more energy than one that distributes their byproduct wet, with zero moisture removal.

The purchased spirits variable is a ratio of purchased and redistilled spirits to total distilled production to determine how much of the product was distilled on-site versus elsewhere. As that ratio increases, the expected energy use would decrease as less mashing, cooking, and fermenting processes are being performed on-site for the same output. The estimated Purchased Spirits coefficient (-1.77) implies that a plant that uses 100% purchased spirits will use only 17% of the energy of a distillery that does all the distillation on-site (i.e., e^-1.77 = 0.17 or 17% of a full plant).

Finally, HDD is normalized to an arbitrary region – in this case, Durham, NC – so that the coefficient represents percent departures above or below that location, given the distillery heats the on-site production or warehouse spaces (reflected in Climate Control variable). With a Climate Control coefficient of 0.067, a one unit increase to the ratio would increase predicted energy by roughly 6.7% to heat the same amount of space. In this example, a one unit increase in the ratio variable means double the HDD of Durham, NC.

Table 4 Distilled Spirits Model Results

Variables	Coefficient	Standard Error	Spline Knot Points	Spline Knot Points (Proof Gallons)
Production Volume 1	0.590***	0.096		
Production Volume 2	0.850***	0.152	13.55	766,814

Variables	Coefficient	Standard Error	Spline Knot Points	Spline Knot Points (Proof Gallons)
Production Volume 3	0.942**	0.378	15.37	4,732,670
Production Volume 4	0.543*	0.301	16.07	9,530,426
Production Volume 5	0.977***	0.210	16.84	20,583,493
Moisture Removal %	0.345	0.219		
Purchased Spirits	-1.77***	0.256		
Climate Control	0.067	0.059		
Constant	2.72	1.19		
Final Regression Results				
Dependent Variable	Total Source Energy			
Number of Observations	59			
R^2	0.954			
Adjusted R^2 value	0.947			
F Statistic	129.18			
Significance (p-level)	< 0.0001			
RMSE	0.435			

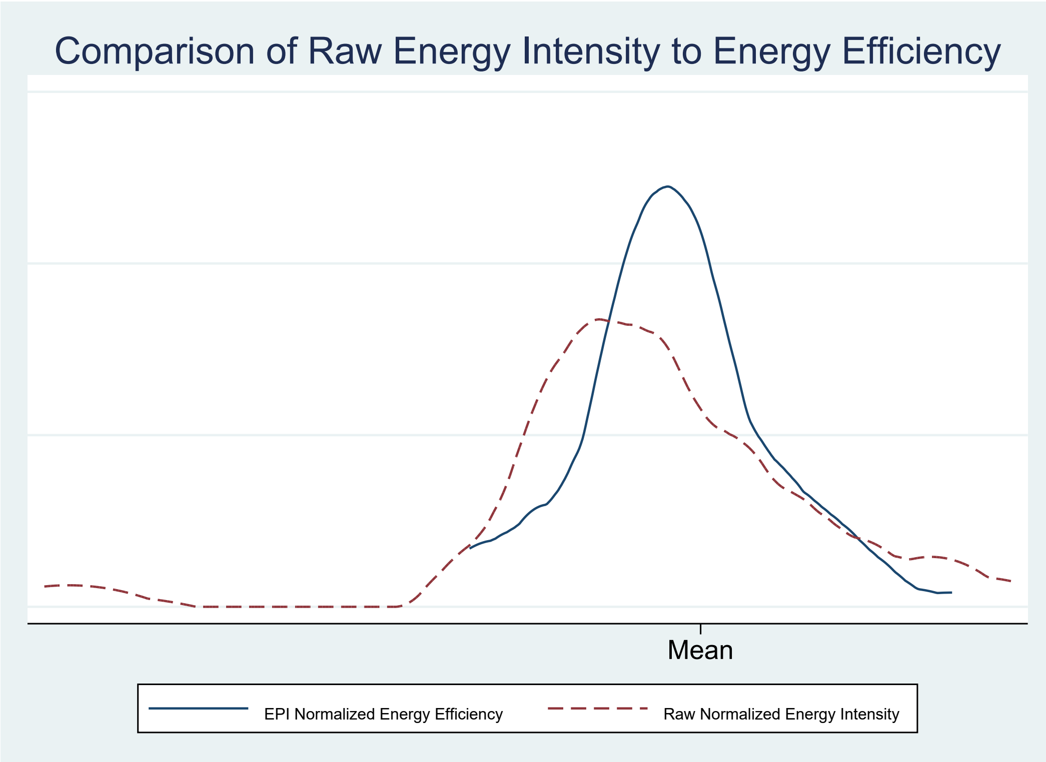
*** Significant at the 99% level; ** Significant at the 95% level; * Significant at the 90% level

4.2 Stylized Results

When examining the raw data on energy intensity (energy/total proof volume) for the distilled spirits industry, the range of performance varies. Tables 2 and 3 show that the energy intensity coefficient of variation – i.e., the standard deviation divided by the mean – is 0.78 and 0.32 for plants that do and do not remove byproduct moisture respectively, while the full sample has a value of 0.66. For plants that remove moisture, the range of performance is much narrower than for those that do not, but these plants also have substantially larger production volumes on average and use more total energy. When comparing the range of intensities to the range of EPI performance where additional factors have been accounted for, the range of efficiency is narrower still. This is consistent with the results of a meta-analysis of EPI studies for other industries (Boyd 2016). In the distilled spirits industry, the difference in total energy consumption between an “average” plant (score of 50) and an “efficient” plant (score of 75) is roughly 18%. These results show a similar spread to some of the light industries and illustrate the opportunity for energy efficiency improvements.

For illustration, the dashed red line in Figure 2 takes the raw energy intensity data and transforms it into the kernel density distribution of plants that lie above or below the average total intensity of 0.059 MMBtu/proof volume represented as a percent difference. The solid blue line, representing the kernel density from the EPI analysis, shows that most of the intensity differences come from differences that are accounted for in the analysis. These results show how comparing distilleries based on energy intensity alone will misjudge the actual energy efficiency of the plants.

Figure 2 Comparing the Distribution of Energy Intensity to Efficiency

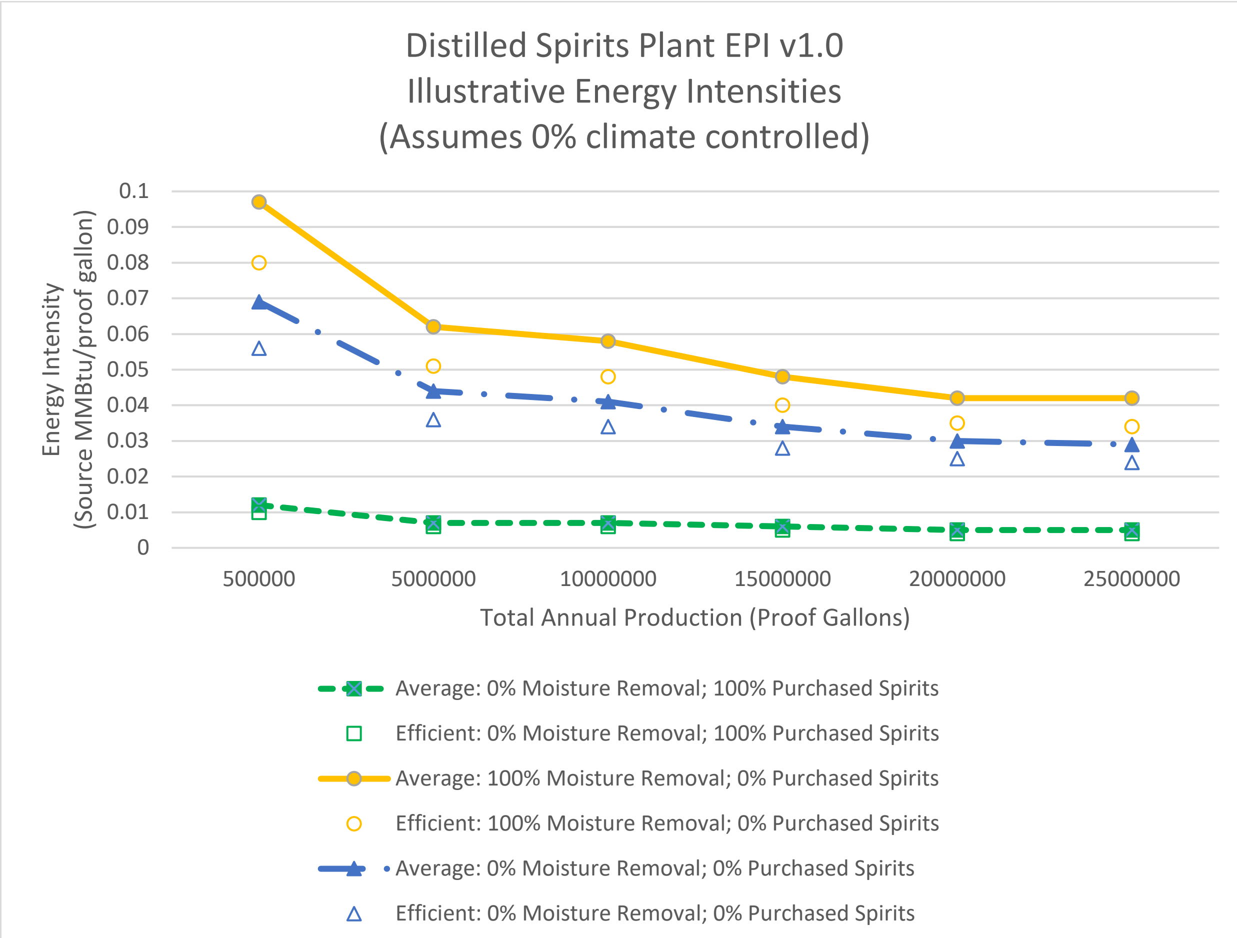


4.3 Visualizing Distilled Spirits Efficiency

The EPI predicts how much energy a distillery would consume based on different production volumes, and adjusts the predicted energy to account for differences in distillery characteristics and processes. These characteristics can include whether the distillery purchases spirits or dries the byproducts from the distillation process. This allows a distillery to get a more precise estimation of energy efficiency based on its unique operating conditions.

Distilleries can use the EPI to see how much energy an “average” or an “efficient” distillery with identical characteristics would be predicted to consume. The graphs below visualize how the energy intensities change based on a range of production volumes. Distilleries can use these graphs to quickly get an idea of how much energy the model predicts average and efficient plants of various production quantities and characteristics would consume.

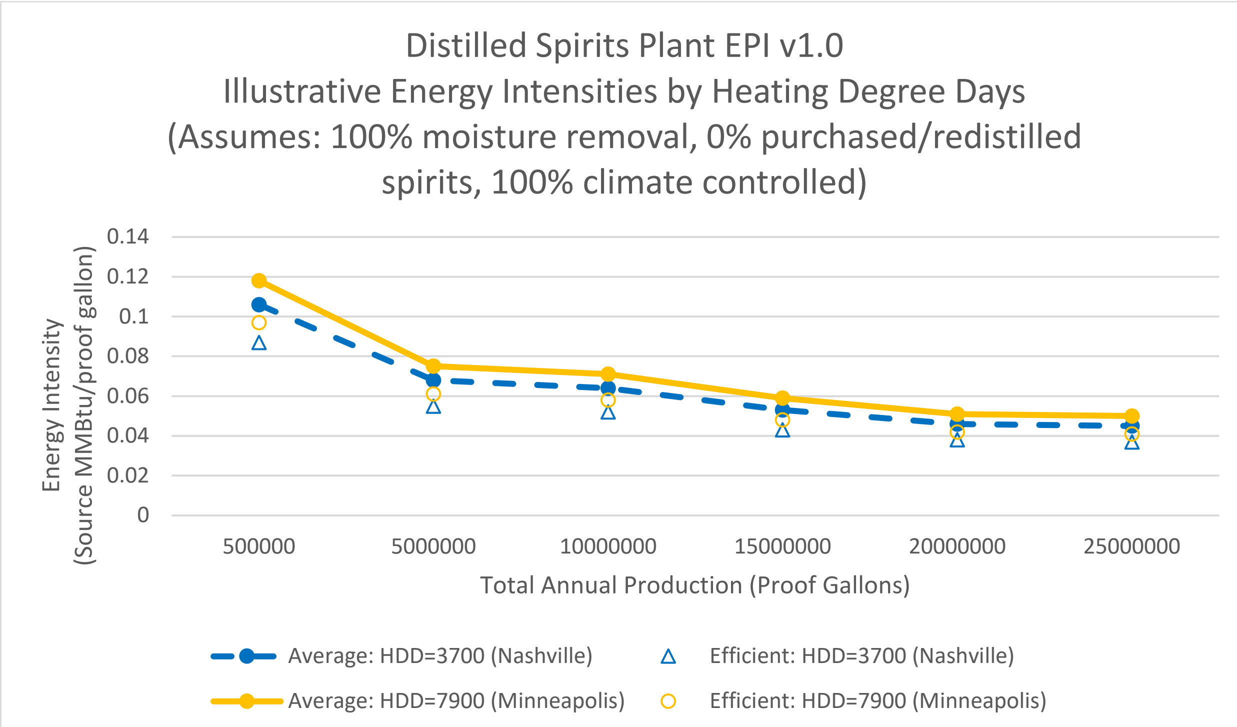
Figure 3 Distilled Spirits Predicted Energy Intensities for a Variety of Plant Characteristics



A plant that distills from a mash or wash (solid and dash/dotted line), which starts with a low alcohol content, uses more energy than one that distills purchased spirits (dashed line), which are at a higher alcohol content. The graph also shows that distilleries that remove 100% of the moisture from the grain and other byproducts (solid line) require more energy per proof gallon distilled than a distillery that does not remove any moisture (dash/dotted line).

In all scenarios, the distilleries operate at a lower energy intensity as total mix production increases; i.e., as production volume increases, the amount of energy required to produce one proof gallon of alcohol typically decreases. However, economies of scale in the distilled spirits industry are not linear. As plants with lower production volumes increase their proof gallon output, they would expect to see larger intensity improvements than a plant whose production volume is already high and increases their output by the same amount.

Figure 4 Distilled Spirits Predicted Energy Intensities Based on Weather



The energy required to operate a distillery also depends on whether parts of the production area and warehouses have space heating. The solid line represents the amount of energy required per proof gallon in a cold climate and the dashed line represents the energy intensity required in a temperate climate. In the scenario above, the colder climate requires ~11% more energy than one in a temperate climate. However, this value is dependent on the climate and the percentage of total space that is heated.

5 SCORING DISTILLED SPIRITS EFFICIENCY

This section describes the spreadsheet tool that was created based on the above analyses, and published as Version 1.0, Release 06/12/2023. Suggestions for how to use the tool and interpret the results are also shown below.

5.1 How the Distilled Spirits EPI Works

The Distilled Spirits EPI scores the energy efficiency of facilities that are primarily engaged in manufacturing or processing distilled spirits, with more than 50% of their product value coming from the distillation of whiskey, grain neutral spirits, vodka, gin, rum, or other liquors, excluding brandy. Facilities whose sales are less than 50% of the listed products that can submeter the relevant production activities would also be eligible. To use the tool, the following information must be available for a plant for a consecutive 12-month period:

- 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
  - Heating degree days (HDD)
- Plant Characteristics
  - Total proof volume of distilled production
    - A proof gallon/liter is one liquid gallon/liter of spirits that is 50% alcohol at 60°F. For example, distilled spirits bottled at 80 proof (40% alcohol) would be 0.8 proof gallons per gallon of liquid. At 125 proof, a gallon of liquid would be 1.25 proof gallons.
  - Amount of purchased and redistilled spirits
  - Percentage of moisture removed from distillation byproducts
    - To calculate, subtract the weight of the byproduct after processing from the pre-processing weight and then divide by the difference of the pre-processing weight and the weight of the byproduct if it were fully dried, i.e.:

$$\frac{\text{Weight: prior to processing} - \text{Weight: post processing}}{\text{Weight: prior to processing} - \text{Weight: if fully dried}}$$

  - Percentage of the production or warehouse space that is climate controlled (heated)

Based on these data inputs, the EPI will report an Energy Performance Score (EPS) for the plant in the designated 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 (i.e., the plant performs better than 75% of plants if the entire industry shared identical characteristics). 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. 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[6] set of plant-level inputs. Energy managers in the distilled spirits industry were encouraged to test the EPI by inputting data for their own plants and then provide comments on the results to the developers. Approximately 73% of companies that provided

---

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

data reviewed each draft version of the model and provided feedback. 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.[7] Example inputs and outputs of the spreadsheet-based tool are shown in Figures 5-6.

Figure 5 Input Section of the Distilled Spirits EPI Spreadsheet Tool

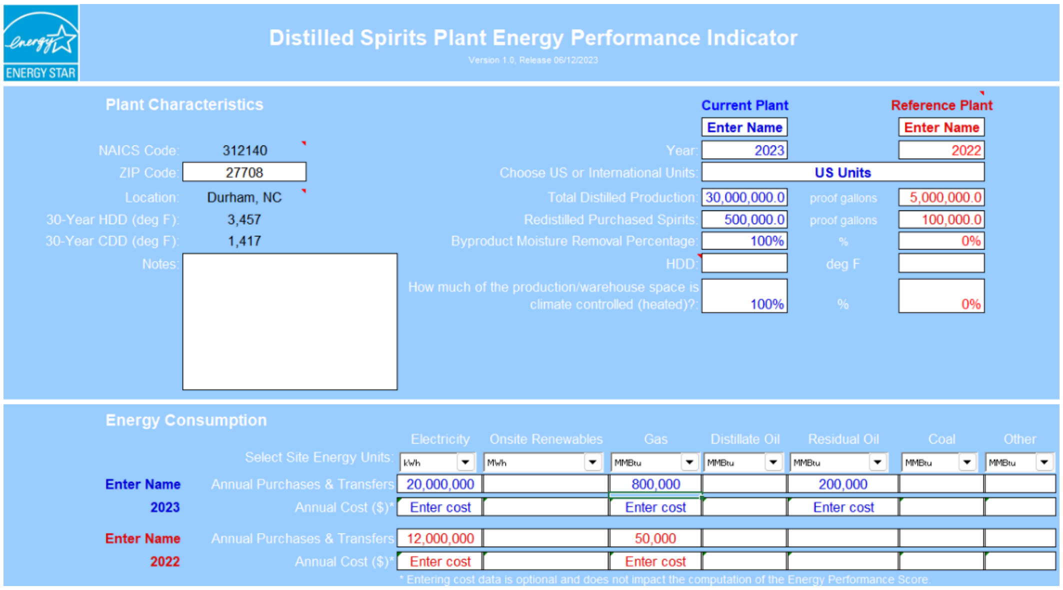
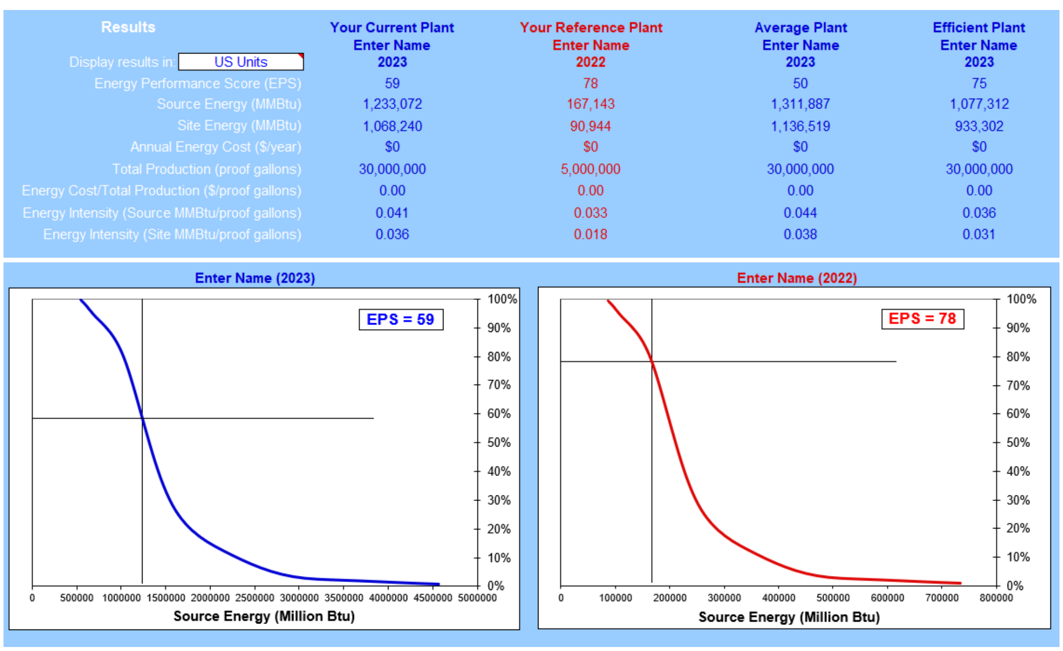


Figure 6 Output Section of the Distilled Spirits EPI Spreadsheet Tool



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

5.3 Use of the ENERGY STAR Distilled Spirits 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 model described in this report is 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 model will need to be updated. EPA plans to improve the model every few years, contingent on newer data being made available and industry use and support of the EPI tool.

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)[8] use of user provided energy consumption data

- 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.

[8] Total Source Energy is the total amount of energy used onsite by the plant (site energy) and the transmission, delivery, and production losses to get the energy to the site.

- TSE per relevant unit of production is also computed.

3. EPI computes the Predicted "Average Practice" 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.
- The lookup table returns a score on a scale of 1-to-100.
- The Predicted TSE for a median and 75^{th} 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

Alcohol and Tobacco Tax and Trade Bureau (2023). *Distilled Spirits Statistics*. https://www.ttb.gov/distilled-spirits/statistics

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. (2006). Development of a Performance-based Industrial Energy Efficiency Indicator for Cement Manufacturing Plants. Argonne IL, Argonne National Laboratory. https://www.energystar.gov/sites/default/files/buildings/tools/ANL-DIS-06-3.pdf

Boyd, G. (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.

Stuckrath, C., and E. Worrell (2022). "Energy Efficiency Improvement and Cost Saving Opportunities for Distilleries." US Environmental Protection Agency. https://www.energystar.gov/sites/default/files/tools/Distilleries_guide_final%2020220202_508.pdf

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.