Key Considerations for Industrial Benchmarking in Theory and Practice

September 2013

By: Kyle Aarons, Solutions Fellow
Center for Climate and Energy Solutions
For the Industry Working Group of North America 2050

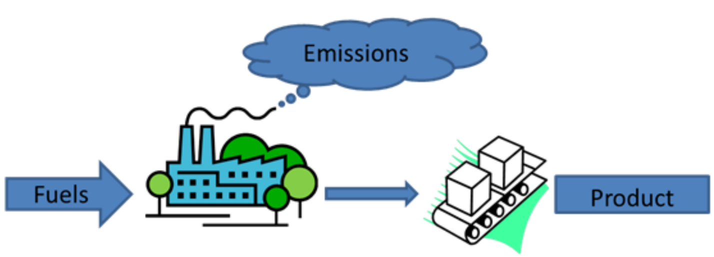
Industry Working Group Co-Chairs:
Craig Golding, Manager, Cap and Trade, Ontario Ministry of the Environment
Stuart Clark, Program Manager, Air Quality, Washington State Department of Ecology

Contents

Purpose

The term "benchmarking" can refer to a wide variety of concepts. For the purposes of this paper, "benchmarking" refers to developing and using metrics to compare the energy or emissions intensity of industrial facilities. Benchmarks are primarily used to compare facilities within the same sector, but can also be used to identify best practices across sectors where common process units, such as boilers, are used. Figure 1 displays a basic example of a benchmark, showing the typical, but not exclusive, representation as tons of emissions per tons of product (e.g., tons of greenhouse gases per tons of steel produced). Several programs exist at multiple levels of government that either encourage or require a reduction of emissions or energy intensity at an individual plant over time. An example is a recognition program that rewards steel mills that reduce their energy consumption per ton of steel by five percent over ten years. This type of program does not necessarily require consistent methodologies across steel mills, and therefore is not the primary focus of this paper. However, useful lessons can still be drawn from these types of programs, which are discussed in the "Benchmarking in Practice" Section.

Figure 1: Emissions benchmark definition



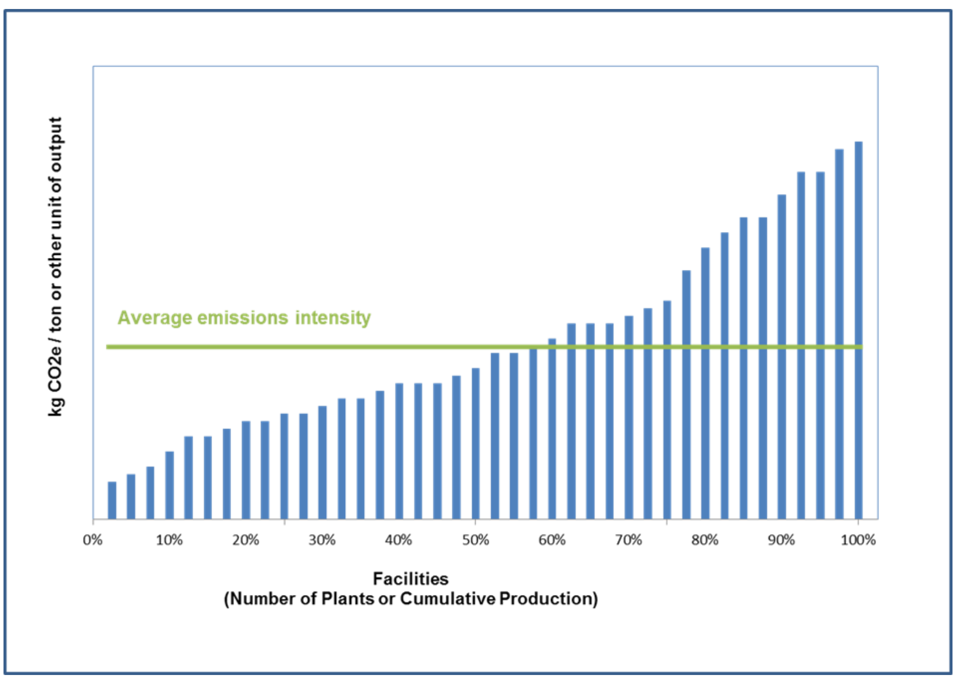
$$Benchmark = \frac{tons\ \ emissions}{tons\ \ product}$$

The intended audience for this paper is government officials in the early stages of considering a new program that involves benchmarking, or the stakeholders interested in the development of such a program. Industry trade organizations or individual companies may also consider establishing voluntary benchmarking programs. However, recommendations made by the Industry Working Group of North America 2050 (Working Group) in this paper are not expressly intended for these industry-initiated programs. Programs established by trade organizations can provide useful lessons for government-run programs and a few of these programs are discussed in the "Benchmarking in Practice" Section.

The goal of this paper is to encourage consistency in benchmarking methodology across programs within a single jurisdiction, as well as across jurisdictions. When facilities are benchmarked using a consistent methodology, it is possible to identify best practices as well as opportunities for improvement across sectors and jurisdictions. For example, two paper mills in neighboring states will only be able to compare their performance if both states use the same data collection methods and metrics. To encourage such consistency, this paper defines and explains key issues that arise when policymakers establish a benchmarking program. It also includes guiding principles recommended by the Working Group based on benchmarking literature and successful programs.

Benchmarking can be valuable because it increases the availability of industrial efficiency data so that businesses and stakeholders can more effectively identify and implement improvements to increase productivity. In this context, benchmarking is used by both industrial trade organizations and governments. Additionally, benchmarking is a tool to push facilities to make efficiency investments they may not otherwise make. Using benchmarking data, policymakers can set a voluntary or mandatory performance target, for example, to achieve efficiency gains. Figure 2 displays a hypothetical benchmarking curve and target. The ultimate purpose of benchmarking has bearing on the benchmarking methodology, as further discussed in the "Use of Benchmarking" subsection, and in this limited context the policy purpose of benchmarking is discussed in this paper. However, this paper primarily focuses on benchmarking itself and only touches on the issues policymakers must take into account when determining where to set the efficiency target, whether it should be voluntary or mandatory, how to enforce the target, and so on.

Figure 2: A benchmarking curve is developed by arranging each facility in order of emissions intensity. Policymakers must then determine the target intensity; an average is one of many options.



This paper addresses both emissions benchmarking (emissions released per unit of production) and energy benchmarking (energy consumed per unit of production). Principles that apply to one of these metrics almost invariably apply to both, but a separate discussion of each is required since the benchmarking programs surveyed address one or the other. For example, allowance allocation for the California cap-and-trade program involves emissions benchmarking, while Energy Star for Industry involves energy benchmarking. Unless explicitly noted, the reader should assume that the issues and recommendations throughout this paper apply to both energy and emissions benchmarking. Volumetric benchmarking is another option for quantifying emissions from industrial processes, which takes the form of parts of pollutant per million parts of total emissions (ppm). Volumetric benchmarking is touched on in the discussion of New Source Performance Standards, but is otherwise not addressed in this paper since it is not generally used for greenhouse gas emissions.

Benchmarking Issues and Recommendations

To ease navigation, each section below begins with the main question that will be answered within the sections and a list of the Working Group's recommendations. These recommendations are collected in the following section: Summary of Working Group Recommendations.

General Policy Considerations

Key Question: What factors influence the success of a benchmarking program aside from the benchmarking methodology itself?

Recommendations:

- There are several critical preconditions to a successful benchmarking program:
  - Common definitions
  - Reliable data
  - Reliable measurement and verification systems
  - Committed contributions
- Targets set based on benchmarks should be SMART:
  - Specific
  - Measurable
  - Appropriate
  - Realistic
  - Timed

Before getting into the specific decisions that must be addressed when creating a benchmarking methodology, it is useful to have a set of general principles to follow. Several conditions must be met before a benchmarking policy should be considered:[1]

- The availability of common definitions, including which greenhouse gases will be covered if an emissions benchmark is being used, and clear standards as to which industrial sectors and facilities are covered under which terms;

- The availability of reliable, consistent data;

- The availability of reliable, consistent measurement and verification systems to ensure confidence in emissions and production data; and

- The contribution of considerable efforts by all stakeholders, including industrial representatives.

Two central elements of a benchmark are that it allows for a fair comparison of facilities regardless of their size, and that it is applicable to a wide range of facilities to ensure only a single benchmarking methodology is required. This paper focuses on the creation of benchmarking curves rather than setting a specific benchmark for a sector, but it may be helpful to consider some target-setting principles. Rietbergen and Blok provide some guidance for this in their paper, Setting SMART Targets for Industrial Energy Use and Industrial Energy Efficiency. This guidance can be remembered with the acronym SMART (Specific, Measurable, Appropriate, Realistic, Timed). These principles are primarily relevant for mandatory or voluntary performance standards:[2]

- Specific: The target must explicitly specify what emissions intensity is to be achieved;

- Measurable: The target must allow for regular measurement of how much progress facilities are making toward the target in order to ensure it is encouraging action;

- Appropriate: Targets must be relevant to the overall aims and objectives of a broader policy;

- Realistic: The target must be achievable within the compliance period, which involves a careful balance. Targets must push operators to make efficiency investments beyond business-as-usual, but must not be so strict that they force facilities to leave the jurisdiction or simply ignore the targets as unachievable; and

- Timed: Targets must have a set timeframe, and should be set in the short or medium term to promote early action.

Data Gathering

Key Question: What data are necessary for a successful benchmarking program and how should they be gathered?

Recommendations:

- Policymakers should consider mandatory reporting regulations prior to the development of a benchmarking program if reliable, consistent data are not already available
- Benchmarking data should come from operating facilities rather than the theoretical literature whenever possible
- Data used for benchmarking should be transparent whenever possible, but trade secret laws may prevent this in some cases
- Policymakers should consider what balance to strike between fostering industry cooperation and designing a strict program

In order to implement an industrial benchmarking program, officials need reliable data on the numerator of the benchmark equation, generally in the form of energy consumed or greenhouse gas emitted, and the denominator, generally in the form of mass of product produced. Reliable, consistent data are critical to a successful benchmarking program, though the level of accuracy required varies by the purpose of the benchmarking program.

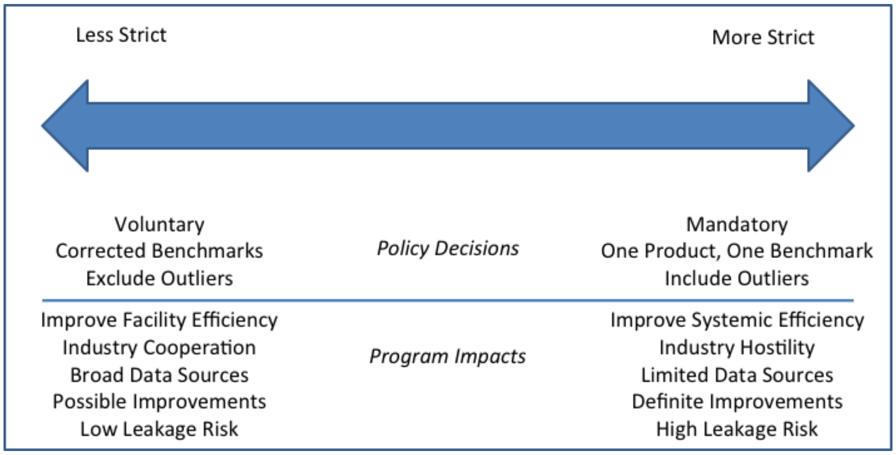
Benchmarks can be derived from different types of sources. It is easier to set benchmarks based on arbitrary targets or the theoretical literature, but these can be unrealistic and lack transparency. It is therefore preferable to base benchmarks on data collected from operating facilities.[3]

Data sources for production and energy consumption at the facility level can be difficult to find. Sources include industry groups and trade associations, such as the Cement Sustainability Initiative, International Aluminum Institute, or Northwest Food Processors Association.[4] Unfortunately industry data cannot always be relied upon because they are typically not publicly available and do not include all industrial sectors.[5] Another possible source for data is government surveys, though confidentiality is also an issue here.[6] These surveys include the Energy Information Administration's Manufacturing Energy Consumption Survey (MECS) and the U.S. Census Bureau's Annual Survey of Manufactures and Economic Census.[7] All publicly available data collected by U.S. Department of Energy (DOE) Industrial Assessment Centers is housed in a central database.[8] These surveys can include production and consumption data, which can be converted to emissions data with standard conversion factors. Air permits held by state or local air agencies may also contain useful data, though the absence of a consistent monitoring methodology across agencies would make it difficult to fairly compare facilities across jurisdictions.

Lack of available emissions data has become less of a barrier in the past few years due to the implementation of mandatory reporting rules for greenhouse gases at both the state and federal levels.[9] California's mandatory reporting rule[10] and the U.S. Environmental Protection Agency's (EPA) reporting program[11] are useful in finding annual greenhouse gas emissions for large industrial facilities. If reliable data on production and emissions (or energy consumption) are not available, officials should consider developing a reporting program before attempting to create a benchmarking program. Mandatory greenhouse gas reporting programs are in place in 18 states, including California, Wisconsin, and Oregon. Greenhouse gas reporting programs are otherwise outside the scope of this paper.

Close cooperation with industry is critical in gathering accurate data. For certain metrics, such as emissions data or energy consumption, cooperation can be compelled through mandatory reporting regulations. However, production data, including quantity and content of production, inputs, and possibly processes, are also necessary for most benchmarking metrics. These data may not be easily accessible, and likely will not be publicly releasable because such a release would reveal trade secrets and impair competitiveness. Thus it may be preferable for policymakers to take a more cooperative approach with industry to encourage facilities to share data voluntarily. Figure 3 shows the decisions that can make a policy more or less strict and the associated effects. Terms used in this figure are defined later in this Section.

Figure 3: Program impacts of making a benchmarking program more or less strict



Policymakers must consider whether to restrict data gathering to their own jurisdiction, or to also go beyond their borders. State-level programs should generally gather data from outside of the state as well.[12] Most states do not have enough facilities from every industrial sector such that a benchmark would be representative of actual efficiency potential. Gathering data from outside of the jurisdiction will enable the identification of new efficiency improvement opportunities that may not exist within the state.[13] Coordinating benchmarking programs across state lines will make this identification much easier. Additionally, many industrial firms operate, or at least market their products, in a variety of jurisdictions and would appreciate consistent information across their business territory.

The Stockholm Environment Institute outlined a number of useful recommendations for data gathering in its whitepaper, Issues and Options for Benchmarking Industrial GHG Emissions. These recommendations include:

- Common guidelines, tools, and methods should be used to measure each factor that goes into a benchmark to ensure consistency across facilities within each sector, and to the extent feasible, across sectors;
- The same protocol should be used in the construction of benchmarks and in monitoring production, emissions, and/or energy consumption after the program begins; and
- Data from more than one year should be used to account for cyclical changes in demand. As an example, California uses a three-year average.[14]

Use of Benchmarking

Key Question: In what policy contexts can benchmarking be used?

Recommendation:

- There are some commonalities in successful benchmarking methodologies across policy contexts (mandatory performance targets, voluntary performance targets, and market based allowance allocation), but some methodology decisions are context-dependent

Benchmarking can be used for three main purposes in a government program that is aiming to improve emissions or energy intensity. These uses are: 1) a mandatory program that requires each facility within a sector to achieve a certain level of performance; 2) a voluntary program that rewards facilities within a sector for achieving a certain level of performance; and 3) a market based program that allocates allowances to facilities based on the benchmark emissions intensity, usually that of a high-efficiency facility within the same sector. Each of these is described in more detail below. A fourth use, comparing an individual facility to itself over time, is only addressed to the extent that lessons can be drawn for the above three purposes.

The first two uses of benchmarking, the creation of mandatory targets or voluntary goals, are very similar in terms of how they are developed. Both can begin with the creation of a benchmark curve (see Figure 2) that plots the emissions or energy intensity of each facility in a defined sector to compare their relative performance. Next, a policy determination is made to choose a target performance level – for example: average, best ten percent, or ninety percent of average intensity. Alternatively, a mandatory or voluntary target can be set based on a level of performance deemed to be achievable by policymakers without surveying existing facilities, for example by setting a target based on the emissions performance of a specific technology. This is the approach generally followed when EPA sets New Source Performance Standards, which is described in the "Benchmarking in Practice" Section.

Facilities are then either required to meet the target (in a mandatory program) or rewarded for meeting the target (in a voluntary program). It should be noted that rewards in a voluntary program can take a variety of forms. Top-performing facilities or companies can simply be recognized through a certification, such as in EPA's Energy Start for Industry Program, or be excused from other regulations such as mandatory efficiency auditing.

The third use of benchmarking, for allocation in a market based system, also requires the creation of a benchmark curve that shows the relative performance of each facility in each defined sector. However, instead of determining a target performance level for each facility to achieve, in the market based context a policy determination is made to set how many allowances each facility will receive. Emission allowances are generally allocated to a facility based on a formula that includes an emissions-per-production benchmark multiplied by its production. For example, in the European Union's Emissions Trading System (EU-ETS), each facility in each defined sector is freely allocated allowances based on the emissions of a facility in the lowest tenth percentile of the emissions intensity distribution.

Since all three uses of benchmarking involve the definition and determination of an energy or emissions intensity metric, in theory all three could share a common methodology. In practice, however, the decisions discussed throughout this paper may have different resolutions depending on the policy context.[15] For example, carefully defining a sector or product is less critical in a voluntary context since a facility that does not easily fit into the defined category can simply choose not to participate and suffer fewer repercussions than if it were a mandatory or market based regulatory context. These types of considerations are addressed where they arise in each of the benchmarking issues discussed below.

Type of Benchmark

Key Question: What metric should be used in a benchmarking program?

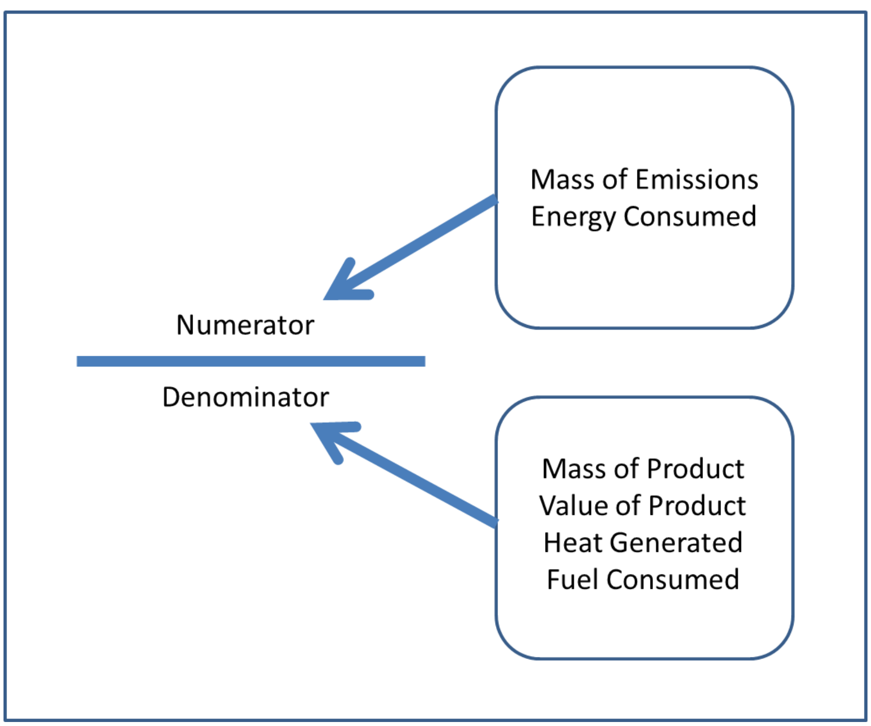
Recommendations:

- Product benchmarks should be used whenever possible
- Product should be measured by mass
- Heat benchmarks should be used when product benchmarking is not possible
- Fuel benchmarks should be used when neither product nor heat benchmarking is possible

Depending on the information being tracked, benchmarks for industrial efficiency can vary in both the numerator and denominator. The numerator is generally either a measure of energy consumption or greenhouse gas emissions, depending on the focus of the program. These two possibilities are largely interchangeable for the purposes of this paper, and the decision to measure one or the other generally has no bearing on the issues discussed. For the sake of simplicity, this paper will refer to emissions-based benchmarks where both emissions-based and consumption-based benchmarks could apply.

Benchmark types are typically designated by what is being measured in the denominator. The most common types of benchmarks are: 1) product; 2) heat; and 3) fuel. A product benchmark measures the amount of emissions per unit of product, which itself can be measured in either mass (e.g., tons of aluminum produced) or value (e.g., price of aluminum produced) terms, discussed further below. These options are summarized in Figure 4.

Figure 4: Possible units for an industrial benchmark



Where possible, a product benchmark should be employed. A product benchmark focuses on emissions released per unit of product. Compared to the other options, focusing on the product allows for the most efficiency improvement opportunities throughout the production process. That is, feedstock choice, facility location, fuel choice, combustion efficiency (where applicable), and process efficiency are all accounted for in a product benchmark. Using a product benchmark therefore encourages facilities to make improvements in all of these factors.

Amount of product can either be measured in mass or economic value. Measuring value can account for differences in output quality between facilities, and therefore properly credit a more emissions-intensive process that results in a higher quality product. However, physical measures of output are generally preferable since product prices and value added can change significantly over time.[16]

When product benchmarking is not possible, for example if a facility does not manufacture a product that has been defined or simply produces steam for use by another facility, the next best option is heat benchmarking. A heat benchmark measures emissions per unit of heat produced. It therefore is only applicable to processes where fuel is combusted to produce an intermediate heat carrier, such as steam, that can be measured and monitored. A heat benchmark encourages facilities to improve the efficiency of fuel combustion, but it does not encourage improvements in other factors, such as feedstock choice, as a product benchmark would.

When neither product nor heat benchmarking is possible, fuel benchmarking is the best option. A fuel benchmark tracks emissions per unit of fuel combusted. It can be used in processes that involve fuel combustion where heat benchmarking is not possible because no intermediate heat carrier is created. A fuel benchmark encourages a facility to combust fuel efficiently, but it does not encourage the maximization of heat production or manufacturing per unit of fuel, which is why it should only be employed when no other option is available.

Unit to Benchmark

Key Question: Should benchmarking focus on the efficiency of an entire facility, or on the process units within each facility?

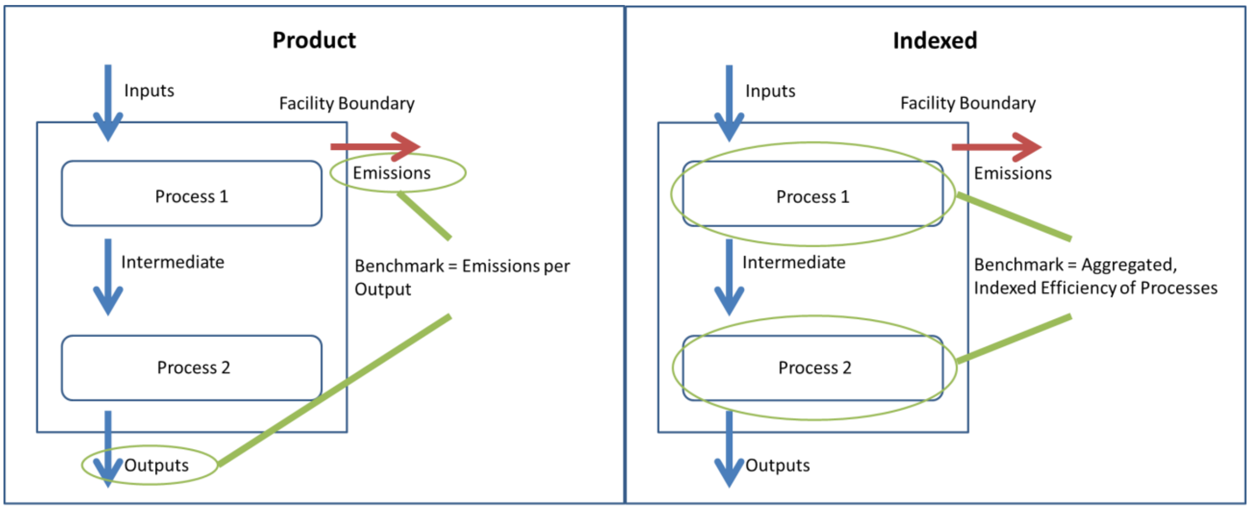
Recommendations:

- Benchmarks should be set at the facility level
- Product benchmarking should be favored over indexed benchmarking in the allowance allocation context
- Indexed benchmarking can also be successfully used in the performance target context, but product benchmarking in these contexts fosters consistency across jurisdictions and programs

Another key issue for officials when developing a benchmarking program is to decide the parameters of the unit subject to the benchmark. For example, a benchmark could apply to the final product of a facility, such as steel from a mill, an individual process within a facility, such as the conversion of iron ore to liquid iron within a mill, or to the combined products of each facility owned by the same company. Since benchmarking programs tend to focus on improving facility efficiency (as opposed to process unit or system efficiency) and since most industrial facilities house multiple processes, benchmarks are generally set at the facility level.

A benchmark can be set at a facility in two major ways: based simply on the amount of final product produced or on an indexed aggregation of the efficiency of each process within the facility. The first option results in an absolute metric, such as tons of greenhouse gas emissions per ton of steel produced. The second option generally results in a relative, unitless metric that compares the efficiency of similar facilities with each other. Figure 5 illustrates the difference between product and indexed benchmarking in a simplified example.

Figure 5: Comparing product and indexed benchmarking



Product benchmarks are generally favored because they account for all aspects of a facility rather than only the efficiency of process units. Indexed benchmarks encourage facility operators to improve the efficiency of process units, but do not address feedstock choice, fuel choice, or overall facility composition. For example, a product benchmark would encourage an operator to switch to a different feedstock that can take advantage of a more efficient type of process unit, whereas an indexed benchmark would not compare different process unit types against each other.

One major difficulty associated with product benchmarks is the potential complexity of accounting required when facilities produce multiple products or export intermediate products. In its simplest form, such as that shown in Figure 1, assigning emissions to a specific product is very straightforward. However, facilities may produce multiple products, in which case it must be determined what proportion of facility emissions to assign to each product. One way to do this would be to assign all of the emissions from any product-specific processes to that product, and divide the remaining emissions proportionally by the mass of each product produced. This approach is discussed further in the "Multi-Product Facilities" subsection.

Additionally, a facility might produce an intermediate product that is used as an input for a separate facility. In this case, it must be determined whether to assign the emissions associated with this intermediate product to the facility that produces it or the one that consumes it. Intermediate products are addressed more fully in the "Intermediate Products" subsection.

Although the difficulties in product benchmarking discussed above are not insurmountable, using an indexed benchmark avoids them altogether. By focusing on process efficiency, indexed benchmarking does not need to account for every product exported from the facility, nor is it relevant whether an intermediate product is used internally or shipped to another facility.[17] Additionally, indexed benchmarking allows for comparison among similar process units regardless of how they are employed.[18] For example, common industrial technologies such as boilers could be benchmarked across sectors. Thus every facility with a benchmarked process would be encouraged to improve the efficiency of that process regardless of what the facility produces. The downsides of such an approach are that facilities without such common processes would be overlooked, and operators of facilities where benchmarked processes do exist are only encouraged to make improvements in those specific processes.

The most common approach to an indexed benchmark is the process-step method. The process-step methodology focuses on individual process units within a facility.[19] Each process step in production is identified and a benchmark is determined for each, either in terms of energy consumed or emissions released per unit of mass produced. Each process within a facility is compared to identical processes in other facilities to determine the maximum achieved efficiency for that specific process. Each process unit is then assigned a score, indexed against the benchmark of the most efficient unit. It is also possible to use a theoretical state-of-the-art process unit as the benchmark rather than an actual unit. Finally, the scores of each process unit within a facility are aggregated to assign a total score to the facility. By focusing on processes rather than final products, the process-step methodology accounts for the heterogeneity of inputs and outputs at facilities that produce the same type of product.[20]

One application of the process-step approach is the Benchmarking and Energy Savings Tool (BEST), as described by Worrell and Price in their article, An Integrated Benchmarking and Energy Savings Tool for the Iron and Steel Industry. BEST calculates an Energy Efficiency Index (EEI) for each plant by determining the difference between actual energy intensity of each process unit and that of a hypothetical state-of-the-art unit and then aggregating units within an entire plant. At the plant level, the EEI identifies top performers and facilities that are lagging. BEST also identifies how an operator can optimize investments in efficiency by identifying which individual process units within a plant are most behind a state-of-the-art unit.

These two types of benchmarks, product and indexed, can be thought of as two ends of a spectrum, with some room to operate in the middle. For example, a program that employs a product benchmark could take on some of the flexibility of an indexed benchmark by including a credit-trading system so facilities where improvements can be made more easily can sell credits to facilities where external factors make improvements more difficult.

The issue of what unit to benchmark is closely related, but not identical, to the issue of how much to control for variables. An indexed benchmark automatically controls for a number of variables, such as fuel choice, by focusing on the efficiency of specific processes. However, an indexed benchmark can be further "corrected" for variables that affect process efficiency, such as, potentially, climate. A product benchmark can also be corrected by defining products more narrowly, for example by setting a separate benchmark for steel produced using coal or natural gas as fuel.

Definition of Product or Sector

Key Question: Which facilities should be included in a single benchmark?

Recommendations:

- The number of benchmarked products should be minimized; related products should only be given separate benchmarks if there is a significant difference in emissions intensity and if each product makes up a large share of total sector emissions
- Outliers should only be removed from the benchmarking curve by transparent consideration of a clear set of criteria that were established prior to the construction of the curve

Defining each product to benchmark is one of the most critical steps in establishing a benchmarking program. Determining if two closely related products should be defined as a single product or two distinct products with unique benchmarks in a consistent manner requires the establishment of a clear set of criteria. For allocation in the EU-ETS, the following factors were considered when determining whether related products should be benchmarked together or separately:

- The difference in emissions intensity between related products. Ecofys, a consulting company that was instrumental in establishing the benchmarking methodology for the EU-ETS, suggested a threshold difference in emissions intensity of 20 percent;

- The share of the emissions from a product group in the total emissions of the sector (a product that is responsible for a larger share of total sector emissions is more likely to be given its own benchmark). Ecofys suggested that benchmarks for a sector should cover 80 percent of its emissions;

- The share of the emissions from a product group in the overall jurisdiction (a product that is responsible for a larger share of emissions in the jurisdiction is more likely to be given its own benchmark); and

- The number of installations producing a product (a product that is produced by a larger number of facilities is more likely to be given its own benchmark).

When defining products to benchmark, one possible example to follow would be that of California, as shown in Table 1. Equivalent benchmarks from the EU-ETS are shown, where available. The full list of EU-ETS benchmarks is much longer and available here.

Table 1: Benchmarked Products in California and the EU-ETS analogs

SECTOR	PRODUCT	**CALIFORNIA BENCHMARK*** (ALLOWANCE PER METRIC TON OF PRODUCT UNLESS OTHERWISE NOTED)	**EU-ETS BENCHMARK*** (ALLOWANCE PER METRIC TON OF PRODUCT UNLESS OTHERWISE NOTED)
Crude Petroleum and Natural Gas Extraction	Heavy Crude Oil Extraction	0.0654 allowance / barrel of heavy crude oil equivalent	N/A
	Light Crude Oil Extraction	0.0100 allowance / barrel of light crude oil equivalent	N/A
Natural Gas Liquid Extraction	Natural Gas Liquid Processing	0.0145 allowance / barrel of natural gas liquids produced	N/A
Potash, Soda and Borate Mineral Mining	Mining and Manufacturing of Soda Ash and Related Products	1.045	0.843
Pulp and Paper	Through-Air-Dried Tissue Manufacturing	1.43	N/A
	Recycled Boxboard Manufacturing	0.550	0.273
	Recycled Linerboard Manufacturing	0.516	0.248
	Recycled Medium Manufacturing	0.434	0.248
*Petroleum Refineries***	Petroleum Refining	0.0465 allowance / barrel of primary refinery products	0.0295 allowance / carbon weighted metric ton of petroleum product
All Other Petroleum and Coal Products Manufacturing	Coke Calcining	0.376	0.376

Industrial Gas Manufacturing	Gaseous Hydrogen Production	8.62	8.85
	Liquefied Hydrogen Production	TBD	N/A
Nitrogenous Fertilizer Manufacturing	Nitric Acid Production	0.385	0.302
	Calcium Ammonium Nitrate Solution Production	0.099	N/A
Glass Manufacturing	Flat Glass Manufacturing	0.519	0.453
	Container Glass Manufacturing	0.291	0.275
Mineral Wool Manufacturing	Fiber Glass Manufacturing	0.434	0.682
Cement Manufacturing	Cement Manufacturing	0.786	0.766 allowance / metric ton of clinker
Lime Manufacturing	Dolime Manufacturing	1.54	1.072
Gypsum Product Manufacturing	Plaster Manufacturing	0.050	0.048
	Plaster Board Manufacturing	0.147	0.131
Iron and Steel	Steel Production using an Electric Arc Furnace	0.199	0.283 (includes emissions from purchased electricity)
	Hot Rolled Steel Sheet Production	0.0929	N/A
	Pickled Steel Sheet Production	0.0139	N/A
	Cold Rolled and Annealed Steel Sheet Production	0.0345	N/A
	Galvanized Steel Sheet Production	0.0556	N/A
	Tin Steel Plate Production	0.0217	N/A

Turbine and Turbine Generator Set Units Manufacturing	Testing of Turbines and Turbine Generator Sets	0.00782 allowance / horsepower tested	N/A

***Products across jurisdictions are defined slightly differently and are thus not directly comparable. This information is included for illustrative purposes only.**

****California will begin using a Complexity Weighted Barrel (CWB) methodology for petroleum refineries starting in 2015. This methodology is very similar to the Complexity Weighted Tonne approach in the EU-ETS.**

Once the products or industries to benchmark are clearly defined, another issue to consider is the treatment of outliers. Two decisions must be made regarding outliers when constructing a benchmarking program: 1) Whether to include outlying facilities in the benchmark curve; and 2) Whether to hold outlying facilities to the established benchmark.

In the development of benchmarking curves in the EU-ETS, no facilities were excluded in the early stages of data collection and calculations. However, in exceptional cases, outlying facilities were removed from the curve, and therefore did not factor into the calculation of the benchmark, for specific and transparent reasons. For allowance allocation and voluntary programs, outlying facilities should be subject to the benchmark because they can purchase additional allowances from other facilities or at auction in the allocation context, or choose to not participate in the voluntary context. In the context of a mandatory performance standard, policymakers can choose to exclude outlying facilities based on criteria set prior to the development of the benchmarking curve.

Outliers are not likely to pose a significant problem in voluntary programs since operators that do not believe the established benchmark is realistic can choose not to participate. Programs that require participating plants to achieve a certain percentage of emissions intensity reduction over time, without comparing facilities to each other, do not have to address the outlier issue.

Finally, policymakers must consider how to weigh each facility in a benchmarking curve. In the most straightforward approach, weight can be assigned on a per-facility basis, such that the emissions intensity of each is weighted equally when determining the industry average (or any other chosen target). Weight can also be assigned based on emissions or production levels. Since benchmarking is intended to look only at emissions intensity and treat large and small facilities equally, weighting should be made on a per-facility basis. This is the approach taken by both the EU-ETS and California.

Controlling for Variables

Key Question: What factors, both within and outside of a facility operator's control, should be accounted for when determining a benchmark?

Recommendations:

- The "one product, one benchmark" approach should be followed to promote systematic efficiency
- Variables such as production technology, fuel choice, feedstock, climate, resource availability, or capacity should not be corrected for when a benchmark is set

Another critical consideration for policymakers is whether, and to what extent, to control for factors other than process efficiency. The "one product, one benchmark" approach, as used in the market based cap-and-trade systems of both California and the EU-ETS, treats all production of a defined product equally. The alternative, which this paper refers to as a "corrected benchmark," controls for factors such as feedstock, capital equipment, and fuel to isolate the efficiency of industrial processes.

Under the "one product, one benchmark" approach, no adjustment is made for variables such as production technology, fuel choice, feedstock, climate, resource availability, capacity, or any other facility-specific factors.[21] For example, this type of approach, as used in California's cap-and-trade program, would have a single benchmark for cement production. Each cement plant receives emission allowances according to the same benchmark, regardless of the type of fuel the plant uses or the composition of the final product. Facilities operating within such a system are therefore encouraged to take every possible step to reduce emissions, including changing fuel, equipment, feedstocks, and outputs, in addition to process efficiency.[22]

This "one product, one benchmark" method is favored for allowance allocation in cap-and-trade programs because it fosters a systematic approach to emission reductions, meaning changes to production methods and fuels are encouraged and rewarded. This method is simpler, more consistent, and more transparent because correction factors, which can be subject to political influence, are not required.

In contrast, a "corrected benchmark" approach isolates the efficiency of individual industrial processes within a facility and excludes other factors from the benchmark. This can be done by only benchmarking similar facilities against each other, for example by setting a separate benchmark for natural-gas fueled Portland cement plants rather than setting a single benchmark for all cement production. This approach encourages facilities to improve process efficiency, but provides less incentive for the systematic restructuring of industrial sectors, such as focusing production near feedstocks or using fuels that result in fewer carbon emissions.[23] Another common method of isolating process efficiency is to use the process-step approach to create benchmarks, which is described in the "Unit to Benchmark" subsection. Figure 7 summarizes how these two sets of options interact.

Figure 7: Interaction between decisions of which unit to benchmark and whether to control for variables

		UNIT TO BENCHMARK	
		Product	Indexed
CONTROLLING FOR VARIABLES	Uncorrected	One product, one benchmark (total emissions over total product)	Aggregated efficiency of process units with a facility, relative to a standardized unit
	Corrected	Total emissions over total product, corrected for certain variables (e.g., climate)	Aggregated efficiency of process units within a facility, relative to a standardized unit, corrected for certain variables (e.g., climate)

There are benefits and drawbacks to each choice, and the purpose for benchmarking must be considered when deciding between these two approaches. The "one product, one benchmark" tends to be favored for the purpose of allocating emission allowances due to its simplicity. This approach also reflects the purpose of benchmarking in a market based cap-and-trade program to result in the fair distribution of allowances, which should solely be based on emissions intensity.

On the other hand, a "corrected benchmark" may be more appropriate for voluntary or mandatory programs that aim to improve process efficiency without necessarily addressing systematic efficiency. The "corrected benchmark" results in more realistic goals because it does not require a facility served with high-carbon grid electricity, for example, to meet a benchmark set to a facility served by renewable grid electricity. This approach is therefore more economically sensitive and would be less likely to lead to leakage or plant closures. This method is also more likely to foster a cooperative relationship between industry and policymakers, which can lead to the increased availability of industry data. However, use of "one product, one benchmark" is encouraged whenever possible to build consistency across programs and jurisdictions.

As with the differentiation of product benchmarks and indexed benchmarks, uncorrected and corrected benchmarks can be thought of as two ends of a spectrum. One methodology that falls in the middle of these extremes is the complexity-weighted ton approach used in the EU-ETS, as well as California in its second phase, to benchmark petroleum refineries.

Petroleum refining does not lend itself to basic product benchmarking due to the heterogeneity of refineries and their products. Refineries produce a variety of petroleum products from crude oil of varying sulfur content and density. Thus each refinery must be configured differently to account for a specific set of inputs and outputs, yet common process units that can be benchmarked against each other are found in many facilities. To account for this, a benchmark based on a "complexity-weighted tonne" (CWT) was developed by CONCAWE and Solomon Associates for the EU-ETS that functions as a product benchmark, but corrects for the necessary heterogeneity among refineries. It is based on a proprietary methodology developed by Solomon Associates that was already in use internally by the refining industry.

Complexity here does not refer to the configuration of a refinery, but to the amount of work necessary to convert the input to a process to the output. For example, even under ideal conditions a fluid coker results in over seven times as much greenhouse gas emissions when converting its inputs to outputs as compared to an atmospheric distillation unit. Thus a simple emissions-per-ton-of-product benchmark for a refinery would reflect the complexity of the process units within each refinery, rather than the efficiency of each unit.

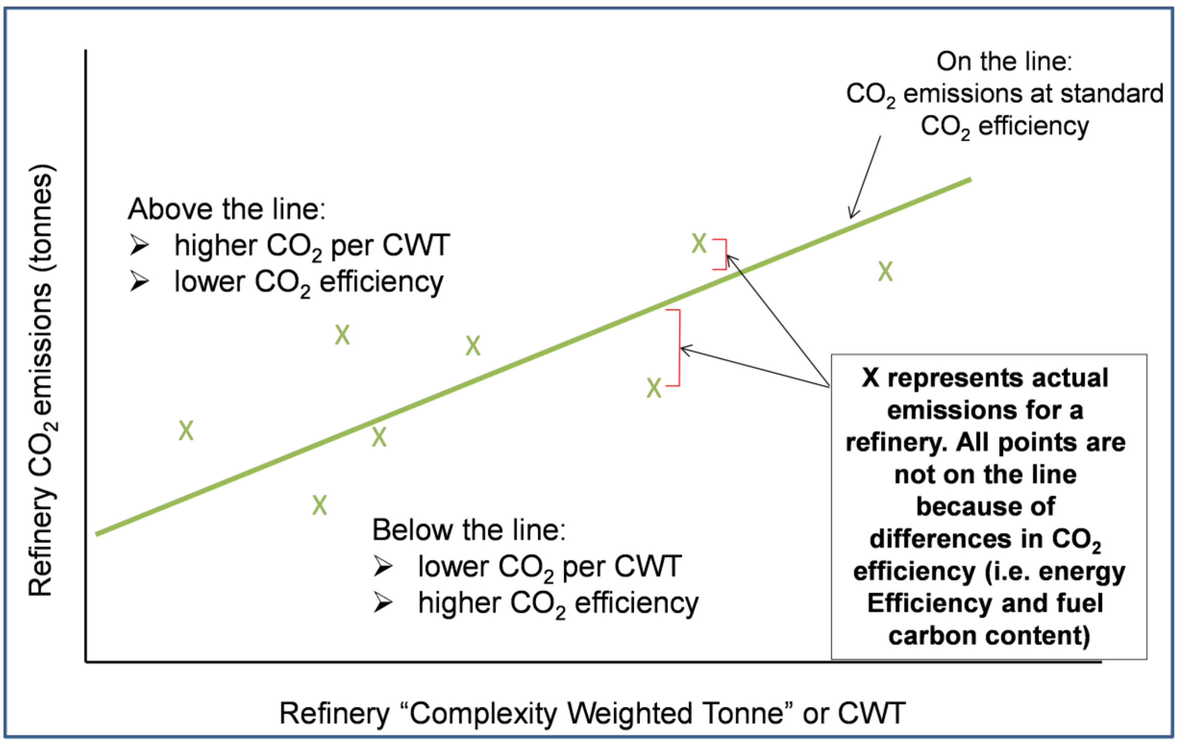
The basic concept is that the CWT is a product benchmark, but the product is weighted by the complexity of the processes within the facility for which the benchmark is being determined. A full explanation of this methodology is included in Box 1. By addressing each process unit separately, the CWT method has elements of an indexed benchmark. However, the CWT tracks the total emissions from a refinery rather than just those associated with process units, and therefore encourages a broader set of efficiency improvements than would a traditional indexed benchmark.

Box 1: Complexity Weighted Tonne (CWT) Methodology

1. Define a list of generic process units, applicable across refineries (51 such process units are defined for the EU-ETS);
2. Allocate each process unit a CWT factor indicating its propensity to emit greenhouse gases at standardized conditions of energy efficiency and a standardized fuel emission factor;
3. Multiply the throughput of each process unit at a refinery by its CWT factor;
4. Total the results from Step 3;
5. Add an "off-sites" allowance for energy consumption at a refinery that is not linked to a specific process unit;
6. The resulting CWT is the total "product" of the refinery – the denominator in a benchmark;
7. The numerator of the benchmark is total refinery emissions.

The CWT allows for benchmarking across two axes – emissions and complexity. Refineries only fall short of the benchmark if the difference in emissions is disproportionately greater than complexity relative to the benchmark, as shown in Figure 6.

Figure 6: Illustration of CWT methodology. The distance between each marker and the central line can be plotted in order to create a traditional benchmarking curve.



Source: Alan Reid, The CONCAWE CWT methodology for allocation of CO2 allowances in refining. Presentation at Industrial Benchmarking Workshop (September 24, 2012, New York City).

Intermediate Products

Key Question: How should a benchmarking program deal with intermediate products that are created and consumed within a single facility in some cases?

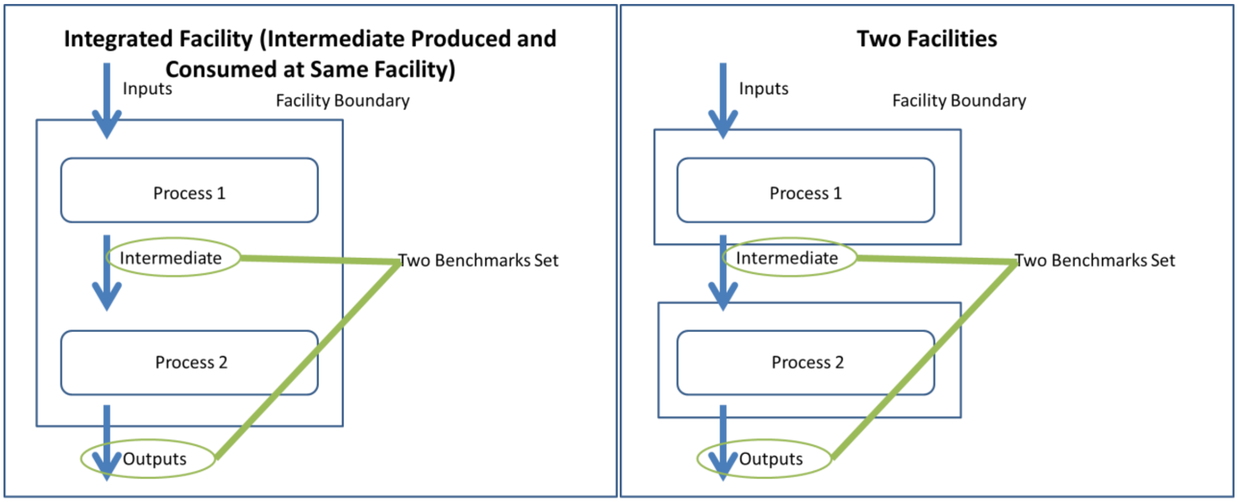
Recommendation:

- Any intermediate product that is ever traded between facilities should be assigned its own benchmark, which would apply in every case of that intermediate product's production

Intermediate products are products that are created through an industrial process but require additional processing before having value to the consumer market. For example, pulp is an intermediate product formed as paper is manufactured from wood. As noted above, intermediate products are not a concern for the development of indexed benchmarks since this approach focuses solely on individual processes. To continue the example above, paper mills (that produce paper from pulp) will only be compared to other paper mills, whether in a stand-along nonintegrated mill or as part of an integrated mill that also features a pulp mill (that produces pulp from wood). Thus whether a mill exports its pulp to another facility or uses it internally does not affect the indexed benchmarking methodology.

Intermediate products can, however, add complications to product benchmarking. There are cases where one facility exports an intermediate product while another facility produces and consumes the same intermediate product in an integrated process, such as in the paper mill example above. Thus, if only exported products were benchmarked, accounting inconsistencies could arise because integrated facilities would not be credited for the creation of intermediate products. One way around this problem would be to assign, in the non-integrated case, emissions from the intermediate product manufacturing to the facility that consumes the intermediate product to produce a final product. However, this would make the second facility responsible for the emissions of the first facility even though it would not have any control over the first facility's operations. Instead, a separate benchmark should be developed for any intermediate product that is ever traded between installations. This benchmark would be used regardless of whether the intermediate product is used at the same facility or in another facility. See Figure 8 for an illustration of this concept. Therefore, an integrated facility that produces and consumes an intermediate product will effectively be treated like a multi-product facility, as discussed below.

Figure 8: Benchmarks should be set for any intermediate that is traded between facilities, even in cases where the intermediate is produced and consumed by an integrated facility



Multi-Product Facilities

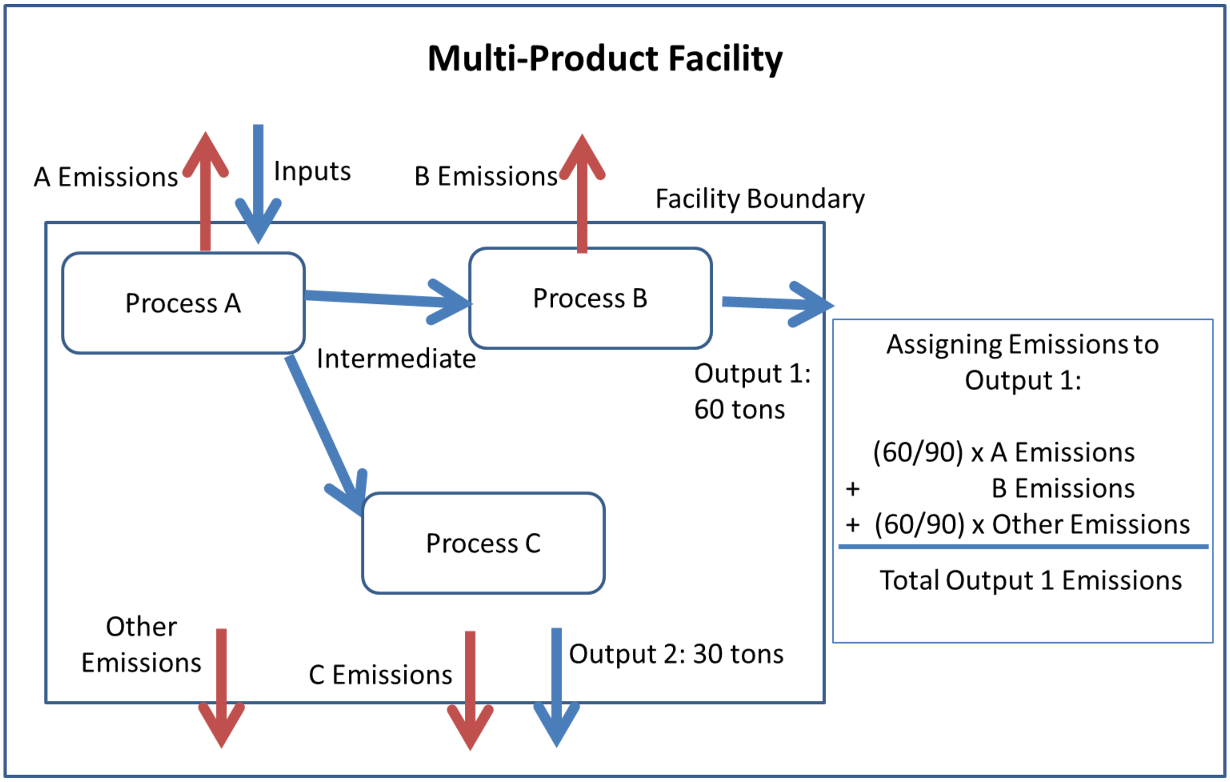
Key Question: How should a benchmarking program assess facilities that produce multiple products?

Recommendation:

- For each product in a multi-product facility, emissions should be quantified by assigning all product-specific process emissions to that product, plus a share of common emissions in proportion with the facility's production of that product (see Figure 9)

Policymakers must determine how to benchmark multiple products that are produced by the same facility. This issue is similar to that of intermediate products. As with intermediate products, multi-product facilities are not a concern when indexed benchmarking is used. Since indexed benchmarking focuses on the efficiency of individual process units, the ultimate mix of final products is not factored into the metric. When a product benchmark is being used, policymakers must decide how to allocate facility emissions to each product. The most common method, as illustrated in Figure 9, is first to separate the processes that are unique to a single product and those that are common to multiple products, including facility emissions not associated with a process. Each product should be assigned all emissions for each process that is only associated with that product, plus a proportion of common process emissions that matches the proportion of production, in terms of mass, of that product at the facility.

Figure 9: Assigning emissions to an individual product in a multi-product facility



Indirect Emissions and Combined Heat and Power (CHP)

Key Question: How should benchmarking programs account for CHP units that increase emissions on-site (direct emissions) while decreasing emissions off-site (indirect emissions)?

Recommendations:

- In the performance target context, indirect emissions should be factored into industrial benchmarks to avoid discouraging investments in CHP systems
- In the allowance allocation context where electricity generation is benchmarked separately, CHP units can be benchmarked separately from the remainder of the facility
- CHP benchmarking must account for heat production and electricity production separately
- In the EU-ETS, a share of emissions is assigned to heat production according to a reference efficiency based on stand-alone steam production from boilers, and the balance of CHP emissions are assigned to electricity production

In mandatory or voluntary programs focused on improving industrial efficiency, the decision of whether to include indirect emissions is relatively straightforward. When only direct emissions are factored into a benchmark, there will not be an incentive for a facility to reduce its electricity consumption because the emissions associated with electricity generation are not accounted for in the benchmark. Worse, facilities may be discouraged from investing in combined heat and power (CHP) systems that increase direct emissions while decreasing indirect emissions even if the result is lower emissions overall. Thus, in programs that are focused on industrial efficiency, indirect emissions should be included in the metric for each facility. As discussed below, this is not necessarily the case in economy-wide programs that cover electricity emissions separately.

When including indirect emissions, policymakers must determine what emissions factor to use for electricity. This factor could be a uniform factor based on the average or marginal emissions factor in a jurisdiction, or it could be an estimate of the actual emissions factor of the electricity used on a per-facility level. The latter approach would allow facility operators to have more precise information when deciding whether to use grid electricity or install cogeneration (also known as CHP). However, the extensive data requirements of this approach are likely prohibitory. Additionally, industrial facilities generally do not have any control over their electricity source, so calculating an indirect emissions factor on a per-facility basis would have very limited benefit. In sum, electricity emissions should be assigned based on an average emissions factor for each jurisdiction.

The issue of indirect emissions can be more complex in an economy-wide program, such as a cap-and-trade program that uses benchmarking for allowance allocation. In these programs, power plants are generally allocated allowances under a different set of rules from industrial facilities. Regardless, power plants must hold enough allowances to account for all of their emissions, so if an industrial facility were also required to hold allowances for electricity-based emissions there would be a double-counting problem. Thus, in the cap-and-trade allocation context, indirect emissions cannot simply be assigned to industrial facilities to avoid the problem of discouraging CHP and another solution is required.

One option to avoid discouraging CHP in a program where indirect emissions are not assigned to industrial facilities is to benchmark CHP units separately from the remainder of the facility.[24] A CHP unit cannot be benchmarked with the same methodology as a traditional power plant because it generates useful heat in addition to electricity. A CHP unit would have to be benchmarked for both useful heat and electricity and receive allocations for both. Thus, an industrial facility would be allocated allowances based on the use of at least three total benchmarks: one for the electricity produced by the CHP unit (using the same methodology used for power plants, not addressed in this paper), one for the useful heat produced by the CHP unit (using a heat benchmark), and at least one for the final product(s) made at the facility (using a product benchmark). It is critical that the emissions from the CHP unit only be counted once for the combined heat and electricity generation of the CHP unit. Where allowances are allocated in different ways for heat and electricity, as in both the EU-ETS and California cap-and-trade program, it is critical to accurately assign shares of emissions to electricity and heat.

The EU-ETS provides an example of how to account for CHP emissions in an economy-wide cap-and-trade program. In this program, CHP units are benchmarked separately from the rest of the facility. CHP emissions are assigned to either the heat production or electricity production aspect of the unit based on a reference efficiency of stand-alone steam production in boilers.[25]

The first step in this calculation is to determine how much fuel would be needed to produce the amount of heat produced by the CHP unit under ideal conditions. This may be easier to understand as a sequence of equations:

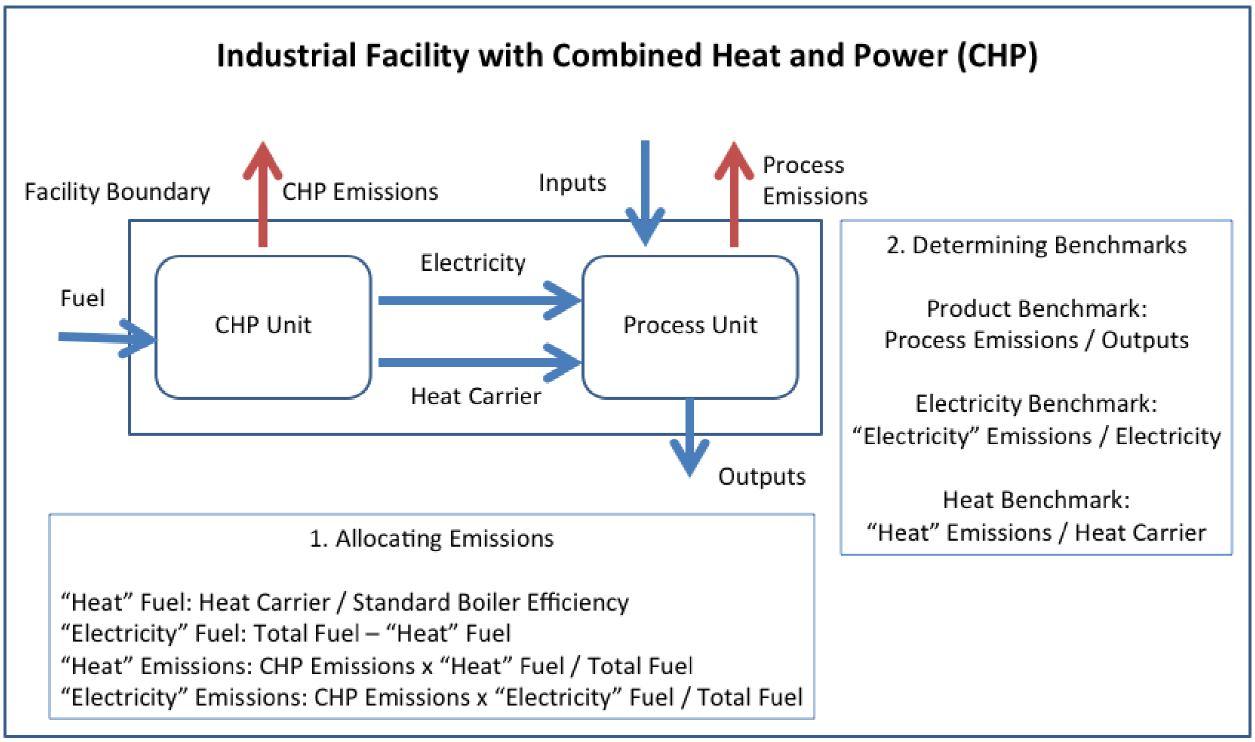
1) Fuel Consumed x Reference Efficiency = Heat Produced

Since the amount of heat produced by the CHP unit can be measured, this equation is rearranged to:

2) Fuel Consumed = Heat Produced / Reference Efficiency

Thus, by dividing the amount of heat produced by the reference efficiency for a boiler, a certain amount of fuel consumed by the CHP unit is assigned to heat production. The emissions associated with this share of fuel consumption are then used to develop the heat benchmark for the CHP unit. The balance of the fuel consumed by the CHP unit is assigned to electricity production, and the emissions associated with this share of fuel consumption are then used to develop the electricity benchmark for the CHP unit. This is illustrated in Figure 10.

Figure 10: Assigning total CHP emissions to either heat production or electricity production



It is also possible to properly encourage CHP in a cap-and-trade program simply based on changes to the price of electricity caused by the carbon price. That is, if the carbon price is accurately reflected in the price of electricity, industrial facilities will be incentivized to install CHP systems even though they will need to hold additional carbon allowances because the reduced electricity payments, due to the efficiency of CHP systems, will outweigh the increased cost of emissions allowances. This is addressed more fully in Stockholm Environment Institute’s White Paper: Issues and Options for Benchmarking Industrial GHG Emissions.

Summary of Working Group Recommendations

General Policy Considerations

- There are several critical preconditions to a successful benchmarking program:
- Common definitions

- Reliable data
- Reliable measurement and verification systems
- Committed contributions
- Targets set based on benchmarks should be SMART:
- Specific
- Measurable
- Appropriate
- Realistic
- Timed

Data Gathering

- Policymakers should consider mandatory reporting regulations prior to the development of a benchmarking program if reliable, consistent data are not already available
- Benchmarking data should come from operating facilities rather than the theoretical literature whenever possible
- Data used for benchmarking should be transparent whenever possible, but trade secret laws may prevent this in some cases
- Policymakers should consider what balance to strike between fostering industry cooperation and designing a strict program

Use of Benchmarking

- There are some commonalities in successful benchmarking methodologies across policy contexts (mandatory performance targets, voluntary performance targets, and market based allowance allocation), but some methodology decisions are context-dependent

Type of Benchmark

- Product benchmarks should be used whenever possible
- Product should be measured by mass
- Heat benchmarks[26] should be used when product benchmarking is not possible
- Fuel benchmarks[27] should be used when neither product nor heat benchmarking is possible

Unit to Benchmark

- Benchmarks should be set at the facility level
- Product benchmarking should be favored over indexed benchmarking in the allowance allocation context
- Indexed benchmarking can also be successfully used in the performance target context, but product benchmarking in these contexts fosters consistency across jurisdictions and programs

Definition of Product or Sector

- The number of benchmarked products should be minimized; related products should only be given separate benchmarks if there is a significant difference in emissions intensity and if each product makes up a large share of total sector emissions

- Outliers should only be removed from the benchmarking curve by transparent consideration of a clear set of criteria that were established prior to the construction of the curve

Controlling for Variables

- The "one product, one benchmark" approach should be followed to promote systematic efficiency
- Variables such as production technology, fuel choice, feedstock, climate, resource availability, or capacity should not be corrected for when a benchmark is set

Intermediate Products

- Any intermediate product that is ever traded between facilities should be assigned its own benchmark, which would apply in every case of that intermediate product's production

Multi-Product Facilities

- For each product in a multi-product facility, emissions should be quantified by assigning all product-specific process emissions to that product, plus a share of common emissions in proportion with the facility's production of that product

Indirect Emissions and Combined Heat and Power (CHP)

- In the performance target context, indirect emissions should be factored into industrial benchmarks to avoid discouraging investments in CHP systems
- In the allowance allocation context where electricity generation is benchmarked separately, CHP units can be benchmarked separately from the remainder of the facility
- CHP benchmarking must account for heat production and electricity production separately
- In the EU-ETS, a share of emissions is assigned to heat production according to a reference efficiency based on stand-alone steam production from boilers, and the balance of CHP emissions are assigned to electricity production

Benchmarking in Practice

This section explores a handful of benchmarking programs, in the allowance allocation, mandatory performance target, and voluntary performance target contexts, to further elucidate the key concepts discussed above. To ease navigation, each section below begins with a list of key characteristics for the program being discussed. These key characteristics are collected and repeated in Table 2.

A brief summary is included for each program to add context, followed by a discussion of any issues that should be especially helpful for policymakers setting out to tackle the same issues for a new program.

Allowance Allocation

Due to the mandatory nature of cap-and-trade programs and the high associated costs to industry if benchmarks are developed improperly, benchmarking in the allowance allocation context has received more academic and practical attention than in other contexts. Thus, the benchmarking methodologies developed for the EU-ETS and California should be especially useful in drawing useful lessons for new benchmarking programs.

European Union Emissions Trading System (EU-ETS)

Program Summary

Program Type: Allowance allocation

Benchmarked Unit

1. Greenhouse gas emissions / Metric ton of product
2. Greenhouse gas emissions / Heat produced
3. Greenhouse gas emissions / Fuel consumed

Comparisons Enabled: Facilities that produce the same product can be compared to each other

In 2005, the European Union (EU) implemented a greenhouse gas cap-and-trade program that included the industrial sector, the Emissions Trading System (EU-ETS). For the first two phases of the EU-ETS, running through 2012, free allocation to industrial facilities was primarily based on historical emissions. However, starting in 2013, allocation to industrial facilities is based on allocation rules that are harmonized among the EU member states. These rules are based on benchmarking wherever possible. Although this paper is generally constricted to domestic examples, it is impossible to ignore the contribution of the EU-ETS in industrial benchmarking due to its comprehensive methodology.

As a leader in greenhouse gas cap and trade, the EU needed to develop a benchmarking methodology to determine allowance allocation without many examples to follow. Ecofys and other consultants worked with industry for several years to develop rules that would equitably distribute emission allowances while taking into account the many complicating factors discussed throughout Section 2. Ecofys developed the principles listed in Box 2 in the development of this methodology, many of which are reflected in the benchmarking policy recommendations made in the "Issues and Recommendations" Section. Most significantly, the EU-ETS follows the "one product, one benchmark" principle that, in most cases, holds all related products to a single benchmark, regardless of additional variables such as resource availability or process type. This paper does not delve into the details of these principles, but a complete explanation is available here.

Box 2: Principles formulated by Ecofys in development of EU-ETS benchmarking methodology

- Base the benchmark level on the most energy efficient technology;
- Do not use technology-specific benchmarks for technologies producing the same product;
- Do not differentiate between existing and new plants;
- Do not apply corrections for plant age, plant size, raw material quality and climatic circumstances;
- Only use separate benchmarks for different products if verifiable production data is available based on unambiguous and justifiable product classifications;
- Use separate benchmarks for intermediate products if these products are traded between installations;
- Do not use fuel-specific benchmarks for individual installations or for installations in specific countries;
- Take technology-specific fuel choices into account in determining benchmarks;
- Use historical production to allocate allowances for existing installations;
- Use product-specific capacity utilization rates in combination with verifiable capacity data to allocate allowances to new installations;
- Use heat production benchmark combined with a generic efficiency improvement factor for heat consumption in processes where no output-based benchmark is developed.

Source: Ecofys, Developing Benchmarking Criteria for CO_2 Emissions. Prepared for the European Commission. (February 2009).

California Cap-and-Trade

Program Summary

Program Type: Allowance allocation

Benchmarked Unit

1. Greenhouse gas emissions / Metric ton of product
2. Greenhouse gas emissions / Fuel consumed

Comparisons Enabled: Facilities that produce the same product can be compared to each other

Like the EU, California maintains a market-based greenhouse gas cap-and-trade program that encompasses industrial sources. A large percentage of emission allowances in the industrial sector are freely allocated, especially in the early years of the program. California's program, which began operation at the beginning of 2013, follows the example of the current phase of the EU-ETS and primarily uses benchmarking for allocation. Additional information on California's cap-and-trade program can be found here.

In setting up its benchmarking methodology, California largely adopted the rules established by the EU-ETS. In terms of the issues discussed throughout the "Issues and Recommendations" Section, the principles followed by California and the EU-ETS are nearly identical. One significant difference is in the type of benchmark to use if a product benchmark is not feasible. In the EU-ETS, a heat benchmark (emissions per energy content of heat-carrying intermediary (such as steam) produced) is the preferred choice, followed by a fuel benchmark (emissions per energy content of fuel combusted). In California, there is only a single fallback to a product benchmark, called an energy-use benchmark. This metric tracks the amount of emissions released per energy content of energy consumed, making it very similar to the fuel benchmark used in the EU-ETS. The California approach is less complex, but in not incorporating a heat benchmark, does not have the capacity to account for the efficiency of facilities that create heat carriers. For example, if a facility creates only steam it would be benchmarked in the EU based on the amount of energy it produces, but in California it would be benchmarked based on the amount of fuel it consumes. Thus the facility in the EU would have an incentive to produce steam more efficiently, whereas the facility in California would not because the steam produced there is not part of the benchmarking equation.

There are also some industry-specific variations to account for differences between the two markets. For example, the cement industry benchmark in California is based on cement as the final product, while a product near the end of the cement production cycle called "clinker" is the benchmarked product in the EU-ETS. Ultimately, the similarities between benchmarking in the two programs vastly outweigh the differences. The many principles they have in common could serve as the foundation for new benchmarking programs.

Mandatory Performance Targets

Mandatory performance targets represent the most straightforward use of benchmarking. An emissions intensity target is set for a category of sources, which are all then legally required to meet that target. The Clean Air Act authorizes a variety of these intensity targets, which are set according to different methodologies, including the creation of a typical benchmarking curve as shown in Figure 2. These targets include New Source Performance Standards for both new sources and existing sources, as well as mercury and air toxics standards for power plants, each of which is summarized below.

New Source Performance Standards

Program Summary

Program Type: Mandatory performance standard

Benchmarked Unit: Varied; includes emissions per mass unit of product

Comparisons Enabled: Facilities within a regulated sector can be compared to each other

One major category of mandatory emission performance standards in the United States are New Source Performance Standards (NSPS), as defined in Section 111 of the Clean Air Act. New, modified, and reconstructed facilities that fall within a category of sources for which an NSPS has been defined must not exceed the emissions threshold set by EPA for that source category. A variety of air pollutants can be covered by an NSPS, but greenhouse gases are not yet included in this program. Since NSPS cover a variety of pollutants from a variety of sources, they are set in a variety of formats. These include intensity targets similar to a product benchmark, such as the standard for sulfuric acid production units: 0.25 grams of sulfuric acid mist per kilogram of sulfuric acid produced.[28] They also include volumetric standards, such as the standard for nitrogen oxides from hospital waste incinerators of 250 parts per million,[29] fuel benchmarks, and operational standards, such as inspection requirements.

Although the NSPS for a specific sector functions as a benchmark that all new, modified, or reconstructed facilities within a covered sector must meet, NSPS are not set based on benchmarking as described elsewhere in this paper. EPA sets a performance standard as the:

> degree of emission limitation achievable through the application of the best system of emission reduction which (taking into account the cost of achieving such reduction and any nonair quality health and environmental impact and energy requirements) the [EPA] Administrator determines has been adequately demonstrated.[30]

Thus, the emissions standard simply must be achievable by a demonstrated technology, known as the Best Demonstrated Technology (BDT) regardless of the performance of existing facilities. To determine the BDT, EPA typically conducts a technology review to identify the available emission reduction systems and how effective they are at reducing air pollution.[31] By determining the level of emission reductions associated with each system of reduction technology, EPA identifies its options for potential emission limits. EPA then evaluates each of these potential limits in conjunction with financial costs, systematic impacts on air pollution (for example caused by additional energy requirements), and other environmental impacts such as solid waste generation. EPA then sets a numeric emissions limit, based on the effectiveness of the determined BDT.

In addition to New Source Performance Standards (NSPS) for new sources, Section 111 of the Clean Air Act requires EPA to set NSPS for existing sources in some circumstances. To date, NSPS have been set for municipal landfills, sulfuric acid producers, and various waste combustors.[32] In general, EPA has relatively little experience setting NSPS for existing sources. Unlike NSPS for new sources, NSPS for existing sources are set at the state level, but must follow emission guidelines set by EPA. EPA sets these emission guidelines for existing sources based on BDT, which is similar to how targets are set for new sources. However, the Clean Air Act allows more flexibility in the regulation of existing sources. States and EPA are allowed to set a less stringent standard or longer compliance period where warranted, considering the financial costs of the emission control technology, the useful life of the facilities, the feasibility of installing the required control equipment, or other factors making less stringent limits or longer compliance schedules appropriate.

There is no need for a benchmarking curve in this process, though other issues discussed in this paper must be addressed. For example, EPA must determine how to define the sectors for which a standard is set, and the unit of the performance standard. Although the standard-setting methodology generally differs from that of a more typical benchmarking program, EPA's NSPS program offers an example of how intensity-based emissions targets can be imposed, monitored, and enforced.

Mercury and Air Toxics Standards for Power Plants

Program Summary
Program Type: Mandatory performance standard
Benchmarked Unit: Varied; generally air pollutants per MMBTU of fuel consumed
Comparisons Enabled: Power plants can be compared to each other

The Mercury and Air Toxics Standards for Power Plants (MATS) are the first-ever national emission standards for power plant air toxics. These standards apply to both new and existing coal and oil-fired power plants. In addition to mercury, the standards cover hazardous air pollutants such as metals, acid gases, and organic air toxics.[33]

Section 112 of the Clean Air Act requires EPA to set intensity-based emissions limits based on the Maximum Achievable Control Technology (MACT). Depending on the specific pollutant, these emission rates are expressed either in terms of product benchmarks (pounds of pollutant released per megawatt-hour (or gigawatt-hour) of electricity produced) or fuel benchmarks (pounds of pollutant released per MMBTU of fuel consumed). Unlike the NSPS, EPA must use benchmarking to set the emission rate standards because the Clean Air Act requires the standards for existing sources to be at least as stringent as average of the best 12 percent of existing units. EPA is authorized to set the emission rate limits for subcategories, as well as separate rates for new and existing unites. In demonstrating compliance, operators are allowed to average different units that are in the same subcategory at the same facility. Certain pollutants are regulated by required work practice standards rather than emission rate limits. EPA has set different standards for new[34] and existing[35] sources.

To set the MACT, EPA begins by gathering data from emission sources within each subcategory that it is regulating. For the MATS, this took place through the 2010 Information Collection Request. EPA is authorized to collect emissions data by Section 114 of the Clean Air Act, but much of the information must be kept confidential. As in a typical benchmarking program discussed throughout the "Issues and Recommendations" Section, EPA then ranks regulated units in each subcategory from highest to lowest emissions. EPA is required to set the MACT floor at the average emission rate of emission units within the best performing 12 percent of units. EPA and states can then consider other factors, including public health benefits and compliance costs, to set the final MACT at a stricter rate than this MACT floor, but the required rate can never be less stringent than the MACT floor.[36] The process of setting the final MACT generally does not involve benchmarking.

Thus the setting of the MACT floor is done in a very similar fashion to how allowance allocation benchmarks are determined in California and the EU-ETS. The process for EPA is relatively less complex because power plants are more homogenous than industrial facilities, though there are still a few subcategories for each fuel where EPA must set distinct benchmarks. EPA's effort is also simplified by its statutory authority to directly collect all necessary emissions data, which states would not have when crafting any industrial benchmarking program.

Multi-Sector Voluntary Programs

Although generally less sophisticated than the benchmarking methodologies developed for California's cap-and-trade program and the EU-ETS, multi-sector voluntary programs might serve as examples for more simplified benchmarking programs that can be adopted at the state or regional level. Most of the programs summarized below are designed to help facilities improve their own performance without making direct comparisons among facilities. In this way these programs differ from the type of program described in in the "Issues and Recommendations" Section, but they still provide useful lessons in terms of data gathering and metric choice and can serve as a basis for new initiatives to improve industrial efficiency.

EPA's Center for Corporate Climate Leadership

Program Summary
Program Type: Voluntary, multi-sector
Benchmarked Unit: Varied; generally greenhouse gas emissions per ton of product
Comparisons Enabled: Facilities can compare themselves to the sector average

EPA's Center for Corporate Climate Leadership was launched in 2012 and provides a variety of resources to assist all companies, including industrial companies, to identify and achieve cost-effective greenhouse gas emission reductions. This center is not focused solely on benchmarking, but one of its major programs, the Climate Leadership Awards (CLA), recognizes companies that have reduced their emissions intensity as determined through benchmarking. The CLA is a co-venture among EPA, The Climate Registry, The Center for Climate and Energy Solutions, and the Association of Climate Change Officers. In addition to a number of other possible awards, companies can be recognized for monitoring and reporting emissions, setting emissions intensity goals, and planning activities to achieve those goals.

For the purposes of this paper, the major takeaway from the Center for Corporate Climate Leadership is its development of a publicly available benchmarking tool called the EPA Corporate GHG Goal Evaluation Model. This model uses data from a variety of sources to calculate fuel consumption and economic output values for common industrial sectors. These results are combined with emission factors to estimate CO_2 emissions intensity per sector. By inputting its own data and goals, a facility can compare its emissions intensity to the industry average, both historically and into the future because the model includes business-as-usual forecasts.

The data sources used by the Corporate GHG Goal Evaluation Model are publicly available. The central data sources are the input-output table projections and historical tables published by the Bureau of Labor Statistics. These tables show the flow of commodities from production through intermediate use by industries and finally as purchases by final users. The model uses information from the Energy Information Administration's (EIA) Annual Energy Outlook, EIA's State Energy Data System (SEDS), and EIA's Manufacturing Energy Consumption Survey (MECS) to estimate the amount and type of energy consumed by each sector based on energy expenditures and energy prices per sector. Thus, the model is able to calculate the quantity of energy and electricity each sector consumed in a given year, which can be combined with fuel-specific emissions factors to determine emissions intensity.

Users of the model specify weighted revenue by North American Industry Classification System (NAICS) code and the model estimates fuel-specific greenhouse gas emissions intensity for a composite NAICS sector with the same revenue-weighting as the specific facility. This allows users to compare their emissions intensity reduction goals with a business as usual scenario for the sector average. Many additional details and capabilities of this model can be found here.

Though the Corporate GHG Goal Evaluation Model is useful for facilities attempting to reduce emissions intensity, it is not useful as a means to compare facilities against each other. Therefore, if state policymakers are seeking to develop a program that identifies top performers within an industrial sector in terms of emissions intensity, this program should not serve as a model. The model does however incorporate data sources that would likely be useful in the establishment of a new program.

EPA's Energy Star Challenge for Industry

Program Summary

Program Type: Voluntary, multi-sector

Benchmarked Unit: Process unit efficiency

Comparisons Enabled: Facilities can compare themselves to others in the sector

Energy Star for Industry employs an Energy Performance Indicator (EPI) to measure the energy intensity of industrial facilities. The EPI is an indexed benchmark that accounts for a number of variables to focus on process efficiency alone. By controlling for variables such as labor, energy and material costs, and output, the EPI allows operators to answer the question "How would my plant compare to everyone else in my industry, if all other plants were similar to mine?"[37] Each participating facility is assigned an EPI that is indexed to the efficiency of a facility that maximizes process efficiency but is otherwise identical.

Data for the EPI are collected by the U.S. Bureau of the Census, including the Census of Manufacturing and the Manufacturing Energy Consumption Survey. Not all data used to determine an EPI are public, and EPA and its contractors operate under strict confidentiality requirements to protect trade secrets and other proprietary information. Thus the EPI is not likely to be copied by policymakers outside of Energy Star due to data confidentiality issues, though EPI values assigned to facilities could be employed directly, especially but a group of states that pool resources to develop such a program.

As a benchmark that excludes factors other than process efficiency, the EPI is an excellent tool to identify which facilities can make significant improvements by investing in their processes. It is also successful at recognizing and crediting facilities that have made significant improvements to their performance. However, the EPI does not identify opportunities to improve systematic efficiency, such as moving facilities closer to feedstock sources. The EPI and Energy Star for Industry provide an excellent example of why benchmarking methodology should be tailored to how the benchmarks will be used.

DOE's Better Buildings, Better Plants

Program Summary

Program Type: Voluntary, multi-sector

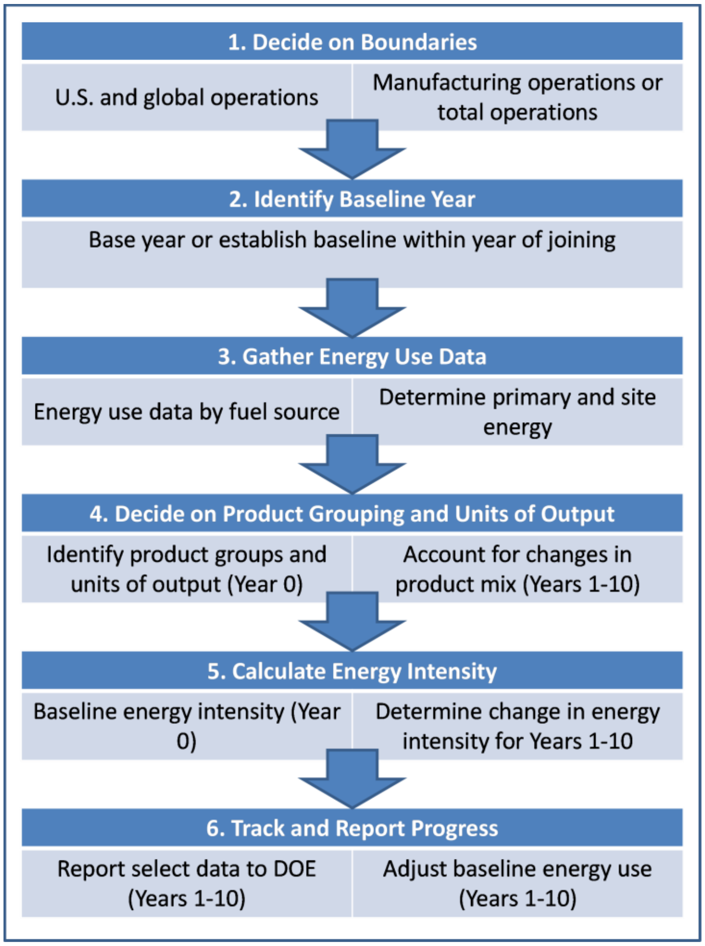
Benchmarked Unit: Energy consumption per ton of product

Comparisons Enabled: A facility can compare its current to its historical performance

The U.S. Department of Energy (DOE) runs two closely related voluntary initiatives to encourage improved energy efficiency in the buildings and industrial sectors: The Better Buildings, Better Plants Program and the more involved Better Buildings, Better Plants Challenge. Both replaced the Save Energy Now program in 2011, and involve a voluntary pledge to reduce energy intensity by 25 percent over 10 years in addition to annual reporting. Partners in the Challenge also agree to transparently pursue innovative approaches to energy efficiency and make a significant, near-term investment in an energy saving project or set of projects. The industrial element of both the Program and Challenge are open to any company in the U.S. manufacturing sector. In addition to technical support, participants in each program receive public recognition upon signing up, and increasing levels of recognition and exposure for achieving various milestones.

After becoming a partner, a facility conducts an initial energy evaluation and develops a general plan for meeting its energy goals. Next, facilities must choose a metric to track and determine their energy intensity baseline according to the steps shown in Figure 11. Better Buildings, Better Plants defines energy intensity as the amount of energy consumer per unit of output, which can be measured in mass, number, volume, or functionality. Physical units of output are preferred in this metric, but financial units are also acceptable in some cases. Facilities then begin implementing energy-saving projects and report their energy data and progress annually.

Figure 11: Better Buildings, Better Plants Steps for Developing an Energy Intensity Baseline



Source: U.S. Department of Energy, Guide for Better Buildings, Better Plants Program Partners (November, 2011).

Relative to EPA's Energy Star for Industry, DOE's Better Buildings, Better Plants uses a much more flexible methodology. The benefit of this is that unique facilities will be able to easily take advantage of Better Plants, Better Buildings. The disadvantage is that without a consistent metric imposed on every facility within a sector, Better Plants, Better Buildings does not enable comparisons across facilities to give operators a sense of their relative performance in terms of energy intensity. Because this program allows comparisons of historical performance but not across facilities, this paper refers to this program type as "internal benchmarking."

More information on DOE's Better Plants Program and Challenge can be found here.

Southwest Energy Efficiency Project (SWEEP) Colorado Industrial Energy Challenge (CIEC)

Program Summary
Program Type: Voluntary, multi-sector
Benchmarked Unit: Varied energy efficiency metrics
Comparisons Enabled: A facility can compare its current to its historical performance

The Southwest Energy Efficiency Project (SWEEP) is a public interest organization promoting greater energy efficiency in Arizona, Colorado, Nevada, New Mexico, Utah, and Wyoming. SWEEP initially focused on utility energy efficiency, but has expanded to include the industrial sector as well. SWEEP leads the Colorado Industrial Energy Challenge (CIEC), which is open to any industrial facility in Colorado with more than $400,000 in annual energy costs. The CEIC provides assistance and recognition to facilities that set a five-year energy efficiency goal. Participation also requires reporting of energy consumption or intensity each year to demonstrate progress toward the goal. Facilities are free to set their own consumption- or intensity-based goals, which range from 5 to 40 percent. Outstanding achievements in the program are recognized annually with an event and a certificate from the Governor's Energy Office.

Of the benchmarking programs described in this Section, the CEIC is one of the most basic. It does not provide uniform metrics to allow for facilities to compare themselves against others within the same sector. However, where minimal resources are available to develop a sophisticated measurement methodology, CEIC provides a good example of a public program that uses energy consumption measurement and public reporting to encourage improvements.

Single-Sector Voluntary Programs

Single-sector voluntary programs are generally designed by trade organizations to help improve efficiency within the organization's sector. They typically assist operators in improving performance as measured against a historical baseline rather than against similar facilities, and thus do not serve as exact models for policymakers to follow in developing the type of benchmarking program discussed throughout the "Issues and Recommendations" Section. However, these programs can still provide valuable lessons in some critical elements of a benchmarking program, such as monitoring protocols.

World Business Council for Sustainable Development Cement Sustainability Initiative

Program Summary
Program Type: Voluntary, single-sector
Benchmarked Unit: Greenhouse gas emissions per ton of product
Comparisons Enabled: Cement facilities can be compared to each other

The World Business Council for Sustainable Development is a coalition of over 100 companies from a variety of sectors around the globe that share a commitment to sustainable development through economic growth, ecological balance, and social progress. The Cement Sustainability Initiative (CSI) is a collaboration of over twenty cement companies from around the globe that was started in 1999. The CSI has a variety of sustainability goals, including a reduction of carbon dioxide emissions from the cement industry. The CSI has approached this goal by developing a carbon dioxide accounting protocol and helping its members set and reach mitigation targets.

The CSI, working with the World Resources Institute, developed the "Cement CO_2 and Energy Protocol: CO_2 and Energy Accounting and Reporting Standard for the Cement Industry" (protocol), now in its third version. The protocol is used by most international cement companies, the EU-ETS, EPA's Mandatory Reporting of Greenhouse Gases Rule, and the California Climate Registry. The protocol is fully explained here.

As discussed above, the protocol addresses direct emissions, the main sources of indirect emissions, and on-site power generation. The protocol allows for the determination of emissions by either making calculations based on inputs, or by directly measuring outputs. The protocol accounts for emissions from raw materials and combustion both within and outside of kilns. To determine indirect emissions, the protocol favors the use of emission factors provided by the electricity suppliers, with factors supplied by the government as a fallback. The protocol recommends that facilities report both absolute emissions and emissions intensity, in the form of mass of carbon dioxide per mass of cementitious material produced.

The CSI has driven significant carbon dioxide emission reductions in cement plants. Fifteen of the CSI's twenty-four member companies have been tracking progress toward their carbon dioxide goals in their annual sustainability reports. As of 2010, 930 facilities were following the protocol, with representation of about twenty-five percent of cement production, including over 70 percent from the Americas and over 95 percent from Europe. Between 1990 and 2010, cement production in participating companies grew by 61 percent while carbon dioxide emissions grew by only 39 percent, equating to a 16 percent reduction in net carbon dioxide emissions per metric ton of cement. More information on the success of the CSI can be found here.

The CSI and associated protocol are an excellent example of a flexible benchmarking methodology that can be adapted to different purposes. The CSI specifically encourages cement plants to benchmark against their own past performance, and thus is another example of internal benchmarking and does not fit directly into the arena of benchmarking programs discussed throughout the "Issues and Recommendations" Section. However, by requiring the use of the comprehensive Cement CO_2 and Energy Protocol, the CSI positions its participants to benchmark themselves against each other and report to government programs such as the EU-ETS.

World Steel Association Climate Action Recognition Program

Program Summary

Program Type: Voluntary, single-sector

Benchmarked Unit: Greenhouse gas emissions per ton of product

Comparisons Enabled: Steel facilities can be compared to each other

The World Steel Association (worldsteel) was founded in 1967 to promote steel and the steel industry to customers, the industry, media, and the general public. It represents about 170 steel producers, which together account for approximately 85 percent of world steel production. Worldsteel operates a confidential

CO_2 data collection program that is open to all steel producers, including those that are not members of worldsteel. Program participants, of which there are 40, collect and report data according to program guidelines, and in exchange receive a report showing a range of emissions against which the participant should compare itself. Program participants are also publicly recognized by worldsteel and can use a Climate Action Member logo on any marketing materials. The program methodology uses a simple benchmark of tons of CO_2 emissions per ton of crude steel production.

The worldsteel Climate Action Recognition Program enables a broad understanding of emissions from the steel sector and provides steel facilities with a consistent methodology to monitor their own emissions. The program is an example of a benchmarking program that compares facilities against each other rather than simply encouraging facilities to compare themselves against their own historical performance. Though data transparency should be an element of a benchmarking program when possible, the worldsteel Climate Action Recognition Program exemplifies how a program can provide participants with valuable information about its competitors without revealing proprietary details.

More information can be found on the worldsteel CO_2 emissions data collection methodology here.

Northwest Food Processors Association Energy Roadmap

Program Summary

Program Type: Voluntary, single-sector

Benchmarked Unit: Varied energy efficiency metrics

Comparisons Enabled: A facility can compare its current to its historical performance

The Northwest Food Processors Association (NWFPA) was organized in 1914 as a canners association, and currently represents the interests of food processors in Idaho, Oregon, and Washington. In order to remain competitive, the NWFPA understands the importance of energy efficiency. In 2008, the NWPFA established an energy vision and long-term energy goals with the support of the Northwest Energy Efficiency Alliance (NEEA) and the DOE. The NWPFA set a goal to reduce member-wide energy intensity by twenty-five percent in ten years and by fifty percent in twenty years. To guide the achievement of this goal, the NWFPA developed an industry-specific Energy Roadmap.

The Energy Roadmap for Members (as distinguished from the Association Energy Roadmap, which provides a path for energy intensity reductions at the association level) charts a pathway for individual food processors to reduce energy intensity. The first step in the Roadmap is to conduct a qualitative self-assessment to determine if a facility is “reactive,” “growing,” or “managed, measured, planned.” This step helps operators identify and adjust their perspective on energy efficiency. The next step is to build the personnel and resources necessary to improve facility performance, such as creating an Energy Team and building a business case.

After the foundation for the benchmarking effort has been set, the Roadmap suggests determining an energy baseline and setting a performance target. All energy inputs should be accounted for in a single measure by converting to BTUs. Production should be expressed in mass to the extent possible for consistency. For the target, the Roadmap recommends a SMART goal, as described in the “General Policy Considerations” subsection. Baselines and targets will vary widely across facilities due to the heterogeneity of the food processing industry. A product benchmark that compared all facilities against each other would thus be impractical in this industry, but all facilities can seek to achieve a similar percentage improvement in performance from their individual baselines. A fuel benchmark that tracks greenhouse gas emissions per unit of fuel consumed would also be possible in this sector.

The Roadmap next guides participating facilities through an energy assessment and development of an energy action plan to identify opportunities for improvement and how those opportunities will be prioritized and captured. The final major stage of the Roadmap is to monitor performance and adjust the energy action plan appropriately to ensure progress is being made toward the facility's goals. The Roadmap places considerable emphasis on reliable measurement, and suggests using an energy monitoring and management system (EMMS), such as the Green Energy Management System, and electrical sub-metering rather than relying on utility bills alone.

As with most other voluntary benchmarking programs, the Energy Roadmap is designed as a tool for plants to improve based on historical performance rather than directly compete against their peer facilities. However, the encouragement of consistent metric use and monitoring methodology across facilities lays the groundwork necessary for a benchmarking program that fairly compares facilities against each other. The Roadmap is also an excellent example of putting benchmarking into the context of a broader energy efficiency program. The Roadmap contains several elements designed to alter the culture at food processing facilities to be more innovative and proactive in regards to saving energy, and therefore cutting costs. Benchmarking becomes a tool designed to help operators achieve goals they believe in, rather than an administrative burden they do not understand. Policymakers should consider following this example and building benchmarking into broader energy efficiency campaigns that encourage operators to be more amenable to improving efficiency.

More information about the NWFPA Energy Roadmap can be found here.

Lessons from Existing Benchmarking Programs

There is a significant amount of experience with internal benchmarking for voluntary programs in the industrial sector in North America. The basic attributes of those described above are summarized in Table 2. There are well over 100 programs in addition to these, including governmental and nongovernmental, that enable and encourage industrial facilities to improve their energy and/or emissions performance compared to historical levels. These programs have led to significant improvements, but ultimately can only drive incremental gains because targets do not necessarily account for the maximum achievable performance within a sector.

Table 2: Comparison of programs summarized in Section 4

PROGRAM	PROGRAM TYPE	BENCHMARKED UNIT	COMPARISONS ENABLED
EU-ETS Allowance Allocation	Allowance allocation	1) Greenhouse gas emissions / metric ton of product 2) Greenhouse gas emissions / heat produced 3) Greenhouse gas emissions / fuel consumed	Facilities that produce the same product can be compared to each other
California Allowance Allocation	Allowance allocation	1) Greenhouse gas emissions / metric ton of product 2) Greenhouse gas emissions / fuel consumed	Facilities that produce the same product can be compared to each other
New Source Performance Standard	Mandatory performance standard	Varied; includes emissions per mass unit of product	Facilities within a regulated sector can be compared to each other
Mercury and Air Toxics Standards for Power Plants	Mandatory performance standard	Varied; generally air pollutants per MMBTU of fuel consumed	Power plants can be compared to each other
EPA's Center for Corporate Climate Leadership	Voluntary, multi-sector	Varied; generally greenhouse gas emissions per ton of product	Facilities can compare themselves to the sector average
EPA'S Energy Star for Industry	Voluntary, multi-sector	Process unit efficiency	Facilities can compare themselves to others in the sector
DOE's Better Buildings, Better Plants	Voluntary, multi-sector	Energy consumption per ton of product	A facility can compare its current to its historical performance
Southwest Energy Efficiency Project: Colorado Industrial Energy Challenge	Voluntary, multi-sector	Varied energy efficiency metrics	A facility can compare its current to its historical performance
World Business Council for Sustainable Development Cement Sustainability Initiative	Voluntary, single-sector	Greenhouse gas emissions per ton of product	Cement facilities can be compared to each other
World Steel Association Climate Action Recognition Program	Voluntary, single-sector	Greenhouse gas emissions per ton of product	Steel facilities can be compared to each other
Northwest Food Processors Association Energy Roadmap	Voluntary, single-sector	Varied energy efficiency metrics	A facility can compare its current to its historical performance

In developing new programs that compare facilities to each other, policymakers can take many lessons from these voluntary, internal benchmarking programs. These include how to communicate the benefits of efficiency investments to industry, how to recognize leaders, and how to gather data. However, several novel and complicated issues arise when developing a benchmarking curve in North America. These include how to define a product category, how to include intermediate products, and how to account for indirect emissions without double counting emissions. For these issues, California has the most experience through its benchmarking work to allocate allowances for its cap-and-trade program. California's benchmarking work should therefore serve as a starting point for policymakers aiming to develop new benchmarking programs.

Conclusion

Policymakers developing new benchmarking programs have many issues to wrestle with, even in the limited context of setting a benchmarking curve. Some of these are purely policy issues that are dependent on the goal of the program. For example, which sectors to cover and whether to limit the program to internal benchmarking are decisions that cannot be guided by technical recommendations. However, there are a variety of technical decisions discussed throughout this paper that can be guided by the available literature and the experience of existing programs. For these issues, we encourage policymakers to follow the lessons learned and presented in this benchmarking whitepaper to avoid repeating research and previous deliberations.

Further Reading

Gregory A. Norris, Jim Sullivan, and Bella Tonkonogy, Evaluating Corporate Climate Performance: A Model for Benchmarking GHG Reductions. ACEEE Summer Study on Energy Efficiency in Industry (2007).

C.K. Lyn Griffth, An Overview of U.S. Federal Government Industrial Energy Efficiency Programs 2005-2011. Energy Pathways LLC. Rockville, MD (2012).

Martijn Rietbergen and Kornelis Blok, Setting SMART targets for industrial energy use and industrial energy efficiency. 38 Energy Policy 4339 (2010).

Gale Boyd, Elizabeth Dutrow, and Walt Tunnessen, The evolution of the ENERGY STAR energy performance indicator for benchmarking industrial plant manufacturing energy use. 16 Journal of Cleaner Production 709 (2008).

Elena Giacone and Salvatore Manco, Energy efficiency measurement in industrial processes. 38 Energy 331 (2012).

Michael Ruth, Ernst Worrell, and Lynn Price, A Process-Step Benchmarking Approach to Energy Use at Industrial Facilities: Examples from the Iron and Steel and Cement Industries, Presented at the 2001 ACEEE Summery Study on Energy Efficiency in Industry (2001).

Stockholm Environment Institute, White Paper: Issues and Options for Benchmarking Industrial GHG Emissions. Submitted to the Washington State Department of Ecology (June 30, 2010).

Ecofys, Developing Benchmarking Criteria for CO_2 Emissions. By order of the European Commission (February 2009).

California Air Resources Board, Cap and Trade Regulation Discussion Draft Appendix B: Development of Product Benchmarks for Allowance Allocation (July 2011).

California Air Resources Board, Cap and Trade Regulation Initial Statement of Reasons Appendix J: Allowance Allocation (July 2011).

Industrial Efficiency Policy Database website: United States (Federal).

Glossary

Allowance Allocation: The methodology used to distribute emission allowances in a cap-and-trade program. There are a variety of methods to allocate these allowances, including free allocation based on an emissions intensity benchmark.

Benchmark: In general, a benchmark is something that serves as a standard by which others may be measured or judged. For the purposes of this paper, it means a standard level of energy or emissions intensity that industrial facilities are compared against, as determined by a set methodology determined by policymakers.

Benchmarking Curve: A visualization that orders industrial facilities within a defined sector according to a certain benchmark. For example, a benchmarking curve might display each steel mill in a certain jurisdiction according to how much greenhouse gas is released per ton of steel produced.

Cap and Trade: A cap-and-trade system sets an overall limit on emissions, requires entities subject to the system to hold sufficient allowances to cover their emissions, and provides broad flexibility in the means of compliance. Entities can comply by undertaking emission reduction projects at their covered facilities and/or by purchasing emission allowances (or credits) from the government or from other entities that have generated emission reductions in excess of their compliance obligations.

Combined Heat and Power System (CHP): Combined heat and power (CHP), also known as cogeneration, refers to a group of proven technologies that operate together for the concurrent generation of electricity and useful heat in a process that is generally much more energy efficient than the separate generation of electricity and useful heat.

Emissions Intensity: The average emission rate of a given pollutant from a given source relative to the intensity of a specific activity; for example tons of carbon dioxide released per ton of steel produced.

Energy Intensity: The average consumption rate of a given fuel of a given facility or process unit relative to the intensity of a specific activity; for example megawatt hours of electricity consumed per ton of steel produced.

Feedstock: The raw material used to fuel or supply an industrial process.

Leakage: A reduction in emissions of greenhouse gases within a jurisdiction that is offset by an increase in emissions of greenhouse gases outside the jurisdiction. For example, if a regulated facility moves across the border to continue operations unchanged rather than reducing its emissions.

Mandatory Performance Standard: A regulatory program that requires all facilities in a covered sector to achieve a specified emissions or energy intensity. These programs may allow for the use of tradable credits for compliance.

Process Unit: A discrete unit that prepares, treats, alters, converts, or otherwise alters a feedstock or intermediate product as part of a system at an industrial facility.

2101 WILSON BLVD. SUITE 550 ARLINGTON, VA 22201 703-516-4146

C2ES.ORG

Endnotes

---

[1] Centre for European Policy Studies, *Benchmarking in the EU ETS: Policy Issues for the EU and Beyond* (Preliminary Version). (Brussels, 2009), available at http://www.ceps.eu/system/files/task_force/2009/07/ExSum%20to%20post%20on%20website.pdf.

[2] Martijn Rietbergen and Kornelis Blok, *Setting SMART targets for industrial energy use and industrial energy efficiency*. 38 Energy Policy 4339 (2010).

[3] Centre for European Policy Studies, *Benchmarking in the EU ETS: Policy Issues for the EU and Beyond* (Preliminary Version). (Brussels, 2009), available at http://www.ceps.eu/system/files/task_force/2009/07/ExSum%20to%20post%20on%20website.pdf.

[4] Stockholm Environment Institute, White Paper: *Issues and Options for Benchmarking Industrial GHG Emissions*. Submitted to the Washington State Department of Ecology (June 30, 2010).

[5] Ibid.

[6] Ibid.

[7] Ibid.

[8] Department of Energy, Industrial Assessment Centers Database: http://iac.rutgers.edu/database/.

[9] Stockholm Environment Institute, White Paper: *Issues and Options for Benchmarking Industrial GHG Emissions*. Submitted to the Washington State Department of Ecology (June 30, 2010).

[10] California Air Resources Board, Mandatory Greenhouse Gas Emissions Reporting, Home Page, http://www.arb.ca.gov/cc/reporting/ghg-rep/ghg-rep.htm.

[11] U.S. Environmental Protection Agency, Greenhouse Gas Reporting Program, Home Page, http://www.epa.gov/ghgreporting/.

[12] Stockholm Environment Institute, White Paper: *Issues and Options for Benchmarking Industrial GHG Emissions*. Submitted to the Washington State Department of Ecology (June 30, 2010).

[13] Ibid.

[14] California Air Resources Board, Cap and Trade Regulation Discussion Draft Appendix B: Development of Product Benchmarks for Allowance Allocation (July 2011).

[15] Stockholm Environment Institute, White Paper: *Issues and Options for Benchmarking Industrial GHG Emissions*. Submitted to the Washington State Department of Ecology (June 30, 2010).

[16] Ernst Worrell and Lynn Price, *An Integrated Benchmarking and Energy Savings Tool for the Iron and Steel Industry*, 3 International Journal of Green Energy 117 (2006).

[17] Stockholm Environment Institute, White Paper: *Issues and Options for Benchmarking Industrial GHG Emissions*. Submitted to the Washington State Department of Ecology (June 30, 2010).

[18] Ibid.

[19] Michael Ruth, Ernst Worrell, and Lynn Price, *A Process-Step Benchmarking Approach to Energy Use at Industrial Facilities: Examples from the Iron and Steel and Cement Industries*, Presented at the 2001 ACEEE Summery Study on Energy Efficiency in Industry (2001).

[20] Ernst Worrell and Lynn Price, *An Integrated Benchmarking and Energy Savings Tool for the Iron and Steel Industry*, 3 International Journal of Green Energy 117 (2006).

[21] Stockholm Environment Institute, White Paper: *Issues and Options for Benchmarking Industrial GHG Emissions*. Submitted to the Washington State Department of Ecology (June 30, 2010).

[22] Ibid.

[23] Ibid.

[24] Stockholm Environment Institute, White Paper: *Issues and Options for Benchmarking Industrial GHG Emissions*. Submitted to the Washington State Department of Ecology (June 30, 2010).

[25] Ecofys, *Developing Benchmarking Criteria for CO_2 Emissions*. By order of the European Commission (February 2009).

[26] As defined as emissions per unit of heat produced

[27] As defined as emissions per unit of fuel consumed

[28] 40 CFR 60.30d-60.32d (2013).

[29] 40 CFR 60.30e-60.39e (2013).

[30] Clean Air Act § 111(a)(1).

[31] Environmental Protection Agency, *Background on Establishing New Source Performance Standards (NSPS) Under the Clean Air Act*, available at http://epa.gov/carbonpollutionstandard/pdfs/111background.pdf.

[32] Franz Litz, Pace Energy and Climate Center, Presentation: *What's Ahead for Power Plants in Minnesota? Greenhouse Gas Performance Standards under Section 111 of the federal Clean Air Act*, available at http://www.pca.state.mn.us/index.php/view-document.html?gid=16094.

[33] Jennifer Macedonia, Bipartisan Policy Center, Presentation: *Summary of EPA Power Sector Regulation: Air Toxics Standards* (January 2012), available at http://www.naruc.org/Publications/BPC%20Handouts%20on%20MATS%20012412JLM%20for%20NARUC.pdf.

[34] 78 FR 24073 (2013), available at http://www.gpo.gov/fdsys/pkg/FR-2013-04-24/pdf/2013-07859.pdf.

[35] 77 FR 9304 (2012), available at http://www.gpo.gov/fdsys/pkg/FR-2012-02-16/pdf/2012-806.pdf.

[36] 76 FR 24976 (2011), available at http://www.gpo.gov/fdsys/pkg/FR-2011-05-03/pdf/2011-7237.pdf.

[37] Gale Boyd, Elizabeth Dutrow, and Walt Tunnessen, The evolution of the ENERGY STAR energy performance indicator for benchmarking industrial plant manufacturing energy use. 16 Journal of Cleaner Production 709 (2008).