The benefits of creating a cross-country data framework for energy efficiency

Alex Katzman, VP of Business Development and Project Manager for SEAD Energy Efficiency Data Access Project, Enervee

Michael McNeil, Deputy Leader, International Energy Studies Group (Acting), Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory

Stephen Pantano, Senior Director of Global Programs, CLASP (Collaborative Labeling and Appliance Standards Program)

Abstract

As manufacturers now sell a similar range of consumer electronics and home appliances to major markets around the world, the task of identifying a product's energy efficiency rating has usually been the responsibility of each country and its respective government agency. This has led to a multitude of energy efficiency testing procedures, ratings, and certifications, resulting in disparate data being captured on identical products. Furthermore, lack of consistent product identification criteria means product energy performance is not easily connected to relevant information about the product such as market availability, price or real world energy consumption.

This paper presents a new data standard for reporting energy performance and related product information that can be adopted internationally. To inform the development of this standard, we explore the existing energy efficiency market data for the two example products of TVs and Room Air Conditioners. This paper discusses current/future use cases of appliance level energy efficiency data across all stakeholders, including consumers, retailers/manufacturers, global standards organizations, third party service providers, and regulatory agencies. It also explains the key benefits of moving to a common international data framework for energy efficiency, such as: 1) a centralized product information repository for comparing energy use, ratings/certifications, and pricing data 2) improved access to relevant consumer electronics and appliance data to facilitate new policy development and harmonization across markets 3) enablement of retailers and other third parties to embed actionable energy efficiency information as part of the consumer experience.

1. Introducing a Cross-country Data Framework for Energy Efficiency

Among the most useful tools in the development of effective consumer electronics and home appliances efficiency programs and promoting consumer awareness of efficiency are data that allow for:

- Linking consumer retail information to efficiency parameters more effectively.
- Tracking efficiency trends over time (monitoring) and comparison of efficiency in different markets and jurisdictions (benchmarking).
- Tracking retail prices of appliances over time, and between countries.

Currently, broad access to consumer electronics and home appliances efficiency and price data and comparability between markets is imperfect. Product energy use labels, in countries that require them, are often not prominently displayed on sales websites; and in many cases it is difficult to access detailed product energy rating information. Policy makers must invest significant time and money in collecting data about the market status of the efficiency of available products when determining appropriate levels for energy efficiency policies. In addition, comparability of attributes across markets is often difficult, even for identical products, due to inconsistencies in model identifiers and efficiency metrics and protocols. A well-defined *data standard* for appliance attributes and efficiency characteristics could potentially solve this and many other product energy information access problems.

Key objectives of the data standard include: (1) development of internationally compatible data specifications for populating, retrieving, and cross-referencing data from web-based product databases and (2) definition of standardized methods for populating these product-specific databases. Data sources may include the use of product certification databases (such as those maintained by agencies that administer energy labeling and minimum performance standards programs); software tools (e.g., “webcrawlers” to gather and regularly update product-specific availability, efficiency and pricing data); and guidelines for the gathering and reporting of actual field (on-site) data on product energy use.

This paper outlines a proposed data framework for consumer electronics and home appliances energy performance and cost data being developed as part of the Super-efficient Equipment and Appliance Deployment (SEAD) Initiative of the Clean Energy Ministerial.[1] The goal of this project is to publish a data standard that can be implemented by relevant organizations that collect and store data about the energy performance, market availability and costs of energy-consuming equipment and appliances. This data standard can be used for sharing data among relevant organizations, and in the development of databases that store and analyze this information.

The data framework relies on two main data source types and anticipates a high degree of data collection automation. Product model lists are collected directly from retail websites, using online shopping application programmer interfaces (APIs). Connection to these online data feeds enables automated downloading of product model numbers, many product attributes and current retail prices. The second dataset type consists of efficiency certification databases, often maintained by regulating bodies responsible for minimum energy performance standards (MEPS) and labeling programs.

The combination of these two datasets provides a wealth of new information. The correlation of models found on retail websites with those in efficiency databases provides a time-dependent picture of the distribution of efficiencies on offer. These data could be updated frequently (perhaps daily), creating a high temporal resolution picture of the evolution of the market. Finally, datasets collected in this way can be used to compare the market for efficiency in different countries. By connecting real time prices with efficiency ratings and market share for the same product sold across markets, there is the potential to compare the cost consumers are willing to pay for efficiency across markets. This step may require robust algorithms for converting efficiency metrics between test procedures used in different countries.

.

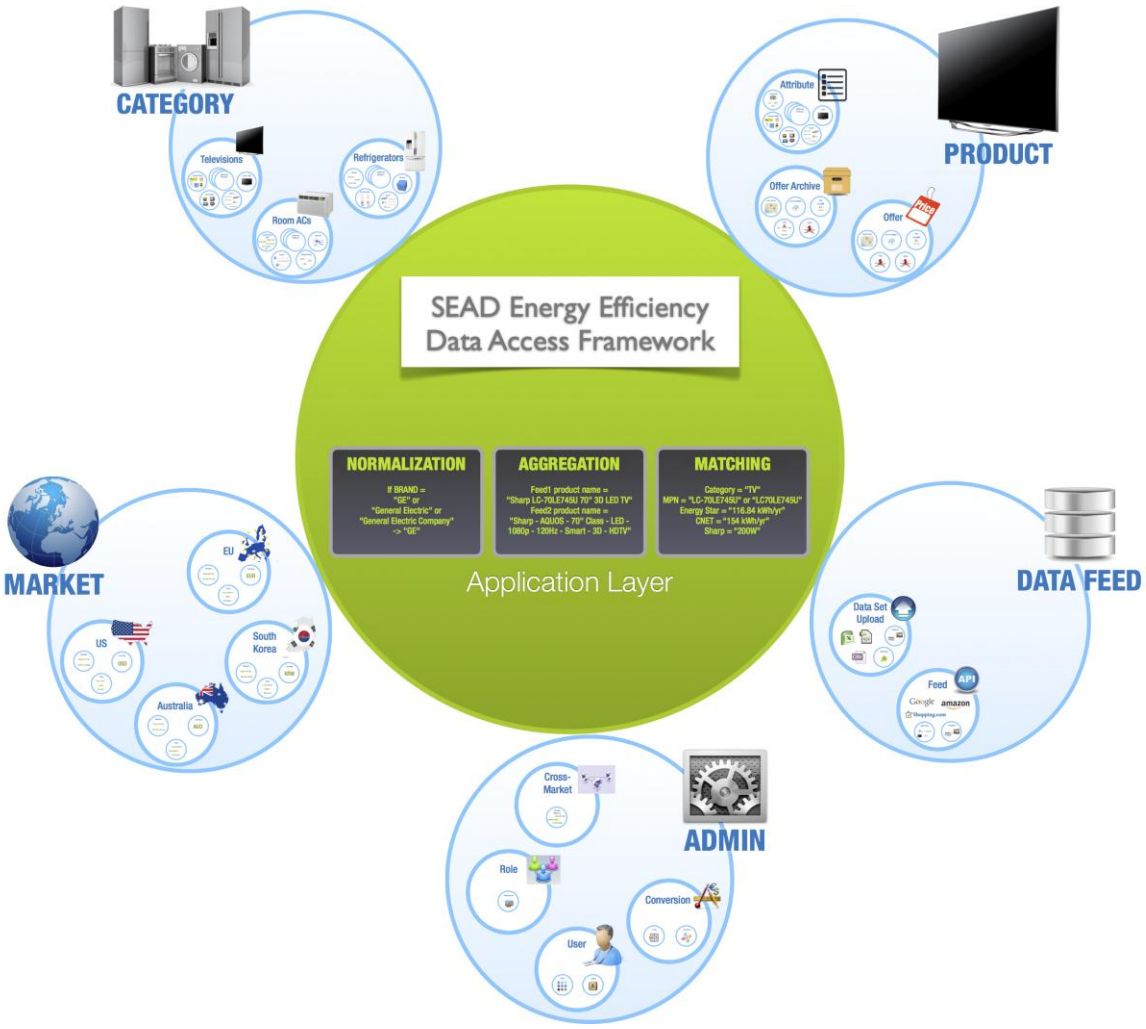
The following sections provide a detailed description of the framework developed as a pilot project supported by the U.S. Department of Energy as part of the SEAD initiative. We then consider two products – room air conditioners (Room ACs) and televisions (TVs) – as a test case to demonstrate the facility of the data standard. The paper concludes by considering applications of a fully implemented framework and discussion of implications.

2 Defining the Framework and Data Standard

2.1. Overview

The SEAD Energy Efficiency Data Access framework was designed to be able to capture and compare energy efficiency product data for the many types of consumer electronics and home appliances that are sold around the world. It uses a product-centric model with an extensible design to ensure that it can meet the requirements of any market or category. The core of the framework standardizes markets, brands, categories, product attributes, and units of measurement with common definitions. An intelligent application layer would connect with the framework to normalize, aggregate, and match data across the disparate energy efficiency certification and retailer data sources. The outcome is a new data standard that defines a methodology for capturing and sharing this data.

Figure 1 – Data Framework Conceptual Overview



2.2. Identifying Use Cases

To ensure that the new framework design could handle a wide range of potential use cases, feedback was gathered through surveys and interviews with the members of the SEAD Standards and Labeling Working Group, which consists primarily of policymakers from the 13 SEAD member economies. The outcome of the surveys and interviews was a prioritized list of use cases.

Table 1 – Prioritized Use Cases

#	Use Case	Need Type	Priority
1	Access to other countries' certifications and ratings for monitoring, verification, and policy design	Policy	High
2	Connect sales and efficiency trends	Policy	High
3	Ability to assess impact of marketing campaigns on consumer adoption	Policy	Medium
4	Enables comparison of product rating methods across countries	Policy	High
5	Facilitates procurement activities	Consumer	Medium
6	Enables development of consumer mobile apps/online comparison tools	Consumer	High
7	Facilitates product energy efficiency awards	Consumer	Medium
8	Facilitates utility incentive programs	Consumer	Medium
9	Enables comparison of testing procedures	Policy	Medium

Expressed needs broke down into two distinct categories: Policy and Consumer. 'Policy' refers to use cases that relate to analyzing, benchmarking, or creating appliance standards. 'Consumer' refers to use cases that enable consumers to have access to and leverage this data in their purchasing decisions. The overwhelming feedback was that this data standard should focus mostly on Policy development but also enable development of consumer mobile apps/online comparison tools.

2.3. Creating a Global Data Standard

An essential part of establishing a data standard is agreeing on a simple, consistent model that can be applied across diverse markets. To accomplish this task, the concept of a Global Category was developed which defines a universal set of product attributes that are gathered irrespective of the category. This is the mandatory information that will be captured for any category, covering the unique product identifiers, market, brand/manufacturer, pricing, energy, certification, and marketing related data.

Global Category Definition

To the greatest extent possible, the Global Category definitions have been based off of the IEA 4E Mapping and Benchmarking[2] and GS1 Global Product Classification (GPC)[3]. These existing definitions were chosen to facilitate harmonization of product category definitions across markets using international standards. Table 2 shows the proposed definition.

Table 2: Global Category Data Standard

Attribute	Type	Notes
Manufacturer Part Number (MPN)	String	Product Model Number stripped of extraneous characters
Brand	String	Normalized Across All Markets
Manufacturer	String	Parent Relationship to Brand
Market	Code	Defined region based on ISO 3166-1 country codes
Global Trade Item Number (GTIN)	String	Universal Product Code (UPC) or International Article Number (EAN)
QR Code	String	Quick-Response Code
Price	Numeric	Min, Max, Mean, Median, and Standard Deviation of Daily Prices
MSRP	Numeric	Manufacturers Suggested Retail Price
Product Energy	Numeric	Operating and Standby Watts, Annual Energy Consumption (from a single data source)
Product Certification	Code, Date	Certifying Authority, Current Test Procedure, Registration Date
Product Images	String	Links to Images of the Product (multiple)
Product Marketing	String	Product Title, Headline, and General Description

Product Category Definition

While the Global Category attributes capture the common data across all products, it is also necessary to capture attributes specific to each product category, since the data required for TVs will vary substantially from that of Room ACs or Refrigerators. These category-specific attributes are the descriptors that bridge regional differences across markets. Below are examples of the TV and Room AC category definitions.

Table 3: TV Data Standard

Attribute	Type	Notes
Screen Technology	String	LCD, LED, Plasma, Projection, CRT
Functions Available	String	Analog, Digital
Screen Size	Numeric	Size in inches, centimeters, meters, or other unit of measurement
Resolution	String	HD or Conventional
Aspect Ratio of Screen	String	4:3, 16:10, 16:9
3D Ready	String	Yes, No

Table 4: Room AC Data Standard

Attribute	Type	Notes
Cooling Capacity	Numeric	BTU, kW, W
Refrigeration Method	String	Electrically driven, Absorption units
Condenser	String	Air cooled, Water cooled

Cooling		
Heat Transfer Fluid	String	Air, Water
Type	String	Unitary, Split units, Multi-split
Cooling/heating	String	Cooling, Heating, Cooling/heating
Variable speed drive/multi-speed compressor	String	Present or Not
Refrigerant	String	Refrigerant Type

2.4. Customization by Market

While the standard sets guidelines for which attributes to capture, it still allows for flexibility when implementing for a specific market. Each market using the framework will have the capability to add new attributes to an existing category or define a new category for products that are not currently covered. These extensions cannot be compared across markets.

Modifying a Category Definition

For example, some countries may want to classify types of Room ACs in more granularity and therefore there may need a 'SubType' attribute added to their Room AC category definition. Additionally, there may be a need to set up an attribute to capture a numeric rating that is assigned by a country's certification program.

New Category Definitions

It is also possible to define an entirely new category that is not currently covered by the data standard. This new category would automatically inherit the base attributes (such as MPN, UPC, or Brand) and would only need to have defined the specific attributes unique to that product category and market.

As more and more countries begin to adopt the data standard, the result will be that the majority of categories will have been standardized across all markets.

2.5. Methodology for Populating the Database Framework

To make the adoption of the data standard as seamless as possible, a comprehensive approach for populating the framework from existing data sources is detailed below. An application layer is required to normalize the disparate data across markets into the new standard.

A system using the framework would create unique products from online shopping APIs such as Google, Amazon, and EBay by pulling in product IDs, marketing details, technical specifications, and prices. Access to Amazon, EBay, and other international online shopping APIs is provided for all organizations that agree to be part of that retailer's affiliate network.

It would then match energy efficiency data based on MPN or GTIN via CSV uploads of certification data sets. An unlimited number of feeds from retailer APIs and certification data sets can be utilized in the framework to ensure strong coverage of each market.

Online Shopping APIs

Online shopping feeds from the retailers with the broadest product coverage will be utilized for creating a list of all current products on the market. One or many APIs per market will be accessed on a daily basis to pull the most up to date products and pricing data. The methodology for doing this as follows:

- Data feed retrieval: establish a connection with the retailer API to select which products should be returned by the data feed.

- Normalization: map the raw data in the retailer data feed to the definitions of the data standard. Data rules will need to be applied to common factor MPN, Brand, and product attributes.

- Aggregation: combine product specifications and prices (Min, Max, Mean, Median, and Standard Deviation of daily offers) from multiple retailer feeds to create and maintain a single, unique product record in the framework.

- Matching: update offers and product specifications on already existing products based on MPN or GTIN. Matching MPNs requires logic to strip extraneous characters and process the many wildcards present in certification records.

Data rules

Rules can be defined per retailer API to control for the variation in data quality coming from online data sources. The power of this data framework is that rules can be configured to reject products based on category attributes, using logical operators such as any, in, not, >, <, >=, <=, and =.

Examples would be: TV Screen Technology in "LCD, LED" or Screen Size > "15 inches" or for Room AC Volume Capacity <= "14 kW". This ensures only data of the highest quality is captured into the framework. Additionally, it will be necessary to rewrite attribute names and values to normalize them across data sources and markets.

Certification Data Sets

Government certification/rating data will be utilized to capture products' energy efficiency information. This data will be imported into the framework as a Comma-separated Value (CSV) file. The framework will then match each record's MPN to an existing product in the database following the methodology below:

- Data set parsing: Certification data sets are uploaded in CSV format to add energy efficiency related attributes to product records.

- Matching: Capture the associated energy efficiency data such as energy consumption, certification level, certifying authority, registration date, etc. on already existing products based on a matched MPN or GTIN code.

If an MPN match is not successful there is also the capability to capture the certification record directly into the database. This will cover the use case whereby certification data sets contain historical products which are not currently available on the market.

2.6. Mapping Products Across Markets

As many manufacturers sell the same products across many markets using different MPNs and GTINs, there is a need to be able to link comparable products. The data framework includes a reference that links an MPN and/or GTIN from one market to another. The power in capturing product attributes in a standard way is to enable automated cross-market mapping.

By using logical rules to compare MPNs and key product attributes such as a product's size, capacity, or energy consumption, the framework could facilitate identification of potential matches. As testing procedures may differ across markets, the challenge will be to identify which conversion factors and thresholds to apply. With the framework linking the test method with each product's energy certification record, it would be possible to compare differences in the actual test procedure conditions (such as AC room temperature or TV viewing mode) between comparable products. Further analysis

by all countries planning to participate in the data standard would be required to identify appropriate conversion factors and thresholds.

A good example would be for Samsung TVs across the US, Europe, South Korea, and Australia. Samsung TVs: UN40EH5000F (US), UE40EH5000 (Europe), and UA40EH5006M (Australia). Notice that the first two letters of the MPN define the market: UN for the US or South Korea, UE for Europe, or UA for Australia. This is the most basic way of identifying comparable products but does not guarantee success.

The next step would be to look at attributes such as Screen Size and Annual Energy Consumption. Using logical rules to first convert these attributes to a common unit of measurement, it will be possible to apply thresholds based on the underlying testing procedures to have the framework automate the identification of potential matches across markets. By continually fine-tuning the attribute conversion and threshold algorithms, the framework can significantly cut down on the amount of manual work required to clearly establish a set of like products across markets in each category.

3. Benefits of a Data Standard

A data standard accomplishes the following primary goal – improving access to data about consumer electronics and home appliances products' availability, pricing, and performance.

3.1. Existing Data

Currently, much of the data to feed into the proposed standard is not provided in a consistent format. For example, currently 26 countries have some type of minimum performance standard or energy performance labeling program.[4] Several of these have multiple programs, and may have different government or non-governmental agencies responsible for administering the program.

Each agency maintains their own qualification system, verifying that products comply with the requirements of their program. Often, these agencies publish lists of qualified products. For instance, the US ENERGY STAR program publishes a list of qualified products on their website.[5] This list contains 37 attributes about each product, such as manufacturer, manufacturer's product number, energy rating, size, and ENERGY STAR model identification number. Meanwhile, the Australian government collects similar information for its categorical labeling program, including 17 fields. Aligning these data formats to make reasonable comparisons across economies requires working directly with the governments that collect this data, as IEA 4E does in its mapping and benchmarking studies.

However, this certification data is further limited in that it doesn't necessarily represent the range of products available in a market, and it is not connected to price information. For instance, data from an endorsement label naturally represents the higher performing products on the market; it would not be possible to use this data to illustrate trends in efficiency of products in that market. However, each day vendors of this equipment are offering this equipment for sale – with associated identifying information such as manufacturer part number/model number (MPN), global trade identification number (GTIN) and manufacturer. By collecting these offers, we can know what fraction of the market is represented by the high performing products and what fraction of products do not have energy information available. When these products are available for sale, these products also have price information. By matching the offers for products at a given price with the energy rating provided by the rating agency, we can estimate average incremental cost for increased efficiency.

3.2. Uses

While a snapshot summary of most of this information can be done with existing tools, it requires significant undertaking on the part of the government or organization doing the research. In many cases, governments must purchase data or do their own research; the high cost of this puts it out of the reach of many organizations and often means that potential uses of the data are not cost effective. A data standard lowers the incremental cost of implementing any particular use by reducing the amount of work needed to reformat and clean the data sources.

Importantly, a data standard also opens up the possibility for continuous, real-time monitoring. Under the conventional approaches, the manual effort involved in collecting and aligning data sources means that information may lag by months or years, and only be available publicly in aggregate summary reports. While with the standard it make take effort to build an initial platform for viewing information, summaries can be updated on demand when source data is updated.

By developing a data standard, this data could be used to feed any number of applications or resources, limited only by what people are willing to design and create. We propose some potential applications of this data that could be developed.

Informing policy

Currently, it is an expensive process to collect data on market status and the availability of products. Most energy performance standards and labeling programs use cost-effectiveness as a test for setting improved target levels. This data standard could enable real-time monitoring of efficiency distribution within a market or region and changes in price among products of different efficiencies. Lowered cost of this process by automating the data collection can put this type of data collection into the reach of some governments that currently can't afford to do thorough market evaluation, and can shorten the revision process for those that already do this.

Cross-market advocacy

Currently, mapping and benchmarking studies by organizations like IEA 4E provide a way for countries to compare the performance of products in their markets. However, these are limited in how frequently they can be updated and how many products can be done. Implementation of a data standard would allow richer data to inform studies such as these, as well as broader product coverage.

Comparison of test methods

In many cases, products are available across markets and are tested to different standards. However, because of variations in test methods, direct comparison of the results is not possible – it is comparing apples to oranges. By using the data standard to identify the same product sold in multiple markets, conversion factors can be estimated.

Consumer mobile applications

Currently, consumers must rely on energy efficiency information that is presented via various types of labeling schemes – usually showing the relative efficiency of a product and an estimated annual energy cost based on average utility rates and usage profiles (i.e. hours of TV watched per day). The data standard could enable developers to build a mobile application that allows consumers to scan a product's bar code to instantly compare personalized energy cost across all products of a specific class. The application could then recommend more efficient products within a specific price range that are available to purchase at local retailers within the consumer's local area. By also including the

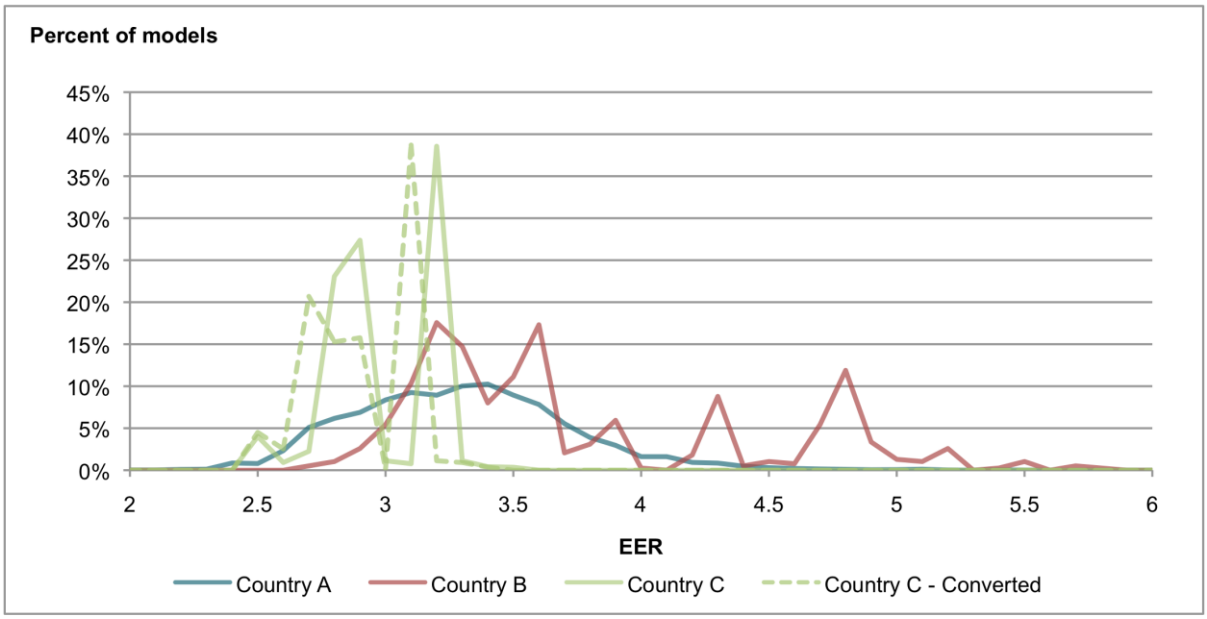
savings available from retailer incentives and utility rebates, it could incentivize consumers to purchase more energy efficient appliances.

4. Example Application

4.1. Efficiency Distribution Across Markets

As part of the development process, we collected sample data sets of room air conditioner performance from several countries. While this is the most basic comparison available, it does illustrate the first possible level of comparison. With the implementation of a data standard, this information could be made available via web portal that is up-to-date with current product information, along with data from additional countries.

Figure 2 – Distribution of Room AC EER



In this example, Country A and Country B follow the same methodology for calculating the energy efficiency ratio (EER). However, Country C uses a different methodology. Previous research assessed the differences in test procedures and demonstrated that the ISO 5151 test conditions could be estimated by dividing the NAFTA test condition EER used in Country C by 1.032. This conversion factor has been applied to the dotted green line in Figure 2. While this conversion approximation is available for EER, many other products do not have such conversions available. Applying analysis of the same products across markets, as described above, could be used to obtain additional conversions; building the test method and conversion factors into the data standards allows for improved comparison in product availability. [6] [7]

4.2. Price-Efficiency Trend Across Markets

A key additional benefit to the data framework is the ability to compare price to efficiency across markets. In the chart below, we show an illustrative example using real data to compare the average EER (in Watts Output/Watts Input), which has been normalized to W/W, and retail price level in US dollars for Room ACs across three different markets. This type of analysis could be automatically generated based on any interval (i.e. daily, weekly, monthly, quarterly) and using any price bucket (i.e. specific dollar amounts as shown below or more general divisions such as low, medium, or high). In addition, it could be filtered by size to look at the price-efficiency relationship for Room ACs of a specific cooling capacity.

Figure 3: Price-efficiency Comparison of Room ACs

5. Conclusion

This paper has presented the first steps of an attempt to develop the framework of an international consumer electronics and home appliances database and demonstrate some of its most important applications. The next stage of development will rely heavily on input from stakeholders within the SEAD initiative and elsewhere, to optimize and populate the database. Establishment of a full set of use cases will drive database design. In addition, the population of the database is in early stages. A major test of the data standard and algorithms will be its ability to associate products for offer in online shopping websites with models listed in certification database. The intended result of this project is a well-populated database that allows access to both efficiency and price data for a wide range of uses and end users.

An equally important outcome of the project will be the establishment of an international data standard for appliances. The natural follow-up to this outcome is to seek wide diffusion and stakeholder input, with the ultimate goal of seeking adherence to the standard from as many countries as possible. As argued above, the establishment and adoption of such a standard would yield rich benefits to policymakers, consumers and researchers concerned with consumer electronics and home appliances energy efficiency.

6. References

[1] Super-efficient Equipment and Appliance Deployment (SEAD) Initiative, http://www.cleanenergyministerial.org/OurWork/Initiatives/Appliances.aspx

[2] IEA 4E Mapping and Benchmarking Annex, http://mappingandbenchmarking.iea-4e.org

[3] GS1 Global Product Classification, http://www.gs1.org/gdsn/gpc

[4] CLASP Global S&L Database, http://www.clasponline.org/en/Tools/Tools/SL_Search

[5] Energy Star - Televisions, http://www.energystar.gov/index.cfm?fuseaction=find_a_product.showProductGroup&pgw_code=TV

[6] E3 Energy Rating – Compare Appliances, http://reg.energyrating.gov.au/comparator/product_types/

[7] Energy Star – Room Air Conditioners, http://www.energystar.gov/index.cfm?fuseaction=find_a_product.showProductGroup&pgw_code=AC