

**State Energy Data Needs Assessment
January 2009**

**Energy Information Administration
U.S. Department of Energy
Washington, DC 20585**

This report was prepared by the Energy Information Administration, the independent statistical and analytical agency within the Department of Energy. Unless referenced otherwise, the information contained herein should be attributed to the Energy Information Administration and should not be construed as advocating or reflecting any policy position of the Department of Energy or any other organization.

Contents

	Page
Executive Summary	ii
Chapter 1. Consultations with Customers, Stakeholders, and EIA Staff	1
Chapter 2. Integrated State Energy Data Programs	3
Chapter 3. End-Use Consumption Surveys: Data Quality and Scope	5
Chapter 4. Supplier Surveys: Data Quality and Scope	14
Chapter 5. Communication, Collaboration, and Outreach	18
 Appendices	
Appendix A. Energy Independence and Security Act of 2007 (P.L. 110-140), Sec. 805. Assessment of Resources	22
Appendix B. Geographic Reporting Levels for EIA Data	23

Executive Summary

Section 805(d) of the Energy Independence and Security Act of 2007 (EISA), Public Law 110-140, requires the Energy Information Administration to assess State-level energy data needs and submit to Congress a plan to address those needs (see Appendix A). In response, this report identifies gaps in EIA's current State-level energy data programs based on stakeholder outreach and internal deliberation, and outlines 34 possible initiatives to close those gaps.

To help users understand in ballpark terms what various enhancements to State-level data might cost, the report provides some information on the potential costs of the possible initiatives. It is not, however, intended as a budget document. In particular, it does not address the prioritization among the different possible initiatives, or their priority relative to improvements in EIA's national or regional data and analysis programs. Moreover, in many cases, the initiatives discussed in this report, such as expanding the sample size for the end-use energy consumption surveys, could be implemented on varying scales and with possible opportunities for cost-sharing, both of which could significantly impact actual implementation costs.

EIA Background. The Energy Information Administration (EIA), created by the Congress in 1977, is the statistical agency of the U.S. Department of Energy. EIA conducts a comprehensive data collection program that covers the full spectrum of energy sources and energy flows, generates short- and long-term domestic and international energy forecasts, and performs analyses on a wide range of energy topics. EIA's policy-neutral energy data and information are designed to meet the needs of Government, industry, and the public for the purpose of promoting sound policy decision-making, efficient markets, and public understanding. EIA disseminates its data products, analyses, reports, and services primarily through its Web site and telephone contact center. Major users of EIA's work products include the Congress, Federal and State government, industry, academia, financial institutions, news media and the public. By law, EIA's products are prepared independently of Administration policy considerations.

Report Organization. The report is organized as follows. Chapter 1 reviews the process EIA undertook to gather information from stakeholders and customers, consistent with the strong emphasis on consultation in EISA Section 805, and from EIA staff and managers. The process was informed by the direction in EISA Section 805(c) to ensure "the quality, comparability, and scope of State energy data, including data on energy production and consumption by product and sector and renewable and alternative sources, required to provide a comprehensive, accurate energy profile at the State level. . . ." Chapters 2 through 4 cover three major categories of gaps: the future of EIA's integrated State energy data programs, the quality and scope of EIA's end-use consumption data, and the quality and scope of data obtained from EIA's surveys of energy suppliers and markets. Each chapter describes user needs and develops initiatives to address the identified data gaps. Chapter 5 describes EIA's current efforts and proposed plans regarding communication, collaboration, and outreach to State energy data users and discusses opportunities and constraints related to the sharing of company-level energy data.

Integrated State Energy Programs. EIA currently collects and publishes data at different geographic levels depending on the type of fuel or energy source (e.g. petroleum, natural gas, coal, uranium, renewables, and electricity) and the type of data being reported, such as prices, reserves, supply, distribution, trade and marketing, consumption, sales, and end use. Only data relevant to a given State are collected and reported for that State because not all States have production plants, refineries, storage facilities, pipelines, ports, etc. Additionally, insufficient sample sizes can prohibit State-level aggregations, and some data are more easily interpreted if they are reported by regions or districts that are associated with established production, marketing, or transmission boundaries. While EIA posts State-level data throughout its Web site, EIA has two State energy data compilation programs to assist data users. The

State Energy Profiles (SEPs) is a data portal to all of EIA's State-level data and fuel-specific State energy profiles, essentially a one-stop shop for users interested in a specific State's energy picture. The *State Energy Data System (SEDS)* fulfills an additional need by providing annual *estimates* of State-level consumption, prices, and expenditures for those seeking time-series data on key energy variables. Both *SEPs* and *SEDS* (which rely essentially on EIA data) would benefit from improvements to EIA's State-level data collection programs but would also benefit from investments aimed specifically at improving the quality of *SEDS* estimates and reducing the time lag between estimation and release. *SEPs* investments could streamline data integration, add on mapping features, and improve timeliness of release.

End-Use Consumption Surveys. A wide range of stakeholders consulted prior to the writing of this report expressed strong interest in a variety of enhancements to EIA's surveys of energy consumption in end-use sectors, including larger sample sizes to increase the availability of State-level data, more frequent end-use surveys, faster processing and issuance of end-use survey data, and expansion of the end-use survey program to additional sectors, such as agriculture and household transportation energy. Options for improving the consumption survey data programs include expanding the current survey designs to cover up to all 50 States. An increase sample size and/or the addition of even a limited number of States would be a vast improvement over current programs because it would permit more complex analysis of key indicators of energy use, publication of more building types, and more accuracy for secondary uses of the data by other Federal agencies. Additional options aimed at improving end-use data and estimates focus on restoring and adding new end-use surveys, conducting feasibility studies to explore alternative methods to update end-use estimates, and decreasing the time between when data are collected and released.

Supplier Surveys: Data Quality and Scope. In addition to data gaps associated with end-use data, stakeholders identified needs associated with EIA's supplier surveys. EIA's supplier surveys cover all aspects of the energy supply picture (petroleum, natural gas, coal, renewables and electric power) and are published at national, regional, and State levels, depending on the survey, the sample, and confidentiality requirements. Stakeholders and customers indicated that EIA supplier data are important for State-level policymaking as well as for responding to emergencies and understanding markets within and affecting their States, and they emphasized the importance of maintaining quality in EIA data. Taking the feedback we received from stakeholders over the years, as well as from workshop participants and others during our 2008 outreach efforts, EIA developed several data quality options aimed at reducing statistical error in EIA surveys. The options include increasing sample sizes to provide better statistics, improving survey coverage by updating frames (the list of survey respondents), and adding new surveys to obtain data not currently collected. Current statistical data edits, nonresponse follow-up methods, and other internal EIA quality checks are covered by strict adherence to EIA statistical standards and controls.

Communication, Collaboration, and Outreach. One of the requirements of Section 805, as noted previously, is "to ensure the quality, comparability, and scope of State energy data, including data on energy production and consumption by product and sector and renewable and alternative sources, required to provide a comprehensive, accurate energy profile at the State level. . . ." EIA's stakeholders and customers note that it takes Federal initiatives to provide data that are comparable across States. Consequently, several options discussed in the report are aimed at expanding EIA's interaction with the States, as well as the U.S. territories, and include hosting conferences and workshops to assess data needs, and options aimed at using collaborative, Web-based tools to share information.

Resource Considerations. EIA's 2009 budget request of \$110,595,000 supports primary data collection, data processing and dissemination, short- and medium-term forecasting and economic and industry

analysis. Costs (both start-up and ongoing) associated with each of the 34 possible initiatives to address State data gaps are *preliminary* estimates and are presented in variety of ways depending on the proposal (all costs include Federal staff and contractor support). For example, some costs are reported as increments to annual costs as presented in EIA's 2009 budget request. This is the case for initiatives that address weekly, monthly, or annual surveys and/or initiatives that are improvements to data quality. For improvements to the end-use (consumption) surveys, which are conducted every 4 years, EIA presents preliminary cost estimates in terms of increments to the projected budget for the next upcoming consumption survey cycle. EIA's end-use consumption survey program is currently funded on a levelized basis over the 4-year cycle required to complete the current set of 3 surveys; annual incremental resource requirements after an initial transition period can be estimated as the cycle cost divided by the cycle length. For initiatives aimed at increasing the frequency of the consumption surveys (to every 2 or 3 years), the costs are per-cycle costs, noting that this also translates into larger *annual* budgets to fund shorter cycles on a levelized basis. In addition, some proposals identify needs that require coordination and resource commitments among statistical agencies. Consequently, the costs of the proposals in this report cannot easily be added together for a grand total. More accurate budget numbers, both start-up costs and per-survey cycle costs, would require more detailed assessments. Furthermore, EIA's budget is part of the broader Federal budget, and EIA's funding must be weighed against other Federal priorities.

In some cases, improvements to EIA's State energy data can be undertaken for a relatively small cost, particularly improvements associated with supplier survey data quality, State data integration programs, and outreach. Expanding the consumption survey programs to include more or all States would be much more costly. For example, EIA currently projects that the 2009 RECS, the 2010 MECS, and 2011 CBECS will cost roughly \$28,000,000 to complete (or about \$7,000,000 on an annual basis). Simply expanding the RECS by 50 percent more sample units would cost about \$3,000,000 more per 4-year RECS cycle. Similarly, increasing the sample size for the quadrennial CBECS by 50 percent would cost about \$7,000,000 more than what EIA currently projects for the 2011 CBECS.

Compiling this report has helped EIA look carefully at gaps in its State-level data collection efforts and develop options to address them. As it is unlikely that EIA would receive the budget to support all the initiatives in this report, it is important for EIA, along with its stakeholders, to consider State energy data gaps within the context of gaps in EIA's other data and analysis programs to be sure funds are directed at the most critical needs. It is equally important to seriously consider options to the expansion or improvement of existing State energy programs. The inclusion of several feasibility studies in this report recognizes the value of fleshing out some of these options.

An old management adage that still holds true today in a variety of energy and other contexts states that you can't manage what you don't measure. Given the current interest in energy and energy efficiency issues at the State level, as well as at the Federal level, it is not surprising that there is considerable interest in more, and more timely, State-level energy data. EIA can respond to this interest in several ways. First, EIA can continue to seek to operate as efficiently as possible to allow the use of a portion of our existing resources to address new priorities, including State-level data. Second, EIA can propose reallocating resources away from existing EIA activities towards work that is judged to have a higher priority. EIA has done this on several occasions over the past 5 years, and in many cases has met strong resistance by users (and/or their representatives) of the existing data proposed for elimination. A third option would be to seek additional resources through the budget process. Decision makers in the Administration and the Congress could then consider our request within the context of two overarching questions. First, what overall level of resources should be devoted to the energy mission of the

Department of Energy, including, but not limited to EIA? Second, what portion of overall energy mission resources should be devoted to energy data programs at EIA?

Chapter 1. Consultations with Customers, Stakeholders, and EIA Staff

Stakeholder Outreach

EIA's first step in responding to the requirements of section 805(d) was to form the State Data Assessment Team consisting of six EIA analysts who conducted the necessary research, gathered input from stakeholders¹, developed options to fill data gaps, and compiled this report. Over the past 15 years, the number of people and organizations seeking information from EIA has increased significantly. Additionally, their interests in and needs for different types of energy information have constantly changed and expanded. EIA's formal stakeholder outreach related to the preparation of this report began with a presentation by the Deputy EIA Administrator to the National Conference of State Legislatures at its July 2008 Legislative Summit in New Orleans. During September and October, the State Data Assessment Team consulted with additional stakeholders. Team members held meetings and telephone conferences with representatives of the National Association of State Energy Officials (NASEO) and the Northeast-Midwest Institute. They attended NASEO's Annual Meeting in September 2008 in Overland Park, Kansas, where they met informally with State officials and gave a presentation that explained EIA's work on Section 805. NASEO members expressed their need to for greater accuracy in EIA's data, more access to company-level data, and more data that would help States track progress towards meeting energy-related goals. For example, two of the goals established by the Environmental Protection Agency's (EPA) *National Action Plan for Energy Efficiency Vision for 2025*² require the establishment of mechanisms for measuring and evaluating a State's progress in achieving greater energy efficiency. States are requesting State-level data on energy consumption that would assist them in measuring their progress toward reaching this goal.

In October 2008, EIA held a workshop inviting attendees to share their perspectives on the scope, accuracy, quality, comparability, and timeliness of EIA's State energy data. Announcements about the workshop were sent to two of EIA's extensive list serves (State Energy Data and Energy Consumption) and to the NASEO Winter Fuels Conference mailing list. Recipients of the announcements represented a wide range of organizations, including the:

- Federal Energy Regulatory Commission
- Environmental Protection Agency
- U.S. Departments of Transportation, Commerce, Labor, Treasury, and Health and Human Services
- State Energy Officials (in all 50 States and the U.S. Territories)

¹ The U.S. Government Accountability Office defines a "stakeholder" as an individual or group with an interest in the success of an organization in delivering intended results and maintaining the viability of the organization's products and services. Stakeholders influence programs, products, and services. Examples include congressional members and staff, representatives of central management and oversight entities such as the Office of Management and Budget; and representatives of key interest groups, including the organization's customers and interested members of the public. EIA's major customer groups include: Federal, State, and local governments; academic and research communities; businesses and industry; foreign governments and international organizations; news media; financial institutions; and the general public.

² National Action Plan for Energy Efficiency (2008). *National Action Plan for Energy Efficiency Vision for 2025: A Framework for Change*. <www.epa.gov/eeactionplan>

- National Conference of State Legislatures
- National Governors Association
- National Association of Regulatory Utility Commissioners
- American Council for an Energy Efficient Economy, the Alliance to Save Energy, the National Association of Home Builders, and
- ExxonMobil, Amerigas, and other energy companies.

Over 80 people registered for the workshop, which was attended by representatives of the Federal and several State governments, including the California Energy Commission and the New York State Energy Research and Development Authority (NYSERDA), non-profit organizations, and private companies. Several workshop participants also provided written comments. Workshop topics included the availability of measures of renewable energy consumption; electricity grid reliability; data on distributed generation, plug-in vehicles, and other new technologies; end-use data on appliances such as gas clothes dryers; and fuel prices for home heating, among many other topics.

EIA In-House Assessments

The State Data Assessment Team initially reviewed EIA's State Data Directory http://tonto.eia.doe.gov/state/SEP_MorePrices.cfm, which helped serve as a guide when considering data gaps. Because EIA's analysts are continually responding to customers, they are uniquely aware of EIA's data gaps, so the Team then canvassed EIA management and analysts to identify gaps and propose options to address the gaps. EIA staff addressed topics such as survey data quality and coverage, State-level data estimation procedures, and information dissemination. Lastly, the Team reviewed public and Congressional requests for State-level information received by EIA during the first 5 months of 2008.

Results of the Assessment

After analyzing all of the feedback received from external and internal stakeholders, the State Data Assessment Team categorized State data needs into four main categories:

- 1) "Integrated State Energy Data Programs," which covers improvements to the dissemination of data and estimates via the *State Energy Profiles (SEPs)* and the *State Energy Data System (SEDS)*;
- 2) "End-Use Consumption Surveys: Data Quality and Scope," which covers data needs to increase the comprehensiveness of EIA's existing and proposed end-use consumption data collection activities;
- 3) "Supplier Surveys: Data Quality and Scope," which covers data needs associated with increasing the comprehensiveness of EIA's existing supply-side surveys; and
- 4) "Communication, Collaboration, and Outreach," which covers actions EIA could undertake to consult and collaborate with stakeholders.

The specific energy data needs identified by EIA and options for addressing them are covered in the next four chapters of this report.

Chapter 2. Integrated State Energy Data Programs

EIA receives comments and questions about State-level data every day. Some users are looking for comprehensive datasets, such as total energy consumption for all 50 States (and U.S. territories) for the past 10 years. Other users are looking for quick access to a comprehensive profile of their State's current energy situation. All users expect to find timely, technically accurate, and complete State-level data on the EIA web site. One participant in EIA's October 2008 State Energy Workshop told EIA, "We need the data for use in making policy. Energy data makes the markets work." Another participant said, "We rely on EIA's data—it's absolutely critical. It's the only source that's reliable State to State."

EIA has two State energy data compilation programs (referred to as "integrated energy" programs) that cover all of the major sources and uses of energy: the *State Energy Profiles (SEPs)* and the *State Energy Data System (SEDS)*. Both programs seek to address the needs of a wide variety of data users. The SEPs program was launched in December 2006 as an innovative web application to improve user access to and understanding of EIA's State-level data. It is a continually updated web-based portal to all of EIA's State energy data; it also includes individual profiles that present key facts and statistics about State energy markets and industries. Because SEPs is a compilation program, it can only include available data. Expanding the scope of SEPs to include, for example, more State-level or U.S. territory end-use data, would require expansion in EIA's data collection programs, discussed in other chapters of this report.

The *SEDS* has provided annual *estimates* of State-level consumption, prices, and expenditures since 1978. In the past few years, an intensive effort has led to significant improvements in timeliness and scope. EIA has begun posting updated estimates for individual energy sources as soon as they are processed. In addition, the time lag from the close of the data year to the release date of *SEDS* integrated totals (all energy sources) has been shortened from 34 months in October 2006 to 23 months in November 2008, with an ultimate goal of an 18-month time lag in the current system. In addition, *SEDS* began providing State-level production estimates in 2008 for all energy sources and States for 1960 forward.

In 2009, EIA will more closely align the programs to gain efficiencies and to provide users with an up-to-date, comprehensive, and accurate State Energy Profile for each State as quickly as possible after the data become available.

Improved Integrated State Energy Data Products

- *Improve Timeliness of the SEDS.* The *SEDS* provides annual estimates of consumption, prices, expenditures, and production for all energy sources and States. Many stakeholders, including the Northeast-Midwest Coalition and the National Association of State Energy Officials, have expressed concern about the timeliness of *SEDS* data, which are used as inputs for forecasting and for the estimation of greenhouse gas emissions. State administrative and energy officials also use the data for planning and analysis. The timeliness of *SEDS* depends on the availability of annual data from various EIA program offices and external sources. It would be difficult to improve timeliness to less than 18 months after the data year with the current schedule of data inputs into *SEDS*, but it would be possible to use estimation procedures to compute a set of *preliminary* *SEDS* estimates by using preliminary data from monthly and quarterly surveys. Using model-based estimates would allow the first set of *SEDS* data (for example, coal consumption) to be released 6 months after the end of the data year, and total energy estimates to be released in 15 months.

Initiative 2.1. Start-up Cost: \$300,000; Annual Operating Cost \$300,000 (increment over EIA's 2009 budget request).

- *Improve Quality of the SEDS.* In order to evaluate total energy use in the SEDS, a complete set of State-level consumption and price data for all energy sources must be compiled. There are gaps, however, in the data inputs needed to calculate total energy use. In some cases, source data come solely from non-EIA sources, and EIA has no control over the timeliness or quality of those exogenous data sources. For example, EIA uses asphalt and road oil consumption data from the Asphalt Institute and liquefied petroleum gas by end-use sector from the American Petroleum Institute. EIA has no control over the data quality, data collection, or publication schedule for those data. In other cases, there are no data available from any source. For example, there are no consumption data for lubricants used in the industrial or transportation sectors. When data are not available from EIA or non-EIA sources, they must be estimated. If the missing data could be filled in with EIA survey data, the quality of the SEDS data would be improved. It is unclear, however, what resources would be required to obtain all of the missing data. Therefore, EIA could conduct research to define the level of effort required to obtain the missing data. *Initiative 2.2. Start-up Cost: \$130,000; Annual Operating Cost: \$130,000 (increment over EIA's 2009 budget).*
- *Create a State Data Application Programming Interface (API).* Upgrading EIA's databases would also make it possible to implement a State Data Application Programming Interface (API), a computer model that makes it easy to access and exchange EIA's State energy data. An API would allow Federal and State agencies, financial markets, research institutions, analysts, and others to quickly access EIA State data and process it into new services beyond the scope and resources of EIA. Such an interface would allow EIA to give customers direct access to EIA data for a variety of purposes. EIA would be able to deliver State energy data in multiple formats in the timeliest manner possible. An API is also the first step to developing a global energy navigator—a database management tool that links multiple databases through a single interface. *Initiative 2.3. Start-up Cost: \$500,000; Annual Operating Cost: \$194,000 (increment over EIA's 2009 budget request).*
- *Develop System to Support User-Generated State Energy Maps.* Computer system upgrades would help make it possible for users to create their own State energy data maps. Congressional staffers, members of the press, and energy analysts routinely call EIA to request high-resolution copies of EIA's State energy maps, which display energy infrastructure sites such as electric power plants, transmission lines, natural gas pipeline flow and hubs, oil ports and refineries, and renewable energy potential at a glance. The maps are unique in that they display geospatial data for several different major energy sources. Currently, EIA's mapping environment does not support the specific geospatial analysis needs of some users. EIA is frequently asked but unable to provide maps showing crude oil and petroleum product pipelines. The addition of biomass potential and more detailed data on solar, wind, and geothermal potential are other common requests. A more robust computing environment would make it possible to include the large amounts of data required to display detailed oil pipelines, energy potentials, and other energy map elements. The analytical value of the maps would also be enhanced if they were published as a system that would allow users to generate their own maps, adding and subtracting map layers and creating customized, high-resolution maps suitable for analysis and publication. *Initiative 2.4. Start-up Cost: \$330,000; Annual Operating Cost: \$150,000 (increment over EIA's 2009 budget request).*

Chapter 3. End-Use Consumption Surveys: Data Quality and Scope

EIA's end-use consumption surveys were among the first fielded by EIA after its creation in 1977 and now cover energy end uses in three areas: the Commercial Buildings Energy Consumption Survey (CBECS) covers the commercial buildings sector, the Residential Energy Consumption Survey (RECS) covers the occupied housing portion of the residential housing sector, and the Manufacturing Energy Consumption Survey (MECS) covers the manufacturing share of the industrial sector.³ These resource-intensive surveys are the only source of data for current estimates of energy end uses in homes and commercial buildings and for key industries within manufacturing. Due to limited resources, they are not currently meeting their statutory requirements for frequency and scope⁴.

Two features limit the geographic level at which consumption data can be provided to the public: 1) sample sizes, which are driven by available resources and 2) data confidentiality laws, which require that EIA protect the identity of individual respondents and establishments. In practice, where firms make up a large share of their industry class or building type, they could be identified at lower levels of aggregation. However, EIA would be prohibited from publishing some statistics at the State level or would need to collapse disparate classes of data to protect its confidentiality. The same legislation, however, does allow certified agents of EIA to make limited use of non-public files for statistical purposes only.

Given current resources and confidentiality constraints, EIA is limited to publishing end-use data for the following geographic areas⁵:

- MECS: National and Census Region
- CBECS: National, Census Region, and Census Division
- RECS: National, Census Region, Census Division, and the four most populous States (California, Florida, New York, and Texas)

Stakeholder Needs

Stakeholders expect EIA to take the lead in providing energy consumption data that meet the quality and scope necessary to monitor topics related to climate, the environment, and energy security and they often request data to help evaluate energy programs and policies that are often written, funded, and implemented at the State level. They need more and new consumption data at lower levels of geography, more frequently, and with less lag time between the period of data collection and the release date. A Federal statistical program that can assess the value of a dollar invested in a particular program, technology, or system is a much broader and more complex data operation than EIA has ever run. Such a program would have profound resource implications for EIA.

Stakeholders report they need data for geographic areas at and/or below the State level—counties, metropolitan areas, or cities—to tie outcomes to specific programs. Stakeholders provide the following

³ Although transportation represents about a third of domestic energy consumption, and residential transportation consumption about two-thirds of that, budget shortfalls have prevented EIA from collecting data on transportation for the past 15 years.

⁴ Public Law 92-275 (Federal Energy Administration Act of 1974) and Public Law 95-91 (Department of Energy Organization Act).

⁵ See Appendix B for a brief discussion on EIA geographic reporting levels.

arguments, among others, for larger sample sizes to improve data quality and to provide new estimates for smaller geographic aggregations:

- EPA reports that it needs a larger sample size for CBECS to produce energy performance benchmarks for more building types. EPA offers an online rating system called Portfolio Manager where commercial buildings can be rated for their energy consumption relative to similar buildings. This tool has been used for more than 78,000 buildings; about 5,600 have achieved an EnergyStar rating. Although this rating system has become the industry standard, EPA can only produce benchmarks for 10 broad building types. An increase in the CBECS sample size would be needed to produce them for more diverse building types and principal activities.
- The American Society of Heating, Refrigerating and Air Conditioning Engineers (ASHRAE) states that the CBECS current sample size is too small to evaluate the impact of critical building labeling programs like EnergyStar by building type *within* climate zones. They also believe that States and cities rely on CBECS information to develop rating systems for “comparable buildings.” Because data at that level are imprecise or absent, ASHRAE is concerned about the accuracy, utility, and impact of the ratings.
- The National Renewable Energy Laboratory (NREL) reports it needs much larger RECS and CBECS sample sizes to perform necessary multivariate analyses to understand the adoption rates and impact of the new technologies, building designs, and energy-efficient equipment they test and promote.
- The National Association of Home Builders (NAHB) indicates it needs larger samples to estimate how much consumption is explained by householder behavior versus that which the builder can control. Although NAHB makes no direct request for State-level estimates, building codes are an important feature in that analysis because they are enacted at the State and local level.
- The American Council for an Energy-Efficient Economy (ACEEE) produces an annual State Energy Efficiency Scorecard, which ranks States according to their adoption and implementation of energy efficiency policies and programs. Without sufficient State-level sample sizes in EIA’s consumption surveys, there is no accurate, direct link between State policies and consumer participation in these programs, which represent large, non-Federal program investments.
- The Department of Housing and Urban Development (HUD) and its local housing authorities use RECS data to calculate utility allowances for subsidized housing by States and localities. State-level estimates would provide a direct source of data closer to what is needed for program management.
- Where utility costs are bundled with rent by landlords, the Bureau of Labor Statistics (BLS) uses RECS data to allocate portions of total rent to housing or utilities. Rents and utility costs vary widely for smaller areas, so BLS would benefit from lower levels of data aggregation.
- The Department of Health and Human Services (HHS) uses RECS to support program needs of the Low Income Home Energy Assistance Program (LIHEAP). Improved State-level estimates would provide marked improvement in allocating funds in this multi-billion dollar grant program to States.

- ACEEE notes that because EIA summarizes consumption data by Census Division, EIA necessarily aggregates statistics in the Mountain Division across a huge area—from Montana on the Canadian border to Arizona on the Mexican border. Program evaluations are not served by combining data across such diverse States and climates. Without additional sampling, EIA could increase breakouts in the Mountain Division only by decreasing the sample in the Pacific Division, which would reduce the accuracy of current estimates for California.
- Stakeholders want EIA to balance sample allocations between new and existing building stock for the RECS and CBECS to improve the accuracy of comparisons among and between buildings of different ages. Although Federal or industry programs offer explicit guidance on energy-efficient building design and practices, program participation is often voluntary. This situation is less true for State and local governments where codes can be tied to building permits. Therefore, State-level data would provide more accurate analysis of the sources of differences in energy consumption for new construction.
- Numerous officials from a wide array of State program offices, consulting firms, utility companies, and equipment manufacturers contact EIA directly for data that are not available without large increases in survey sample sizes: State-level energy intensities within consumption sectors including breakouts by industry, building type or activity, fuel, end use, and equipment type; and ‘typical’ intensities for specific buildings types, industry codes, or residential classes.

There is a recurring dilemma when trying to optimize sample allocations across such overlapping geographic variables as physical geography and climate zones. A good sample can achieve sufficient coverage by States or by climate zones but not both without increasing the likelihood that some buildings or establishments could be individually identified. For reasons of data confidentiality, EIA will probably need to continue to aggregate some data to mask the location of particular buildings.

Options for Improving Consumption Data Surveys

The consumption surveys draw samples from large, heterogeneous populations, and the survey methodologies necessary to produce accurate results mirror that complexity. Therefore, for RECS and CBECS, EIA samples, lists, and enumerates units selected from area clusters at rates that have to balance sampling error across key characteristics: geography, building size and type or activity, main fuel used for space heating, and so on. Area-probability field studies are the most expensive but provide the best frame coverage for complex populations, the highest item and unit response rates, and the ability to collect key data in physical form (e.g., household square footage measurements and interviewer observations of buildings characteristics). MECS is considerably less expensive per sampled unit because it is drawn from a known list of establishments with routine updates maintained by the Census Bureau and utilized for other economic data collections, and can be conducted by mail and the Internet.

Options for improving the consumption survey data programs are described below. The first options expand the current survey designs to improve data quality without State-level estimates, other than for the four most populous States in the RECS. Programs with expanded survey designs would be a vast improvement over current programs because they would permit more complex analysis of key indicators of energy use, publication of more building types, and more accuracy for secondary uses of the data by other Federal agencies. Short of a 50-State estimates program (or a 50-States and U.S. territories program), also described below, an intermediate option within each of these designs is to add some but not all States. Proposals are also included that describe data collection needed to produce baseline

measures for 50 States, although such a program would require considerably more resources. Additional options focus on restoring and adding new end-use surveys, conducting feasibility studies to explore alternative methods to update end-use baselines, and decreasing the time between data collection and release.

For each initiative or proposal, EIA estimates preliminary costs, which are presented in different ways depending on the initiative. Costs are presented either in terms of increments to EIA's 2009 budget request (for wholly new initiatives); increments to the projected budget for an upcoming survey cycle (in the case of initiatives that increase an ongoing survey's sample size or data scope to improve data quality); or relative per-cycle costs (for initiatives that address frequency of surveys). More accurate budget numbers (both start-up costs and per-survey cycle costs) require knowing, *inter alia*, which States/Regions would be added, the homogeneity of key characteristics within them, and the level of accuracy desired.

Enhance Data Quality by Adding Data Elements and Increasing Sample

There are many ways in which EIA could improve the value and accuracy of consumption surveys. By far, the most important is to improve the accuracy of end-use estimates. Managing the sample sizes and allocations among key characteristics is equally important as adding questions to capture new phenomena that effect consumption patterns and levels. Changes in sample coverage below address two stakeholder needs: 1) produce sufficient sample size to be able to perform needed multivariate analyses (i.e., having the statistical power to understand the relative contribution of three or more characteristics) and 2) produce accurate State-level energy intensities and end-use estimates with which States can monitor and evaluate energy programs.

- *Obtain More Geographic Detail on Fuel and Nonfuel Uses of Fossil Fuels for Manufacturing (MECS).* Doubling the quadrennial MECS sample size to 31,000 sample units would allow EIA to provide estimates for manufacturing energy consumption for industry groups by Census Region, improve the statistical accuracy of national analyses, estimate energy efficiency in this sector, and calculate changes in carbon emissions over time that result from structural change. This initiative would serve essential missions of many Federal, State and industry energy, environmental and commercial interests. *Initiative 3.1. Cost per 4-year Cycle: \$2,800,000 (increment over EIA's projected cost for 2010 MECS).*
- *Enhance the Quality of the RECS.* Increase the quadrennial RECS by 50 percent to 9,750 sample units and add questions to improve the accuracy of multivariate data analysis and end-use estimates. For example, collect data on: the share of remodeling that is done to incorporate energy efficient equipment, systems, and designs; the degree of compliance with building energy codes (and which version or source) for new construction; and building operation, because behavior can often explain more variation in consumption than can technology and equipment. Conduct a periodic sub-metering study on a subsample of RECS households to measure actual energy use by refrigerators, hot water heaters, televisions, and computers and their peripherals. Compare actual consumption by end use and behavior to estimates based on non-linear models developed by EIA staff. Such a comparison would help EIA adjust for bias that might occur in an infrequent survey that lags in accounting for new technology, standards, and growing plug loads. *Initiative 3.2. Cost per 4-year Cycle: \$3,160,000 (increment over EIA's projected cost for 2009 RECS).*

- *Enhance the Quality of the CBECS.* Increase the sample size for the quadrennial CBECS by 50 percent to 17,250 sample cases (which include buildings and establishments) and target specific building types that are big energy users, such as data centers, laboratories, convention centers, and arenas. Add questions to improve the accuracy of multivariate data analysis and end-use estimates. For example, EIA could collect new data on: the degree of compliance with building energy codes (and which version or source) for new construction; building operation, to isolate energy management practices from fixed factors; and building type, to explain more variation in energy consumption, e.g., linear shelf feet rather than number of refrigerators for groceries, size of the eating area for food service, volume of transactions/sales for retail, number of beds or rooms for dorms/hotels/hospitals, and number of service bays for auto repair shops. EIA could conduct a quality study on a subsample of the CBECS buildings in detail and compare the end-use energy use of these buildings to the estimates developed by EIA non-linear models. EIA could validate the information collected during the field interviews and also provide an independent estimate of the energy consumption by end use. This information could lead to improvements in future CBECS questionnaires, as well as measuring the quality of current CBECS estimates.
Initiative 3.3. Cost per 4-year Cycle: \$6,880,000 (increment over EIA's projected cost for 2011 CBECS).

Restore and Add New Surveys to Fill Data Gaps

In addition to increasing the geographic scope of consumption surveys to provide a neutral source of data for Federal, State, and local energy policy, EIA could increase coverage across consumption sectors to improve forecasts of short- and long-term energy demand. Data gaps are sometimes addressed by increasing sample sizes to permit publication of more subclasses, or by adding questionnaire items or surveys. To increase the coverage and accuracy of the amount and sources of growing energy demand, particularly for electricity and petroleum-based fuels, EIA would need to restore and add data collections for transportation and agriculture. The transportation sector is going through a dynamic era of technological, fuel, and industry change in response to energy prices and global economic and climate concerns. Measuring transportation energy consumption poorly, infrequently, or not at all has implications across all sectors of the economy. With the emergence of biofuels, energy and food policy are now more closely linked.

- *Collect Data on End Uses of Energy by the Residential Transportation Sector.* Restore the Residential Transportation Energy Consumption Survey (RTECS), which was discontinued after the 1994 data year due to insufficient funds. EIA could expand the current RECS to include a follow-on study of residential transportation end uses, which account for two-thirds of the entire transportation sector and most of motor gasoline consumption. Many stakeholders value a reliable, policy-neutral source of data to understand on-road fuel economy, price elasticities, vehicle miles traveled, commuting behavior, and vehicle purchases relative to new energy policies and technologies. *Initiative 3.4. Start-up Cost: \$1,000,000; Cost per 4-year Cycle: \$4,048,000 (increment over EIA's 2009 budget request).*
- *Collect Data on the End Uses of Energy by the Non-Residential (Truck) Transportation Sector.* To increase coverage of the transportation sector, EIA could sponsor, in part, the collection of data on the physical and operational characteristics of the Nation's private and commercial truck populations. Until 2002, the Census Bureau conducted a truck survey with its quinquennial economic census. A similar truck survey, with additional questions about fuels used, energy end-uses, and costs would be beneficial. The series would produce national and State-level estimates

of the total number of trucks and their end uses. Expanding the survey to the State level could require data coordination and cost-sharing with Federal, State, and local agencies. These data could serve missions of Federal energy and transportation agencies by providing a comprehensive data set for assessing energy efficiency and the environmental impact of the Nation's truck fleet. Cost estimates only assume EIA's role adding fuel-related questions. *Initiative 3.5. \$3,000,000 (increment over EIA's 2009 budget request).*

- *Collect Data on End Uses for the Agricultural Sector.* EIA currently collects data on industrial sector energy end uses only for the manufacturing portion. Adding an agricultural survey would improve the industry coverage in EIA's State Energy Profiles. For example, fuel rates vary significantly by scale and type of operation, but EIA now assumes that commercial rates prevail. An agricultural series would provide the only baseline to: measure opportunities for new energy-efficient technologies and practices; allow a more uniform evaluation of the impact on production agriculture of Federal and State energy policies, such as fuel tax abatements, efficiency incentives, and alternative fuel use; and produce estimates of the share of greenhouse gases resulting from different enterprises, production practices, and technologies. Two agencies would be served: the Department of Agriculture (USDA) could measure the farm-level response to changes in energy prices and supply (food security), and the Department of Energy could study the flow of crops into biofuels compared with other uses (energy security). Cost estimates only assume EIA's role in a new agricultural survey. *Initiative 3.6. Start-up Cost: \$200,000; Cost per 4-year Cycle: \$1,500,000 (increment over EIA's 2009 budget request).*

Produce End-Use Data for All 50 States

Developing and operating a 50-State consumption data program (or a 50-State and U.S. territories data program) would require significantly more resources than EIA's current program and significantly more than any of the proposals discussed so far. For example, a preliminary estimate for adding States to RECS would be \$750,000 for each additional State. For the CBECS, each additional State would add about \$1,200,000 to the total survey budget. State selection criteria would vary according to measurement goals, accuracy and confidentiality requirements, and costs relative to sampling efficiencies for other States. For example, if States were selected according to population or building rank, coverage would quickly increase for one Division in the Midwest⁶ (Illinois, Ohio, and Michigan in the East North Central Division) and one in the South (Georgia and North Carolina in the South Atlantic Division). Population or the number of buildings as the main selection criteria would yield no data improvements for the Mountain or West North Central Division States and might not meet the goals of States with aggressive energy policies, stakeholders such as the ACEEE, and many Federal agencies with little or no data to monitor programs. The criteria for adding States would need to be carefully developed, with stakeholders' input, and clearly communicated.

- *Provide Residential Energy Consumption Data for All 50 States (maintaining current 4-year cycle).* Increasing the RECS sample to cover 50 States would ensure that new (often State) policies can be monitored for their impact on fuel type used, intensities, and end uses. Accuracy would improve for national, regional, and division-level estimates so that analysts could isolate the effect of such factors as new efficiency standards, building technologies, and program participation from factors over which consumers have no control, such as weather. EIA would have sufficient sample counts to produce estimates for small appliances and home electronics—a growing portion of residential

⁶ See Appendix B for a map of the U.S. Census Regions and Divisions.

consumption, a source of greenhouse gases, and an opportunity for technology change and innovation. Expanding the program to include all States would serve essential missions of many Federal, State, and energy industry, environmental, and commercial interests. *Initiative 3.7. Start-up Cost: up to \$8,000,000; Cost per 4-year Cycle: up to \$26,460,000 (incremental cost over EIA's projected cost for 2009 RECS).*

- *Provide Residential Transportation Data for All 50 States (maintaining current 4-year cycle).* If residential transportation data were deemed necessary to collect via a follow-on survey to the 50-State RECS design, it would require additional funds beyond funds required for the 50-State RECS. *Initiative 3.8. Start-up Cost: \$1,000,000; Cost per 4-year Cycle: \$15,660,000 (incremental cost over EIA's projected cost for 50-State RECS).*
- *Use Current Population Survey to Collect State-Level Residential Transportation Data.* Some State-level residential transportation data could be collected by adding questions to the Current Population Survey (CPS), such as the vehicle identification number and the current odometer reading for all vehicles held by members of the sampled household. The CPS, which is conducted by the Bureau of the Census for the Bureau of Labor Statistics, is a large monthly survey that would yield about 15,000 interviews per month. Estimates would permit State-level estimates of fuel use by month for large States and 6- or 12-month averages for smaller States. The data could potentially be combined with other social variables, such as household characteristics, employment, and income. *Initiative 3.9. Start-up Cost to EIA: up to \$4,000,000 (over 3 years); Cost per Annual Cycle: up to \$900,000 (incremental cost relative to EIA's 2009 budget request).*
- *Provide CBECS Data for All 50 States (maintaining current 4-year cycle).* Increasing the CBECS sample would allow EIA to publish State estimates for some major building types and dramatically increase the number of types that could be published at the national, regional, and division level. A larger sample would also allow EIA to undertake energy efficiency analysis; calculate changes in carbon emissions over time; and monitor the adoption of new building design, equipment technologies, and energy management tools and practices. *Initiative 3.10. Start-up Cost: up to \$13,000,000; Cost per 4-year Cycle: up to \$42,640,000 (incremental cost over EIA's projected cost for 2011 CBECS).*

Increase Frequency of Energy End-Use Data

- *Collect End-Use Data More Frequently (maintaining current sample sizes).* Note that per-cycle costs are roughly the same regardless of whether the cycle is 2, 3 or 4 years but shorting the current cycle time would require additional resources on a per-year basis.
 - *Conduct the MECS biennially.* With the current 4-year cycle, EIA cannot accurately describe how energy consumption in manufacturing relates to changing energy market conditions, to the cost and availability of capital for investment in new technologies and energy management practices, and to structural shifts in demand for its products. Because energy intensities and fuel-switching capacities vary considerably between industry classes, so would their response capacity. Features such as the lag time between a market signal and demand response, whether it is temporary or permanent, and the relative impact by industry sector are not measureable or are missed by infrequent data collection. Enormous changes have occurred in the structure of American industry, where labor and the supply and cost of energy are key factors in global competitiveness. There is a significant value to energy policymakers,

the Bureau of Economic Analysis, the Federal Reserve Board, and the Bureau of Labor Statistics in having a clearer understanding of the flow of energy, capital and labor. MECS is conducted primarily via the Internet, so the marginal burden to respondents and cost to EIA of updating this series biennially would be minimized. A more frequent MECS could integrate policy-relevant topics into the survey in a more timely manner. *Initiative 3.11. Cost per 2-year Cycle: \$3,987,700.*

- *Conduct CBECS, RECS, and RTECS triennially.* Adhering to a 3-year cycle would ensure that EIA could best account for factors that are important in forecasts of national consumption, and in identifying trends and structural shifts caused by changes in policy, technology, and behavior. EIA would realize some efficiencies in survey management with more frequent data collection, more staff in more specialized roles, and new resources to improve data, management and release processes. *Initiative 3.12 (CBECS). Cost per 3-year Cycle: \$13,763,000. Initiative 3.13 (RECS): Cost per 3-year Cycle: \$9,551,000. Initiative 3.14 (RTECS): Cost per 3-year Cycle: \$4,740,000.*

Increase Timeliness of End-Use Data Releases

- *Improve Timeliness of Data Releases.* To increase the availability of end-use data, EIA would need to reduce the time lag between data collection and data release in part by making better use of the Internet for data release. In the past 10 years, three major computing changes occurred while budget (in real terms) and staff resources declined: EIA migrated from a centrally controlled mainframe to a distributed local-area network processing environment, data collection moved from paper-administered forms to computer-assisted interviews or Internet data collection instruments, and reports moved from physical publications to electronic media on the Web. The processing environment requires more skilled coordination, management, and documentation. The data collection technology requires more pre-survey preparation, which backs into the previous survey cycle. Web publications require ongoing, specialized support functions to meet user needs and expectations. Although EIA has begun work to improve performance on all these fronts, additional resources could help attain consistency and efficiency across the consumption surveys. *Initiative 3.15. Start-up Cost: \$500,000; Cost per 2, 3 or 4-year Cycle: \$800,000 (increment over EIA's 2009 budget request).*

Conduct Feasibility Studies for Alternative Sources of End-Use Data

- *Feasibility Studies for Alternative Sources of Data for End-Use Estimates.* Evaluate alternative Federal, State, and commercial data sources for meeting the statutory requirements of EIA's end-use program under budgetary constraints. Identify other surveys and administrative records that may prove valuable for benchmarking, modeling, or filling data gaps in the program. Evaluate the methods used and quality produced relative to EIA's needs. Identify the potential for interagency data and cost sharing, collaborative data collections, and value-added analyses to meet the challenges of scarce Federal resources, economic disruptions from the global recession, climate change, and health or national security events. *Initiative 3.16. Annual Operating Cost: up to \$500,000 (increment over EIA's 2009 budget request).*

Additional Challenges to Improving Energy Consumption Data

Where there are gaps in the scope, frequency, timing, and/or publication of EIA consumption data, users increasingly resort to *ad hoc* means to address them. For example, EPA is assisting businesses, user

groups, and trade associations in collecting their own consumption data to augment the CBECS program. In other situations, States are independently collecting data to produce more localized benchmarks. These efforts are vulnerable to funding cuts, which would likely cause quality challenges similar to or worse than EIA's. Furthermore, EIA is subject to OMB's statistical standards, whereas State data collections are not. Although resources vary considerably, States are relying more on smaller data collection firms and less expensive, less accurate modes of data collection.

In another example, the Department of Housing and Urban Development produces a Utility Schedule Model using RECS, the basis for calculating utility allowances for various Federal programs. Because RECS is conducted infrequently, users are left to make their own idiosyncratic or no adjustments to the model for non-RECS years. Adjusting a model in rapidly changing energy markets is beyond the scope, resources and expertise of most of these users. As a result, EIA analysts receive direct requests from metropolitan housing authorities (or their consultants) for inter-survey estimates for geographic areas smaller than in the current RECS. Trade associations for commercial buildings and for manufacturing firms that are trying to work around Federal data gaps also request *ad hoc* advice and support. Such requests suggest that the methods used to define and evaluate efficiency and other energy program targets are idiosyncratic and will diverge until EIA can provide more and better consumption data for smaller geographical divisions and analytic subclasses.

Stakeholders and other data users report that the demand for high-quality State-level data will continue to grow. EIA terms, definitions, and survey methods then become the *de facto* basis for benchmarking and assessing energy policies led or enacted by States, as well as by other Federal statistical agencies. Improving the consumption survey program would add considerable value and coherence to data that are central to policy and other decision makers.

Chapter 4. Supplier Surveys: Data Quality and Scope

EIA's supplier surveys cover all stages of the energy supply picture. For crude oil and petroleum products, the program covers reserves, production, transportation, refining, blending, imports, exports, and storage. Much of the data collected is reported at the national, regional, and State levels. For natural gas, EIA collects and publishes data on the operation of natural gas markets at the national and State levels to provide information about natural gas supplies (from domestic and foreign sources), natural gas movement and storage, natural gas consumption by major end-use sectors, and prices along the path from wellhead to end user. EIA surveys cover the electric power industry by collecting data at the individual facility level and made available at the national, State, sector, and facility level (subject to confidentiality requirements) and are designed to capture data from traditional utilities as well as emerging participants (e.g., power marketers, and other entities engaged in the production, sales, or distribution of electricity). EIA also collects extensive data on the coal industry and on the nuclear power and the uranium mining and milling industries. EIA collects and disseminates a limited amount of data from the manufacturers of solar thermal collectors, solar photovoltaic cells and modules, and geothermal heat pump equipment. These data along with information collected by EIA on its electric power surveys and the MECS provide statistics on the growth in renewable energy in the United States, assess the effectiveness of incentive programs, and provide a basis for projections of renewable energy consumption. Finally, EIA collects and disseminates data on the production of alternative-fueled vehicles and hybrid-electric vehicles by suppliers, and on the use and fuel consumption of alternative-fueled vehicles by Federal, State, and fuel- provider fleets, and local transit companies.

Many stakeholders and customers have said that EIA supplier survey data are important for State-level policymaking, as well as for responding to emergencies and understanding markets within and affecting their States. Stakeholders and customers emphasize the importance of EIA's ability to maintain quality in its data. They have pointed out room for improvement in EIA's data, mentioning issues such as occasional discrepancies between weekly and monthly data and gaps in what EIA is currently able to collect and publish. Taking the feedback received from stakeholders over the years, as well as from workshop participants and others during our 2008 outreach efforts, EIA has developed several data quality options aimed at reducing statistical error in EIA surveys. Statistical error is inherent in all survey data, regardless of their source and the care and competence of data collectors. There are various potential sources of statistical error, such as the following:

- **Sampling.** A data collection may be based on a sample of the population rather than on a complete enumeration. Variation occurs by chance because a particular sample is surveyed. This variation decreases with increased sample size. A preferred sample size is determined in order to keep key statistics under a specified threshold of sampling error, but sometimes the actual sample size is solely determined by cost.
- **Coverage.** The frame (the list of those surveyed) may not be complete or may contain companies that are out of scope or no longer in business. Complete and up-to-date frames are important in reducing statistical error due to lack of coverage in EIA surveys. Frame improvements are particularly important in the deregulated electricity sector, which has undergone a tremendous number of mergers, acquisitions, divestitures, and company name changes over the past several years.
- **Target population.** Reporting thresholds can be used to determine a survey frame and thereby exclude smaller companies from reporting. Exclusions result in a frame that does not completely

cover the target population of companies of all sizes. Also, a frame may be incomplete and not cover the parts of the target population that are more difficult or more expensive to reach.

- **Respondent.** Respondents may commit errors in reporting their data. Respondents frequently report statistics in units other than those requested by EIA. For example, respondents will report data in gallons rather than in barrels.
- **Nonresponse.** Not all of the units that are surveyed respond (unit non-response), and some respondents may not provide all of the information requested (item non-response).
- **Processing.** Errors may occur from transcribing data incorrectly.
- **Concept.** The data collection elements may not measure the items that they were intended to measure, or the estimation methodology may provide inaccurate results.

Most of the options discussed below focus on the first three sources of survey error identified above, either by increasing sample sizes to provide better statistics or by improving survey coverage by updating frames. Other quality initiatives include improving the scope of several EIA surveys by increasing the sample size to obtain State-level estimates for all 50 States and better estimates for current statistics, or by adding new surveys. Current statistical data edits, nonresponse follow-up methods, and other internal EIA quality checks are covered by strict adherence to EIA statistical standards and control the last four sources of errors; therefore, no initiatives are proposed for any of these sources of error at this time.

Upgrade Frames

A comprehensive frame is essential to conducting a statistically valid survey. Frames-related initiatives include:

- *Develop Frame for Weekly Retail On-Highway Diesel Price Survey (EIA-888).* Construct an outlet frame of truck stops and service stations to improve the quality of EIA's weekly diesel price data. EIA is currently unable to provide average prices for more detailed geographic areas than national, Petroleum Administration for Defense (PAD) Districts, sub-PAD regions, and California. *Initiative 4.1. Start-up Cost: \$250,000; Annual Operating Cost: \$50,000 (increment to EIA's 2009 budget request).*
- *Update Frame for Weekly Retail Gasoline Price Survey (EIA-878).* The frame of motor gasoline outlets was last updated in 2001. Currently, EIA provides select State and city prices in addition to national, PADD, and sub-PADD average prices. The updated frame could be used to select a new, larger sample that could produce State-level data for all States. *Initiative 4.2. Start-up Cost: \$260,000; Annual Operating Cost: \$50,000 (increment to EIA's 2009 budget request).*

Expand Scope

Additional initiatives would expand the number of data elements on current EIA survey forms or would increase the sample size in order to obtain more accurate State-level data⁷. Others would increase the

⁷ Expanding sample sizes to include some or all U.S. territories would require additional funding for each initiative.

sample size in order to reach additional States. Still others propose new surveys to collect information on new energy data elements, including renewable energy. This group of initiatives includes the following:

- *Expand Natural Gas Production Survey (EIA-914)* to collect data from all natural gas-producing States and to add crude oil and lease condensate production from all oil-producing States. Expanding the survey would promote improved and timelier data on production trends nationally and particularly for a number of States where oil and gas production is rapidly changing. *Initiative 4.3. Start-up Cost: \$400,000; Annual Operating Cost: \$325,000 (increment to EIA's 2009 budget request).*
- *Expand Scope of the Monthly Gas Marketer Survey (EIA-910)* to obtain data on sales by marketers for more States. The current scope of the survey includes only the 12 States that offer natural gas retail choice programs with the highest concentration of natural gas sold by marketers. Currently, there are 22 States with residential sales by marketers and 49 States with commercial and industrial sales by marketers. *Initiative 4.4. Start-up Cost: \$600,000; Annual Operating Cost: \$500,000 (incremental to EIA's 2009 budget request).*
- *Expand Power Plant Operations Survey (EIA-923)* to collect fossil-fuel receipts from smaller electric plants (those with capacities between 1 megawatt and 50 megawatts). An expanded survey would complete the picture of the electric power industry and match the reporting threshold by which fuel consumption and electricity generation data are collected. *Initiative 4.5. Start-up Cost: \$50,000; Annual Operating Cost: \$65,000 (incremental to EIA's 2009 budget request).*
- *Expand Winter Heating Fuels Telephone Survey (EIA-877)* from 24 to all 50 States and from its current data collection period of October through mid-March to a year-round survey. This survey collects weekly data on retail prices of No. 2 heating oil and propane. These data are used to assess hardships experienced by heating oil and propane users during periods of critical short supplies. The current survey is a cooperative data collection effort between EIA and 24 States. (New initiatives for joint data collection are discussed in Chapter 5.) *Initiative 4.6 Start-up Cost: \$750,000; Annual Operating Cost: \$750,000 (incremental to EIA's 2009 budget request).*
- *Develop and Launch Weekly State Inventory Survey for Selected Petroleum Products* similar to EIA's current monthly collection of inventory data for selected petroleum products such as distillate fuel oil, motor gasoline, kerosene, residual fuel oil, and propane. EIA would use the weekly data to better monitor any pending supply issues. *Initiative 4.7. Start-up Cost: \$1,000,000; Annual Operating Cost: \$500,000 (incremental to EIA's 2009 budget request).*
- *Develop and Launch New Survey to Collect Biomass Energy Consumption* using a frame developed by EPA of all forest-product facilities. This survey would permit EIA to publish State-level biomass data. *Initiative 4.8. Start-up Cost: \$190,000; Annual Operating Cost: \$105,000 (incremental to EIA's 2009 budget request).*
- *Add Methane Data Elements to Existing Survey (EIA-895A), "Annual Quantity and Value of Natural Gas Production Report,"* which currently collects vented and flared natural gas as a single value, despite the fact that vented natural gas is a methane emission and flared natural gas results in carbon dioxide emissions. The distinction is important because methane is a more potent greenhouse gas than carbon dioxide. This initiative would require respondents to provide separate values for vented and flared gas and would lead to improved accuracy of State-level

greenhouse gas emissions. *Initiative 4.9 Start-up Cost: \$70,000; Annual Operating Cost: \$70,000 (incremental to EIA's 2009 budget request).*

Chapter 5. Communication, Collaboration, and Outreach

One of the requirements of Section 805 is to establish guidelines to “ensure the quality, comparability, and scope of State energy data, including data on energy production and consumption by product and sector and renewable and alternative sources, required to provide a comprehensive, accurate energy profile at the State level. . . .” EIA’s stakeholders and customers note that it takes Federal initiatives to provide data that are comparable across States. Section 805 also instructs EIA to “assess any existing gaps” in EIA’s data and to consult with State officials and others on a regular basis.

As a recent independent study noted, “EIA has a wide variety of customers and a long history of collecting feedback in order to better understand the needs of the public and to improve the quality of its products.”⁸ Current consulting and collaborating activities include the following:

- *Consultations with Other Government Agencies.* EIA consults with other State and Federal agencies. For example, EIA consults with the Federal Energy Regulatory Commission (FERC) on a regular basis concerning new and existing survey forms, particularly in conjunction with seeking the necessary periodic approval from the Office of Management and Budget for EIA survey forms.
- *Participation with State Organizations.* One way EIA stays informed about changes in energy markets that affect States in particular is through participation in workshops and conferences with groups such as the National Governors Association, the National Conference of State Legislatures, the National Association of Regulatory Utility Commissioners, and others. The National Association of State Energy Officials (NASEO) is a good example. EIA staff regularly attend NASEO conferences and co-sponsor conferences and workshops with NASEO every year. EIA’s communications with NASEO are intended to assess changes in interest for State energy data, to identify emerging energy issues and trends among the States, and to share solutions to State energy market or distribution problems. EIA staff often provide presentations or suggest and recruit energy industry conference speakers at NASEO’s request.
- *Joint Data Collection.* One of EIA’s best known collaborations with the States is the State Heating Oil and Propane Program (SHOPP), a cooperative data collection program between State energy officials and EIA that operates during the winter months of October through March. EIA provides funding to the States through grants to cover 50 percent of their costs for participating in the weekly heating oil and propane price survey. EIA selects the sample for each State, edits and verifies the incoming data, aggregates the data, and disseminates the data on the EIA Web site. The States make weekly phone calls to the companies in their sample and transmit the prices to EIA via an Internet data collection system. SHOPP is important in terms of outreach. It provides EIA with a contact person in each State Energy Office, enabling a communication network with the participating State Energy Offices. The cooperative program promotes the States’ ability to monitor market conditions for heating oil and propane during the winter season.

⁸ *Challenges, Choices, Changes: An External Study of the Energy Information Administration*, May 2006, p. 21.

In addition to continuing activities such as those described above, EIA could expand its interaction with the States through the following initiatives, which also could involve the U.S. territories, as feasible and appropriate:

- *Develop a Plan To Evaluate and Address Emerging State and U.S. Territory Data Needs.* New data needs related to emerging State and territorial energy issues are of particular concern to them. In order to collect and publish data that are responsive to those needs and also comparable across States and territories, EIA proposes to initiate an ongoing plan. In each cycle, the first step would be to identify emerging issues and to evaluate the data needs through iterative interactions with State and territorial officials and others. The second step would be to interview EIA analysts and to research current data collections. The third and final step would be to write recommendations about the ways in which EIA could address the needs by expanding the scope and coverage of existing surveys, by adding new surveys, or by other means. EIA's recommendations to fill the gaps could help ensure that current, comparable datasets are available for State- and territorial-level energy analysis. *Initiative 5.1. Start-up Cost: \$130,000; Annual Operating Cost: \$65,000 (increment to EIA's 2009 budget request).*
- *Conduct Biannual State and U.S. Territory Energy Data Workshops.* EIA could conduct workshops to educate users of State and territorial energy data about current and new EIA initiatives to enhance the relevancy, scope, quality, comparability, and timeliness of EIA's data, as well as to seek feedback. The workshops would be aimed at a broad audience. Each workshop would result in a record of written and oral comments from workshop participants and proposed follow-up actions which, if funded, would address the data needs identified in the workshop. *Initiative 5.2. Start-up Cost: \$65,000; Annual Operating Cost: \$65,000 (increment to EIA's 2009 budget request).*
- *Consult with Other Government Agencies on State and U.S. Territory Data Needs.* EIA could regularly consult with Federal and State agencies and organizations about the relevancy, scope, quality, comparability, and timeliness of EIA's data in order to gain information about State and territorial data needs. EIA would write annual reports describing the data needs and recommending changes in EIA's data collection activities which, if funded, would meet the needs. *Initiative 5.3. Start-up Cost: \$130,000; Annual Operating Cost: \$130,000 (increment to EIA's 2009 budget request).*
- *Investigate Options for Supporting Independent State and U.S. Territory Data Collection.* Currently, EIA publishes a set of statistical and methodological guidelines on EIA's Web site. One option for cooperative data collection between EIA and the States and territories would be for EIA to work with them to assess their data needs, plan an approach to collecting the data, and develop detailed, individualized guidance on how to collect the targeted data. With that approach, State and Territory data collection efforts could mesh with EIA's, yielding comparable datasets. *Initiative 5.4. Start-up Cost: \$130,000; Annual Operating Cost: \$130,000 (increment to EIA's 2009 budget request).*
- *Establish a "Wiki" to Allow for Collaboration in a Common Workspace.* EIA understands the mutual benefits of good communications with the States and other stakeholders who are interested in State- and territory-level information, and EIA communicates with stakeholders in a variety of ways. Some electronic media currently in use are "one-way," in that EIA posts content to its Web site or sends list-serve notices about new products. EIA also uses interactive approaches, such as those that allow users to create customized graphs and tables. There are also

options for two-way electronic communications, including the use of wikis (Web sites, or parts of a Web site, using software that allows for content to be edited and revised by all users.) The use of a wiki would allow EIA's data users to collaborate in a common workspace on data projects. The main benefit of a wiki would be quick and efficient communications during collaborations and a public record of data projects. *Initiative 5.5. Start-up Cost: \$50,000; Annual Operating Cost: \$45,000 (increment to EIA's 2009 budget request).*

Sharing of Company-Level Data

Section 805 directs the Administrator to "share company-level data collected at the State level with each State involved, in a manner consistent with the legal authorities, confidentiality protections, and stated uses in effect at the time the data were collected, subject to the condition that the State shall agree to reasonable requirements for use of the data, as the Administrator may require. . . ." Information collected by EIA falls into three groups with respect to how it is shared:

- *Confidential* information is collected under the pledge of confidentiality pursuant to the *Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA)*.
- *Protected* information is collected under a pledge that EIA will protect information to the extent it satisfies exemption 4 of section 522(b) of the Freedom of Information Act (FOIA). Title 15 Section 796(d) of the United States Code also contains a provision that requires EIA to protect certain types of information that, if made public, would reveal methods or processes that are entitled to protection as trade secrets. Broadly stated, information is considered *protected* by EIA if the release of respondent-identifiable information would be likely to cause harm to the respondent's competitive position or impair EIA's ability to obtain similar data in the future. EIA staff and contractors are subject to the penalties (imprisonment, fines and termination) cited in Title 18 Section 1905 of the United States Code for disclosing trade secret information or confidential statistical data in any manner not authorized by law.
- *Public Information* is information that is collected with the notice to respondents that the information may be publicly released in company- or individually-identifiable form.

Respondent-level confidential and protected data may be shared under certain conditions as outlined in CIPSEA and other authorizing legislation pertinent to EIA, including the DOE Organization Act, the Federal Energy Administration Act, and the Energy Supply and Environmental Coordination Act of 1974. Requests for data sharing are reviewed on a case-by-case basis. EIA formalizes the sharing of confidential and protected data with a Data Sharing Agreement signed by EIA's Administrator. Several factors are evaluated when a request is received, including: 1) the legal authority controlling the sharing of requested data; 2) the resources involved in providing the requested data, such as if any special file formatting or re-programming is required; 3) what direct identifiers, if any, need to be removed from the data file before its release; 4) the duration of the sharing activity, i.e., is it one-time or on-going; 5) can the agency requesting the data appropriately safeguard it; and 6) will any tabular data satisfy the requesting agency's data needs or is respondent-level data necessary for its use and purpose for accessing the information.

Confidential data, or data that is collected under CIPSEA, can only be shared for statistical purposes and CIPSEA attaches stringent data protection and eligibility requirements. CIPSEA allows EIA to designate other agents, besides contractors, who may have access to identifiable information collected under CIPSEA as long as the agent's access is properly controlled by EIA. These agents, such as researchers,

function under the agency umbrella, i.e., under the supervision of EIA employees through contracts or special agreements and subject to the same limitations and penalties that are described in CIPSEA. The fact that CIPSEA allows EIA to share information collected under CIPSEA with approved agents does not create a researcher's right to such information. If EIA chooses to designate an agent, it still remains responsible for all confidential information protected under CIPSEA. EIA must approve the researcher and the researcher must agree in writing to comply with all terms and conditions required by EIA prior to any access.

Protected data can be shared for official use by other DOE components, other Federal agencies, the General Accounting Office, and any Committee of the Congress. Official use of EIA data by other Federal agencies may include both statistical and non-statistical uses. A court of competent jurisdiction may obtain protected information in response to a court order. With respect to the States, Title 15 Section 779(a)(3) of the United States Code requires the Administrator to provide, upon request, to State governments all relevant information he/she possesses concerning the status and impact of energy shortages and the extent and location of available supplies of crude oil, petroleum products, natural gas, and coal, within the distribution area serving that particular State. EIA has, in fact, worked closely with the States over many years to share relevant information in its possession concerning the status and impact of energy shortages. However, because Section 779(a)(3) explicitly states that such information is to be provided in accordance with other provisions of this Code chapter, which includes Section 796(d) cited previously, the sharing of protected information remains subject to confidentiality requirements.

As noted, existing laws also permit the sharing of protected or confidential information for statistical purposes related to the EIA mission provided that the entity seeking access to the data has the ability to protect it from disclosure. EIA's implementation of data-sharing agreements provides State entities or agencies with an avenue to access company-level data for statistical purposes.

Appendix A. Legislation

ENERGY INDEPENDENCE AND SECURITY ACT OF 2007 (P.L. 110-140)

Sec. 805. ASSESSMENT OF RESOURCES

(a) 5-Year Plan-

(1) ESTABLISHMENT- The Administrator of the Energy Information Administration (referred to in this section as the 'Administrator') shall establish a 5-year plan to enhance the quality and scope of the data collection necessary to ensure the scope, accuracy, and timeliness of the information needed for efficient functioning of energy markets and related financial operations.

(2) REQUIREMENT- In establishing the plan under paragraph (1), the Administrator shall pay particular attention to—

- (A) data series terminated because of budget constraints;
- (B) data on demand response;
- (C) timely data series of State-level information;
- (D) improvements in the area of oil and gas data;
- (E) improvements in data on solid byproducts from coal-based energy-producing facilities; and
- (F) the ability to meet applicable deadlines under Federal law (including regulations) to provide data required by Congress.

(b) Submission to Congress- The Administrator shall submit to Congress the plan established under subsection (a), including a description of any improvements needed to enhance the ability of the Administrator to collect and process energy information in a manner consistent with the needs of energy markets.

(c) Guidelines-

(1) IN GENERAL- The Administrator shall--

- (A) establish guidelines to ensure the quality, comparability, and scope of State energy data, including data on energy production and consumption by product and sector and renewable and alternative sources, required to provide a comprehensive, accurate energy profile at the State level;
- (B) share company-level data collected at the State level with each State involved, in a manner consistent with the legal authorities, confidentiality protections, and Stated uses in effect at the time the data were collected, subject to the condition that the State shall agree to reasonable requirements for use of the data, as the Administrator may require;
- (C) assess any existing gaps in data obtained and compiled by the Energy Information Administration; and
- (D) evaluate the most cost-effective ways to address any data quality and quantity issues in conjunction with State officials.

(2) CONSULTATION- The Administrator shall consult with State officials and the Federal Energy Regulatory Commission on a regular basis in--

- (A) establishing guidelines and determining the scope of State-level data under paragraph (1); and
- (B) exploring ways to address data needs and serve data uses.

(d) Assessment of State Data Needs- Not later than 1 year after the date of enactment of this Act, the Administrator shall submit to Congress an assessment of State-level data needs, including a plan to address the needs.

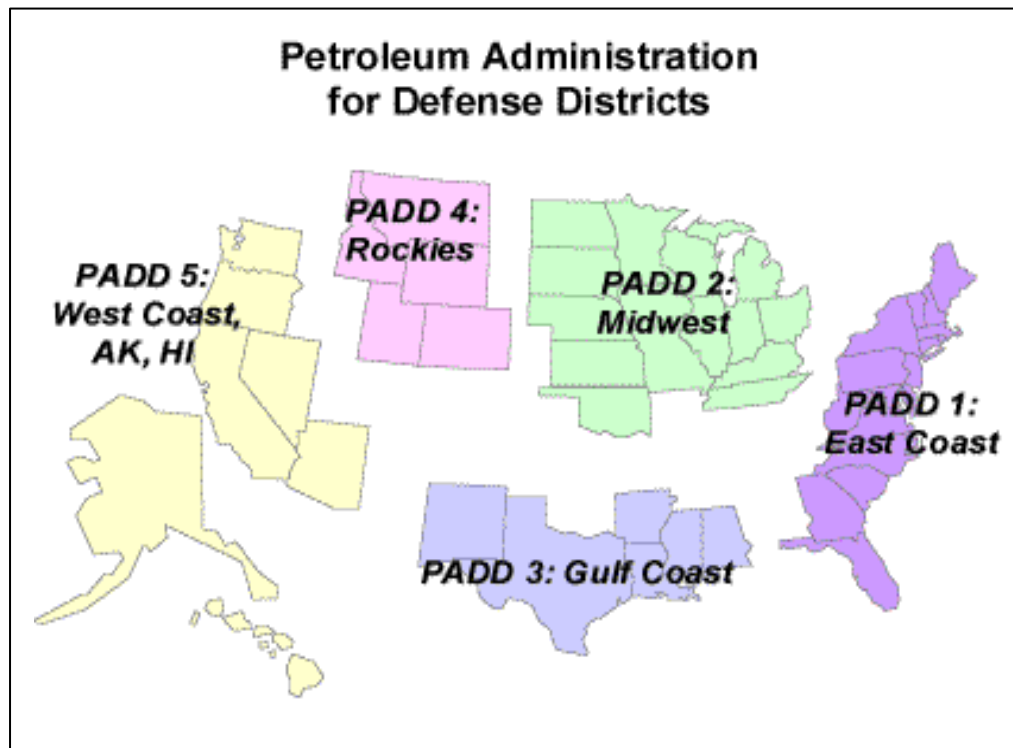
(e) Authorization of Appropriations- In addition to any other amounts made available to the Administrator, there are authorized to be appropriated to the Administrator to carry out this section--

- (1) \$10,000,000 for fiscal year 2008;
- (2) \$10,000,000 for fiscal year 2009;
- (3) \$10,000,000 for fiscal year 2010;
- (4) \$15,000,000 for fiscal year 2011;
- (5) \$20,000,000 for fiscal year 2012; and
- (6) such sums as are necessary for subsequent fiscal years.

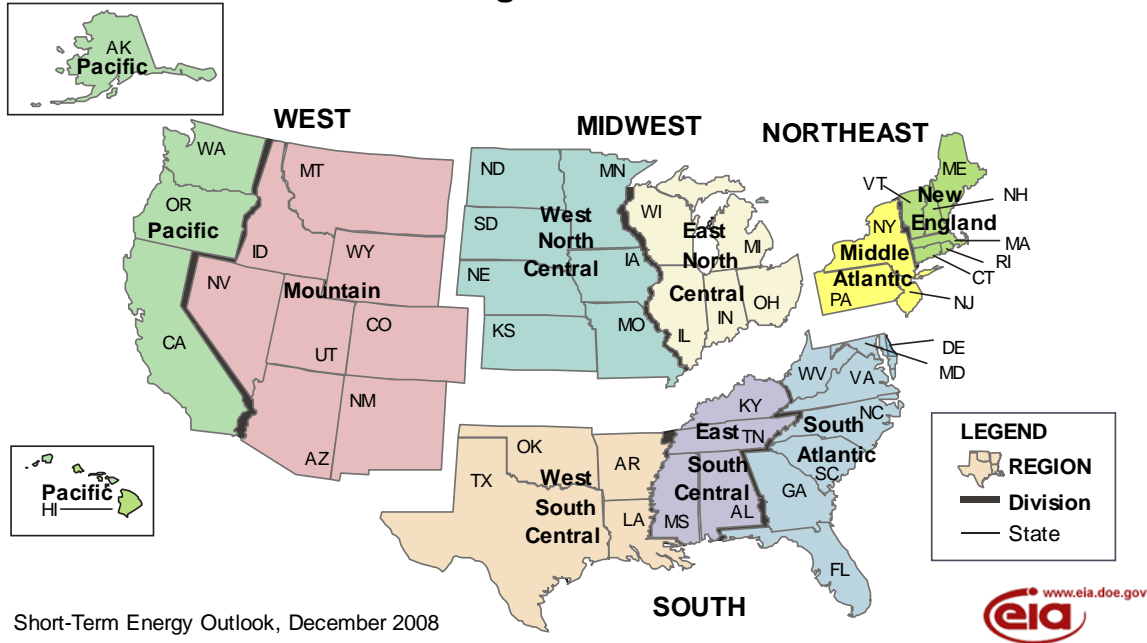
Appendix B. Geographic Reporting Levels for EIA Data

EIA currently collects and publishes data at different geographic levels depending on the type of fuel or energy source (e.g. petroleum, natural gas, coal, uranium, renewables, and electricity) and the type of data being reported, such as prices, reserves, supply, distribution, trade and marketing, consumption, sales, and end use. Only data relevant to a given State are collected and reported for that State because not all States have production plants, refineries, storage facilities, pipelines, ports, etc. Additionally, insufficient sample sizes can prohibit State-level aggregations, as discussed in Chapter 5, and some data are more easily interpreted if they are reported by regions or districts that are associated with established production, marketing, or transmission boundaries. Some data are available by both State and by region, other data are available by region alone. For example:

- Much of EIA's petroleum data are reported at the five Petroleum Administration for Defense Districts (PADD) in addition to U.S and State levels (see map below).
- EIA's Energy Consumption Surveys are reported primarily at the Census Divisions and Regions level (see map below).
- EIA's *Electric Power Annual* reports summer and winter data at the NERC (North American Electric Reliability Council) region level (see map below).

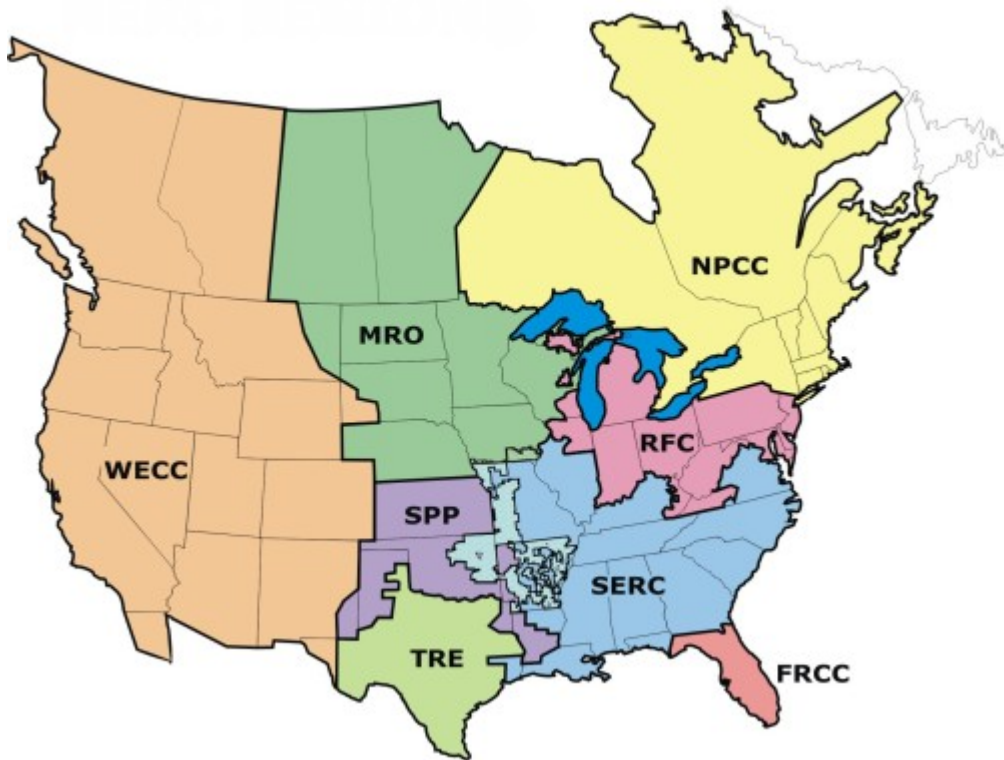


U.S. Census Regions and Census Divisions



Short-Term Energy Outlook, December 2008

North American Electric Reliability Corporation (NERC) Regions



[FRCC](#) - Florida Reliability Coordinating Council [SERC](#) - SERC Reliability Corporation
[MRO](#) - Midwest Reliability Organization [SPP](#) - Southwest Power Pool, RE
[NPCC](#) - Northeast Power Coordinating Council [TRE](#) - Texas Regional Entity
[RFC](#) - ReliabilityFirst Corporation [WECC](#) - Western Electricity Coordinating Council