



Independent Statistics & Analysis

**U.S. Energy Information
Administration**

Assistant Administrator for Energy Statistics | Office of Electricity, Renewables & Uranium Statistics

Supporting Statement for Survey Clearance: Electric Power and Renewable Surveys

Part B: Collection of Information Employing Statistical Methods

OMB Number 1905-0129

FORM EIA-63B, Photovoltaic Module Shipments Report

FORM EIA-411, Coordinated Bulk Power Supply Program Report

FORM EIA-826, Monthly Electric Utility Sales and Revenue Report with State Distributions (discontinued form to be replaced by Form EIA-861M)

FORM EIA-860, Annual Electric Generator Report

FORM EIA-860M, Monthly Update to the Annual Electric Generator Report

FORM EIA-861, Annual Electric Power Industry Report

FORM EIA-861S, Annual Electric Power Industry Report (Short Form)

FORM EIA-861M, Monthly Electric Power Industry Report (replaces EIA-826)

FORM EIA-923, Power Plant Operations Report

FORM EIA-930, Hourly and Daily Balancing Authority Operations Report

April 15, 2016



TABLE OF CONTENTS

B	STATISTICAL METHODS.....	1
B.1	Respondent Universe.....	1
B.2	Statistical Methods.....	4
B.2.1	Cutoff Sampling Methods.....	5
B.2.2	Relative Standard Error as a Measure of Sample Accuracy.....	6
B.2.3	Electric Power Surveys Sampling and Estimation.....	6
B.3	Maximizing Response Rates.....	11
B.4	Test Procedures and Forms Consultations.....	13
B.5	Statistical Consultations.....	13

B STATISTICAL METHODS

B.1 Respondent Universe

Form EIA-63B, "Photovoltaic Module Shipments Report"

Form EIA-63B is a mandatory monthly survey in addition to being a mandatory annual survey of companies engaged in photovoltaic module manufacturing, shipping, importing, or exporting within the United States, its territories, and possessions. The new monthly survey, like the existing annual version, will collect information on volume and value of shipments, technical characteristics of the modules shipped, and employment data. The survey frame for the Form EIA-63B will contain approximately 76 respondents; 16 will report monthly and 60 will report annually. The monthly frame will include only 'large' producers with the intent of capturing at least 90% of peak kilowatts (kWp) shipped. Respondents reporting total shipments of at least 100,000 kWp during the previous year will be surveyed monthly, and the published monthly data will reflect only the sampled units. The annual survey will collect data from all known U.S. producers using a shorter version of the monthly form. Potential respondents will be identified from U.S. Department of Energy (DOE) and U.S. Energy Information Administration (EIA) databases, industry/manufacturer directories, and trade publications.

Form EIA-411, "Coordinated Bulk Power Supply Program Report"

The Form EIA-411 survey collects a subset of the electric power system reliability information collected by the North American Electric Reliability Corporation (NERC) in the execution of its responsibilities as the Electric Reliability Organization for the United States.¹ The data collected by the Form EIA-411 include information about regional electricity supply and demand projections for a ten-year advance period and information on the characteristics and reliability of the transmission system and supporting facilities.

The Form EIA-411 is mandatory for those entities required to report. With the exception of Schedules 7 and 8, the form is completed by each of the eight Regional Entities of NERC. Each Regional Entity compiles the responses from data furnished by utilities and other members within their Region and then provides it to NERC headquarters. NERC then compiles and edits these data and provides an extract to EIA. Schedules 7 and 8 data for each Regional Entity are provided by NERC from, respectively, its Transmission Availability Data System and Generating Availability Data System databases.

There are nine respondents to this survey, the eight NERC Regional Entities and NERC headquarters.

Form EIA-826, "Monthly Electric Utility Sales and Revenue Report with State Distributions"

EIA will discontinue this survey and replace it with the new Form EIA-861M, "Monthly Electric Power Industry Report." See the discussion below under Form EIA-861M.

Form EIA-860, "Annual Electric Generator Report" and Form EIA-860M, "Monthly Update to the Annual Electric Generator Report"

The mandatory Form EIA-860 annual survey collects data on the status of electric generating plants and associated equipment (such as environmental control systems) that are connected to the U.S. power

¹ NERC was designated as the official Electric Reliability Organization by the Federal Energy Regulatory Commission pursuant to the power system reliability provisions of the Energy Policy Act of 2005. EIA has had a long-standing relationship with NERC and its predecessor for the collection of the Form EIA-411 data.

transmission grids. The target population comprises existing and proposed electric power plants, including the following:

- A. All existing plants that have a total generator nameplate capacity (sum for generators at a single site) of 1 megawatt (MW) or greater; and where the generator(s), or the facility in which the generator(s) resides, is connected to the local or regional electric power grid and has the ability to draw power from the grid or deliver power to the grid; and
- B. All proposed plants that: 1) have an expected total generator(s) nameplate capacity of 1 MW or greater; 2) expect the generator(s), or the facility in which the generator(s) resides, to be connected to the local or regional electric power grid and expected to be able to draw power from the grid or deliver power to the grid; and 3) expect to commence commercial operation within 10 years in the case of coal, petroleum coke, nuclear, and hydroelectric (both conventional and pumped storage) units, or within 5 years for all other units.

Companies complete the form for all the plants they operate. Over 4,700 entities are required to file the Form EIA-860. They currently operate and/or propose to operate about 8,700 facilities, containing over 24,700 generators. The respondents to this survey form the basis of the EIA electric power entity frame, from which samples for other surveys are drawn. Respondents are either self-identified or identified through industry publications and data services to which EIA subscribes.

The Form EIA-860M monthly survey collects data from the Form EIA-860 population when a respondent has any of the following:

- A new generator scheduled to begin commercial operations within the next 12 months
- An existing generator scheduled for retirement within the next 12 months
- An existing generator undergoing modifications resulting in changes in capacity or other major modifications that are scheduled to be completed within 1 month

The survey respondents are generally identified through information submitted on the Form EIA-860 survey. In some cases information from other sources, such as industry publications, is also used.

Respondents are the operators of the power plants where these new generators and existing generators are located. Based on recent experience, EIA estimates that Form EIA-860M will collect data from approximately 200 respondent entities on average each month.

Form EIA-861, "Annual Electric Power Industry Report;" Form EIA-861S, "Annual Electric Power Industry Report (Short Form);" and Form EIA-861M, "Monthly Electric Power Industry Report"

The approximately 4,000 entities that constitute the universe for these three mandatory surveys are all participants in the electric power industry who are involved in the generation, transmission, or distribution of electricity in the United States and its territories. Target population members include electric utilities, wholesale power marketers (registered with the Federal Energy Regulatory Commission [FERC]), energy service providers (registered with the states), and electric power producers.

The Form EIA-861 annual survey collects a range of information related to electricity sales, customers, and demand response and energy efficiency activities. The Form EIA-861 is administered to a threshold sample, where the largest of the 3,405 entities in the survey universe (approximately 2,295 entities) are

sampled, based on a threshold of annual retail sales and other factors explained below. The remaining smaller entities, about 1,110 of the potential respondents, account for only about 1 percent of total U.S. electricity sales and rarely engage in any other activities outside of basic electricity sales and distribution.

In 2013, EIA implemented a burden reduction plan by creating a short form for the smallest 1143 utilities representing approximately between 1% - 2% of the annual national retail sales. These smaller size utilities currently report on Form EIA-861S. Respondents to Form EIA-861S report only a selected subset of data elements that are collected each year on Form EIA-861. All Form EIA-861S respondents are currently required to complete the long form in lieu of the short form once every five years to account for changes in total revenues and sales over time. Consequently, these smaller entities are required to provide only a limited amount of sales, revenue, and customer count data on Form EIA-861S and, for certain respondents, they also report data on time-based rate customers and advanced meter reading. Once every five years and in lieu of the short form, all entities that normally complete Form EIA-861S are currently required to complete Form EIA-861.

EIA disseminates monthly totals of both reported and estimated electricity sales and revenues data in its monthly publication, the *Electric Power Monthly* (EPM). The survey form EIA-826 (note: *Form EIA-861M replaces Form EIA-826 in the proposals of this ICR*) is the data collection instrument for monthly data on electricity sales and related end-use categories. Since monthly survey reports are not required by all utilities, a model-based imputation process is used to estimate for non-sampled entities and for nonresponses. Only larger utilities report monthly sales and revenues data on Form EIA-826, as well as annually on Form EIA-861. Data reported by these utilities form the bases for EIA's model-based imputations. This enables EIA to publish monthly estimates of *total* electricity sales and revenues by state, region and nation.

Smaller entities need to complete all parts of the annual Form EIA-861 on a periodic basis to maintain the accuracy of the statistical imputation procedure used to generate key estimates from data not collected on Form EIA-861S. The results from research by EIA, using historical data, showed that requiring Form EIA-861S respondents to complete Form EIA-861 less frequently, once every eight years, will result in a significant reduction in respondent burden without compromising the accuracy and quality of the model based estimates. Based on this research, EIA will extend the time frame for EIA-861S respondents to complete Form EIA-861 in lieu of Form EIA-861S from once every 5 years to once every 8 years.

The Form EIA-861S is completed by all electric utilities with annual retail sales in the prior year of 100,000 megawatt hours (MWh) or less, except when of any one of the following conditions hold:

- The respondent has retail sales of unbundled service
- The full set of data is required from the respondent to ensure that statistical estimates for a state or business sector are of acceptable quality
- The respondent instead reports in aggregate under the Tennessee Valley Authority (TVA) or WPPI Energy²
- The company is part of the sample for the monthly form.

² TVA and WPPI Energy (the latter a consortium of public power utilities in the Midwest) consolidate responses for their wholesale customers and deliver the information to EIA. Extracting a subset of utilities from the joint filings of TVA and WPPI would be administratively cumbersome and would defeat the purpose of the proposed change to the Form EIA-861 frame, which is to reduce the cost of managing the survey for both EIA and respondents.

Utilities for which any of these exceptions apply must complete the Form EIA-861 survey. Note that in any given reporting period respondents can only complete either the Form EIA-861 or the Form EIA-861S, but not both.

The monthly Form EIA-861M survey will collect data from a sample of the universe of entities that report on the Form EIA-861 and Form EIA-861S surveys. Cutoff sampling will be used to select the survey frame for the Form EIA-861M. Based on the number of entities that are currently reporting on the Form EIA-826, the Form EIA-861M frame will include approximately 620 entities. The survey respondents will be generally identified through information submitted on the annual Form EIA-861. New respondents to the Form EIA-861M survey frame update the annual Form EIA-861 census survey. All respondents to the monthly Form EIA-861M also report on the annual Form EIA-861/861S).

Form EIA-923, "Power Plant Operations Report"

The target population for this mandatory annual and monthly survey comprises all electric plants in the United States that are connected to the electric power grid and have a generating capacity of 1 MW or greater. The survey frame is established by the Form EIA-860 survey and is identical to the set of operational and standby power plants in the frame of that survey.

There are over 7,300 power plants for which data are collected through Form EIA-923. The data collected includes electric power generation, fuel consumption, fossil fuel stocks, delivered fossil fuel cost, combustion byproducts, operational cooling water data, and operational data for NO_x, SO₂, particulate matter, mercury, and acid gas control equipment. Not all respondents answer all questions on the survey instrument; for example, questions on fuel consumption are inapplicable to hydroelectric, wind, and certain other types of power plants.

A survey form is completed for each power plant in the survey. If a single entity operates several power plants, it will complete a separate form for each plant and each form is treated as a unique response for estimation and burden calculation purposes.

EIA is proposing to reduce the current monthly sample using a more efficient model-based cutoff design. It will reduce the number of monthly respondents (from approximately 2,108 to 1,333 plants). These plants will report monthly data on electricity generation, fuel consumption, and, in some cases, the cost and quality of certain fossil fuel deliveries.³ At the end of the year, most of the monthly respondents (approximately 1,064 of the 1,333 monthly plants) will also file a supplemental form that provides annual information on nonutility power sales and the operation of environmental control equipment.⁴

The approximately 6,000 plants not in the monthly sample file data annually only. This includes data on generation and fuel consumption and, where applicable, data on fuel cost and quality and environmental equipment performance.

Form EIA-930, "Hourly and Daily Balancing Authority Operations Report"

The mandatory Form EIA-930 is a census survey of hourly electric power operating data from Balancing Authorities in the contiguous United States. The data collected include:

- Hourly demand,

³ Plants report fuel cost and quality data only if they meet the following criteria: 200 MW or greater nameplate capacity for plants fueled by natural gas, petroleum coke, distillate fuel oil, and residual fuel oil, and/or have coal-fueled capacity of 50 MW or greater.

⁴ Plants file data on environmental control system performance only if the plant has a minimum of 10 MW of nameplate capacity using combustible fuels.

- Hourly next-day demand forecast,
- Hourly net generation,
- Hourly actual interchange with each interconnected Balancing Authority.

Balancing Authorities are generally either Regional Transmission Organizations or electric utilities that have transmission grid management responsibilities. There are currently 66 Balancing Authorities in the contiguous United States and they constitute the census frame for this survey.

B.2 Sampling Methodology and Estimation Procedures

Four of the surveys in this Information Collection Request use sampling methodology:

- The three surveys of electric power sales and related data: the annual Form EIA-861, the annual Form EIA-861S (Short Form), and the monthly Form EIA-826 survey, which will be discontinued and replaced by the Form EIA-861M survey.
- The Form EIA-923 monthly and annual surveys of power plant operations.

These surveys are discussed below.

B.2.1 Cutoff Sampling Methods

To reduce reporting burden for respondents and the agency's own workload, EIA utilizes cutoff samples for the survey frames of four of its electric power surveys: the monthly Forms EIA-861M (formerly the Form EIA-826) and EIA-923; and the annual Forms EIA-861, EIA-861S, and EIA-923. These cutoff samples are comprised of all units with measures of size capacity and/or production variables larger than predefined thresholds, taking into account the need for data on multiple variables of interest within survey instruments.

For the Form EIA-861 series of surveys, EIA uses a monthly version of the annual survey Form EIA-861 to collect the same data from a smaller set of industry participants. This is done to better track the change over time of electric utility sales and revenues while reducing burden both for respondents and the agency. The monthly version is the Form EIA-826 (which will be discontinued and replaced by the Form EIA-861M survey) and its respondents are drawn from the Form EIA-861 annual frame using cutoff sampling.

The cutoff sample for Form EIA-861M consists of all units with sales and revenues of electricity larger than predefined thresholds which are specific to key estimates for various publication cells but are generally upwards of 80 percent of total sales or revenues.

Cutoff sampling eliminates the monthly reporting burden for smaller industry participants. Because smaller units have been responsible for a high percentage of reporting errors historically observed for Form EIA-861, cutoff sampling may also reduce the levels of non-sampling error affecting the published estimates. Section B.2.3 below provides details about the sampling and estimation methods used for Form EIA-861M.

On the Form EIA-923 series of surveys, EIA uses a monthly version of the annual survey Form EIA-923 to collect the same data from a smaller set of industry participants. This approach allows EIA to better track the change over time of power plant operations data while reducing burden both for respondents and the agency. The monthly version is the Form EIA-923M and its respondents are drawn from the Form EIA-860, "Annual Electric Generator Report" frame using model-based cutoff sampling.

The current monthly sample design starts with a convenient core sample of power plants, calculates relative standard errors (“RSEs”—a measure of reliability of estimates), and then adds more sample units as needed in publication cells where the RSE is larger than desired. In order to reduce the number of power plants in the monthly sample from approximately 2,200 to 1,333, a new sample methodology will be implemented by first determining the minimum sample size needed to achieve a target RSE reliability constraint for each publication cell. The new cutoff sample for Form EIA-923 will consist of all units with a capacity size that is larger than predefined thresholds which are specific to key estimates for various publication cells but are generally upwards of 80 percent of total electricity generation.

Cutoff sampling eliminates the monthly reporting burden for smaller industry participants. Because smaller units have been responsible for a high percentage of reporting errors historically observed for Form EIA-923, cutoff sampling may also reduce the levels of non-sampling error affecting the published estimates. Section B.2.3 below provides details about the sampling and estimation methods used for Form EIA-923.

B.2.2 Relative Standard Error as a Measure of Sample Accuracy

EIA uses the relative standard error (RSE) measure to evaluate the reliability of an estimated statistic. Relative standard error is based on the standard deviation about the mean of a statistic in a sample. Standard deviation is a measure of the spread of data values in relation to the mean value. Standard error normalizes this measure in terms of sample size or the number of samples, and relative standard error expresses this result as a percentage of the mean.

B.2.3 Electric Power Surveys Sampling and Estimation/Imputation

Form EIA-861M

Note: This clearance is proposing to discontinue the Form EIA-826 survey and replace it with the Form EIA-861M survey. The Form EIA-861M will employ the same sampling and estimation methodology as the Form EIA-826.

The Form EIA-861M will be a monthly survey of a sample of electric power utilities and marketers that sell or deliver electric power to end users. The sample will include electric power entities providing bundled electric service and will include all power marketers.⁵ The survey will collect information on retail electric sales and revenue by operating company, end-use sector (residential, commercial, industrial, transportation, total), and state. These data will be the monthly equivalent to the corresponding annual data reported on the Form EIA-861.

The Form EIA-861M will use the annual Form EIA-861 as its sampling frame and estimates for the nonsampled members of the universe, or for any monthly data that is missing or fails edits. The sample will basically be a cutoff sample composed of those companies that typically sell most of the electricity in each customer class in each state. Respondents include:

- All investor-owned utilities (IOUs), except for a few small IOUs in Alaska

⁵ Bundled service, which includes the supply of energy and delivery, is typically provided by traditional electric utility companies. In states where retail choice for electric power service is available, the energy may be supplied by a power marketer (also referred to as an energy service provider) and the utility provides delivery service only.

- All retail power marketers
- All federal utilities
- All entities selling electricity to the public transportation sector
- A sample of the municipal and cooperative utilities

The Form EIA-861M sample design and estimation procedures will employ a weighted linear regression model to represent the relationship between a respondent's past annual data and current monthly data. For example, the model used for estimating monthly data of a company in January 2017 would be:

$$\hat{y}_{i_{2017}} = \hat{\beta}_1 x_{i_{2015}} + w_i^{-1/2} e_i \quad (1)$$

where 1 represents January, and x represents 2015 annual data.

These model parameters are estimated for all respondents having valid data for both the current month on the Form EIA-861M and the most recent annual Form EIA-861 data. The model is not used when prior annual data are unavailable, as is the case for respondents that are new to the target population (although this is uncommon). For a new respondent, the respondent's monthly data are used in estimating totals.

The model parameters are estimated separately each month using three independent variables (sales volume, revenues, and number of customers) by end-use sector and geographic region. The models are then applied to estimate data for entities not in the Form EIA-861M monthly sample but with valid annual Form EIA-861 data. The same model is used to impute for non-respondents and cases where submitted data fails EIA's data review. Data for which EIA has a complete census each month – that is, retail power marketers and the electric utilities that provide distribution services for the marketers⁶ -- is added to the sampled and estimated values to obtain monthly estimates for the entire universe.

The monthly cutoff sample thresholds for the Form EIA-861M will be selected based on the criterion of having estimated relative standard error (RSE) values less than 1 percent for all data published at the state level by end-use sector. The RSE is a percentage measure of the precision of a survey statistic and is used in part as one way to measure error introduced by using model-based predicted monthly values in place of missing and non-sampled data for the quantities of interest (revenues, sales, and number of customers). Threshold values for the cutoff sampling have been adjusted over time to maintain low RSEs for the published estimates.

In past years, EIA had Form EIA-861 data for the universe of respondents that were one or two years old. However, with the 2013 change in the Form EIA-861 from a survey of the respondent universe to a sample, EIA had data for the respondent universe that was one-to-six years old. As noted in section B.2.1 of the EIA clearance package submitted in October 2012, EIA research indicated that this would not adversely impact the accuracy of the estimates.

As already described above, the entities excluded from reporting on the Form EIA-861 that subsequently became the respondents to the Form EIA-861S accounted for approximately 1 percent of total retail

⁶ If imputation is necessary due to nonresponse of either census group, the most recent census monthly data would be used to impute for the missing data.

sales, and it was determined that the model-based estimation methodologies then (and presently) employed for Form EIA-861's monthly version, Form EIA-861M, provide sufficiently accurate imputed/predicted values for these entities. The Form EIA-861M is a smaller cutoff sample of the Form EIA-861, and its model-based estimation methodologies are generally weighted least-squares regressions. The retail sales, revenue, and customer counts reported on the most recently completed Form EIA-861 census cycle are utilized as regressor data.

The present proposal to further decrease the frequency of the need for Form EIA-861S respondents to report on Form EIA-861 from every five years to every eight years is a result of EIA's extensive simulations using historical EIA-861 data. A sequence of estimates were made for 2014 data (dependent variable) using each historical year (2004-2013) as the independent variable. EIA tested and confirmed the validity of the described sampling and estimation strategy by using less recent but relevant regressor data. These simulations assumed that the total census frame is maintained every year on Form EIA-861S to prevent over/under respondent imputation.

For a general description of cutoff sampling as applied at EIA, see "Model-Based Sampling, Inference and Imputation" at <https://www.eia.gov/electricity/data/methodology/eiawebme.pdf>.

Form EIA-923

Note: This clearance contains a proposal to revise the Form EIA-923 cutoff sampling methodology.

The Form EIA-923 is a monthly and annual survey of electric power plants. The survey collects information that includes electric power generation, energy source consumption, end of reporting period fossil fuel stocks, as well as the quality and cost of fossil fuel receipts.

A revised and more efficient model-based cutoff sampling methodology is being proposed for the Form EIA-923. EIA utilizes cutoff sampling to reduce the burden on smaller capacity plants and EIA's own workload. EIA collects monthly data on the Form EIA-923 from respondents (i.e., plants) with a large energy capacity and collects annual totals at the end of the reporting year on the Form EIA-923 from plants with a small energy capacity. The monthly sample is selected from the Form EIA-860 sampling frame of electric power plants. With the combination of monthly and annual respondents, the Form EIA-923 achieves a full census of operational and standby power plants with generation capacity of at least 1MW.

Since a cutoff sample is used, EIA predicts values for the smaller plants using weighted linear regression models depending on the module estimated. EIA can estimate values for out-of-sample plants in current time, because there is a census of the previous year's data; i.e., regressor data. See Equation (1) above for details.

Nameplate capacity was used as the measure of size in the original sample selection process and its high correlation with other data elements on the Form EIA-923 ensured good coverage results for other reported values. Since then, sample validation studies were conducted on fuel consumption, receipts, costs, and stocks, and the sampling methodology was adjusted accordingly.

The sampling strata are defined by facility type, energy source, and geographic region. For instance, one stratum is identified as electric utilities burning coal in the South Atlantic Census Division. There are four facility types: electric utilities, independent power producers, commercial facilities, and industrial

facilities; 14 energy source categories, which correspond to the energy source classifications used in the *Electric Power Monthly* (EPM); and 10 geographic groups, which correspond to the 10 modified Census divisions published in the EPM.

The original Form EIA-923 monthly sample was selected with the 2008 data collection cycle. By the inception of the 2015 collection cycle, nearly 1,750 plants have been added to the monthly sample to maintain the targeted sample coverage ratios. These sample additions are necessary as new large plants become operational. EIA intends to conduct future sample validation studies as the universe of power plants changes to reassess sample size.

Form EIA-923 Sample Selection Criteria

The Form EIA-923 sampling methodology provides accurate results while minimizing the burden on the industry and the federal government. The following five steps are used in selecting plants for the monthly sample:

1. Select preliminary cutoff samples based on nameplate capacity values
2. Add sample units, where necessary, based on generation, consumption, and stocks
3. Add sample units, where necessary, to provide adequate sample counts for estimation groups
4. Add sample units, where necessary, to reduce relative standard errors (RSEs) of key estimates to acceptable levels
5. Add other facilities, based on special-case criteria

The first three steps were designed to ensure adequate coverage of the target population by including all of the largest contributors to key data elements. The fourth step helps ensure that the published estimates meet reasonable reliability standards, which is the key goal, given acceptable resource expenditure. The final criterion covers special cases, as described below.

Facilities in the target population that meet any one of the sample selection criteria applied at any of the five steps are included in the final sample. Further, any additional prime movers and energy sources used by a sample facility are also included in the sample, even if they individually did not meet any of the sample selection criteria.

Each sample facility reports data each month for all combinations of prime mover and fuel source. All nuclear and pumped storage facilities are included in the monthly sample. The remainder of this section provides further detail on the sampling steps.

Step 1: Select Cutoff Samples Based on Nameplate Capacity.

Pre-determined capacity coverage percentages are tested to ensure a certain proportion of operational Form EIA-860 capacity is covered within each stratum. Stand-by generators are not included in the operational capacity totals when data are aggregated to the level of prime mover, and only the largest consumed fuel source for each generator is used in identifying the stratum.

Different target coverage percentages are selected for each facility classification using plants with nameplate capacity over 25 MW and are applied to all regions and energy sources within each facility classification. Otherwise, the percentages of capacity included in the sample are, as follows:

1. Electric utilities – 70 percent
2. Independent power producers – 70 percent
3. Commercial facilities – 50 percent
4. Industrial facilities – 50 percent

Step 2: Add Units Based on Generation, Consumption, and Stocks.

Facilities accounting for large percentages of actual past reported gross generation, fuel consumption, or fuel stocks, are added to the sample, even if their nameplate capacities fall below the percentage cutoff for capacity coverage.

Step 3: Add Units to Ensure Adequate Sample in Estimation Groups.

Instead of fitting models separately for each sampling stratum, the cases are collapsed to form “estimation groups” that have data as homogenous as possible while also having the largest number of observations. This will provide better fitting models than if the models were fit separately on each sampling stratum. These groups are re-formed over time as changes in the data occur. Units below the threshold value were added to any estimation group with fewer than 10 usable observations, until the usable count was brought up to 10.

Step 4: Add Sample to Meet Reliability Standards.

Weighted linear regression equations, identical to those currently employed in the Form EIA-923 imputation system, are run, and relative standard error (RSE or Coefficient of Variation) estimates are calculated for each publication group by month. An additional diagnostic measure, the relative standard error for a super population (RSESP), is calculated to indicate the adequacy of the regression model fit. Limits for both measures (RSE and RSESP) are set individually for each facility classification and applied to all energy sources for the U.S. total for each classification.

If one or both of the error measures fall outside of the limits, the next largest facilities, ranked by gross generation, are included until the estimated RSE/RSESP values are brought into the desired range. If only the RSESP estimate (and not the RSE estimate) is out of range, then it is difficult to lower the estimate of RSESP based on sampling, alone. In these cases, a change in modeling may be necessary. The RSE/RSESP data quality limits are outlined, as follows:

- Electric utilities – RSE less than 5 percent and RSESP less than 20 percent
- Independent power producers – RSE less than 5 percent and RSESP less than 20 percent
- Commercial facilities – RSE less than 10 percent and RSESP less than 30 percent
- Industrial facilities – RSE less than 10 percent and RSESP less than 30 percent

Step 5: Add Special Cases.

Lastly, additional facilities are added to the sample, as necessary. These include storage-only facilities (used in estimating stocks); new facilities for which the EIA has no prior-year's annual data for use in regression imputation; and any new large facilities that the survey staff identifies for inclusion in the sample.

As already described above, the electric power plants currently reporting on a monthly basis on the Form EIA-923M account for 94 percent of total electricity generation. It was determined that the current model-based estimation methodologies employed for the remaining population of power plants that submit data annually on the Form EIA-923 provide sufficiently accurate imputed/predicted values. However, it was determined that in order to reduce burden for respondents and the agency, decreasing the number of power plants in the monthly sample should be employed. The reduction to the monthly sample and new sample methodology was developed so that an accurate imputation of data could still be performed for power plants submitting data annually, despite the reduction of power plants submitting data via the monthly sample.

The electricity generation and energy source consumption reported on the most recently completed Form EIA-923 census cycle are utilized as regressor data (independent variables used in the estimation methodology for imputing for out-of-sample and non-respondent power plants). As stated previously, the new model-based cutoff sample methodology then determines the minimum sample size needed to achieve a target RSE reliability constraint for each publication cell.

The present proposal of implementing this new sample methodology in order to decrease the population of power plants reporting monthly is a result of EIA's extensive research into how a lower monthly sample size will affect the ability to impute/predict values for power plants reporting on an annual basis. EIA tested and confirmed the validity of the described sampling and estimation strategy by using less recent but relevant regressor data. These simulations, however, presuppose that the total census frame is maintained every year on the Form EIA-923 to prevent over/under respondent imputation.

Form EIA-930

The purpose of the EIA-930 imputation procedure is to account for missing data or anomalous data reported by respondents that would prevent the calculation of complete or reasonably approximate regional and national totals, which are published by EIA. The imputed values are only used in the calculation of regional and national aggregate values; the values for each respondent Balancing Authority (BA) are published as reported.

The data processing system currently performs imputations for demand and net generation values only. Imputed values are flagged in the database as either missing or anomalous. Once corrected data are received, the corrected actual reported values are used to re-calculate regional and national totals, which are then updated.

Demand values that will be imputed are identified as follows:

- Blank/missing values
- Zero values
- Negative values
- Values that are equal to or greater than 1.5 times a maximum value specified for each BA

Net generation values that will be imputed are identified as follows:

- Blank/missing values
- Zero values, except for a small number of BAs for whom zero net generation is a valid value
- Negative values, except for a small number of “generation-only” BAs for whom a negative value between -11 and 0 is valid
- Values that are equal to or greater than 1.5 times a maximum value specified for each BA

The maximum values specified for each BA are based on their maximum reported historical demand and net generation values excluding likely anomalous values and, for demand, a comparison to FERC-714 monthly peak demand values if comparable.

Imputation for demand values is performed as follows:

1. Use the BA's reported demand forecast value for the hour in question. (This step is skipped for a small number of BAs due to the non-comparability of their demand and demand forecast values.)
2. If the previous step's value is missing, use the previous hour's demand value.
3. If the previous step's value is missing or imputed, use the demand value for the same hour of the prior day.
4. If the previous step's value is missing or imputed, use the last available daily average demand for that BA.

Imputation for net generation values is performed as follows:

1. Use the previous hour's net generation value.
2. If the previous step's value is missing or imputed, use the net generation value for the same hour of the prior day.
3. If the previous step's value is missing or imputed, use the last available daily average net generation for that BA.

B.3 Maximizing Response Rates

The response rates for the EIA electric power and renewable surveys have historically been at or near 100 percent. Recent response rates are shown below in Table 1.

Table 1. Recent Survey Response Rates (2015 Data)

Frequency of Survey	Survey	Survey Frame	Number of Responses	Response Percentage	Non-Responses	Non-Response Percentage
Annual Surveys Collection of 2014 Data*	EIA-63B	171	171	100%	0	0%
	EIA-411	9	9	100%	0	0%
	EIA-860	3160	3154	99.8%	6	0.2%
	EIA-861	2264	2264	100%	0	0%
	EIA-861S	1111	1110	99.9%	1	0.1%
	EIA-923	5600	5593	99.9%	7	0.1%
Monthly Surveys Collection of January 2016 Data in March 2016**	EIA-826	533	533	100%	0	0%
	EIA-860M	188	187	99.5%	1	0.5%
	EIA-923	2108	2108	100%	0	0%
Hourly and Daily Collection of April 21, 2016***	EIA-930	66	1650	97%	49	3%

*The most current, complete annual surveys collection was for 2014 data, completed and published in January 2016. The annual surveys collection of 2015 data is currently being processed and prepared for final publication in late 2016.
The latest monthly data collection was for January 2016 data. The data, which are considered preliminary, were published in late March 2016. *This day was chosen as a representative day. Hourly data are posted within one hour and daily data are posted the next day.

To maximize response rates, EIA survey forms and instructions are designed and written for clarity and conciseness. Data that are not expected to change from year-to-year or month-to-month are pre-populated on the forms. Notifications are emailed early to maximize the time that respondents have to complete the surveys.

As noted in Part A, EIA's Internet data collection system makes forms available on-line as soon as respondents obtain a secure ID and password. Given the high Internet use rate among respondents to these surveys (approximately 95 percent), most online respondents will log on in the next data collection period and access their required forms. Form due dates are the same each period so that respondents can schedule their completion activities.

Any non-respondents are contacted by email, telephone, and letter to request data submission, until an insignificant or zero non-response rate is obtained. Follow-up email messages citing sanctions for failing

to file the required form(s) are sent to all non-respondents. If the follow-up email messages do not result in a response, additional correspondence requesting immediate submission is sent to the supervisor of the primary contact and, if necessary, to higher-level management officials at the non-responding entity. These letters are sent from the EIA Office Director or (rarely) from the Assistant Administrator or Administrator of EIA.

Respondents who file via the Internet system are given the opportunity to either correct or explain unusual data during their submission. These explanations are reviewed by EIA staff. Respondents are contacted if further clarification is needed. For those respondents that do not file via the Internet, but rather on a hard copy of the form, email messages are sent and/or telephone calls are made to confirm corrections or clarifications of any data suspected to be in error.

Changes in plant ownership and/or contacts have contributed in the past to non-response. To address this issue, EIA developed an improved centralized frame system for the electric power surveys. This system affords all survey staff almost immediate knowledge of changes in entity ownership and/or contacts. This frame system is integrated with the EIA's Internet Data Collection system so that access can quickly be given to new owners and/or contacts.

B.4 Test Procedures and Forms Consultations

The survey designs are reviewed by EIA cognitive specialists and modified as necessary to improve clarity and reduce burden. Draft surveys are also made available for review as part of the 60-day and 30-day Federal Register Notice (FRN) comment periods. When the FRNs are published, emails are sent to over 5,000 monthly and annual survey contacts, 66 Balancing Authorities, and a list of about 250 stakeholders concerning the survey changes with links to the proposed forms. In addition, the draft forms and instructions are posted on EIA's website for the general public. When comments are received on form design and content, the comments are evaluated and changes made as necessary.

For additional information on forms review, please see Part A, section A.8., "Summary of Consultations Outside of the Agency." For a summary of the comments received and EIA's responses, see the Public Comment File submitted with this OMB package.

B.5 Statistical Consultations

For additional information concerning this proposed information collection, please contact Stan Kaplan at 202-586-5114, or at Stan.kaplan@eia.gov.