



EIA STANDARDS MANUAL

[Foreword](#)

Interagency Standards

1. [2002-1](#), Office of Management and Budget Requirements

Systems Standards for Information Collection, Processing, and Dissemination

2. [2002-2](#), Information Technology (IT) Systems
3. [2002-3](#), Information System Documentation

Information Collection and Processing Standards

4. [2002-4](#), Survey Collection/Processing Planning, Design, and Testing
5. [2006-5](#), Frames Development and Maintenance
6. [2002-6](#), Respondent Contact Records (RCRs)
7. [2002-7](#), Response Rates and Imputation
8. [2002-8](#), Measuring Response Rates
9. [2002-9](#), Edit Procedures
10. [2006-10](#), Survey Data Evaluation
11. [2002-11](#), Data Quality Measures

Information Dissemination Standards

12. [2002-12](#), Policy for Releasing Information
13. [2006-13](#), Revisions
14. [2002-14](#), Dissemination of Information Based on Reported and Derived Data (Estimates)
15. [2002-15](#), Rounding
16. [2002-16](#), Codes, Abbreviations, Acronyms, and Definitions
17. [2002-17](#), Information Utility
18. [2002-18](#), Information Integrity
19. [2002-19](#), Accuracy Measures of Data and Estimates
20. [2002-20](#), Quality Assurance Reviews
21. [2007-21](#), Data Protection and Accessibility
22. [2002-22](#), Nondisclosure of Company Identifiable Data in Aggregate Cells
23. [2002-23](#), Reproducibility
24. [2002-24](#), Documentation for Public-Use Electronic Products
25. [2002-25](#), Statistical Graphs

Model Standards

26. [2002-26](#), Model Documentation
27. [2002-27](#), Model Archival
28. [2002-28](#), Proprietary Models

Public Comments on Agency Compliance with Quality Guidelines Standard

29. [2002-29](#), Public Comments on Compliance with Information Quality Guidelines

Business Process Documentation, Continuity of Operations, and Records Management Standards

30. [2002-30](#), Business Process Documentation
31. [2002-31](#), EIA Continuity of Operations
32. [2002-32](#), EIA Records Management

Additional Materials

1. [Standard 2002-4 Supplementary Materials](#), Survey Form Design Checklist
2. [Standard 2002-10 Supplementary Materials](#), Developing A Survey Data Evaluation Plan and Suggested Approaches for Evaluating Different Types of Survey Data Nonsampling Error
3. [Standard 2002-11 Supplementary Materials](#), Additional Suggested Data Quality Measures, and Periodic Quality Reviews
4. [Standard 2002-15 Supplementary Materials](#), Guidelines on the Standard for Rounding
5. [Standard 2002-20 Supplementary Materials](#), Quality Assurance Review Guidelines
6. [Standard 2002-22 Supplementary Materials](#), Guidelines for Implementation of a Disclosure Limitation Rule
7. [Standard 2002-25 Supplementary Materials](#), Guidelines for Graphs
8. [Standard 2002-26 Checklist A](#): Explanatory Model Documentation Components
9. [Standard 2002-26 Checklist B](#): Supplementary Model Implementation Documentation Components
10. [Standard 2002-26 Guidelines](#) for Mathematical Specifications in Model Documentation

NOTE: Some Related Information used in conjunction with EIA's Standards is available only on EIA's Intranet site accessible to persons working for EIA.

FOREWORD

Standards are used by the Energy Information Administration (EIA) in support of EIA's [Information Quality Guidelines](#) to help ensure and maximize the quality, objectivity, utility, and integrity of information disseminated by EIA. Standards document the professional basis upon which EIA expects to be judged by our stakeholders and the level of quality and effort expected in all our activities, including those of our contractors. Standards provide a means for and assurance of consistency among and within activities conducted by EIA. Finally, standards provide users of EIA products information on the methods and principles employed in the collection, analysis, and dissemination of information.

The standards in this Manual must be followed by all EIA staff and contractors and are effective as of the approval date indicated in each standard. This *Manual* supersedes the previous Standards Manual issued by EIA.

This *Manual* contains 32 EIA standards designed for general application to EIA models, surveys, data systems, and information products. EIA conducts over 50 surveys, utilizes over 100 data systems, operates over 25 models for analysis and forecasting, and disseminates information in numerous electronic and printed products. The application of standards over such a wide diversity of activities, of course, requires judgment.

In those rare instances where the strict application of a standard is impractical or infeasible, an affected EIA

program office should consult with the Statistics and Methods Group (SMG) to consider alternative methods of achieving the objective of the standard or to request an exemption from the standard.

The SMG Director will decide whether an exemption should be granted based on a review of the circumstances as well as consultation with other members of EIA Senior Staff, as appropriate. If an exemption is not granted by the SMG Director, the affected program office has the option of requesting an exemption granted by the EIA Deputy Administrator.

Date Issued: September 29, 2002

Guy F. Caruso
Administrator
Energy Information Administration

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-1

Title: Office of Management and Budget Requirements

Superseded Version: 88-02-03

Purpose: To identify Office of Management and Budget (OMB) requirements that are mandatory for Federal information programs and activities.

Applicability: All EIA information collection and dissemination activities.

Required Actions: The following OMB requirements must be adhered to:

- [Metropolitan Statistical Areas](#) for use by all Federal statistical agencies to establish nationally consistent definitions for collecting, tabulating, and publishing statistics for a set of geographic areas
- [North American Industry Classification System](#) (NAICS) for use by all Federal statistical agencies to classify business establishments (NAICS replaced the Standard Industrial Classification [SIC])
- [Data on Race and Ethnicity](#) for use by all Federal agencies to establish a minimum standard for maintaining, collecting, and presenting data on race and ethnicity for all Federal reporting purposes
- [Standard Occupational Classification \(SOC\) User Guide](#) for use by Federal statistical agencies to classify workers into occupational categories for the purpose of collecting, calculating, or disseminating data
- [Federal Statistical Confidentiality Order](#) applies to statistical agencies or units as defined in the Order (within EIA the Order applies only to the Energy Consumption Division which the Order refers to as the Energy End Use and Integrated Statistics Division); the Order clarifies, and make consistent, government policy protecting the privacy and confidentiality interests of individuals or organizations which furnish data for Federal statistical programs
- [Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies](#) (February 22, 2002, Volume 67 of the *Federal Register* [FR], pages 8452-8460); requirement for Federal agency information quality guidelines
- [Controlling Paperwork Burdens on the Public](#) (August 29, 1995, 60 FR pages 44978-44996); OMB approval requirements for an information collection (e.g., survey) involving 10 or more persons in a 12-month period
- [OMB Circular A-130, Transmittal Memorandum #4, Management of Federal Information Resources](#) (November 28, 2000); requirement for all Federal agencies to establish a policy for the management of Federal

information resources

- [OMB Circular A-16, Coordination of Geographic Information and Related Spatial Data Activities](#) (August 19, 2002); provides direction for Federal agencies that produce, maintain or use spatial data (i.e., information about places or geography traditionally shown on maps) either directly or indirectly in the fulfillment of their missions.

Related Information: None

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-2

Title: Information Technology (IT) Systems

Superseded Version: Standards 88-02-01, 88-03-02, 88-03-03, and 88-03-04

Purpose: To ensure that the EIA IT systems are developed, maintained, and secured using best IT practices and that they adhere to established policies, standards, and procedures.

Applicability: All IT systems at EIA.

Background: EIA has standards for developing and documenting computer systems. These standards ensure that systematic procedures are used when new IT systems are developed and existing systems are revised; that these systems meet business and functional requirements; and that the systems can be operated, modified, and backed up. With the advent of the Internet, the EIA IT standards have been expanded to address data access and security issues arising in the Web environment.

Required Actions: All EIA and contractor staff involved in developing and maintaining IT systems are required to follow applicable Federal Information Processing (FIPS), Department of Energy, and EIA-specific IT system standards, policies, guidelines, and procedures listed below.

Related Information:

1. [Federal Information Processing Standards \(FIPS\)](#)
2. [Department of Energy IT Standards](#)
3. [2002-30](#), Business Process Documentation
4. [2002-31](#), EIA Continuity of Operations
5. [2002-32](#), EIA Records Management
6. [EIA Information Systems Security Policy](#)
7. [EIA Password Generation, Protection, and Use Procedures](#)
8. [EIA Operations Security Program Manual](#)
9. [EIA Software Systems Life Cycle Standard](#)
10. [EIA Web-Based Data Collection Systems Standard](#)
11. [EIA Security Architectural Standard For Web-Based Data Collection Systems](#)
12. [EIA Desktop PC Standard](#)

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-3

Title: Information System Documentation

Superseded Version: Standards 88-03-02 and 88-03-03

Purpose: To ensure that EIA information systems are documented adequately to allow personnel unfamiliar with them to become knowledgeable about them and to operate them, if necessary.

Applicability: All EIA information systems.

Required Actions:

1. If an EIA information system is a modeling system, it must follow the requirements of EIA [Model Documentation Standard 2002-26](#).
2. All other EIA information systems (e.g., survey systems, secondary data systems, and other systems supporting other key EIA business practices) must have documentation of all operations (both automated and manual) necessary to operate, maintain, and update the systems.
 - All new information systems should have up-to-date *Users* and *Developer's Reference Manuals* (see section 7.8 of [DOE Systems Engineering Methodology](#)).
 - All information systems created prior to the approval date of this standard should have up-to-date *Operations* and *Programmer's Maintenance Manuals* as required by the predecessor [EIA Standards Manual](#).
 - When documenting an information system that is not a modeling system, the following topics should be covered, if applicable: 1) an overview of integrated manual and automated operations, workflow, interfaces, and personnel requirements; 2) frames development and updating; 3) sampling design and methodology; and 4) a detailed description of each step in collection/processing/dissemination, including distribution of survey materials, establishment of access to on-line reporting options, data collection methodologies, respondent contact handling, non-response procedures, data entry procedures, data editing procedures, error correction/adjustment procedures, estimation methodology, quality comparison procedures, withholding procedures, revision procedures, dissemination, and back-up and archiving procedures. This information may be incorporated into the existing documentation or written as a separate document.

Related Information:

1. [2002-30](#), Business Process Documentation
2. [2002-31](#), EIA Continuity of Operations
3. [2002-32](#), EIA Records Management

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-4

Title: Survey Collection/Processing Planning, Design, and Testing

Superseded Version: Standard 88-04-01

Purpose: To ensure that appropriate procedures are followed in the design, development, and testing of EIA survey programs and to obtain EIA and Office of Management and Budget (OMB) approvals.

Applicability: All EIA survey programs, including existing programs undergoing major revisions. Major revisions include the following: change from census to sample survey or vice versa; change in sample size of over 25% from the size of the previous survey collection; change in the type of reporting entity from

which information is collected; major sample redesign; or substantive changes in the information collected.

Required Actions:

1. **EIA Review:** Before beginning the EIA approval process, a survey sponsor should consider the data needs, collection methods, processing system, dissemination plans, and estimated resource requirements. A proposal for a new survey or a major revision of an existing survey should then be developed for review and approval by the Sponsoring Office Director and the Statistics and Methods Group (SMG) Director prior to submission of the proposal to OMB. The proposal is typically documented in a supporting statement that includes:
 - The energy subject-matter background and rationale for a new/revised effort, including a discussion of uses of the data/data products.
 - An explanation of the reasons why the new/revised effort is needed at this time (e.g., Congressional request or directive, information requests from other public or private data users, EIA management decision, etc.).
 - A brief discussion of survey methodology, including the universe being covered, the type of reporting entity being considered for data collection and analysis, sampling approach and sample size if the collection is to be a sample survey, frequency of collection, mode(s) of collection, topics of information to be collected, etc.
 - Information on how the data will be disseminated, including intended release times to the public and data dissemination media.
 - Initial estimates of the per-year resources (staff and contract dollars) and annual respondent burden of the proposed survey.

2. **OMB Review:** Any new survey collecting information from 10 or more persons in a 12-month period or any major revision to an existing survey generally requires OMB approval. In addition, a renewal of OMB approval is required at least once every three years, even if no changes will be made to an existing survey. The sponsoring office should coordinate with SMG to ensure all OMB requirements are satisfied. The actions required to obtain OMB approval include:
 - Consulting with stakeholders (e.g., survey respondents and data users) who may be affected by the proposed action. At a minimum, OMB requires publication of a *Federal Register* notice requesting public comments. The *Federal Register* notice will be included on EIA's Web site. Consultations with stakeholders may also be expanded to include any other means for collecting comments, such as individual meetings, focus groups, presentations at conferences/workshops, cognitive testing, and pretests/pilot tests.
 - Preparing a clearance package for submission to OMB. A clearance package contains a supporting statement (written according to OMB's requirements), draft respondent letter, survey form, and instructions.
 - Reviewing the clearance package within EIA prior to submission to OMB.
 - Submitting the clearance package to OMB. Simultaneously, issuing a second *Federal Register* notice requesting that public comments be sent to OMB.

Related Information:

1. [Controlling Paperwork Burdens on the Public](#) (8/29/95, 60 *FR* 44978-44996); OMB approval requirements for an information collection (e.g., survey) involving 10 or more persons in a 12-month period
2. [Standard 2002-4 Supplementary Materials](#), Survey Form Design Checklist

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2006-5

Title: Frames Development and Maintenance

Superseded Version: Standard 2002-05

Purpose: To require plans to ensure that necessary steps are taken to develop and maintain frames, and that a frame's coverage is evaluated and documented.

Applicability: All EIA surveys.

Required Actions for Creating New Frames:

A plan for constructing a frame file should be developed and implemented when a new frame needs to be created. The plan should include the following descriptions:

- The frame's major applications and frequency of use.
- The extent that the new frame is expected to reflect the population(s) of interest.
- Clear, standardized, and comprehensive definitions of terms and codes.
- Data items necessary to support unique and efficient identification of elements in the frame.
- Data items necessary for sampling purposes such as:
 - (a) Information referencing a geographic area;
 - (b) Measures of size and/or other sampling attributes;
 - (c) Information relating to company structure that shows how frame elements are organized such as identification of sites, facilities, plants, outlets, subsidiaries, joint ventures and other corporate structures; and
 - (d) Important industrial, technical, legal, or historical relationships between or within frame elements.
- Identification, description, and evaluation of any sources, including any or all parts of an existing frame that may be used to construct a new frame.
- What measures will be used to evaluate the quality of the new frame (under/over coverage, accuracy of sampling and identification information, timeliness, etc).
- Describe any important checks, counts, totals, or statistics that will be calculated from the data fields for the frame elements and the performance measures that will be produced to measure and monitor the frame with respect to quality (see Standard 2002-11 and supplement to 2002-11). Use the measures to evaluate the frame to provide comparisons and/or some trend analysis between the old and updated frame.
- The frequency and conditions for updating the new frame, the methods and sources that will be used, and a projected schedule and resource budget for this task (see Required Actions for Maintaining and Updating Existing Frames).
- Any manual or automated matching/merging/un-duplication methods used.

Required Actions for Maintaining and Updating Existing Frames:

Plans for maintaining and updating existing frames must be documented, implemented, and revised as needed. Maintenance includes:

- Identifying and adding new frame elements (births).

- Identifying and coding frame elements that are no longer in business (deaths).
- Checking and accurately coding whether frame elements are in or out of scope.
- Checking and accurately coding whether frame elements are active or inactive, or other operational status classifications.
- Revising the ownership, name, address, contact, etc. of any data field for a frame element.
- Incorporating suggestions to improve frame quality.
- Incorporating cross-reference information between surveys or survey systems.
- Resolving conflicts between frame elements and other sources of information.
- Recording the date of changes to the content of an element and documenting the person who made the change.

All frames should be maintained and updated based on a predetermined schedule. Updates may also occur anytime as described in the pre-determined frame update plan as information is discovered prior or after the pre-determined scheduled update. The update process should be examined to consider the sources of the update information, the timing, and how the update is performed to avoid any potential source of bias to the survey. Plans to update under a pre-determined schedule should document:

- The frequency or the necessary conditions for updates.
- The projected schedule and resource budget for an update.
- Comprehensive searches that examined the most current information sources and how information from these sources will be incorporated into the frame
- How sources of information exogenous to the frame will be evaluated and used, the timeliness and data limitations of any exogenous information sources, and their associated Internet links if available.
- How updates will be incorporated into all appropriate files, mailing lists, and other survey systems that use the frame.
- The measures that will be used to evaluate the updated frame for quality. These include under/over coverage, accuracy of sampling and identification information, and performance measures that will be produced to measure and monitor quality.
- Descriptions of any important checks, counts, totals, or statistics that will be calculated from the data fields for the frame elements. Use the measures to evaluate the frame to provide comparisons and/or some trend analysis between the old and updated frame.

Required Actions for Frames Retention:

To ensure that the most recent information is retained, the record for each element of the frame must have appropriate transactional variables such as the most current known status of the frame element and the last date the record was modified. Before each major update of a frame, the current version must be archived electronically or be preserved in a manner that allows its recreation with minimal time and effort. Historical versions of frames should be retained in accordance with each survey office's record retention schedule and include:

- all respondents whether active or inactive when the historical version was archived for retention;
- all information about an individual frame element, including all transactional and descriptive data needed for sampling or survey control.

A respondent's record should not be deleted from a frame file unless it was added to the frame because of either a factual mistake or administrative error. If a survey office determines that an inactive respondent should no longer be included as part of the current frame file, then the survey office should first consult with the users of the frame to insure that this action results in no adverse consequences to users of the frame or to future frame updates, or any conflicting data requirements from other survey files. Respondents no longer considered part of the frame should be appropriately coded in the status code field on the frame file so that the record does not appear in future archives of the frame.

Related Information: None

Approval Date: March 22, 2006

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-6

Title: Respondent Contact Records (RCRs)

Superseded Version: Standard 88-04-03

Purpose: To ensure that all contacts with survey respondents are recorded.

Applicability: All EIA surveys.

Required Actions:

1. A Respondent Contact Record (RCR) form must be completed for each communication (e.g., Fax, telephone, mail, e-mail, or any other form of contact) with a respondent. An RCR must be prepared regardless of who initiated the communication or the reason.
2. The Respondent Contact Record should provide the following information:
 - Survey form number
 - EIA Respondent Identification Number
 - Responding entity's reporting name
 - Responding entity's contact person information
 - a) Name
 - b) Telephone number
 - c) Mailing address
 - d) Fax number (if applicable)
 - e) Internet e-mail address (if applicable)
 - Date of contact
 - Reporting period (if applicable)
 - Survey data collection staff member's name
 - Reason for the contact (questions/issues addressed)
 - Information received from respondent
 - Actions taken, if any.

Related Information:

1. [EIA Operations Security Program Plan](#) for information on retention of survey records

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-7

Title: Response Rates and Imputation

Superseded Version: Standard 88-04-02

Purpose: To maximize survey response and minimize the combined variance and bias.

Applicability: All EIA surveys.

Background: Data disseminated by EIA are to be based on surveys with acceptable response rates. The minimum desired level for the unweighted response rate (see [Standard 2002-8](#)) is 80 percent of the eligible respondents. The minimum desired level for a response coverage rate (see [Standard 2002-8](#)) is 80 percent for “key” national-level totals as identified by the survey manager. Imputation methods (including weight adjustment) are to be used to account for the data not submitted.

Required Actions:

1. Systematic effort must be taken to maximize response rates, subject to resource constraints. For a survey not meeting the minimum desired response level, the manager must consider steps to help ensure high response rates in the future:
 - Communicate ahead of time. Notify respondents that they have been selected to report on a survey prior to the first collection period. Provide advance notice so that selected respondents are prepared to report.
 - Use a cover letter signed by a high-ranking executive within the EIA survey office explaining the purpose of the survey, the value to the participant, and how the survey will be used.
 - Increase follow-up communications after the questionnaire is distributed by use of mail (e.g., letter or postcard), phone, e-mail, and fax to stimulate participation.
 - Develop and use a series of letters for late respondents and nonrespondents.
 - Develop lists of high-level executives in respondent companies to contact if other efforts are not successful.
2. Summarize in the survey documentation any imputation methods (including weight adjustment) used to account for survey nonresponse. Monitor response rates over time. If response rates deteriorate, analyze the reasons for nonresponse to assess the effects on the survey and to improve future survey operations. Also, analyze the imputation results to determine the changes in the variances due to imputation.
3. If a response rate is below 80%, consider variance estimates and bias before disseminating the information

Related Information:

1. [Standard 2002-8](#), Measuring Response Rates
2. [Standard 2002-11](#), Data Quality Measures

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-8

Title: Measuring Response Rates

Superseded Version: None

Purpose: To ensure that uniform procedures are followed to measure response rates.

Applicability: All EIA surveys.

Required Actions:

The following general formulae should be used to measure response rates:

- Unweighted unit response rate = R / N
- Response coverage rate (sometimes called weighted response rate) =

$$\frac{\sum \omega_i t_i}{\sum \omega_i}$$

R N

where

R = the number of eligible units completing the survey and responding,

N = the number of eligible units in the survey,

Additionally for the response coverage rate formula:

- The numerator represents a weighted total contributed by eligible responding units.
- The denominator represents the corresponding weighted total for all eligible units in the survey.
- ω_i = the (sampling) weight for the ith unit.
- t_i = a measure of size or a “key” variable for the ith unit.
- For sample surveys $N = n$, the number of sampling units.
- For census surveys $\omega_i = 1$ for all i .

In the above formulae,

- The denominator includes all original survey units that were identified as being eligible, including units with pending responses with no data received, new eligible units added to the survey, and an estimate of the number of eligible units among the units of unknown eligibility. The denominator does not include units deemed out-of-business, out-of-scope, or duplicates.
- The numerator includes all survey units that have submitted sufficient information (based on criteria established by the survey staff) to be considered complete responses for the survey period.

The unweighted response rate is used to indicate the proportion of eligible reporting units that responded to the survey, while the response coverage rate is generally used to indicate the (sampling-weighted) proportion of an estimated national total that is contributed by respondents. The response coverage rate provides a measure of the survey’s coverage of key variables.

Related Information:

1. [Standard 2002-7](#), Response Rates and Imputation
2. [Standard 2002-11](#), Data Quality Measures

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-9

Title: Edit Procedures

Superseded Version: Standard 88-04-05

Purpose: To ensure that survey data are subjected to appropriate edits, edit procedures are applied consistently, and edits are evaluated in a systematic and quantitative manner.

Applicability: All EIA surveys.

Required Actions:

1. The survey documentation must describe procedures for handling records that fail edit checks. The documentation must explain the purposes of the edit checks, the decision procedures for resolving edit messages, and how edit message records are maintained. Edit checks should address the following indicators of data quality (as applicable):
 - a. Consistency
 - b. Range
 - c. Completeness
 - d. Arithmetic/calculation accuracy
 - e. Comparability with other sources
 - f. Validity of codes.
2. For key edits as identified by the survey staff, maintain measures for the number of:
 - a. Edit messages generated
 - b. Edit messages resulting in revisions of the originally submitted data.
3. Document the procedures for evaluating the performance of the edits and the results of any evaluations.

Related Information:

1. [Standard 2002-2](#), Information Technology (IT) Systems
2. [Standard 2002-7](#), Response Rates and Imputation
3. [Standard 2002-11](#), Data Quality Measures

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2006-10

Title: Survey Data Evaluation

Superseded Version: Standard 2002-10

Purpose: To periodically evaluate survey data quality and limitations.

Applicability: All EIA surveys.

Required Actions:

EIA surveys are subject to periodic assessments of data quality. An assessment, conducted either by the survey office or by the Statistics and Methods Group, should focus on one or more of the following topics:

1. Identify and discuss potential sources of nonsampling error - i.e., coverage, measurement error, nonresponse, data collection and processing, survey methodology, imputation, estimation, and revision. Describe how those sources may affect the survey results and assess those effects, if any, over time. (See [Standard 2002-10](#)

[Supplementary Materials](#), Developing A Survey Data Evaluation Plan and Suggested Approaches for Evaluating Different Types of Survey Data Nonsampling Error.)

2. Compare the survey data with similar data available from other sources. Assess differences, and identify and discuss specific reasons - e.g., coverage, survey methodology, definitions, time periods, etc.- for any differences. Also, identify areas of research for resolving any unexplained differences that persist over time. Document the data sources, including Internet links, and any adjustments to the other data series for comparison with the survey data.
3. Identify any changes to the survey over time. Discuss how those changes affect the estimates. Assess the effects on the data from those changes, including evaluating the effectiveness of those survey changes on the accuracy and quality of the data.

Related Information:

1. [Standard 2002-10 Supplementary Materials](#), Developing A Survey Data Evaluation Plan and Suggested Approaches for Evaluating Different Types of Survey Data Nonsampling Error
2. Federal Committee on Statistical Methodology, Subcommittee on Measuring and Reporting the Quality of Survey Data. [Statistical Policy Working Paper 31](#), Measuring and Reporting Sources of Error in Surveys, Washington, D.C., June 2001.
3. Examples of EIA Survey Data Evaluation Materials:
 - Energy Information Administration. [Accuracy of Petroleum Supply Data 2001](#). “Petroleum Supply Monthly” Feature Article, DOE/EIA-0109 (2002/09), Washington, D.C. September 2002.
 - Energy Information Administration. [Comparisons of Independent Petroleum Supply Statistics](#). “Petroleum Supply Monthly” Feature Article, DOE/EIA-0109 (1999/12), Washington, D.C., December 1999.
 - Energy Information Administration. [Energy Consumption Series: Residential Energy Consumption Survey Quality Profile](#). DOE/EIA-0555(96)/1, Washington, D.C., March 1996.

Approval Date: March 22, 2006

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-11

Title: Data Quality Measures

Superseded Version: Standard 88-04-06

Purpose: To collect information for use in evaluating and improving the quality of EIA survey data.

Applicability: All EIA surveys.

Required Actions:

Each survey should collect the measures below to support EIA-wise measures or should have a plan with milestones for collecting those measures not currently available.

1. Frames - Frame size, volatility (e.g., births, deaths, mergers), and estimated coverage.
2. Response Rates - Unweighted unit response rate and response coverage rate (as defined in [Standard 2002-8](#)).
3. Revisions of Disseminated Data - Percent difference between first disseminated data and the final disseminated data for key data series.

4. Timeliness - Time from the close of the survey reference period until the dissemination of key data series.
5. Data Edits - For key edits (as identified by the survey staff), the number of survey responses with that edit flag and the percentage of flagged responses that are revised.
6. Sample Surveys - Sampling variances (or confidence intervals or relative standard errors).
7. Imputation - Item imputation rates for key data items.

Related Information:

1. [Standard 2002-11 Supplementary Materials](#), Additional Suggested Data Quality Measures and Periodic Quality Reviews
2. [Standard 2002-5](#), Frames Development and Maintenance
4. [Standard 2002-8](#), Measuring Response Rates
5. [Standard 2002-9](#), Edit Procedures
6. [Standard 2002-13](#), Revisions
7. [Standard 2002-19](#), Accuracy Measures of Data and Estimates

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-12

Title: Policy for Releasing Information

Superseded Version: None.

Purpose: To ensure that information intended for public release is released according to a dissemination plan that provides fair access to customers.

Applicability: All EIA information products.

Required Actions:

1. To provide fair access to customers, major information products should follow publicly available release schedules.
 - Establish the schedule and mode for the release of information products.
 - Provide security procedures to prevent unauthorized premature release of the information.
2. Sensitive information must be protected against any unauthorized pre-release and may be released only according to an established release schedule. Information is considered sensitive by EIA if it substantively affects financial or energy markets.
 - Prior to the official release time, sensitive information should not be read or discussed in public, or otherwise exposed on airplanes, restaurants, public transportation, or in other public places and/or in any media.
 - The office responsible for producing the sensitive information must follow strict security procedures to protect the information from any unauthorized release prior to the official release schedule. The security procedures for protecting sensitive information against an unauthorized release include:
 - a. Establish and adhere to office lock-up procedures to physically secure the information and to restrict access to it. Depending upon the circumstances, these security procedures may include placing locks on desk drawers, filing cabinets, and office room doors.
 - b. Maintain a list of authorized individuals who have access to the information and designate a primary

responsible party for safeguarding the information prior to its official release.

3. Information products that are non-sensitive may be released prior to their release schedule if available earlier

Related Information:

1. [DOE/EIA Operations Security Program Manual](#), including Section 3, "Fair Practice Data Disclosure," Policy Reference No. 3, "News and Data Releases," and Appendix F, "Embargoed Data Products and Release Guidelines"

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-13

Title: Revisions

Superseded Version: Standard 2002-13

Purpose: To provide EIA customers with information about revisions in disseminated data.

Applicability: All EIA information products.

Required Actions:

1. Establish a policy for anticipated revisions. Show the date the policy became effective, and make it available electronically to users. Program offices should document any change to their revision policy and the effective date for each change.
2. The first dissemination of a data value in an information product should be identified as "preliminary" if revisions are anticipated in a subsequent dissemination. Scheduled revisions to these values should be identified as "revised" (or "final," where appropriate) when a revised value is disseminated.
3. Preliminary and revised data must be identified through the following:
 - A. data value labeling such as:
 - (i.) data marked "P" for preliminary, and
 - (ii.) data marked "R" for revised); or,
 - B. text in the
 - (i.) product title,
 - (ii.) table titles,
 - (iii.) headers or footnotes; or,
 - C. other text accompanying the release of the information product.

Historical databases accessed by the public should identify preliminary data or any other data that are subject to revision either by appropriately marking the data or by text in the product title, header or footnotes of the output table that shows the data. After the revised data are released with the appropriate notice to users, no additional data labeling is required for final or revised data that are archived in an historical database.

4. When unscheduled revisions of data that have not been released as "final" are required due to previously unrecognized errors or respondent resubmissions, clearly identify the revisions and communicate the reasons for revising the data in a notice to data users (see item 3 above).

5. Data previously released as “final” should not be revised without consulting with and notifying all other program offices within EIA. The program office seeking to revise data previously released as “final” must contact the Revisions Coordinator in EIA’s Statistics and Methods Group (SMG) ([Renee Miller](#) as of March 2006) or, in his/her absence, the Director of SMG regarding the proposed change, reason(s) for the change, and the procedures for notifying representatives of the other offices. A program office may proceed with its proposed change if no office objects to the revision within three weeks following notification. If there is any disagreement between the program office and another EIA office concerning a proposed revision, an inter-office team will meet to discuss the proposed change and resolve the differences.

6. Do not disseminate information if “errata sheets” are anticipated.

Related Information:

1. [Standard 2002-7](#), Response Rates and Imputation

Approval Date: March 22, 2006

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-14

Title: Dissemination of Information Based on Reported and Derived Data (Estimates)

Superseded Version: Standard 88-05-03

Purpose: To ensure that EIA customers are aware of how information is produced; i.e., whether it is prepared solely from data reported on EIA surveys, is based on data from non-EIA sources, or is derived using models, and the methodology used to produce the information.

Applicability: All EIA information products.

Required Actions:

1. For information based solely on reported data from surveys:
 - If the data reflect imputation or other adjustment methods, the methods must be described and be accessible from the data product.
 - For sample surveys, a brief discussion of sampling error and its potential effects on the data must be accessible from the data product.
 - Any data item that is a balancing item (i.e., it is not itself reported but is computed as the remainder after other components are taken into account) must be identified.
 - The sources and known limitations of any information from sources outside of EIA must be described.
2. For products with information derived using models:
 - Forecasts and derived estimates must be clearly identified (e.g., in the title, header, footnote, or text of the information product). If only selected cells in a table are estimates, the cells may be marked “E”. A description of the forecasting model or derivation procedure must be accessible from the product along with any available evaluation of its accuracy. Any total or subtotal comprised partially or wholly of estimates must be labeled as well.
 - Any number originally disseminated as a projected value must be identified in subsequent releases.

Related Information:

1. [Standard 2002-7](#), Response Rates and Imputation
2. [Standard 2002-19](#), Accuracy Measures of Data and Estimates
3. [Standard 2002-26](#), Model Documentation

4. [Standard 2002-27](#), Model Archival
5. [Standard 2002-28](#), Proprietary Models

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-15

Title: Rounding

Superseded Version: 88-05-07

Purpose: To ensure consistency in rounding.

Applicability: To all data released by EIA.

Required Actions:

1. To round a number to n digits (decimal places), add one unit to the nth digit if the (n+1) digit is 5 or larger and keep the nth digit unchanged if the (n+1) digit is less than 5.

Note: This is a simplified description of what should be done when rounding. See [Guidelines on the Standard for Rounding](#) for further details on how to round, including rounding negative numbers.

2. All calculations should be carried out prior to rounding. In particular, tabulations to produce summary data and computations performed for purposes of estimating standard errors should be done on data as collected; i.e., no rounding should take place prior to the completion of these kinds of tabulations.

3. Sums of the column (row) values in a table should be derived by using the unrounded column (row) values, with appropriate rounding of the column (row) total being done after its derivation. To handle the problem of the rounded column (row) values not adding to the rounded column (row) total, a footnote should be used such as, "Totals may not equal the sum of their components because of independent rounding."

Related Information:

1. [Standard 2002-15 Supplementary Materials](#), Guidelines on the Standard for Rounding

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-16

Title: Codes, Abbreviations, Acronyms, and Definitions

Superseded Version: 88-02-02

Purpose: To ensure consistency in the use of codes, abbreviations, acronyms, and definitions throughout EIA.

Applicability: All EIA information products.

Required Actions:

1. Codes, Abbreviations, and Acronyms
 - In general, avoid using codes, abbreviations, and acronyms in EIA information products. When codes, abbreviations, or acronyms must be used, define them in footnotes, text, or link to a glossary definition the first time they appear in a product.
 - Use the approved codes, abbreviations and acronyms from the following sources as applicable:
 - a) [Preferred Metric Units for General Use by the Federal Government, Federal Standard 376B](#)
 - b) [The EIA Publishing Style Guide](#)
 - c) [Government Printing Office \(GPO\) Style Manual](#)
 - d) [Official Postal Service Abbreviations](#) (for States and street addresses)
 - e) [Federal Information Processing Standards](#) (FIPS PUBS).
2. Definitions
 - EIA has a [Glossary](#) of established definitions for commonly used energy terms. When using a term in the Glossary, use the established definition.
 - If you have questions, would like to create a definition for a new energy term, or would like to propose a revision to an established definition, contact the Glossary Coordinator, [Grace Sutherland](#) in EIA's Statistics and Methods Group for procedures on coordinating with affected EIA staff and customers.

Related Information: None

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-17

Title: Information Utility

Superseded Version: None

Purpose: To maximize utility (i.e., usefulness to intended users) of energy information products disseminated by EIA.

Applicability: All EIA information products.

Required Actions:

1. Each program office must ensure the continuing usefulness of the information products it disseminates by:
 - Monitoring information needs of its customers and responding to customer feedback.
 - Reviewing on a regular basis the frequency and mode of access of information products.
 - Assessing customer input.
 - Developing new information sources and release methods or revising existing information collection methods, models, and information products, when appropriate.
2. EIA must conduct periodic customer surveys to evaluate how effectively information meets users' needs.
3. Users of EIA information must have access to explanatory materials on possible sources of error, data definitions, collection methodology, estimation and imputation, and any data adjustments to assist in the

understanding and interpretation of information disseminated by EIA.

Related Information:

1. [Standard 2002-4](#), Survey Collection/Processing Planning, Design, and Testing
2. [Standard 2002-20 Supplemental Materials](#), Quality Assurance Review Guidelines

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-18

Title: Information Integrity

Superseded Version: None.

Purpose: To ensure that EIA data, analyses, and forecasts are: 1) secure from unauthorized access or revision or destruction, and 2) not compromised through corruption or falsification.

Applicability: All data submitted to EIA by survey respondents or obtained by EIA from secondary sources and all data, analyses, and forecasts releasable to the public by EIA.

Required Actions:

1. EIA will ensure that:
 - Individually identifiable EIA survey data are protected.
 - EIA data systems and electronic products are protected from unwarranted intervention.
 - Only EIA authorized personnel can access EIA information resources.
 - EIA data files, network segments, servers, and desktop PCs are electronically secure from malicious software and intrusion using best available information resource security practices, which are periodically monitored and updated.
2. Options for respondents to submit forms must be consistent with the Draft EIA Policy on Security of Sensitive Unclassified Survey Data Transmitted to EIA.
 - EIA will provide secure electronic methods for its survey respondents to use when they transmit sensitive unclassified survey data to EIA.
 - EIA will inform its survey respondents about the incoming data transmission methods that EIA can accommodate and which are secure.
 - EIA Office Directors may request waivers to this policy. Waiver requests will be reviewed by the Information Technology Council and decided by the Administrator.
3. EIA will ensure controlled access to data sets so that only specific, named individuals working on a particular data set can either read, write, or both read and write that data set. Data set access rights are to be periodically reviewed by the project manager responsible for that data set in order to guard against unauthorized release or alteration.
4. EIA will ensure that all changes to EIA data are documented and that they are both warranted and follow EIA established data revision procedures.

Related Information:

1. [DOE/EIA Operations Security Program Plan](#)

2. [EIA Cyber Security Program Plan](#)
3. [EIA Information Systems Security Policy](#)
4. [EIA Password Generation, Protection, and Use Procedures](#)
5. [EIA Security Architectural Standard for Web-Based Data Collection Systems](#)
6. [Standard 2002-13](#), Revisions
7. [Standard 2007-21](#), Data Protection and Accessibility
8. [Standard 2002-30](#), Business Process Documentation
9. [Standard 2002-31](#), EIA Continuity of Operations

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-19

Title: Accuracy Measures of Data and Estimates

Superseded Version: None.

Purpose: To provide users with information concerning the accuracy or reliability of survey data and estimates.

Applicability: All EIA information products.

Required Actions:

1. Revision and sampling error information, when applicable, must be available for an EIA product.
 - Revision error (applicable to all disseminated data and estimates when they are subsequently revised) is the difference (often expressed as a percentage) between an initial release of data/estimates and its corresponding final disseminated data/estimates for key data series. The office disseminating data/estimates should provide revision error information designed to help users better understand the variability between initial key data/estimates and final key data/estimates. Some ways to present revision error information include the average revision error, the maximum revision error, or the distribution of revision errors during a specified time period. Also, when presenting a measure of revision error, an indication should be made with respect to what data/estimates the error applies (e.g., the difference between the first estimate and the final data, the difference between preliminary and final data).

If revision error for a key data series shows an initial release is an unreliable indicator of the final data/estimate, consider whether publishing the data/estimate with a measure of revision error or withholding the initial data/estimate is the best way to serve EIA's customers.

- Sampling error (applicable only to data/estimates derived from a sample surveys). Include in products the sampling errors or tools to calculate these errors together with a description of sample selection and estimation procedures.

If the sampling error for a key data series shows the data may be an unreliable measure, consider whether publishing the data with sampling error information or withholding the data is the best way to serve EIA's customers.

2. Nonsampling errors affect all data/estimates and may occur for a number of reasons including incomplete coverage of the units of interest, nonresponse, respondent difficulties in understanding/reporting, mistakes in recording/coding data, and other errors of collection, response, coverage, and estimation. While nonsampling error is generally difficult to measure, a discussion of nonsampling error and its potential impact must be available for an EIA information product.

3. If the revision error, sampling error, or nonsampling error discussion is too complicated or extensive to include in an information product, a reference or link to a Web product with the required information should be provided.

Related Information:

1. [Standard 2002-7](#), Response Rates and Imputation
2. [Standard 2002-11](#), Data Quality Measures
3. [Standard 2002-13](#), Revisions
4. [Standard 2002-14](#), Dissemination of Information Based on Reported and Derived Data (Estimates)

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-20

Title: Quality Assurance Reviews

Superseded Version: None.

Purpose: To ensure pre-dissemination quality assurance reviews of all EIA information products.

Applicability: All EIA information products.

Required Actions:

1. Before public dissemination, an EIA information product must undergo a quality assurance review within the originating office. (For guidelines on the pre-dissemination review, see the [Standard 2002-20 Supplemental Materials](#).)
2. The originating office may submit an information product for review by other EIA offices before dissemination.
3. As needed, a program office may use EIA's Independent Expert Review (IER) program to provide technical advice from highly qualified, nationally and internationally recognized subject matter experts. For additional information about the program, contact EIA's IER Coordinator, [William Weinig](#), of EIA's Statistics and Methods Group.

Related Information:

1. [Standard 2002-20 Supplemental Materials](#), Quality Assurance Review Guidelines
2. [Review, Clearance, and Dissemination of EIA Products](#)
3. [Standard 2002-16](#), Codes, Abbreviations, Acronyms, and Definitions
4. [EIA Publishing Style Guide](#)
5. [EIA Style and Standards for Electronic Products](#)

6. [EIA Guidelines for Statistical Graphs](#)

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2007-21

Title: Data Protection and Accessibility

Superseded Version: 2002-21

Purpose: To ensure that survey information collected by EIA is properly safeguarded

Applicability: All EIA survey information.

Definitions

1. EIA survey information collected under a pledge of confidentiality pursuant to the Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA) is the only survey information that may be considered “**confidential**” and should be protected in accordance with those statutory requirements. Confidential information covered under CIPSEA is exempt from release pursuant to a request made under the Freedom of Information Act. For surveys that collect information outside of CIPSEA the information is considered either “protected” or “public information.” .
2. Survey information collected under a pledge that EIA will protect such information, but not collected under CIPSEA, is considered “**protected**,” regardless of the statutory authority used to collect the information. EIA protects such survey information by pledging not to release identifiable information to the extent it satisfies exemption 4 of section 522(b) of the Freedom of Information Act (FOIA). If the information is submitted voluntarily, it must be the type of information that is not customarily disclosed to the public. If the information is submitted pursuant to EIA’s mandatory data collection authority, Exemption 4 applies if releasing the respondent identifiable information is likely to cause substantial harm to the respondent’s competitive position or impair the agency’s ability to obtain similar data in the future.
3. EIA survey information is considered “**public information**” if it is collected with notice to the respondents that the information may be publicly released in company or individually identifiable form, and will not be protected from disclosure in identifiable form.

Required Actions to Protect Confidential and Protected Information

4. For confidential and protected survey information, the program office shall establish procedures and mechanisms to ensure the information is safeguarded during the production, use, storage, transmittal, and disposition of the survey data in any format (e.g., completed survey forms, electronic files, printouts, and information products). Survey managers are responsible for establishing and maintaining a list of persons who are authorized to access such information, and for safeguarding the accessibility to the survey information. The EIA [Information Systems Security Officer](#) shall insure that the EIA computer system infrastructure complies with DOE and EIA cyber security requirements for protecting data and limiting access to users authorized by the appropriate survey manager.
5. Each EIA and contractor employee shall ensure that confidential and protected survey information is not disclosed to unauthorized persons and shall follow applicable EIA and DOE cyber security, records management,

technical and management requirements, standards, and policies for safeguarding the collection, processing, transmitting, storing of records, indexing for retrieving records, and release of such information.

Conditions for Sharing Confidential and Protected Data and Required Actions:

6. By law, in response to a proper request, EIA may share protected survey information when properly requested for official use by other DOE components, other Federal agencies, the General Accounting Office, and any Committee of Congress. EIA's preference is to share such information only when it will be used for statistical purposes. 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. The process for sharing protected survey data is:

a. A request from a federal agency to share protected EIA survey data shall be made in writing by an Officer (at least equal in rank to the EIA Administrator) of the organization requesting the information and the request must be directed to EIA's Administrator.

b. The sharing of protected survey data must be formalized with a Data Sharing Agreement signed by EIA's Administrator. (Contact EIA's [Standards Officer](#) for additional details on procedures for preparing and signing data sharing agreements.)

c. The Data Sharing Agreement shall stipulate that the receiving organization will provide the same level of protection to the data as EIA provides.

d. Individual respondent information shared under a Data Sharing Agreement shall not be made publicly available by the requesting organization except as required by law.

7. A request to share information collected under CIPSEA shall satisfy all of EIA's terms and conditions for designating the requestor as an agent of EIA and must follow the procedures outlined in Section VII of the EIA [CIPSEA Guidance Manual](#) for granting access. EIA has discretion whether to approve a request for accessing data protected under CIPSEA.

8. A public request received under the Privacy Act or Freedom of Information Act for confidential CIPSEA information or other protected survey information shall follow the EIA and DOE FOIA/Privacy Act Procedures. Confidential information is not released in response to a request made under FOIA because it is considered a nonstatistical purpose under CIPSEA and is exempt from release to the public. Under CIPSEA, nonstatistical use is any use of identifiable data that is not statistical, including any administrative, regulatory, law enforcement, adjudicatory, or other purpose that affects the rights, privileges, or benefits of a particular identifiable respondent. Protected survey information not collected under CIPSEA is not released in response to a request made under FOIA to the extent it satisfies exemption 4 of FOIA.

Related Information:

1. [Standard 2002-18](#), Information Integrity
2. [Standard 2002-22](#), Nondisclosure Of Company Identifiable Data In Aggregate Cells
3. [EIA CIPSEA Guidance Manual](#)
4. [Policy on Security of Sensitive Unclassified Survey Information Transmitted to EIA](#)
5. [DOE/EIA Operations Security \(OPSEC\) Program Plan](#)
6. [EIA FOIA Procedures](#)
7. [DOE FOIA/Privacy Act Program](#)
8. [EIA Information Systems Security Policy](#)
9. [EIA Security Architectural Standard for Web-Based Data Collection Systems](#)
10. [EIA Password Generation, Protection, and Use Procedures](#)
11. [DOE Requirement: Sensitive Unclassified Information and Personally Identifiable Information, TMR22](#)
12. [DOE O 243.1 Records Management System](#)

Approval Date: December 20, 2007

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-22

Title: Nondisclosure of Company Identifiable Data in Aggregate Cells

Superseded Version: Standard 88-05-06

Purpose: To ensure that sensitive data cell values are suppressed (i.e., withheld from public release) for the protection of confidential survey data. A sensitive data cell value is one that, if publicly disseminated, may be used to closely estimate the confidential information that an individual survey respondent reported to EIA.

Applicability: Data disseminated by EIA that are collected with a pledge of confidentiality. **EXCEPTION:** For information products where data historically have been disseminated without suppression of sensitive cells, a waiver to this Standard may be obtained if a *Federal Register* notice announcing that such dissemination will continue does not meet with substantive negative response. This must also be explained on the collection form/instructions sent to respondents and in any materials sent to OMB.

Required Actions:

1. Determine what cells are sensitive by using the following criteria:
 - a. If the nonzero cell value is the reported value from either one or two respondent(s), the cell is considered sensitive.
 - b. If the nonzero cell value is the sum of the reported values from more than two respondents, use the PQ rule alone or in conjunction with another rule that has subadditive properties to determine whether the cell value is sensitive. (See Related Information for details on the PQ rule.) Before using a rule other than the PQ rule, consult with the SMG Director, [Nancy Kirkendall](#), regarding the implications.
2. Withhold sensitive cell values from any dissemination and apply complementary suppression to other cells to ensure that the sensitive values cannot be reconstructed from disseminated data.
3. Use the symbol "W" to denote when a data value has been suppressed, along with a footnote explaining that "W" represents "Withheld".
4. Do not publicly reveal parameters associated with disclosure limitation rules used to protect confidentiality. However, the rules must be documented and the documentation available within EIA.

Related Information:

1. [Standard 2002-22 Supplementary Materials](#), Guidelines for Implementation of a Disclosure Limitation Rule
2. [Standard 2007-21](#), Data Protection and Accessibility
3. [Statistical Policy Working Paper 22](#), Report on Disclosure Limitation Methodology
4. [Checklist on Disclosure Potential of Proposed Data Releases](#) (prepared by the Federal Committee on Statistical Methodology's [Confidentiality and Data Access Committee](#))

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-23

Title: Reproducibility

Superseded Version: 88-05-08

Purpose: To ensure that EIA is able to reproduce any publicly released information product.

Applicability: All EIA electronic product files and/or associated data files used to generate EIA energy information products, i.e. databases for on-line queries, data tables, graphs, publications and reports.

Required Actions:

1. All electronic product files must be archived and retained until no longer needed for current business. Complete information products, whether paper or electronic, representing a specific continuing publication product or one-time report, must be permanently archived.
2. The data files and/or databases (at the most disaggregated level), which are used to generate publicly released information products, must be archived and retained until no longer needed for current business.
3. System and model documentation and computer software/programs used to generate an EIA information product must be archived and retained until no longer needed for current business.

RELATED INFORMATION:

1. [2002-30](#), Business Process Documentation
2. [2002-31](#), EIA Continuity of Operations
3. [2002-32](#), EIA Records Management
4. [EIA Record Schedules](#)
5. [EIA Records Management 2000-Policy and Procedures Manual](#)
6. [EIA Styles and Standards for Electronic Products](#), Appendix (see Web Site Records Retention)

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-24

Title: Documentation for Public-Use Electronic Products

Superseded Version: 88-05-09

Purpose: To ensure that electronic information is released to the public with standard documentation to facilitate the understanding and use of the product.

Applicability: All EIA electronic information products.

Required Actions:

All electronic information products should be disseminated with the following:

1. Data description and/or file format and description, if applicable.
2. Source information, such as EIA survey form number and description of methodology used to produce the information or links to the methodology.
3. Time period covered by the information and units of measure.
4. Point of contact in EIA to whom further questions can be directed.
5. Software or links to software needed to read/access the information and installation/operating instructions, if applicable.
6. The date the product was last updated.

Related Information:

1. [EIA Styles and Standards for Electronic Products](#)

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-25

Title: Statistical Graphs

Superseded Version: Standard 88-05-05

Purpose: To ensure the utility (usefulness to intended users) and objectivity (accuracy, clarity, completeness, and lack of bias) of energy information presented in statistical graphs.

Applicability: All EIA information products.

Required Actions:

1. Graphs should be used to show relationships, whereas tables should be used to show individual values. Tables are a better format than graphs for presenting data values.
2. Graphs should contain sufficient information to be understood with information on the printed/web page alone or with additional information in the printed product or accessible from the Web page.
3. Graph titles and axis labels should be clear with no unexplained or undefined acronyms or industry jargon.

4. Both axes of a graph should be labeled with the names of variables, except where the x-axis label “years” is obvious.
5. When using time intervals, spacing should be equidistant only if the intervals are equidistant.
6. In general, the vertical axis of a line or bar graph should include a zero point. The axis should start with zero for positively-valued quantities or have a zero in the middle of the vertical scale for quantities that may demonstrate both positive and negative changes. For line graphs showing relative quantities such as an index, zero is meaningless. In such graphs, the natural basis (such as 100 for a percentage index) should be used with placement in the middle of the scale.
7. A break in the a scale is permissible to accomplish a zero point. A break is indicated by marks (e.g., diagonal lines) on the scale. However, do not connect lines or bars across breaks in the scale.
8. For compliance with accessibility guidelines ([Section 508](#)), an EIA graphic must be understandable when printed or viewed in black and white. Web graphs also need the appropriate alternative text “tags” for use by screen readers.

Related Information:

1. [Standard 2002-25 Supplementary Materials](#), Guidelines for Graphs
2. [EIA Style and Standards for Electronic Products](#) (section on Charts and Graphs)
3. [EIA Guidelines for Statistical Graphs](#)
4. [EIA Publishing Style Manual](#)

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-26

Title: Model Documentation

Superseded Version: 96-01-03

Purpose: To ensure that the procedures, equations, and assumptions which define EIA models are publicly available.

Applicability: All models developed and maintained by EIA or its contractors. (In a large model or system, separable components or modules may be considered as individual models for documentation purposes.)

Required Actions:

1. Model documentation should correspond to a specific archived version of the model.
2. The documentation should include the required components listed in [Checklists A](#) and [B](#) and the optional components as the program office deems appropriate. For items in [Checklist B](#), options are offered for presenting the materials either separately or along with the materials in [Checklist A](#).
3. Documentation should include discussions of the theoretical basis for the model, empirical support for the model, critical assumptions (as judged by the modeler), mathematical specifications, and other information that the program office believes is relevant to potential users.

4. The program office determines whether documentation will be available in one or multiple volumes, the format (hard copy or electronic), and the organization of the materials.
5. In cases where minor documentation revisions are necessary, the program office may produce a new version of the documentation or a supplemental document that describes only the revisions that have taken place in the model.
6. All documentation should undergo a quality assurance review by the program office prior to dissemination. In addition, new or significantly revised model documentation may undergo reviews by other EIA offices and/or independent expert reviewers from outside EIA when appropriate.
7. When a product using model outputs is sent to the Administrator for approval, the responsible Office Director should specify what documentation is available. If up-to-date documentation is not available, the office should provide a schedule for completing the documentation within 90 days of the product's release or should request an exception.
8. The most current model documentation should be available on EIA's Web site (see [Recent Energy Model Documentation](#) main page) or from the model manager.

Related Information:

1. [2002-27](#), Model Archival
2. [2002-28](#), Proprietary Models
3. [Standard 2002-26 Checklist A](#): Explanatory Model Documentation Components
4. [Standard 2002-26 Checklist B](#): Supplementary Model Implementation Documentation Components
5. [Standard 2002-26 Guidelines](#) for Mathematical Specifications in Model Documentation
6. [Review, Clearance, and Dissemination of EIA Products](#)
7. [2002-30](#), Business Process Documentation

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-27

Title: Model Archival

Superseded Version: 91-01-04

Purpose: To ensure that EIA model calculations are reproducible.

Applicability: All models used by EIA.

Required Actions:

1. A model archive package should be prepared by the program office when model outputs are used in an EIA product that is publicly disseminated.
2. A model archive package should contain:
 - a. All source code and program control files needed to compile, link, and execute the reference or base case scenario of the model. If alternate source code versions were used to run the model for other scenarios cited in the same product, provide these versions or create a scenario-specific archive.
 - b. Input files used by the model for the reference or base case cited in the model, along with other file versions

that would be needed to run the model for other scenarios cited in the same product.

- c. Primary output (as opposed to debugging or trace files) from the reference or base case scenario used in the product. It is not necessary to provide all output files for all scenarios cited in the product, as long as those outputs can be regenerated from runs of the archived model and the primary outputs from the model can be verified against disseminated results.
 - d. Instructions for compiling and running the model and comparing the results to disseminated results. A description of changes needed to run the alternative scenarios published in the report or to create scenario-specific archives should be included.
 - e. The source of the proprietary data and software should be included in the archive instructions. However, an archive package should exclude proprietary data and software used in the model.
3. The program office should create and verify an archive package within sixty days of disseminating an EIA information product utilizing model outputs.
 4. The archive package must be retained until no longer needed for current business. Consult with the National Energy Information Center regarding records retention requirements.
 5. (Optional) Develop a policy for public dissemination of the archive that addresses such topics as:
 - a. The model's transportability.
 - b. Additional software required (such as proprietary embedded models) that is not provided by EIA with the model.
 - c. EIA's expectations for how a public user will identify the model if the model has been modified by a user from outside EIA.
 - d. Limits on EIA support.
 - e. Any other issues pertinent to outside use.

Related Information:

1. [2002-26](#), Model Documentation
2. [2002-28](#), Proprietary Models

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-28

Title: Proprietary Models

Superseded Version: 91-01-05

Purpose: To permit the use of proprietary models in EIA modeling systems and in conjunction with EIA products.

Applicability: To all models available to EIA through license, purchase, or subscription.

Required Actions:

1. Every agreement for the acquisition or use of a model should provide for:
 - a. Publicly available documentation of the model's design, theoretical basis, empirical implementation, and objective capabilities.
 - b. A means for EIA to replicate model calculations for a period of three years after each application in an EIA

product.

2. For an active EIA model:
 - a. The model documentation should be available to the public.
 - b. The model version and archive of all model inputs and outputs associated with a disseminated information product must be identified so the results can be replicated.
 - c. All changes EIA makes to the model should be documented and archived to EIA standards.
 - d. A proprietary model should not be embedded in an EIA modeling system unless the model is commercially available.

Related Information:

1. [2002-26](#), Model Documentation
2. [2002-27](#), Model Archival

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-29

Title: Public Comments on Compliance with Information Quality Guidelines

Superseded Version: None

Purpose: To ensure that users of EIA information have administrative procedures to request correction of information not in compliance with information quality guidelines established by the Office of Management and Budget (OMB), the Department of Energy (DOE), and the Energy Information Administration (EIA).

Applicability: All information prepared for public dissemination and disseminated on or after October 1, 2002, regardless of when the information was first disseminated.

Required Actions:

1. A person has a right to request, where appropriate, timely correction of information maintained and disseminated by EIA that does not comply with information quality guidelines issued by OMB, DOE, and EIA. A person requesting correction should be directed to [EIA's Information Quality Guidelines Web site](#) for details on information quality guidelines and the correction request procedure.
2. A requestor must specifically identify the information in question, explain with specificity the reasons why the information is inconsistent with applicable quality guidelines, the need for correction, and the type of correction requested. The correction process is designed to address the genuine and valid needs of EIA and its customers without disrupting EIA processes.
3. DOE's Chief Information Officer (CIO) is the point of contact for all public requests for correction of information disseminated by any DOE component. The DOE CIO will forward any request for correction of EIA information to EIA's Statistics and Methods Group (SMG). SMG will forward the request to the EIA office that disseminated the information. The program office must respond to a request within 60 days and provide a copy of the response to SMG and the DOE CIO. EIA's responses will be included in an annual report to OMB.
4. When considering a request for correction, EIA staff should determine what action, if any, is necessary.
 - a. If the information is in compliance with applicable information quality guidelines, the requester should be so informed.
 - b. If the information is not in compliance, the program staff will determine the appropriate level and method of

correction, considering the nature and timeliness of the information involved, significance of the error on the use of the information, the magnitude of the error, and the cost-effectiveness of undertaking a correction. The requestor should be informed of the action to be taken.

c. Upon deciding that information requires correction, the program staff shall provide notification of the intention to correct.

5. If a requestor does not agree with EIA's response to the request for correction, the person has a right to an independent administrative appeal process. A person requesting an appeal should be directed to [DOE's Web site](#). The DOE CIO is the point of contact for all appeal requests.

6. SMG will review and respond to appeals. If SMG was directly involved in responding to the initial request for correction and thus can not be independent in responding to an appeal, the appeal will be handled by EIA's Deputy Administrator.

Related Information:

1. [OMB's Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies](#)
2. [DOE's Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Department of Energy](#)
3. [EIA's Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Energy Information Administration](#)

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-30

Title: Business Process Documentation

Superseded Version: None

Purpose: To ensure that EIA's essential business processes are adequately documented so they can be operated efficiently and effectively by anyone with the necessary skills, information input and technical resources.

Overview:

Documentation requirements for specific portions of EIA software systems, surveys and models are delineated in several existing EIA Standards. This Standard covers the documentation requirements for the remaining portions of these processes and for EIA's other business processes not covered by existing standards (e.g., administrative, financial, customer service). The documentation will consist primarily of manuals that provide instructions and reference materials to individuals learning how to perform the function. These manuals would be used not only by staff newly assigned to the function, but also by veteran employees who must perform the function when the usual performers are unavailable. This might occur when the usual performer is absent from work, retires, or is unavailable during post-disaster recovery operations.

Applicability: All business processes that are essential to EIA's [Continuity of Operations Plan](#). There are no waivers to this Standard.

Required Actions:

This Standard requires the following:

- Designation of those business processes that are essential to EIA's [Continuity of Operations Plan](#).
- Creation and maintenance of Process Performance Manuals (may also be called User Manuals, User Guides, Operations Manuals, Process Manuals, Instruction Manuals, etc.) for all business processes designated as essential.
- A Process Performance Manual (or its equivalent) will contain the following information:
 1. A narrative description of the process, its purpose, operating environment, components, information inputs (and sources) and information outputs (and destinations);
 2. An overview chart showing the connectivities of the process components among themselves and with outside entities;
 3. A set of instructions suitable for an untrained but skilled individual that explains how each process component is performed;
 4. An optional troubleshooting guide and set of frequently-asked-questions that provide help to those encountering difficulty in performing the process; and
 5. Names of individuals to contact who can provide assistance in the performance of the process.

Related Information:

1. EIA Data Systems Standards, [2002-2](#), [2002-3](#)
2. EIA Models Standards, [2002-26](#), [2002-27](#), and [2002-28](#)
3. [EIA Software System Life Cycle Standard](#)

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-31

Title: EIA Continuity of Operations

Superseded Version: None

Purpose: To establish the minimum requirements for planning for continuing operations of the Energy Information Administration in case of emergency or catastrophe.

Overview:

The EIA [Continuity of Operations Plan](#) designates essential functions, provides for the emergency delegation of authority to carry out tasks normally performed by the Administrator, assigns management responsibility for operations under various contingencies, outlines alert and notification procedures and a communications plan, and categorizes vital records necessary to maintain operations or reconstitute agency programs. Lower level organizational planning must be consistent with the requirements of the EIA COOP Plan.

Applicability: All Offices in EIA.

Applicable Laws and Regulations:

- [Executive Order 12656](#), Assignment of Emergency Preparedness Responsibilities (statutory authorities include the National Security Act and the Federal Civil Defense Act)
- DOE Headquarters Continuity of Operations Plan

- EIA [Continuity of Operations Plan](#)
- 36 U. S. C., Chapter 12, Part 1236, Management of Vital Records
- National Archives and Records Administration Regulations, [Vital Records and Records Disaster Mitigation and Recovery](#)

Required Actions:

Each Office will create an implementation plan. These plans shall include, at a minimum, the following basic elements:

- Standing delegations of authority to act for management positions in the organization in the absence of managers.
- Designation of an assembly point for staff in the case of a required building evacuation, and the maintenance of a staff roster to enable the rapid identification of any missing staff.
- A method for the alert and notification of staff of the existence, nature, and required response actions appropriate for the emergency situation.
- An inventory of vital records, updated on a regular basis, to insure their availability as a part of the EIA vital records program.
- A description of the processes the Office will implement to accomplish its mission objectives in various emergency scenarios.

The Office of Information Technology will create a disaster recovery plan to reconstitute EIA's IT infrastructure in the event of an emergency or disaster, in support of EIA's essential functions.

The Office of Resource Management will create a plan to provide necessary administrative support (personnel, procurement, and finance) for the level of EIA operations conducted under emergency conditions.

These plans will be tested regularly and updated as necessary to insure their operability should an emergency situation arise.

Related Information: See applicable laws and regulations above.

Approval Date: September 26, 2002

ENERGY INFORMATION ADMINISTRATION STANDARD 2002-32

Title: EIA Records Management

Superseded Version: None

Purpose: To ensure that procedures and systems are implemented and maintained to assure efficient and effective records management.

Overview:

As the basic administrative tool by which the Government does its work, records are a component of each agency's information resources. Like other resources, records must be managed properly for the agency to function effectively and comply with Federal regulations. This document establishes required standards for activities with respect to records creation, maintenance and use, and disposition, which are essential to sound management of EIA business lines.

Applicable Laws, Regulations, Policies, And Procedures:

- 44 U.S.C. Chapter 31, [Records Management by Federal Agencies](#)
- 36 C.F.R. Part 1222, [Creation and Maintenance of Federal Records](#) (National Archives and Records Administration Regulations)
- [EIA Records Management Policies and Procedures Manual](#)

Applicability: This standard applies organization-wide to both program and administrative records. There are no waivers to this Standard.

Required Actions:

This standard requires the following actions:

1. EIA will establish a planned coordinated set of policies, procedures, and activities to manage the organization's records.
2. EIA will designate a Records Liaison Officer to provide overall policy guidance in records management on an organization-wide basis.
3. Each EIA office will designate an individual or individuals to act as Records Management Contacts and forward the routing symbol and telephone number of the designees to the National Energy Information Center, EIA-30.
4. Each EIA office will create and maintain a records disposition program designed to provide for (1) the removal from valuable office space all records and nonrecord materials which are no longer essential for current operations and (2) the protection of records that should be kept for long periods.
5. EIA will develop a policies and procedures manual governing the maintenance and disposition of all EIA records, including electronic media.
6. The EIA Records Liaison Officer will coordinate the retirement and retrieval of all EIA records to or from Federal Archives and Records Centers.
7. When planning and implementing all technology applications, EIA offices will include records management requirements in order to maximize the organization's ability to document its activities.
8. EIA offices will develop a standardized filing system or structure, which provides a records management mechanism that facilitates ease of records use, access, and disposition.

Related Information:

1. This standard is consistent with the [EIA Software System Life Cycle Standard](#).

Approval Date: September 26, 2002

[Standard 2002-4](#) Supplementary Materials, Forms Design Checklist

A. Data Requirements and Pretesting

1. Each data element is needed to support program goals and does not unnecessarily duplicate information available elsewhere.
2. Items requested can be readily reported by all or most respondents.
3. Sufficient response time is allowed for respondents to supply reliable data.
4. Methodology is appropriate (e.g., mail, telephone, electronic, or personal interview) to complexity and amount of data requested, and funding available.
5. Electronic information collection techniques are considered as a means of reducing respondent burden and government cost.
6. Pretesting is conducted when prior experience cannot be used to determine feasibility of the form.
7. The units (companies, households, etc.) for which data are to be reported are clearly established through appropriate preprinted items, questions and/or at the beginning of the form.

B. Question Content and Wording

1. Question structure (e.g., open-ended, multiple choice) is the most suitable for information requested.
2. Questions do not lead to biased responses.
3. Wording used is based on understanding of the respondents.
4. Questions are as self-explanatory as possible, reducing the need for additional instructions.
5. Double or compound questions are not used.
6. Only approved abbreviations (see [Standard 2002-16](#)) and units of measure are used.
7. Applicable conversion factors are included.
8. Reference periods or dates are clearly specified with frequent reminders provided.
9. Overall logical sequence and transitional phrases are used.
10. Questions on same general topic are grouped together.
11. All items on form are numbered.
12. Skip patterns ensure only appropriate questions are asked of each respondent.
13. Respondent should not be required to code a large (excessive) number of items.

C. Forms

1. Legal authority as the basis for the data collection is cited in on the first page of the form.
2. Filing requirement (voluntary, required to obtain or retain benefits, or mandatory) is cited on the first page of the form.
3. Statement regarding data confidentiality is included on the first page of the form. (This may be a reference to a detailed discussion in the instructions for the form.)
4. Office of Management and Budget (OMB) control number and expiration date are placed in upper right-hand corner of the first page of the form.
5. EIA's estimate of the average burden hours per response is provided. In addition, include a request for comments concerning the accuracy of the burden estimate and any suggestions for reducing the burden. Contact EIA's Clearance Officer, [Jay Casselberry](#), for the wording to be used and additional information regarding the placement on the form. (This may be included in the instructions instead of on the form.)
6. Form number, edition date, and superceded notice are located in the upper-left corner of each page of the form. (These items need not appear on the reverse side of a page.)
7. Request contact person information and signature, as appropriate.
8. Certification may be used consistent with the following EIA policy:
Surveys conducted to collect data intended to be used for regulatory purposes and surveys where the underlying statute(s) require certification of the information reported shall include certification sections. For any other survey, certification is not necessary unless the EIA Administrator approves an exception to this practice. If need is perceived for exception(s) to this policy by EIA Office Directors, each will notify the Administrator in writing of any surveys deserving certification and why. The Administrator will determine which of these

requires certification.

9. Each EIA form and instructions should inform respondents of the implications of reporting false, fictitious, or fraudulent information to a U.S. government agency. The suggested wording (to be printed in bold font) is: **Title 18 U.S.C. 1001 makes it a criminal offense for any person knowingly and willingly to make to any Agency or Department of the United States any false, fictitious, or fraudulent statements as to any matter within its jurisdiction.**
10. Skip patterns are clearly identified.
11. As much as possible, include item instructions next to the item or elsewhere on form, as opposed to appearing on a separate sheet or instructions.
12. Form should be available in an electronic format.
13. Sufficient space is provided for responses and for coding, as required.
14. Form design facilitates data entry without reducing ease of completion by respondents.
15. Use of check boxes and preprinted data is maximized to reduce respondent reporting burden.
16. Paper color, type size, and ink maximize readability.
17. When designing forms, 8 1/2" x 11" paper is preferable for both respondent and EIA use. Also, that size is the only acceptable size in all Federal Courts, effective January 1, 1983.
18. Complex skip patterns should be avoided, especially for self-administered forms.
19. Bar codes on survey forms may be used to facilitate the log-in and initial processing. Those surveys with bar codes should use the Code 39 (also known as 3 of 9) method unless there is a compelling reason to do otherwise.
20. Options for respondents to submit forms to EIA are consistent with the Draft EIA Policy on Security of Sensitive Unclassified Survey Data Transmitted to EIA
 - EIA will provide secure electronic methods for its survey respondents to use when they transmit sensitive unclassified survey data to EIA.
 - EIA will inform its survey respondents about the incoming data transmission methods that EIA can accommodate and which are secure.
 - EIA Office Directors may request waivers to this policy. Waiver requests will be reviewed by the Information Technology Council and decided by the Administrator.

D. Instructions

1. OMB control number and expiration date are placed in the upper-right corner of the first page of the instructions.
2. Form number, edition date and superceded notice appear in the upper-left corner on each sheet of instructions. (These items need not appear on the reverse side of a page.)
3. When placed on a separate page, aligning instructions in two or more narrow columns rather than on full-width lines and as numbered items rather than paragraph style make it easier for respondents.
4. Purpose of form is clearly stated.
5. Description of how data are to be used is included in purpose.
6. Clear, specific statement of when the form should be filed and where it is to be sent are included.
7. Clear, specific statement of who is required to file the form is included.
8. Clear, specific statement of what to file is included.
9. Legal authority, filing requirement (voluntary, required to obtain or retain benefits, or mandatory), and data confidentiality provisions are clearly stated.
10. Phone number and e-mail address are provided for respondent inquiries.
11. All needed definitions are provided. When possible, definitions should be grouped together in one section of the instructions.
12. Unless a waiver has been granted, all terms and definitions will be consistent with the [EIA's Energy Glossary](#).
13. Instructions are provided for each item on the form where necessary to avoid ambiguities.
14. Instructions for specific questions are numbered to correspond to the questions they explain.

15. The interview format clearly distinguishes (a) questions for respondents, (b) instructions or transition phrases to be read by/to respondents, and (c) instructions to interviewers.

16. The interview format for questions with response categories are shown on the form: There should be an instruction to the interviewer as to whether (a) these categories should be read or shown to the respondent; or (b) the question is to be left open and the interviewer is to assign the response to the appropriate category.

17. Each EIA form and instructions should inform respondents of the implications of reporting false, fictitious, or fraudulent information to a U.S. government agency. The suggested wording (to be printed in bold font) is:

Title 18 U.S.C. 1001 makes it a criminal offense for any person knowingly and willingly to make to any Agency or Department of the United States any false, fictitious, or fraudulent statements as to any matter within its jurisdiction.

18. EIA's estimate of the average burden hours per response is provided. In addition, include a request for comments concerning the accuracy of the burden estimate and any suggestions for reducing the burden. Contact EIA's Clearance Officer, [Jay Casselberry](#), for the wording to be used and additional information regarding the placement on the form. (This may be included in the instructions instead of on the form.)

[Standard 2002-10](#) Supplementary Materials, Developing A Survey Data Evaluation Plan and Suggested Approaches for Evaluating Different Types of Survey Data Nonsampling Error

Developing A Survey Data Evaluation Plan

Consider the following actions with respect to developing a survey evaluation plan as required in [Standard 2002-10](#):

1. Review past surveys with similar data to determine what survey evaluation data have been collected and any issues related to the comparison. This review should document what is known about the sources and magnitude of nonsampling error and issues that should be addressed in future survey evaluations.
2. Plan how each priority survey evaluation issue will be addressed. This should include assessing:
 - a. What data will be used (both from within the survey and from outside sources).
 - b. How any outside source data will be obtained.
 - c. What comparisons will be made.
 - d. What other types of evaluations of the survey will be conducted.
3. During the design phase for a survey, identify each potential source of nonsampling error and, where possible and appropriate, develop methods for bounding or estimating the error from each source.
4. For recurring surveys, develop an error profile document itemizing all sources of error. Where possible, estimates or bounds on the magnitudes of these errors should be provided, the total error model for the survey should be discussed, and the survey should be assessed in terms of this model.
5. In addition to identified quality issues, consider other issues arising during the course of the survey and, where appropriate, collect and analyze data for assessment.
6. Where feasible and appropriate, analyze information from any evaluation prior to or concurrent with the analysis of the survey data so that the results of the evaluation can be taken into account when analyzing, interpreting, and disseminating the survey data.
7. Provide a brief summary of any evaluation's findings periodically to users.
8. As part of the evaluation activity, prepare written recommendations on what should be considered for future evaluations.

Suggested Approaches for Evaluating Different Types of Survey Data Nonsampling Error

The following approaches may be useful in survey data evaluation topics as discussed in [Standard 2002-10](#):

Coverage - adequacy of frame

- Matching studies to earlier versions of the same data source or to other data sources
- Comparison of a very reliable external source for some subset of the current population to the current frame
- Analysis of survey returns for deaths, duplicates, changes in classification, and out-of-scope units
- Research, such as reviews of trade association lists, trade publications, telephone listings, Web sites, and State government lists
- Comparison of estimated totals with estimated totals from another source.

Measurement

- Pretests to determine the efficacy of devices to improve response (e.g., reporting options, nonresponse follow-up methods, incentives)
- Pretests to compare alternative ways of collecting or processing data
- Comparison of final design effects with estimated design effects used in survey planning
 - In a longitudinal survey, establishment of a small independent sample for use as a test group to evaluate the effect(s) of changes
 - Discussions (e.g., cognitive interviews) with respondents to assess survey comprehension and task (i.e., completing the survey) performance
 - Site visits to examine administrative records, to compare information in respondents' records to survey responses, and to discuss discrepancies with respondents
 - Comparison with outside data sources measuring the same concept
- Examination of changes in response over repeated questioning
- Agreement among responses to different surveys (e.g., weekly versus monthly, or monthly versus annual)
 - Agreement of statistics derived from different sections of the questionnaire or different questionnaires.

Nonresponse

Unit Nonresponse:

- Description of characteristics of units not responding. Examine changes in response rates over time and/or conduct a follow-up of a sample of nonrespondents to ascertain nonresponse bias
- Assessment of method of handling nonresponse, including imputation, estimation, or weighting procedures

Item Nonresponse:

- Description of items not completed, patterns of partial non-response, characteristics of companies failing to respond to certain items or groups of items.

Imputation and Estimation

- Analysis of the imputation process, including frequency of imputation, initial and final distributions of the variables
- Examination of the choice of estimator/design
- Possibility of fitting survey distributions to known distributions from other sources to reduce variance and bias
- Re-estimation using alternative techniques, such as alternative outlier treatments, alternative imputation procedures, and alternative variance estimation techniques
- Use of generalized variance curves
- Effect of changes in data processing procedures on survey estimates.

Revisions

- Examine effect of revisions (number of times data are revised and the magnitude of the revisions).

[Standard 2002-11](#) Supplementary Materials, Additional Suggested Data Quality Measures, and Periodic Quality Reviews

Additional Suggested Data Quality Measures

The following suggested data quality measures may be useful for evaluating and improving the quality of EIA survey data in conjunction with [Standard 2002-11](#). As appropriate and based on available resources, the survey staff should consider collecting information to calculate the following measures:

1. Response rates by respondent characteristics (e.g., size, geography)
2. Respondent-level reporting problems (i.e., nonresponse, late response, and/or large amounts of data needing revision)
3. Editing impact (i.e., net impact of editing on aggregate disseminated data)
4. Suppression (i.e., number of data items suppressed)
5. Data comparisons (i.e., comparison of survey data to external data (e.g., comparison with data collected by other public or private entities) and comparison of respondent-level data from different surveys (e.g., comparison of monthly versus annual reports for the same data items)).

Periodic Quality Reviews

In addition to data quality measures, an additional action to monitor survey quality is to conduct a *Periodic Quality Review*. The purpose is an appraisal of a survey. Although qualitative information is not always easy to summarize, it may include narrative descriptions on such topics as the following:

1. Frame coverage (i.e., assess the coverage of the frame, including the procedure used to estimate the coverage along with when and how the frame is updated; it may not be possible to estimate the coverage for all surveys)
2. Staffing (i.e., assess the experience level of the staff and their training)
3. Procedures (i.e., assess the data collection procedures, any cognitive testing or pretests conducted, and the dates of the steps in the data collection process)
4. Methods (i.e., assess sampling, editing, imputation, and estimation procedures, including any known biases. Also, consider any poststratification, any unusual definitions, or coding issues).

[Standard 2002-15](#) Supplementary Materials, Guidelines on the Standard for Rounding

1. Rounding - definition of significant digits.

For numbers over one, significant digits are defined as the number of digits kept after rounding. For example, if rounding to the nearest thousand, the number 22,000 has 2 significant digits. When rounding to the nearest tenth, 5,243.4 would have 5 significant digits. For numbers less than one, the number of significant digits is the number of

digits reported to the right of the last zero digit following the decimal. For example, .000041 has 2 significant digits.

2. Rounding - effect of multiplying rounded numbers.

Suppose two rounded numbers contain a different number of significant digits, say n_1 and n_2 digits where n_1 is less than n_2 . Then the product of the two numbers can only be assured of being accurate to n_1 digits.

For example, suppose data are collected on the number of gallons of oil that a company delivered in a week. These data were rounded to the nearest gallon and contained 5 significant digits; e.g., 26,579 gallons. If 26,579 is multiplied by a weight such as 21, which was rounded to the nearest ten thousandth whole number, the product 558,159 should be stated with 2 significant digits; i.e., as 560,000. This is because 21 was the rounded number with the fewest number of significant digits in the multiplication (2 significant digits). If accuracy is needed beyond 2 significant digits, such as 5, use a weight which has 5 significant digits; e.g., 20.983.

Conversely, if the weight 21 is an exact number, the product can be stated with the number of significant digits in the approximate number (26,579) which is 5; i.e., the product should be stated as 558,160.

3. Rounding - how much?

When considering all sources of error that may be involved, it is unlikely that any energy datum at the national level is accurate to more than three significant digits. Nevertheless, it is often necessary to present more digits in totals to avoid having empty cells in the body of the table. The problem of implying a spurious degree of accuracy must be balanced against the need for consistency within and between tables.

4. Rounding - mechanics of applying standard. Use the following procedure to round to the nearest 10^n :

- a) Divide the number by 10^n
- b) Add .5 to the number obtained above, for positive numbers, and add -.5 to the number obtained above for negative numbers
- c) Truncate the decimal part of the number obtained in step 2
- d) Multiply by 10^n .

Note: for decimals, n is negative; e.g., to round to 2 decimal places (hundredths), $n = -2$. For whole numbers, n is positive; to round to thousands, $n = 3$.

Standard 2002-20 Supplementary Materials, Quality Assurance Review Guidelines

Quality Assurance Reviews ensure the objectivity, utility, and integrity of information products. The author should conduct a quality assurance review before asking for reviews by other persons and/or prior to dissemination. These provide an opportunity for a thorough product review in order to ensure quality including conformance to standards.

Statistical Review

The statistical review covers a check of the completeness and accuracy of the description of the survey design, including a description of the collection process and form(s); sample design, if applicable; the mathematical formulae reflecting the estimation procedures for both sampling and non-sampling errors, as available; and the data processing procedures and quality control procedures in data processing. The statistical

review also checks all statistical inferences made in the text, the proper application of the guidelines for disclosure avoidance, and the proper use of rounding. Items that should be reviewed include:

Background and Explanatory Information Review

1. Basic Definitions and Terminology
 - a. The time period(s) to which the data pertain are specified.
 - b. The time period during which the data were collected are specified.
 - c. All terms and concepts critical to understanding the information are defined.
 - d. Codes, abbreviations, acronyms, and definitions conform to Standard 2002-16.
2. Sample Design
 - a. The target population is described.
 - b. The name of the organization that collected (supplied) the information is provided.
 - c. The survey frame is described with regard to source of the frame, reference date, number of units, and an assessment of the frame's quality.
 - d. The units selected for the sample at each stage are described.
 - e. Any stratification and clustering procedures are described.
 - f. The number of sampling units selected at each stage is provided.
 - g. The procedures for allocating the sample size at each stage are described.
 - h. The sample selection process is described, i.e., the selection method at each stage.
 - i. Measures of size stated for sampling with "probability proportionate to size" and types of estimates that are likely to have improved precision as a result are discussed.
3. Data Collection Process
 - a. A general outline of the methods used to collect the data, including mailing and/or interviewing, editing, and other procedures is provided.
 - b. Relevant portions of the data collection forms and their instructions are provided or are available on EIA's web site.
 - c. The quality control procedures used in the collection process are described.
4. Data Reduction, Estimation, and Error
 - a. The derivations of the base weights and any special weights are specified.
 - b. A description of the procedures used to adjust for nonresponse is provided.
 - c. Descriptions of any other adjustments and all imputations are provided.
 - d. Computing algorithms and a specification of the measures of reliability used are provided.
 - e. Descriptions of nonsampling error and other potential biases are provided.
 - f. Sampling error can be calculated for all estimates in the product. For example, if tables show estimates at the State level, the user can calculate sampling error at the State level.
5. Appraisal of the Data
 - a. Response bias and measurement error are discussed, including the response rate or information that could be used to compute it.
 - b. Conceptual and other limitations of the data are pointed out.
 - c. Explanations are given of how the limitations of the data are dealt with in the publication.
 - d. There are discussions of the results of data evaluations and other quality assurance activities. (This may include references to other publications, reports, technical papers, etc.)
6. Revisions
 - a. The process for revising data and/or estimates is described, including the schedule for revisions and any thresholds for unscheduled revisions.
 - b. There is a discussion of how future planned revisions could affect data interpretation.
 - c. Where the difference between old and revised (or new) data series causes discontinuity, the difference is discussed and quantified, if possible.
 - d. The time period covered by the revision is specified as well as how far back the data are revised.
 - e. If there is a difference in period-to-period change as the result of revision, it is fully explained.

Text Review

1. The text should present the subject matter in a well-organized and logical manner.
2. Information in the text, tables, and graphs is consistent. If not, there is an explanation.
3. Units of measure are specified.
4. The numbers and conclusions should be consistent with those in earlier EIA products, unless the present information represents a revision of these earlier products.
5. The text statements should be supported by the data. For example, for data based on sample surveys, statements should be written in terms of confidence intervals or have been tested for statistical significance.

Table and Graph Review

1. In the tables, totals should equal the sum of components within the accuracy appropriate for independent rounding.
2. The same data elements should be consistent between tables.
3. All tables should be self-sufficient and capable of being abstracted directly from the product and so must cite the source of the data, whether collected by EIA or not.
4. All tables and graphs that display data collected on EIA survey form(s), whether the data are previously published or not, cite the EIA form number(s). (Reference to another table or graph in another publication is not a proper citation.)
5. If a publication is based on several survey forms, the reader should be able to promptly and clearly determine which forms correspond to which tables and graphs.
6. All tables and graphs that display data not collected by EIA, whether the data are previously published or not, cite the source of the data.
7. The numbers should be consistent with those in previously disseminated EIA products or the differences should be highlighted.
8. The text, tables, and graphs clearly distinguish among actual data, estimated data, forecasts, projections, preliminary data, revised data, and final data.
9. Units of measure should be specified.
10. Codes, abbreviations, acronyms, and definitions in tables and graphs should satisfy the requirements in [Standard 2002-16](#).
11. Presentation should satisfy the requirements in [Standard 2002-17](#).
12. Graphs should satisfy the requirements in [Standard 2002-25](#).

Analytical Reports and Feature Articles Review

Analytical reports and feature articles require the following additional steps for review:

1. The executive summary of the analysis report must accurately reflect the findings found in the body of the report.
2. The report's purpose, assumptions, methods, findings, and uncertainties must be stated clearly and precisely.
3. The analytical results must support the findings and conclusions.
4. The analytical results may need to be verified; thereby, specifications must be either accessible from the product or in the user's report.
5. For those products whose analytic findings are supported by information from models, the results of sensitivity analyses, if any, conducted for important model parameters should be accessible from the product
6. Data sources and model documentation should be referenced. The model documentation should be consistent with the version of the model used and meet EIA standards for model documentation.
7. An analytical report should satisfy the requirement for reproducibility; i.e., the information is capable of being

substantially reproduced subject to an acceptable degree of imprecision. Substantially reproduced means that an independent analysis of the original or supporting data would generate similar analytic results, subject to an acceptable degree of imprecision or error. To ensure that the reproducibility requirements are satisfied, the author should ensure that the appropriate documentation is available and the necessary data and program files are archived.

Standard 2002-22 Supplementary Materials, Guidelines for Implementation of a Disclosure Limitation Rule

These guidelines assist in understanding and implementing disclosure limitation procedures using the PQ rule, as specified in [EIA Standard 2002-22](#).

- [Section 1](#) presents a background introduction to disclosure limitation.
- [Section 2](#) describes the PQ rule and how to use it.
- [Section 3](#) presents a discussion of the impact of parameter selection and describes the use of a combination rule in special situations.
- [Section 4](#) provides a discussion of complementary suppression.
- [Section 5](#) provides guidelines for applying disclosure limitation when imputation for nonresponse is performed.
- [Section 6](#) describes disclosure limitation when reported data are negative.
- [Section 7](#) describes procedures to use when an item is the difference between two positive reported items.
- [Section 8](#) describes procedures to use when an item is the weighted average of positive reported quantities.
- [Section 9](#) discusses the complexities of disclosure limitation if the data to be disseminated are based on a survey where respondents could waive their pledge of confidentiality.
- [Section 10](#) provides additional details on the PQ rule.
- [Section 11](#) is an illustration of use of the PQ rule.

1. Introduction

Statistical disclosure is the release of confidential information by the cross-tabulation of microdata. This information is in the form of an attribute which may be uniquely identified with a particular responding unit.

Release may be exact or approximate. Two examples:

- If a cell was published based on the survey response from only one company, a user would be able to determine what the company reported to EIA.
- If a cell was published based on survey responses from two companies, each would be able to use the published total to exactly determine the other's value.

Generally, when more than two companies contribute to a cell, the total merely allows a company to form estimates for the other companies' values.

There are two types of disclosure possible, internal and residual. Internal disclosure occurs when members of a cell use their own data and the cell value to obtain confidential information about others in the same cell.

Residual disclosure involves the mathematical manipulation of the data from other cells to obtain the confidential information in the sensitive cell.

Sensitive cells are cells which, if released, would disclose confidential information. Such cells are usually identified by the use of a sensitivity rule. Sensitivity rules determine whether or not internal disclosure would occur and the extent of the disclosure.

In EIA, the procedure for preventing disclosure is to withhold sensitive cells from release. Cells which have been identified as being sensitive by use of a sensitivity rule and are suppressed for that reason are called **primary suppressions**. If totals are to be released and any cell requires primary suppression, then certain other

cells in the table must also be suppressed to prevent residual disclosure. This type of cell suppression is called **complementary suppression**.

The appropriate application of disclosure limitation requires identification of cells which require primary suppression, and additionally, identification of a set of appropriate cells for complementary suppression.

2. The Recommended Primary Suppression PQ Rule and Its Use

The primary suppression rule which is recommended for use in EIA is the PQ rule, one of the so called "priori-posteriori ambiguity rules." It was selected because it satisfies the required property of subadditivity, and provides protection regions. Subadditivity guarantees that if two cells are both nonsensitive, their sum is also nonsensitive. This sensitivity measure is applicable only when publishing totals of nonnegative reported values and is necessary when the reported values are volumetric data, such as production, stocks, or sales volumes. Volumetric data have the greatest potential for disclosure because they can show the greatest diversity among the reported values and, hence, the total may be close to the reported value of a single company or set of companies.

Follow the steps below to use the PQ rule for primary suppression.

- a. Choose the value of q/p . ([Section 3](#) discusses selection of the value for q/p .)
- b. Apply the following formula to each non-zero cell value to determine $S(x)$, the sensitivity measure:

$$S(x) = x_1 - (q/p) * (T - x_1 - x_2)$$

- x_1 is the largest reported respondent value included in the total cell value.
 - x_2 is the second largest reported respondent value included in the total cell value (x_1 can be equal to x_2).
 - T is the total cell value. In a census survey it is the sum of the values reported by respondents; in a sample survey it is the weighted sum of the reported values.
- c. The cell is sensitive if $S(x)$ is nonnegative. If a cell is sensitive, complementary suppression will also be necessary. See [Section 5](#) for a discussion of complementary suppression.

An illustration of the use of PQ rule is provided in [Section 11](#).

3. PQ Selection Criteria and Uses of a Combination Rule

PQ, the input sensitivity parameter, represents the maximum permissible gain in information when one company uses the published total and its own value to create better estimates for other companies' values.

PQ must be greater than one, but need not be an integer. The particular choice of PQ is to be made by the office disseminating the data. The particular value selected and applied should be considered as confidential.

When selecting a value of q/p , take into consideration that as the value of q/p increases, less data cells will be classified as sensitive and withheld from publication. However, as the value of q/p increases, the potential for a user to more closely approximate the information reported by a survey respondent increases and threatens

EIA's promise of confidentiality. This is demonstrated in [Section 10](#).

Larger values of q/p permit release of cells where x_1 or $x_1 + x_2$ represent larger percentages of the total. That is, larger values of q/p permit more information to be gained by releasing the total, T . As an example, if x_1 accounted for 85 percent of the total, x_2 accounted for 5 percent of the total, and $q/p = 9$, then $S(x) = -.05 * T$ and the cell would not be classified as being sensitive.

If an office disseminating information prefers to use a large value of q/p , the PQ rule alone may permit too much disclosure when the cell is dominated by one large company, as in the above example. However, a combination rule may be drafted to overcome this problem.

A combination rule is defined to be the maximum of the sensitivity measure defined for the PQ rule and the sensitivity measure defined for some other subadditive rule. An example of another subadditive rule is the n,k rule. In the n,k rule, the cell is sensitive if the largest n companies account for k percent or more of the total. If $n=1$ and $k=80$, for example, the sensitive measure would be:

$$S_1(x) = x_1 - .8 * T,$$

and, as before, the cell is sensitive if $S_1(x)$ is nonnegative.

The combination rule is subadditive since the maximum of two subadditive sensitive measures is subadditive. In this example the sensitivity measure would be:

$$S^*(x) = \max(S(x), S_1(x)).$$

In the above numerical example, $S(x) = -.05 * T$, and $S_1(x) = .05 * T$. Thus, the maximum is $.05 * T$, and the cell is sensitive.

4. Complementary Suppression

Determining the optimal pattern of cells for complementary suppression is a complicated procedure because it potentially requires a search over all possible cells to select the fewest number of cells, with the smallest possible total value, which adequately protects the cells requiring primary suppression.

An office disseminating data may choose to perform complementary suppression manually, using industry knowledge, and maintaining the same or similar patterns of suppression from one release to the next.

Implementation of a disclosure limitation procedure should be augmented by an audit feature which permits an evaluation of proposed patterns of suppression to assure that there is no residual disclosure.

5. Treatment of Imputation for Nonresponse in Disclosure Analysis

There are two general cases for the application of disclosure analysis when imputation procedures are used to adjust for nonresponse:

- 1 . Imputed values are based on the other respondents' data, as in adjusting weights or "hot decking," and
2. Imputed values are based on data submitted by the nonresponding company in a previous time period.

In the first case, there can be no disclosure of values associated with the nonresponding companies. Only the reporting companies are at risk. To assess this risk, the imputed values are included in the total, T , but the

imputed values are not counted as reported values for identification of the largest two companies. (Recall that in applying the sensitivity measure to sample surveys, T is calculated based on the weighted total, and x_1 and x_2 are defined based on actual unweighted survey responses.)

In the second case, the theoretical justification for the imputation procedure is that there is a high correlation between values reported by the same company in different time periods. Hence, this type of imputed value contains sensitive data for that nonresponse. Thus, the imputed data should be treated as reported data for purposes of disclosure analysis. That is, the imputed values should be included in the total, T , and should be considered as reported values when selecting the largest two values in the cell.

6. Negative Reported Values

If all reported values are negative, apply the PQ rule to the absolute values of the reported data. Complementary suppression must be applied if marginal totals are released.

7. Differences of Positive Reported Values

If the published item is the difference between two positive reported quantities (e.g., net production equals gross production minus inputs), then apply the PQ rule as follows:

- If the resultant difference is generally positive, as is the case for net production of distillate fuel oil, apply the PQ rule to the first item (gross production in the above example). Suppress cells for which the first item is sensitive.
- If the resultant difference is generally negative, as is the case for net production of an item which is used primarily as inputs, apply the PQ rule to the second item (inputs in this example). Suppress cells for which the second item is sensitive.
- If the difference can be positive or negative, and is not dominated by either, ensure that there are at least three nonzero respondents contributing to each cell.

Complementary suppression (see [Section 5](#)) must be applied if marginal totals are released.

8. Weighted Averages

If a released item is the weighted average of two positive reported quantities (such as volume weighted price), apply the PQ rule to the weighting variable (volume in this example).

Suppress the average cell if the weighting variable is sensitive. It is not necessary to use complementary suppression on the average variable.

Both primary and complementary suppression must be applied to the weighting variable, if it is also a released item.

9. Disclosure Limitation for Data Based on A Survey Using Confidentiality Waivers

It is possible for a survey to allow respondents to sign a form waiving the pledge of confidential treatment for information reported to EIA. Disclosure limitation for data from such a survey is complicated because an aggregate value may be based on a combination of information, including reports by respondents not waiving confidentiality as well as respondents that waived confidential treatment. When using disclosure limitation techniques to determine whether a cell value is sensitive and must be withheld, consideration is given to the respondents that have not waived their confidentiality. If waivers are used, consult with the Director of the Statistics and Methods Group, [Nancy Kirkendall](#), on the application of disclosure limitation methods.

10. Further Discussion of PQ Rule

Although the PQ rule does not exactly correspond to any of the rules based on the contribution of the largest company or the sum of the largest two companies, it is possible to calculate the percent contributions of x_1 and $x_1 + x_2$ for which suppression is guaranteed by different values of q/p . We use the following expression for the region where $S(x)$ is positive (the region of sensitivity).

$S(x)$ will be positive and disclosure will occur for any x_1 and x_2 such that:

$$(x_1 + x_2) * q/p + x_1 > T * q/p.$$

Since $(x_1 + x_2) > x_1$ for all x_2 , we know that if

$$x_1 * q/p + x_1 > T * q/p$$

then $(x_1 + x_2) * q/p > T * q/p$, and the cell is sensitive. Thus, if the largest cell contributes a larger proportion to the cell total than $(q/p)/(1+q/p)$, the cell will be sensitive.

Similarly, the smallest value of $(x_1 + x_2)$ which guarantees disclosure, regardless of the contribution of the other companies, is found at the intersection of the disclosure boundary with the line $x_1 = x_2$. Thus, substituting $x_2 = x_1$ into the inequality and solving for x_1 yields $x_1/T > (q/p)/(1+2*q/p)$. Rewriting this in terms of $(x_1 + x_2)$ gives:

$$(x_1 + x_2)/T = 2 * x_1/T > 2*(q/p)/(1 + 2*(q/p)).$$

Thus, if $(x_1 + x_2)$ contributes a larger fraction to the total than $2*(q/p)/(1+2*q/p)$, then the PQ rule will determine that the cell is sensitive.

As the value of q/p chosen as the input sensitivity measure increases,

- less data cells will be classified as sensitive and withheld from publication, and
- the potential for a user to more closely approximate the information reported by a survey respondent increases.

To demonstrate this, the following table summarizes results for selected values of q/p .

EXAMPLE – Based on different values of q/p , a cell will be suppressed by the PQ rule if x_1 , or $(x_1 + x_2)$ exceeds the following percentage of the total:		
If the input sensitivity parameter q/p is chosen as:	A cell would be withheld if the respondent with the largest value (x_1) exceeded this percentage of the cell total:	A cell would also be withheld if the two respondents with the largest values ($x_1 + x_2$) exceeded this percentage of the cell total:
2	66.7%	80.0%
3	75.0	85.7
6	85.7	92.0
9	90.0	94.7

Note that larger values of q/p permit release of cells where x_1 or $x_1 + x_2$ represent larger percentages of the total.

That is, larger values of q/p permit more information to be gained by releasing the total, T . As an example, if x_1 accounted for 85 percent of the total, x_2 accounted for 5 percent of the total, and $q/p = 9$, then $S(x) = -.05 * T$ and the cell would not be classified as being sensitive.

If an office disseminating information prefers to use a large value of q/p , the PQ rule alone may permit too much disclosure when the cell is dominated by one large company, as in the above example. However, a combination rule may be drafted to overcome this problem. (See [Section 3.](#))

11. Illustration of PQ Rule Using Hypothetical Survey Data

A table has the cell shown below:

Coal Production by State (in thousand short tons)	
State	Production
Kansas	317

The 317 thousand short tons of coal production in Kansas represent the following values reported on an EIA survey:

Company A	230
Company B	62
Company C	11
Company D	10
Company E	4

The two largest companies (A and B) are responsible for over 92% of the Kansas coal production of 317 thousand short tons.

Assume the value chosen for q/p is 3.

Applying the PQ rule.

$$S(x) = x_1 - (q/p) * (T - x_1 - x_2)$$

$$S(x) = 230 - (3) * (317 - 230 - 62)$$

$$S(x) = 230 - (3) * 25$$

$$S(x) = 230 - 75$$

$$S(x) = 155$$

Because $S(x)$ is nonnegative, the total coal production for Kansas would be considered sensitive and would be withheld from dissemination. The cell in the EIA information product would appear as:

Coal Production by State (in thousand short tons)	
State	Production
Kansas	W

W = Withheld to avoid disclosure of individual company data.

Complementary suppression would also be necessary.

[Standard 2002-25](#) Supplementary Materials, Guidelines for Graphs

Titles

1. Use a self-explanatory title that includes the message of the graph.
2. Place the title flush left or centered.

Layout

1. Avoid clutter that does not add anything to the message (e.g., too many arrows, bubble boxes, grid lines, extra tick marks, and other non-data features).
2. Avoid using line graphs with more than five lines and pie charts with more than six slices.
3. When possible, order values from highest to lowest or lowest to highest if the data are not a time series.
4. For color graphs, use different colors or different patterns for different data series, but make sure the colors can be differentiated when viewed in black and white. See [Colors](#) below.

Labels, Legends, and Lines

1. Include data and labels directly on the graph when possible and avoid legends, except for use on maps. Labels make legends unnecessary and it is much easier for a reader to associate the lines/bars/slices with the associated text.
2. When possible,
 - Put data values beside, above, or inside bars.
 - Label individual slices of pie charts and show values.
 - Label lines on line charts.
3. Lines should be greater than .216 points wide to be readable (especially electronically).
4. Use different weights (thickness) or colors of lines for different data series.
5. Do not use large symbols or patterns to draw lines.

Scales

1. Within an information product, use the same scale whenever appropriate so that graphs can be compared.
2. In general, avoid comparing items measured in different units in the same graph and use only one vertical scale for plotting. This does not preclude the insertion of a second scale that is a linear derivative of the first; e.g., where

[1]

the same entity can be expressed in different units, such as miles and kilometers. Testing has shown users have difficulty understanding graphs with a dual y-axis.

3. For a dual y-axis graph, labeling is extremely important. Users participating in tests had difficulty associating data lines with the appropriate axes, particularly when the two y-axes represented the same variable but at different

[2]

orders of magnitude. However, users found the dual y-axis graphs to be useful in conveying the relationship between price and volume when the graphs were carefully constructed and labeled. If a dual y-axis graph will be used, consider:

- Using different chart types for the variables, such as bars for volumes and lines for prices.
- Using different colors for each variable and associate that with the color used for the axis label to help users determine which y-axis to use. For example, graph price with a red line and show the price axis in red while graphing the volume line in blue and showing the volume axis in blue.

Cumulative (stacked) graphs

1. Avoid using cumulative (stacked) line graphs or stacked bar graphs. When using cumulative bar graphs, include the data values or percentages for each bar segment.
2. Stacked graphs are primarily used when the focus is only on the “total change over time” and
 - The components do not exhibit marked irregularities,
 - The components do not exhibit seasonality, and
 - The components do not exhibit any sharp upward or downward trends.
3. With stacked graphs, users tend to add the various bar segments when trying to estimate the value, rather than visualizing the trends of the stacked bars separately. Discerning trends is much easier with grouped (side-by-side)

bar charts than with stacked bars.

Three-dimensional graphs

1. Avoid using 3-D graphs when they might impede the user's understanding of the information. When a 3-D graph is used, include the data values with the bars/pie slices.

Underlying Data

1. The data for the graph should be available in a corresponding table, whenever possible. For example, a graph on EIA's Web site could have a link at the bottom to the data.

Colors

1. To ensure that graphics are understandable and any colors can be differentiated when printed on black and white printers, check each graph before including it on EIA's Web site.
2. Use white or very light colored backgrounds.
3. Use dark colors for titles and labels.
4. Avoid light pastels.
5. Refer to the additional color specifications in the [EIA Style and Standards for Electronic Products](#).

Software

1. The preferred graphic software for EIA web products is Microsoft Excel.

Standard 2002-26 Checklist A. Explanatory Model Documentation Components

Elements are required unless "optional" is specified. Materials need not be presented in the order discussed here.

- a) A reference to the appropriate archive package.
- b) Model overview: A concise description of the model, its purposes and uses, how it generates forecasts, critical assumptions, and a discussion of any significant departures from accepted theory or practice.
- c) Process flow diagram (optional): A flowchart showing the sequencing of the data inputs, calculations (processes), and outputs of the model.
- d) Mathematical specifications: The equations of the model, or, for a linear programming model, the objective technology matrix and constraint vector. The relevant equations include those used to transform input data and parameters into model data and parameters, as well as equations that characterize the solutions of algorithms. Refer to the [Guidelines for Mathematical Specifications in Model Documentation](#).
- e) Variable and parameter definitions: All variables and parameters used in the documentation and their units of measurement. If this is prepared with the mathematical specifications, then it is not necessary to present these separately. (It is not necessary to include definitions of computer-code variables that have no exact counterpart in the model documentation, such as temporary computer-code variables, or variables used only in debug statements.)
- f) Inputs, outputs, and data: All should be defined, and their units and sources specified. It should be made clear which data are used as direct model input and which data are used for estimation of parameter values. If data were transformed or manipulated prior to use, an explanation should be provided. (The material in this section need not be presented as a unit. The information on data used for estimation could well be presented with the discussion of "model estimation procedures." The discussion of data transformations could well be included with the "Mathematical Specifications" section, see item b above.) See [Checklist B](#), item a, for guidelines for presenting values of data used as direct input into the model run. The program office may elect to include the data sources for

this direct-input portion of data as descriptive comments within the input file included in the archive package, rather than having it in the explanatory portion of the documentation.

g) **Model estimation procedures:** The methods and data sources used to estimate parameters and other quantities in the model should be identified. Enough information about the estimation technique should be given to allow an expert to exactly reproduce the estimation results. This includes a precise citation of data sources, the data series used, and an exact description of which portions of the data series are used for each estimation. (This material need not all be presented in the same section of the documentation; e.g., the data sources could be presented with the other items listed in item f above.)

h) **Existence and uniqueness of solutions (optional):** For iterative or optimization problems, in cases where the existence or uniqueness of a solution has been demonstrated by analytical means, a description may be presented. In other cases, computer runs should be conducted to test whether the solution methodology is dynamically stable. Also, tests from different initial value conditions should be conducted to provide evidence of the uniqueness of solutions.

i) **Sensitivity analysis (optional):** Tests should be performed to determine whether changes in key model inputs cause key model outputs to respond in a sensible fashion. If available, include the most recent sensitivity analysis in the documentation or reference an information product with the analysis.

Standard 2002-26 Checklist B. Supplementary Model Implementation Documentation Components

For each of these items, options are provided for presenting the material on a computer file, or alternatively, with the other documentation materials.

a) **Direct input data:**

When data are used as a direct input to the model, the values should be given and clearly labeled, possibly on an accompanying computer file or files (e.g., ASCII, spreadsheet, or database files may be used). Another option is to reference the archive package with the direct input data. When data are presented in a computer file, identify each data series including labels identifying each column to ensure the meaning of each row is clear. (If the input files included in the model archive package are labeled in this manner, then this automatically satisfies the requirement.) In cases where the archived input files are labeled to satisfy this requirement, the program office furthermore has the option of including descriptive comments to identify sources of the data. When this is done, it is not mandatory to include this information in the explanatory portion of the documentation ([Checklist A](#), item f).

b) **Parameter estimates from regressions:**

In cases where parameters are estimated by a statistical regression or econometric package, the parameter estimates, R^2 , t-statistics, and other key information which the program office chooses to report, can be presented in one of two ways:

1. A discussion may be included in the documentation, or
2. An output listing from a statistical or econometric package may be presented. Enough information should be included to allow a reader to understand which equation in the documentation is being referred to, and how the parameter names in the package output correspond to names used elsewhere in the documentation.

Explanatory lines within the package outputs may be added for explanation, where necessary. In such cases, a distinctive mark, such as ">>>" (or whatever a writer prefers to use) should be placed at the beginning of the comment, so that a reader can distinguish the added comments from the output of the package.

c) **Correspondence between variable names appearing in the documentation and those in the computer code:** When possible, the variable names in the documentation should match the corresponding variable names used in a

model's computer source code. In cases where the name in the code differs from the name in the documentation, provide a cross-reference list. (Computer-code variables which have no exact counterpart in the documentation, such as temporary variables, would not be cross-referenced.) Cross-referencing can be accomplished in one or a combination of two ways:

1. A cross-reference table may be presented (included either in an electronic file or in the documentation), or
2. When there is a declaration in the computer code of a variable which is the equivalent of a documentation variable, a comment in the code would indicate the name of the equivalent variable in the documentation, in the distinctive form $\{vn\}$ where vn represents the variable name in the documentation. The distinctive form makes it easy to locate the name electronically. (If another distinctive form is used, then the documentation should state the form.) Comments could alternatively be placed next to variable names (e.g., a FORTRAN COMMON block) or another location in the code that the program office chooses.

If the program office elects not to update the variable cross-reference table further, then for all subsequent variable declarations where there is an equivalent variable in the documentation, the documentation name should be identified in a code comment. If the program office elects to use the computer-comment approach throughout the entire code for all of the documentation variables, then publication of the cross-reference table is optional.

If the variable cross-reference table is no longer being updated, the documentation should state this and should explain the format for identifying documentation variable names in code comments.

Standard 2002-26 Guidelines for Mathematical Specifications in Model Documentation

The mathematical specification of the model should be an unambiguous formal statement of the modeler's concept of the methodology and structure to represent real world phenomena. Given the description of the model's methodology and structure along with knowledge of the models inputs and transformations, in principle, one should be able to form an understanding of the specific interrelationships represented in the model, the underlying assumptions, and the essential model outputs. In practice, an interested person may investigate the documentation and computer code as well as discuss the model with the manager to more fully understand the intricacies of the model.

This requirement obviously does not mean that the documentation needs to present one equation in the documentation for each equation in the computer code. Often a single mathematical equation in the documentation is the mathematical equivalent of several computer code statements. The more compact representation would be used in the documentation. This means that certain computer "temporary variables" that are used to temporarily store values (in code) need not be discussed at all in the documentation. In some cases, words rather than an explicitly written equation can be used to specify the model, but only if it is clear how the stated idea would be expressed mathematically.

In cases where an optimization problem is stated, or an iterative process is used to find a solution of some mathematical problem, it is important to state precisely what optimization problem is being solved, or to state precisely, in mathematical terms, the solution obtained when the iterative process is complete. Where iterative processes are used, the basic solution methodology should be described. If there are key parameters, such as a damping parameter, that are involved in the iterative process, then these should be described and the relevant equations containing those parameters given. References to published literature should be given, where applicable, but it is not necessary to describe the iterative process used in the code in complete detail. The mathematical specification in the documentation need not deal rigorously with the possible problems of non-convergence or multiple solutions. However, if an iterative procedure in the code has a default condition which allows the computation of the program to continue even if convergence is not attained, then this condition should be stated.

In some circumstances, one equation in the model documentation can appear several times in the computer code when parallel relationships are used in several separate sectors. In such cases, the documentation should clearly describe the several sectors to which this general equation applies, and relevant subscripts or arguments for representing the various sectors should be clearly defined. (It is not necessary to repeat the equation in the documentation for each instance in which it appears in the code, as long as the documentation makes it clear how the specific instances of the equation fit in to the model structure.)

If a linear program problem is of such a size and complexity that it is impractical to include the complete specification, provide a description of the structure of the problem in as much detail as practical. In addition, provide information needed by an expert to access electronically any desired coefficient or constant in the problem, to determine the dimensions of the objective function vector and constraint matrix, identify which constraints involve strict or non-strict inequality, and to obtain all information necessary to generate the results of the optimization problem.

Portions of the mathematical specifications can be included in an appendix or appendices if the program office chooses to do so. The text and the appendices jointly should provide a complete specification of the model. (If a complete mathematical specification is provided in a report, it is not necessary to repeat these same equations in an appendix, unless the program office chooses to do so.)

In the variable cross-references ([Checklist B](#), item c), it is possible that one documentation variable name might correspond to several computer variable names. In some circumstances, the computer code might use two different variables which have the same definition but are in different units. If the preference is to use one variable name to represent both variables, include a special notation or explanation in the variable cross-references in cases where the correspondence is among variables with the same definition but measured in different units.

[1]

Graphic Standard: Web Graphics that Translate Well in Black and White, Herbert Miller, *Update and Results of Cognitive Testing of EIA Graphs*, paper presented at semi-annual meeting of the American Statistical Association's Advisory Committee on Energy Statistics (Washington, DC, April 19, 2001).

[2]

An example of where users had difficulty occurred when the two y-axes represented the same variable but at different orders of magnitude; e.g., both y-axes were in billions of barrels but of different orders of magnitude; e.g., both y-axes were in billions of barrels but one had a scale of 5, 10, 15 and the other had a scale of .5, 1.0, 1.5. Users did not notice that changes in one variable were 10 times the magnitude of changes in the other variable. Also, percent change is not a good alternative to dual y-axis graphs because users were confused as to whether the graph presented percent change from a base year or from the previous year.