

## **B. Collections of Information Employing Statistical Methods**

### **B1. Description of Respondent Universe and Sampling Plan**

The universe consists of all operators of producing wells in the United States (excluding Alaska) that produce natural gas, including Offshore wells. A natural gas operator forms the responding unit. A cut-off sample of well operators is selected monthly. It is refined to give the targeted coverage in all areas with the minimum amount of operators. For the 2010 survey, EIA met its quality goal with about 240 operator respondents. The cut-off sample is designed to provide about 90 percent coverage at the lower-48 level and an adequate percent coverage in seven areas. The seven areas include the Gulf of Mexico Federal Offshore, Louisiana, New Mexico, Oklahoma, Texas, Wyoming, and Other States (defined as all remaining States excluding Alaska).

The cut-off sample is selected from a current commercially available data source (HPDI LLC) and the EIA-23 frame. These provide the most recent and accurate available data. All estimates are based on the simple ratio estimate method with the exception of Other States which is currently determined by applying the ratio of 2008 total gross natural gas production published in the EIA Natural Gas Annual to the 2008 EIA-914 survey production. The 2010 EIA-914 frame consisted of roughly 8,300 operators who produce at least 1,000,000 cubic feet of gas per day.

### **B2. Data Aggregation Procedures**

Currently, well operator respondents are asked to submit monthly reports of their natural gas production for seven areas. The respondent data submitted for each area are processed by EIA using its Standard Energy Processing System (STEPS), a general-purpose survey processing system originally developed by the Bureau of Census. EIA has an established procedure for both follow-up of non-respondents and for verification of data filed. Data filed on the EIA-914 are aggregated in STEPS and undergo a series of mathematical checks for reasonableness and accuracy. EIA uses documented techniques to estimate total monthly natural gas production volumes for the total US and the seven areas (techniques described in [EIA-914 Monthly Gas Production Report Methodology](#)).

#### **EIA-914 Monthly Gas Production Report Methodology**

EIA implemented the methodology described here in April 2010 and applied it historically to all of 2009 as well as to the current months. Fundamentally there are two parts to the process: the sampling and the estimation. Both represent changes to previous methodologies, and each is described separately below.

## **Sampling Methodology**

The EIA-914 survey collects natural gas production volume information on a monthly basis from a sample of well operators (companies). Production volumes are requested specifically for Texas, Louisiana, Oklahoma, Wyoming, New Mexico, Federal Offshore Gulf of Mexico, and all Other States (except Alaska). Sampling occurs every month and is called a monthly refresh of the sample. The latest available HPDI monthly data is used to select companies to add to the sample group. (HPDI is a commercial vendor of production data.) The sample group of companies changes by 1 or 2 every month. This keeps the sample current and avoids a major change in the sample caused by less frequent updating. A cutoff sample based on company production rates is used.

### ***Data Preparation***

The HPDI database is used for both the sampling and the estimation processes. HPDI acquires well or lease level data from State agencies, places it in their own database format, and sells it. A new HPDI database is acquired every month. HPDI data for five of the smaller producing States is missing or inadequate. For Illinois, Indiana, Kentucky, Pennsylvania, and Tennessee, annual production data from the EIA-23 survey (Annual Survey of Domestic Oil and Gas Reserves) is used to supplement the HPDI database. Hereafter, references to HPDI data include supplemental data from the EIA-23 survey for these five States.

The monthly production data are split into two parts: the group of companies that comprise last month's sample and the group of non-sampled companies. The non-sampled group is sorted largest to smallest based on production from the lower-48 States. A second sorting for Oklahoma only is also done (more on this below).

HPDI data for the most recent months are usually significantly incomplete. A six month average natural gas production level is calculated by company for the most current 6 month period where data are or are nearly complete. The non-sampled companies are sorted by this 6 month average.

### ***Cutoff Sample***

The sampling process uses cutoff criteria of 20 MMcf/d by company for Lower 48 production. For the State of Oklahoma a cut off of 10 MMcf/d is used to increase the coverage in Oklahoma. (Oklahoma has an abundance of small companies, so a lower cut off helps keep the percent coverage up in Oklahoma.)

### ***Adding and Dropping Companies***

Each month, companies whose surveyed production rate falls below the sample cutoff point are reviewed to determine whether or not they should remain in the sample. If a company is below the cutoff (20 MMcf/d in the lower 48 or 10 MMcf/d in Oklahoma) for 6 consecutive months it is contacted and asked why. If the decline in production is not reversible or repairable in the near future the company is dropped from the survey.

When adding a company, the sorted non-sampled HPDI data is used. The highest production level in the non-sampled group of companies is about 20 MMcf/d. Data for the largest companies in the non-sampled group can be reviewed quickly and any company in the non-sampled group that produces more than 20 MMcf/d for 4 months in a row (previous to the last few months of incomplete data) is a candidate to be added to the sample. Production is double checked and confirmed for these identified companies. Once identified, a selected company is informed with a phone call and a survey information packet is sent. The same process is applied to Oklahoma production data with a 10 MMcf/d cutoff.

### ***Other Ways Companies are Added or Dropped***

Mergers and acquisitions, or buying and selling properties can cause a company's production level to move above or below the sample cut off value. An attempt is made to accommodate the larger events in the sample as soon as possible after they occur. These larger events usually appear in news reports, newsletters, press releases, industry trade journals, etc. Small events involving small companies or small volumes of production are ignored if they involve only companies in the non-sampled group. Most smaller mergers and property sales are unknown.

### ***Potential Sources of Errors***

Unknown, deficient reporting of, or improperly handled mergers and property sales are likely the largest cause of sample errors. These events are continuous and make the sample calibration data (HPDI data) a very dynamic data set. The company production in the historical HPDI data set must be merged to match the reported sample data every month. The unknown or missed events are usually small and probably don't contribute large errors, but it is still possible to miss a larger event. The historical sample data used to make estimates can be missing production or have too much production and therefore adversely affect the production estimates because of name changes, multiple name spellings, companies that report under multiple names in the HPDI data set, the lag between the time of a merger and the time of its appearance in the HPDI data set, past multiple mergers, and the potential to improperly assign EIA operator codes. It is impossible to account for all of the mergers and property sales.

### ***Estimating Methodology***

The Simple Ratio Method (SR) is used for all the individual States in the EIA-914 report. The SR method allows the use of the most current historical data available to determine a straightforward ratio. The short lag times mean that any changes in the sample over the shorter lag will be as small as possible and can normally be neglected.

For the Other States group of states, the ratio of the EIA-895A survey (Annual Quantity and Value of Natural Gas Production Report) to the EIA-914 survey annual volumes in the previous calendar year is applied to the current monthly

EIA-914 volumes to calculate the estimate. This is the same process used previously for the Other States group.

***Simple Ratio Method***

The SR method uses a ratio of the total production to the current sample’s production at some point in history. This ratio is then applied to the current reported sample volume to estimate the current total production volume. The ratio is a 6-month average ratio calculated at some lag time that varies by State. The time frame for the 6 month average ratio calculation is moved ahead one month every month so that the lag time is a constant over time. Lag times vary from 6 to 18 months for the different states. Lags are necessary because the HPDI data is incomplete in current months. Some States require a longer lag than others to get back to a time when the data is complete to calculate the 6-month average ratio. Currently a 6-month lag is used for Wyoming, New Mexico, and Louisiana, 9-months for Texas, 12-months for the Federal Gulf of Mexico, and 18-months for Oklahoma. At these lag times the reported production in HPDI should be less than 0.5% different than the final reported production. Recent changes in HPDI’s data collection in Oklahoma may allow a shorter lag time in the future. The equations are as follows:

$$Avg\ SR_{i-L} = \left( \frac{\sum_{i-L}^{i-L-5} \frac{TP_i}{SP_i}}{6} \right)$$

$$TP_{esti} = S_i * Avg\ SR_{i-L}$$

where:

- Avg SR = Simple Ratio, 6-month average
- TP = Total Production, from HPDI
- SP = Sample Production, current sampled group of companies historical production, from HPDI
- L = Lag time in months
- TPesti = Total Production estimate for the current month
- Si = Sampled production for the current month
- i = Current or estimation month.

**B2.1 Revisions**

Revisions to aggregated data are published when respondents submit revised data for prior months or late data for the current month according to the [EIA-914 Revision Policy](#).

**B2.2 Security**

Several steps have been taken to assure confidentiality and security of respondent data and estimates during data processing and report preparation. One example is that respondent data can be submitted to EIA using the secure file transfer (SFT) system. SFT is based on the secure hypertext transfer protocol (HTTPS), an industry standard method to send information over the web using a secure, encrypted process. All information is protected by 128-bit encryption to maintain the privacy and confidentiality of transmitted data. See Section A10 for a discussion on confidentiality of the data.

### **B3. Methods to Maximize Response Rates**

Form EIA-914 is collected under EIA's mandatory data collection authority and its response rate has been good. For example, the 2009 production weighted response rate for the eight areas averaged 99 percent. To maximize response rates, the form was designed for ease of completion, and instructions were written in a clear and concise manner to make them easily understood. Potential respondents initially tested the form and instructions, which were modified based on their feedback. Furthermore, small operators who usually have low response rates are not sampled.

Forms are e-mailed as early as possible to maximize the time that respondents have to complete the surveys; the forms and instructions are available on the EIA Website. Survey non-respondents are contacted by telephone to discuss the requirement to file and any problems or questions that are delaying filing. Follow-up letters regarding the failure to file are also mailed to non-respondents. Specific schedules are followed for telephone calls and letters to non-respondents. Every effort is made to assist respondents in completing the survey and submitting it in a timely manner.

### **B4. Contacts**

For additional information concerning the survey design, please contact Gary Long of the Department of Energy, Energy Information Administration 214-720-6182. For more information regarding the approval request, please contact Jason Worrall 202-586-6075.