Independent Statistics & Analysis
U.S. Energy Information Administration

February 2016

Types of Possible Survey Errors in Estimates Published in the *Weekly Natural Gas Storage Report*

The U.S. Energy Information Administration (EIA) collects and publishes natural gas storage information on a monthly and weekly basis. The Form EIA-191, *Monthly Underground Natural Gas Storage Report*, is a census survey that collects field-level information from all underground natural gas storage operators in the United States known to EIA. EIA also collects weekly information about working natural gas levels from a sample of natural gas storage operators completing the Form EIA-912, *Weekly Underground Natural Gas Storage Report*. The EIA-191 and EIA-912 surveys are the basis for the EIA's *Weekly Natural Gas Storage Report* (WNGSR).

This document lists types of potential errors in EIA estimates published in the WNGSR. Survey errors are an unavoidable aspect of data collection. Error is inherent in all collected data, regardless of the source of the data and the care and competence of data collectors. The type and extent of error depends on the type and characteristics of the survey.

The analyses in this document is based on techniques found in the statistical literature. Groves et al. (2011), Biemer and Lyberg (2003), and Weisberg (2009) all presented concepts and issues involving evaluations of total survey error. Possible survey errors in estimates published in the WNGSR may include sampling error, coverage error, nonresponse error, adjustment error, measurement error, and processing error.

The monthly survey data from the *Monthly Underground Natural Gas Storage Report* (EIA-191), are aggregated and disseminated by EIA in the *Natural Gas Monthly* (NGM), about 60 days after the end of the survey reporting period. The data collected from the EIA-191 include information about working gas storage and base inventories, flows into and out of storage facilities, and design capacities of natural gas storage fields. These data play an integral role in the sample design and estimation of EIA weekly natural gas storage published in the WNGSR.

For the weekly survey (EIA-912), EIA selects a sample of operators for this survey from the frame of all underground natural gas storage operators in the Lower 48 states that are included in the EIA-191 data collection. Compared with the EIA-191, the EIA-912 data collection is simplified, requesting information only pertaining to working gas levels. The WNGSR is usually disseminated about six days after the close of the weekly reporting period. For example, data for the week ending February 26, 2016, were disseminated on March 3, 2016.

There are several reasons for conducting sample surveys instead of census surveys. The smaller number of respondents in the weekly sample means the EIA can collect, review, and disseminate the information faster. This also means lower overall burden on the operators who have to respond, as well as lower data collection and processing costs. Also, the tradeoff between sampling error and reporting error is a key consideration in choosing to conduct a sample, because the increased reporting associated with a census can result in increased reporting error.

Sampling error

Sampling error is the error caused by observing a sample instead of the entire survey frame. Statistics based on the sample (such as totals, means, and medians) generally differ from statistics on the entire frame because the sample includes only a subset of the frame. For example, if EIA measures the weekly underground storage inventory estimate for the East region from a sample of natural gas storage operators, the total stocks estimated from the sample may not be the same as the overall total stocks from the full frame of natural gas storage operators in the East region. Sampling error is the difference between an estimated value based on the sample and the frame value, if the entire frame had been surveyed under the same conditions as the sample.

Statisticians use measures such as standard error and coefficient of variation to measure the sampling error. The standard error (SE), which is measured in the same units as its corresponding survey estimate, is a measure of the sampling variability of the estimate based on all possible samples that could have been selected using the chosen sample design. The coefficient of variation (CV), which may also be referred to as the *relative standard error*, is the standard error expressed as a fraction of the estimate and is usually displayed as a percentage. The glossary in this report provides the defintions of standard error and coefficient of variation as well as other defintions of statistical terminology used in this report.

Standard error – a measure of the sampling variability of an estimate based on all possible samples that could have been selected using the chosen sample design.

Coefficient of variation (or relative standard error) – a measure of the sampling variability of an estimate in which the standard error is expressed as a fraction of the estimate and is usually displayed as a percentage.

Simple estimators have standard equations for measuring sampling errors. However, because the WNGSR estimator is complex, there is no standard formula to compute measures of sampling error. Vartivarian and Kasprzyk (2006) documented the research by Mathematica Policy Research Inc. and Abacus Inc. to develop bootstrap variance estimates for WNGSR estimates that were based on three geographic regions. Using the WNGSR data from April 2015 to November 2015, which are based on the 2015 sample and the new five-region breakout (http://ir.eia.gov/ngs/notes.html), EIA computed estimates of standard errors for both the published estimates of weekly working gas storage and the net changes in weekly working gas storage. The new breakout did not change the companies that are included in the sample, but it may have affected the region (or regions) in which operators reported their data. These estimated standard errors were computed using the bootstrap method, in which 1,000 samples (or replicates) were drawn with replacement from the operators on the sampling frame that were not selected in the sample with certainty. The variability of the estimates of the population values that were produced from the 1,000 replicates was used to calculate the standard errors of the WNGSR estimates.

The glossary in this report provides the mathematical formulation that was used to produce the standard errors of the WNGSR estimates based on the bootstrap method.

Bootstrap – method of estimating measures of sampling error based on the variability of population estimates that are produced by combining data for certainty units with data for replicate samples of the noncertainty units. From Wolter (2007), an estimate of the variance of the estimator $\hat{\theta}$ is:

$v(\hat{\theta}) = (1-\bar{f})\frac{1}{A-1}\sum_{\alpha=1}^{A}\left(\hat{\theta}_{\alpha}^{*} - \hat{\bar{\theta}}^{*}\right)^{2}$, where *A* is the number of replicates, $\hat{\theta}_{\alpha}^{*}$ is the estimator of the population value based on replicate α, $\hat{\bar{\theta}}^{*}$ is the mean of the estimators of the population value over the *A* replicates, and $(1-\bar{f})$ is an approximate finite-population correction factor for reducing the overestimation of variance because the uncorrected bootstrap method acts as if the sample were selected with replacement. An estimate of the standard error of the estimator $\hat{\theta}$ is, therefore, $se(\hat{\theta}) = \sqrt{v(\hat{\theta})}$.

Table 1 in the Appendix gives coefficients of variation for the published estimates of weekly working gas storage from April 2015 to November 2015, based on the 2015 sample for the WNGSR and the new five-region breakout. For the Lower 48 states, the coefficients of variation averaged about 0.9%, which is well below the threshold of 5% used for the sample design. Figure 1 in the Appendix shows the 95% confidence band for the Lower 48 states based on the published estimates and coefficients of variation given in Table 1. The standard errors are directly related to the published weekly working gas estimates. Regionally, coefficients of variation ranged from 1.4% to 4.1%, with the Pacific region posting the highest coefficients of variation.

Table 2 in the Appendix gives standard errors for the published estimates of net change in weekly working gas storage from April 2015 to November 2015, based on the 2015 sample for the WNGSR and the new five-region breakout. The average standard error for the estimates of these net changes by region was approximately 0.5 billion cubic feet (Bcf) for the East region, 0.6 Bcf for the Midwest region, 0.3 Bcf for the Mountain region, 0.8 Bcf for the Pacific region, and 1.2 Bcf for the South Central region (0.8 Bcf for both salt and nonsalt fields). For the Lower 48 states, the average standard error for these estimates of net change was approximately 1.8 Bcf. Figure 2 in the Appendix shows the 95% confidence band for the Lower 48 states based on the published estimates and standard errors given in Table 2. The largest standard error was approximately 2.5 Bcf, so each of these changes (15 Bcf is the smallest in magnitude) is significantly different from 0 when tested individually at the 5% level.

Coverage error

Coverage error can occur when the list of operators in the sampling frame is not complete or when it includes units that are not in the survey's target population. The difference between a value based on the frame and the true population value, which is typically unknown, is the coverage error. EIA mitigates the risk of coverage error in its storage program by monitoring developments in the storage industry and developments in the trade press, including permitting of new storage projects, as well as expansions or sales of existing facilities.

For the WNGSR, the sampling frame used to select the 2015 sample had 131 operators for which data are collected in the EIA-191. EIA publishes field-level storage data by year for the operators for which data are collected in the EIA-191.

Nonresponse error

Some survey units may not respond, or they might provide only partial information. Because nonresponse error is typically present in sample surveys and censuses, EIA follows up with nonrespondents to minimize nonresponse. Nonresponse to the EIA-912 is typically resolved with a form submission within a couple days of publication. Response rates for both the monthly and weekly natural gas storage surveys typically exceed

95%. For the 2015 sample used for the WNGSR, the response rate for the weeks ending April 10, 2015, through January 1, 2016, was 100%, except for four weeks when the response rate was at least 96% (Table 3). For a given response rate in Table 3, the denominator was based on the number of operators in the sample, which had 73 operators that were broken out by the region(s) in which the operators reported. An operator that reported for more than one region was counted more than once in both the numerator and denominator of the response rate.

Occasionally, nonrespondents may differ from respondents. So, to the extent that nonrespondents differ from respondents, the survey data will not accurately represent the entire target population. The difference between an estimated value based on respondents only and an estimated value based on the full sample (where every respondent could have provided the information) is termed nonresponse error.

For the WNGSR, data that are missing or that are of insufficient quality at the time of publication are approximated mainly through imputation, by treating the nonrespondent as an operator not reporting on the weekly survey or by using operator-specific linear extrapolation methods based on the most recent data. If the imputed value for a given operator is too large or too small, then errors may be introduced into the corresponding estimate.

Adjustment error

After survey data are collected, it is sometimes necessary to perform certain adjustments as part of the estimation method to reduce the effect of nonresponse and coverage errors. Although post-survey adjustments are performed to help correct or reduce overall error in survey estimates, they may introduce errors of their own if the original assumptions used to make the adjustments no longer hold. For example, the weekly estimates of natural gas storage are based on a model that estimates the volume of natural gas stored by an operator that did not report to the survey, using a 12-month moving average of reported monthly stock volumes. If the modeled value for a given operator is too large or too small, then errors may be introduced into the corresponding estimate.

The new regions used for the WNGSR are believed to provide more accurate estimates because the new groupings better align with actual operating conditions and practices unique to each region. For 2015, a comparison of the weekly estimates with the monthly census values in the *Monthly Underground Natural Gas Storage Report* suggests that the five-region estimates are closer to the monthly values than the three-region values. The sample stratification used for estimation of nonsampled storage operators' volumes is largely based on the geographical storage regions, with the exception of the salt/nonsalt substrata in the South Central region and the former Producing region. Changing the stratification of the WNGSR resulted in changing the application of the estimation methodology for nonsampled operators, as well as the estimates for these operators. EIA tests of these methodological changes included running the new stratification in parallel with the former stratification during each week of 2015. Using the three regions and the five regions, differences in levels for the Lower 48 states observed during this period ranged from -8 Bcf to +10 Bcf, with the absolute difference averaging about 3 Bcf, or 0.14%. Several of the differences exceeded 4 Bcf—the newly adopted publication revision threshold—in absolute value.

Measurement error

Measurement error is the difference between a respondent's true value and the value the respondent provides during data collection. Sources of measurement error are many, including, but not limited to, reporting errors such as a respondent misinterpreting some of the words in the survey questions; (b) respondent records not corresponding exactly to the data that EIA may be requesting (for example, some respondents may have records only on natural gas flows or scheduled nominations, rather than on measured

inventory levels); and (c) a respondent committing reporting errors such as using the wrong units (for example, million cubic meters rather than million cubic feet) or incorrectly classifying base gas as working gas. Statisticians and operations staff are often able to catch reporting errors and either make the necessary corrections or contact respondents for verification.

There were no published revisions from April 2015 to November 2015 under consideration. However, there were submissions of data received from survey respondents after the publication deadline that were too small to meet the 4-Bcf revision threshold. These revisions typically differed from the published working gas estimate by less than than 1 Bcf (Table 4). Several of the revisions resulted from the resolution of nonresponse. The largest revision during the period totaled about 2 Bcf. Although these revisions typically would not be published because they fall below the revision threshold for the WNGSR, they were included in the publication of the five-region history that accompanied the first release of the report.

EIA collects data from respondents in several ways. Respondents typically provide data electronically, but EIA can also collect data over the telephone or by fax, when needed. The type of submission also could introduce small biases in the reported values. Although EIA conducts preliminary tests of new and revised survey questionnaire forms to minimize the likelihood of misinterpretation of survey questions and different recordkeeping practices, these problems may still occur when collecting data.

Processing error

Although rare, mistakes can happen in processing the data received from a survey. When these mistakes are identified, EIA corrects the errors and revises the published data as quickly as possible. There were no known errors in post-collection operations, such as data entry or tabulation, during the period from April 2015 through November 2015.

Appendix

Table 1. Published estimates of weekly working gas storage (in billion cubic feet) and coefficients of variation (in percents) for April 2015 to November 2015

Week ending	Lower 48 states		East		Midwest		Mountain		Pacific		South Central					
	Stocks	CV	Stocks	CV	Stocks	CV	Stocks	CV	Stocks	CV	Stocks for region	CV	Stocks for salt	CV	Stocks for nonsalt	CV
4/10/15	1,539	1.2	250	2.2	263	1.7	117	3.2	273	4.1	637	1.8	171	3.3	466	2.1
4/17/15	1,628	1.1	271	2.1	282	1.7	118	3.1	276	4.0	681	1.7	191	3.0	490	2.1
4/24/15	1,711	1.1	294	2.1	296	1.7	118	3.1	281	3.6	721	1.7	209	2.9	512	2.0
5/1/15	1,785	1.0	310	2.1	311	1.6	119	2.9	287	3.6	758	1.6	223	2.7	535	2.0
5/8/15	1,897	1.0	341	1.9	339	1.6	123	2.8	292	3.4	802	1.6	243	2.8	559	2.0
5/15/15	1,989	1.0	371	1.9	364	1.6	125	2.7	298	3.4	832	1.6	251	2.7	581	2.0
5/22/15	2,101	1.0	404	1.9	396	1.5	129	2.6	307	3.3	866	1.6	261	2.6	605	2.0
5/29/15	2,233	1.0	440	1.9	431	1.5	136	2.4	316	3.2	910	1.7	277	2.6	633	2.1
6/5/15	2,344	1.0	470	2.0	464	1.5	142	2.3	322	3.2	946	1.7	291	2.7	655	2.1
6/12/15	2,433	1.0	499	2.0	493	1.5	147	2.2	327	3.3	968	1.7	295	2.9	673	2.1
6/19/15	2,506	0.9	524	1.9	517	1.5	150	2.2	333	3.2	982	1.6	295	2.8	686	2.0
6/26/15	2,579	0.9	552	1.9	546	1.5	155	2.1	333	3.2	993	1.6	292	2.8	700	2.0
7/3/15	2,666	0.9	584	1.9	582	1.5	158	2.0	329	3.2	1,013	1.6	297	2.9	716	2.0
7/10/15	2,764	0.9	609	1.9	612	1.5	164	1.9	334	3.3	1,045	1.7	305	3.2	739	1.9
7/17/15	2,823	0.9	627	1.8	636	1.5	167	1.8	337	3.3	1,056	1.6	304	3.2	752	1.9
7/24/15	2,872	0.9	643	1.7	659	1.6	171	1.8	340	3.2	1,059	1.6	300	3.3	759	1.9
7/31/15	2,910	0.9	658	1.7	683	1.6	174	1.8	341	3.2	1,054	1.6	292	3.4	762	1.9
8/7/15	2,975	0.9	681	1.6	713	1.6	176	1.8	344	3.0	1,061	1.7	291	3.5	769	1.9
8/14/15	3,027	0.9	704	1.6	740	1.6	179	1.9	345	2.9	1,059	1.7	286	3.7	773	1.9
8/21/15	3,094	0.9	725	1.5	770	1.7	183	2.1	345	2.8	1,071	1.7	287	3.8	784	1.8
8/28/15	3,190	0.9	751	1.4	815	1.7	188	2.2	345	2.7	1,091	1.7	290	3.9	801	1.8
9/4/15	3,262	0.9	769	1.4	848	1.8	191	2.3	349	2.6	1,105	1.7	293	4.0	811	1.8
9/11/15	3,336	0.9	787	1.4	883	1.8	194	2.3	347	2.7	1,125	1.7	298	4.0	828	1.7
9/18/15	3,441	0.9	812	1.4	918	1.8	197	2.3	352	2.6	1,162	1.6	312	3.9	850	1.7
9/25/15	3,537	0.9	837	1.4	952	1.8	201	2.3	355	2.7	1,192	1.6	321	3.9	870	1.7
10/2/15	3,634	0.9	863	1.4	984	1.8	206	2.2	359	2.9	1,222	1.6	331	3.9	891	1.7
10/9/15	3,731	0.9	880	1.4	1,016	1.8	210	2.1	367	2.8	1,258	1.6	346	3.8	912	1.6
10/16/15	3,813	0.8	899	1.4	1,046	1.8	213	2.0	367	2.9	1,288	1.5	357	3.5	930	1.6
10/23/15	3,875	0.8	905	1.4	1,072	1.8	216	1.9	373	2.9	1,309	1.5	366	3.5	944	1.6
10/30/15	3,931	0.8	916	1.4	1,097	1.9	216	1.9	377	2.8	1,325	1.5	371	3.3	955	1.6
11/6/15	3,985	0.8	929	1.4	1,117	1.9	217	1.9	382	2.8	1,340	1.5	373	3.3	967	1.6
11/13/15	4,000	0.8	934	1.4	1,124	1.9	214	1.6	381	2.7	1,347	1.5	377	3.4	970	1.6

Table 2. Published estimates of net change in weekly working gas storage (in billion cubic feet) and standard errors (in billion cubic feet) for April 2015 to November 2015

	Lower 48 states		East		Midwest		Mountain		Pacific		South Central					
Week ending	Change	SE	Change	SE	Change	SE	Change	SE	Change	SE	Change for region	SE	Change for salt	SE	Change for nonsalt	SE
4/17/15	89	1.4	21	0.2	19	0.4	1	0.1	3	0.6	44	1.2	20	0.9	24	0.8
4/24/15	83	1.9	23	0.6	14	0.4	0	0.1	5	1.2	40	1.3	18	0.9	22	1.0
5/1/15	74	1.3	16	0.4	15	0.4	1	0.3	6	0.3	37	1.1	14	0.7	23	0.9
5/8/15	112	1.9	31	0.4	28	0.5	4	0.2	5	0.4	44	1.8	20	1.4	24	1.2
5/15/15	92	1.7	30	0.3	25	0.4	2	0.3	6	0.4	30	1.6	8	0.5	22	1.5
5/22/15	112	2.0	33	0.7	32	0.2	4	0.2	9	0.9	34	1.6	10	0.4	24	1.5
5/29/15	132	2.4	36	1.3	35	0.5	7	0.2	9	1.0	44	1.6	16	0.5	28	1.5
6/5/15	111	1.7	30	1.0	33	0.5	6	0.1	6	0.5	36	1.2	14	1.0	22	0.7
6/12/15	89	1.6	29	0.8	29	0.6	5	0.1	5	0.6	22	1.1	4	0.8	18	0.7
6/19/15	73	1.3	25	0.3	24	0.5	3	0.1	6	0.2	14	1.2	0	0.7	13	0.9
6/26/15	73	1.5	28	0.4	29	0.5	5	0.1	0	0.6	11	1.1	-3	0.9	14	0.7
7/3/15	87	1.7	32	0.7	36	0.8	3	0.2	-4	0.5	20	1.2	5	1.1	16	0.5
7/10/15	98	1.9	25	0.3	30	0.4	6	0.3	5	0.6	32	1.6	8	1.5	23	0.6
7/17/15	59	1.2	18	0.3	24	0.4	3	0.4	3	0.5	11	0.9	-1	0.2	13	0.9
7/24/15	49	1.2	16	0.4	23	0.5	4	0.2	3	0.3	3	1.0	-4	0.2	7	0.9
7/31/15	38	1.9	15	0.5	24	0.7	3	0.4	1	1.2	-5	1.2	-8	0.6	3	1.0
8/7/15	65	1.8	23	0.5	30	0.7	2	0.5	3	1.3	7	0.9	-1	0.6	7	0.7
8/14/15	52	1.6	23	0.3	27	0.7	3	0.5	1	1.1	-2	0.8	-5	0.5	4	0.6
8/21/15	67	2.0	21	0.5	30	0.7	4	0.8	0	1.5	12	0.8	1	0.4	11	0.7
8/28/15	96	2.0	26	0.6	45	1.1	5	0.5	0	1.2	20	0.7	3	0.5	17	0.6
9/4/15	72	2.1	18	0.4	33	1.1	3	0.4	4	1.6	14	0.7	3	0.5	10	0.4
9/11/15	74	1.8	18	0.4	35	1.1	3	0.1	-2	1.2	20	0.8	5	0.5	17	0.6
9/18/15	105	1.7	25	0.6	35	0.8	3	0.2	5	0.9	37	1.0	14	0.9	22	0.4
9/25/15	96	1.6	25	0.6	34	0.8	4	0.2	3	0.5	30	1.1	9	0.9	20	0.6
10/2/15	97	2.5	26	0.7	32	0.6	5	0.3	4	1.3	30	1.9	10	1.8	21	0.6
10/9/15	97	1.7	17	0.4	32	0.5	4	0.3	8	0.7	36	1.4	15	1.2	21	0.8
10/16/15	82	1.9	19	0.4	30	0.6	3	0.5	0	0.9	30	1.5	11	1.3	18	0.8
10/23/15	62	1.2	6	0.4	26	0.7	3	0.4	6	0.3	21	0.9	9	0.5	14	0.7
10/30/15	56	2.3	11	0.6	25	1.2	0	0.1	4	0.8	16	1.7	5	1.5	11	0.8
11/6/15	54	2.3	13	0.2	20	0.5	1	0.3	5	1.5	15	1.7	2	1.6	12	0.4
11/13/15	15	2.0	5	0.3	7	0.2	-3	1.8	-1	0.3	7	0.7	4	0.6	3	0.4

Table 3. Response rates to the EIA-912, *Weekly Natural Gas Storage Report* for April 2015 to November 2015 (red indicates weeks where the response rate was less than 100%)

Date	Forms received for publication	Operators in sample	Response rate
17-Apr-15	81	81	100%
24-Apr-15	81	81	100%
1-May-15	81	81	100%
8-May-15	81	81	100%
15-May-15	81	81	100%
22-May-15	80	81	99%
29-May-15	81	81	100%
5-Jun-15	81	81	100%
12-Jun-15	81	81	100%
19-Jun-15	80	81	99%
26-Jun-15	81	81	100%
3-Jul-15	81	81	100%
10-Jul-15	81	81	100%
17-Jul-15	81	81	100%
24-Jul-15	81	81	100%
31-Jul-15	81	81	100%
7-Aug-15	81	81	100%
14-Aug-15	81	81	100%
21-Aug-15	81	81	100%
28-Aug-15	81	81	100%
4-Sep-15	81	81	100%
11-Sep-15	81	81	100%
18-Sep-15	81	81	100%
25-Sep-15	81	81	100%
2-Oct-05	81	81	100%
9-Oct-15	81	81	100%
16-Oct-15	81	81	100%
23-Oct-15	80	81	99%
30-Oct-15	81	81	100%
6-Nov-15	81	81	100%
13-Nov-15	81	81	100%
20-Nov-15	78	81	96%

Table 4: Differences in working gas levels from original publication (in million cubic feet) resulting from revisions of data for the EIA-912, *Weekly Natural Gas Storage Report* for April 2015 to November 2015

Period	South Central	East	Midwest	Mountain	Pacific
4/17/2015	0	381	0	0	0
4/24/2015	1	0	0	0	0
5/1/2015	0	0	0	0	0
5/8/2015	10	0	0	0	0
5/15/2015	0	0	0	990	0
5/22/2015	248	0	0	0	0
5/29/2015	0	0	0	0	0
6/5/2015	0	0	0	0	0
6/12/2015	0	0	0	0	0
6/19/2015	-4	0	0	0	0
6/26/2015	2,179	0	1	0	0
7/3/2015	0	0	0	0	0
7/10/2015	0	0	0	0	0
7/17/2015	0	0	0	0	0
7/24/2015	0	0	0	0	0
7/31/2015	0	0	0	0	0
8/7/2015	0	0	0	0	0
8/14/2015	274	891	0	0	0
8/21/2015	0	0	0	0	0
8/28/2015	0	0	0	0	0
9/4/2015	48	0	0	0	0
9/11/2015	0	0	0	0	0
9/18/2015	0	0	0	0	0
9/25/2015	0	0	0	0	0
10/2/2015	0	0	0	0	0
10/9/2015	0	0	0	0	0
10/16/2015	0	0	171	0	0
10/23/2015	895	0	0	0	0
10/30/2015	0	0	0	0	0
11/6/2015	0	0	0	0	0
11/13/2015	0	0	0	0	0

Figure 1: Confidence band of published weekly working natural gas storage compared with published estimates (April 10, 2015–November 13, 2015)

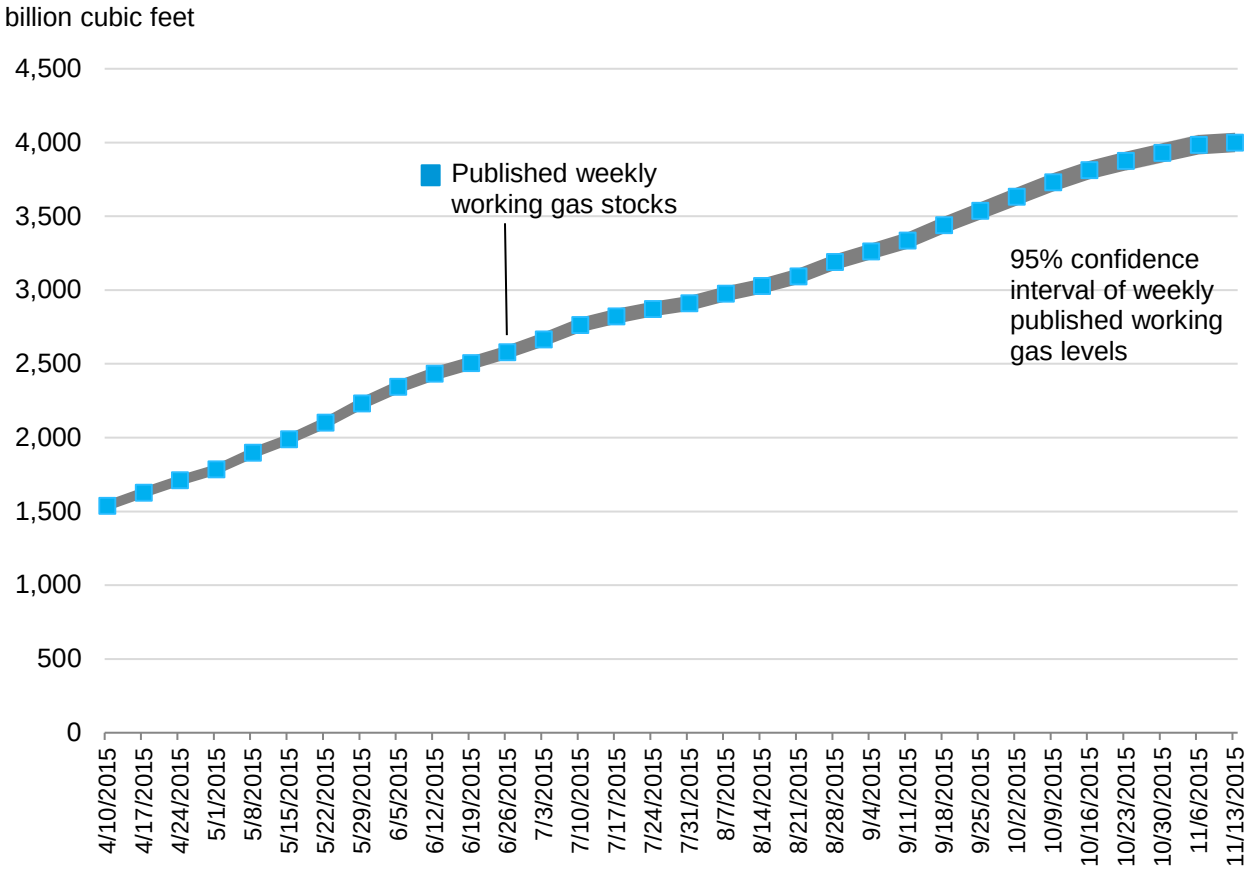
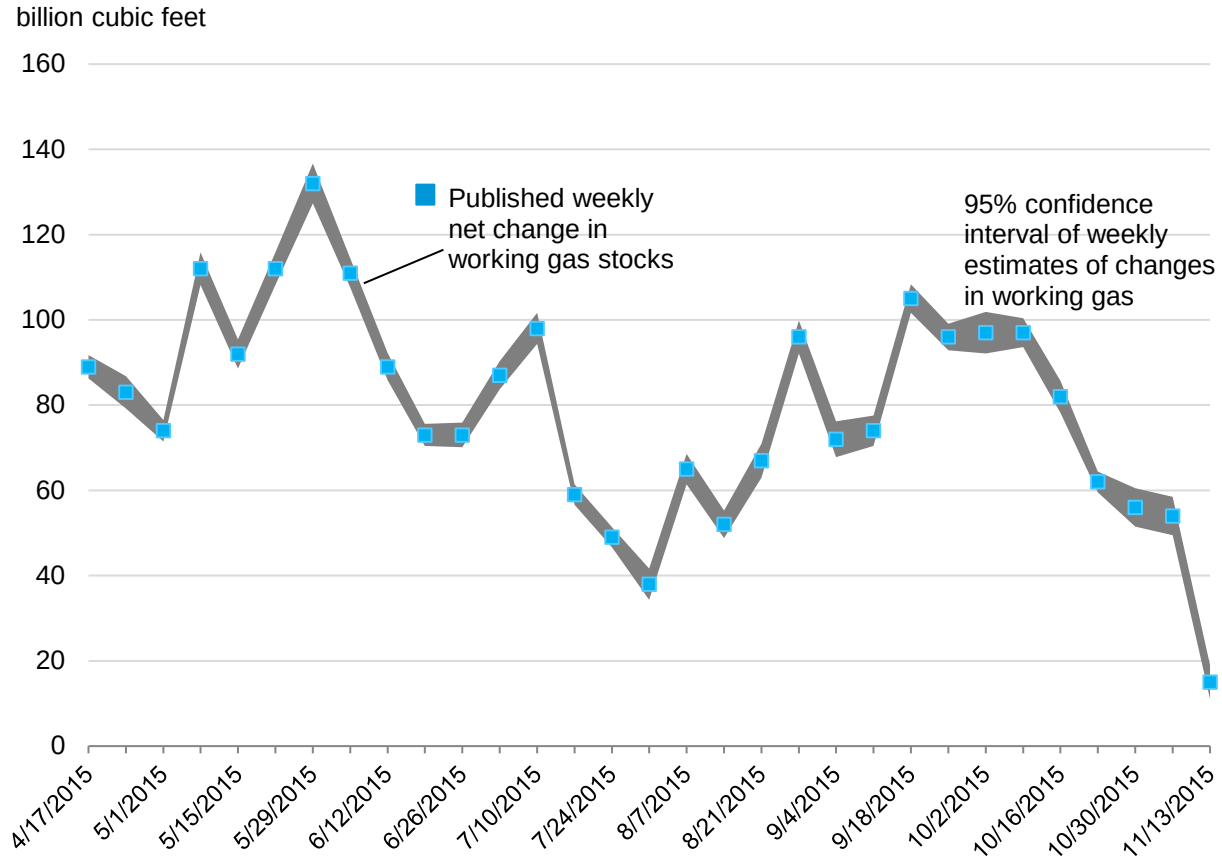


Figure 2. Confidence band of weekly published net changes in natural gas storage (April 17, 2015–November 13, 2015)



Glossary

Adjustment error – error introduced by post-survey adjustments is made as part of the estimation method to reduce the effect of nonresponse and coverage errors.

Bootstrap – method of estimating measures of sampling error based on the variability of population estimates that are produced by combining data for certainty units with data for replicate samples of the noncertainty units. From Wolter (2007), an estimate of the variance of the estimator $\hat{\theta}$ is:

$v(\hat{\theta}) = (1 - \bar{f}) \frac{1}{A-1} \sum_{\alpha=1}^{A} \left(\hat{\theta}_{\alpha}^{*} - \hat{\bar{\theta}}^{*}\right)^{2}$, where *A* is the number of replicates, $\hat{\theta}_{\alpha}^{*}$ is the estimator of the population value based on replicate α, $\hat{\bar{\theta}}^{*}$ is the mean of the estimators of the population value over the *A* replicates, and $(1 - \bar{f})$ is an approximate finite-population correction factor for reducing the overestimation of variance because the uncorrected bootstrap method acts as if the sample were selected with replacement. An estimate of the standard error of the estimator $\hat{\theta}$ is, therefore, $se(\hat{\theta}) = \sqrt{v(\hat{\theta})}$.

Coefficient of variation (or relative standard error) – a measure of the sampling variability of an estimate in which the standard error is expressed as a fraction of the estimate and is usually displayed as a percentage.

Confidence band – a line of varying width that is produced by connecting a series of individual confidence intervals of a specified level.

Confidence interval – an interval (range) estimate, typically centered about a point estimate, with width determined by the level of confidence and the point estimate's standard error.

Coverage error – the difference between a value based on the sampling frame and the true population value, which is typically unknown.

Measurement error – the difference between a respondent's true value and the value the respondent provides during data collection, which could be because of the results of reporting errors or errors in the transmission of the data.

Nonresponse error – the difference between an estimated value based on respondents only and an estimated value based on the full sample, where every respondent could have provided the information.

Processing error – rare mistakes that happen in processing the data received in a survey.

Sampling error – the difference between the estimated value based on a sample and the sampling frame value, if the entire frame had been surveyed under the same conditions as the sample.

Standard error – a measure of the sampling variability of an estimate, based on all possible samples that could have been selected using the chosen sample design. For complex estimators, standard formulas often cannot be used to compute measures of sampling variability such as the standard error.

References

Biemer, P. and Lyberg, L. (2003). *Introduction to Survey Quality*. John Wiley & Sons.

Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2011). *Survey Methodology (Vol. 561)*. John Wiley & Sons.

Vartivarian, S. and Kasprzyk, D. (2006). "Project Summary for EIA-912 Variance Estimation in the Energy Information Administration Weekly Natural Gas Storage Report". Internal Memorandum *EIA-6145*.

Weisberg, H. F. (2009). *The Total Survey Error Approach: A Guide to the New Science of Survey Research*. University of Chicago Press.

Wolter, K. (2007). *Introduction to Variance Estimation*. Springer Science & Business Media.