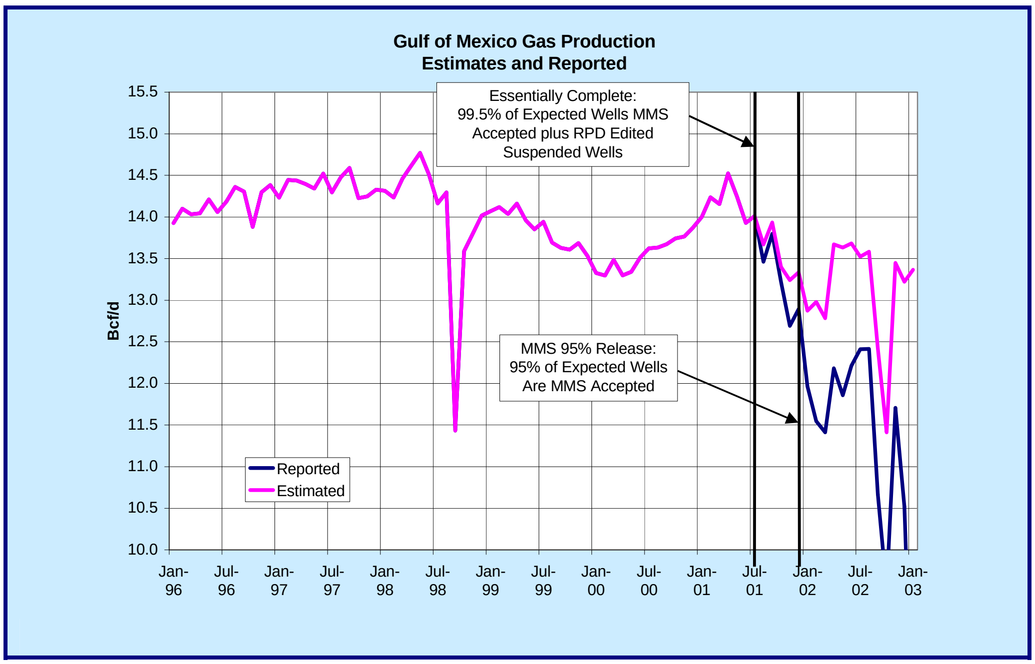
How EIA Estimates Natural Gas Production

February 2004



Office of Oil and Gas
Energy Information Administration

Table of Contents

Summary

Background

The Energy Information Administration (EIA) publishes estimates monthly and annually of the production of natural gas in the United States. The estimates are based on data EIA collects from gas producing States and data collected by the U. S. Minerals Management Service (MMS) in the Department of Interior. The States and MMS collect this information from producers of natural gas for various reasons, most often for revenue purposes. Because the information is not sufficiently complete or timely for inclusion in EIA's *Natural Gas Monthly (NGM)*, EIA has developed estimation methodologies to generate monthly production estimates that are described in this document.

Using the EIA-895 survey, "Monthly and Annual Quantity and Value of Natural Gas Production Report," natural gas producing States submit monthly State-level production information to EIA on a voluntary basis. The States' annual EIA-895 data submissions are appended to their December monthly submission; they also may contain revisions for the 11 previous months' data. MMS collects monthly well-level oil and gas production data from operators on Federal lands, including the Gulf of Mexico (GOM). Of the 32 producing States in the US, 15 States submit at least partial EIA-895 monthly data to EIA. The Natural Gas Division (NGD) in EIA's Office of Oil and Gas prepares monthly production estimates for 10 States that submit EIA-895 monthly surveys, retrieves Wyoming data from the Wyoming website and Colorado data from the Colorado website. NGD also estimates production for the 15 producing States not reporting on a monthly basis. Using MMS, State, EIA-895 and other data, the Reserves and Production Division (RPD) in the Office of Oil and Gas estimates monthly production in Texas, Louisiana, Oklahoma and the Gulf of Mexico.

Total US natural gas marketed wet production in 2001 was 20.6 trillion cubic feet (tcf).[1] Figure S-1 shows the percentage contributions of the major producing states to that total. The sources of the data that EIA uses in creating the estimates are shown in Table S-1. In general, States receive production data from producers on an ongoing basis; production values for a given month are revised many times over a period of years as new well level data come in from producers. Half the gas produced in the United States comes from the Gulf of Mexico and Texas, which each account for about a quarter of total US production (Figure S-1.) Part 1 describes the methodology RPD uses to estimate monthly production in the Gulf of Mexico, using MMS well-level production data posted on their website or provided directly to RPD. Part 2 describes the methodology RPD uses to estimate monthly production in Texas, based on production data posted on the Texas website.

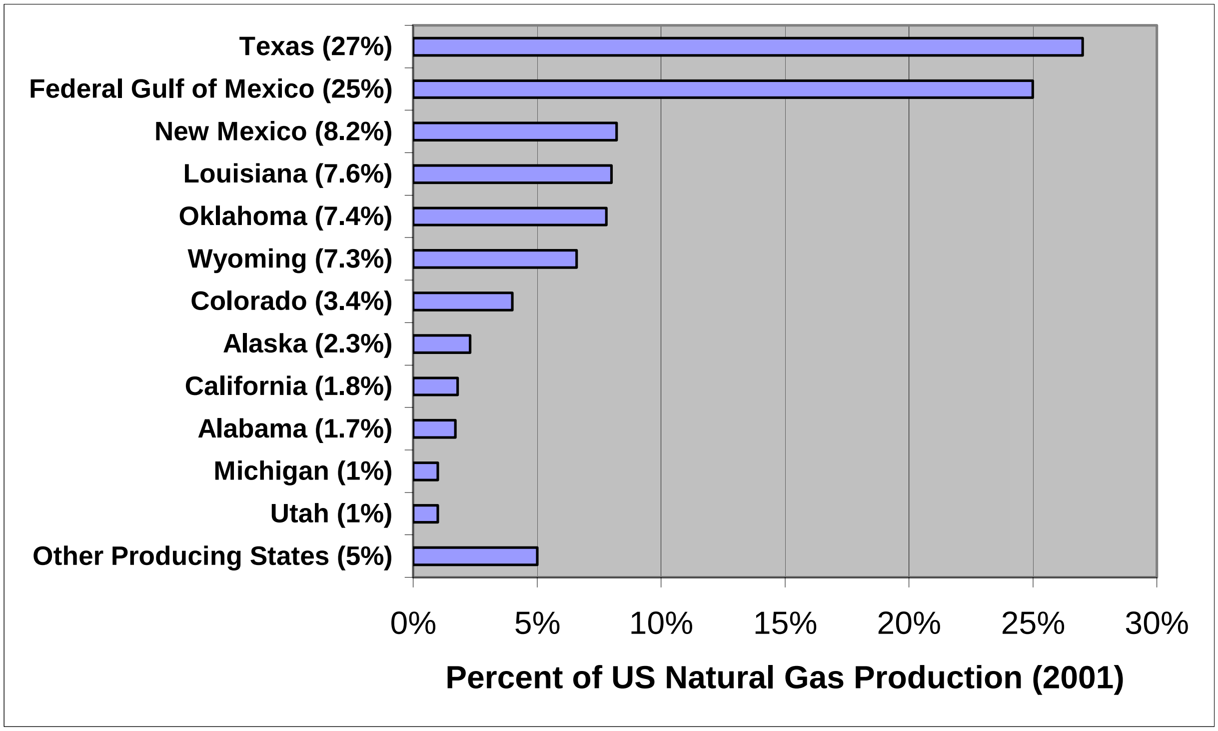
The first estimate for natural gas production in Texas is available within 60 days of the close of the report month, but this number is revised monthly for about 24 months, and there may be revisions even after two years. Reported production volumes generally start out low and, as more reports come in, approach their final values following a relatively stable revision pattern (curve). EIA's estimation model for Texas uses a multinomial statistical methodology and the model for the Gulf of Mexico uses historical data patterns, both of which estimate final

---

[1] Source: Energy Information Administration website, *Natural Gas Navigator, Production* (updated 7/1/03.)

production values from early production data. The revision patterns can be determined from history.

Figure S-1. Major Natural Gas Producing States (2001)



`Table S-1. Sources of Data Used in Estimating Monthly Production

States and Gulf of Mexico	Data Sources
Texas	Texas website (timely and complete well-level data)
Federal Gulf of Mexico	MMS website (monthly well-level data, partial, not timely)
Oklahoma	Texas monthly data and Oklahoma EIA-895 annual data (Oklahoma EIA-895 monthly data are not timely)
Louisiana	Louisiana EIA-895 annual data and bi-annual vintage reports (Louisiana EIA-895 monthly data are not timely)
Wyoming	Wyoming website monthly data (timely, incomplete)
Colorado	Colorado website (timely, incomplete)
Alaska, Alabama, Mississippi, Arizona, North Dakota, Nevada	EIA-895 monthly and annual data (complete and timely)
New Mexico, Utah, Montana, Oregon	EIA-895 monthly data (complete but late) and annual data
California, Michigan, Kansas, Florida	EIA-895 monthly and annual data (partial, often late)
Remaining Producing States	Some EIA-895 monthly data, mostly EIA-895 annual data

Part 3 describes the methodology RPD uses to estimate production in Louisiana. Vintage natural gas production reports from Louisiana became available twice a year (every February and August) beginning in 1998. These data are used to calculate a six-month median ratio of preliminary production values to final values. A preliminary reported value is multiplied by the appropriate six-month median ratio to yield an estimate of the final volume.

Part 4 describes the methodology RPD uses to estimate production in Oklahoma. EIA receives no monthly EIA-895 monthly data from Oklahoma in time for publication in the *NGM*, but relatively stable and close-to-final production data are usually (but not always) available in time for publication in the *Natural Gas Annual* (*NGA)*. To estimate monthly production in Oklahoma, RPD relies on the fact that the production patterns in Oklahoma are similar to those in Texas. EIA computes the month-to-month percent change in estimated Texas gas production (the Texas slope), and applies it to the estimate of production in Oklahoma last month to calculate the current month's estimate.

Part 5 describes how NGD creates monthly production values for the 10 producing States that submit monthly EIA-895 data and for the 18 producing States that do not. Of the 10 States providing monthly data, six States submit timely and complete EIA-895 data. Four States submit timely but incomplete EIA-895 data for which imputation methodologies are used. Colorado and Wyoming data are retrieved directly from their websites and estimation methodologies are used for the remaining producing States that submit no monthly EIA-895 data.

EIA measures the accuracy of its estimates as the difference (i.e., revision) between the first release of a monthly natural gas production estimate for a month and the final estimate for that month. The percent revisions for the five largest producing States and the total United States during 1999-2000 are shown in Appendix 4. The changes from EIA's preliminary estimate to its current best estimate average under three percent at the national level.

The methodologies described in this document are those that EIA is currently using. Other methodologies have been used in the past (some of which are described in Appendices 2 and 3) and new methodologies may be used in the future. Because it is likely that timely natural gas production data will continue to be difficult to obtain, EIA plans to refine its estimation methodologies over time to continually improve the published estimates.

The Natural Gas Product Stream, Terminology and Definitions

The natural gas product stream begins with gross withdrawals of gas from wells and is reduced along the way by various disposition and purification processes until the resulting product stream is total dry marketed production ready for consumption, as depicted in Figure S-2. The bold blue lines represent the major flow of natural gas ultimately reaching consumers. Terms are defined in Table S-2. As the figure shows, in addition to obtaining data from the EIA-895 form, data from the EIA-816 survey, "Monthly Natural Gas Liquids Report," and the EIA-64A survey, "Annual Report of the Origin of Natural Gas Liquids Production,"[2] are also used. Appendix 1 provides a more detailed description of the physical flow of the natural gas product stream and the associated production activities.

[2] The monthly EIA-816 survey gathers information from operators of facilities that extract liquid hydrocarbons from a natural gas stream (natural gas processing plants) or facilities that separate a liquid hydrocarbon stream into its component products (fractionators). The information provided to EIA describes the balance between the supply (beginning stocks, receipts and production) and disposition (input, shipments, fuel use and losses and ending stocks) of natural gas liquids. The EIA-64A annual survey gathers information from all natural gas plant operators on the amount of natural gas processed, natural gas liquids produced, the resultant shrinkage of the natural gas and the amount of natural gas used in processing.

Figure S-2. Natural Gas Product and Activity Stream, EIA Data Sources

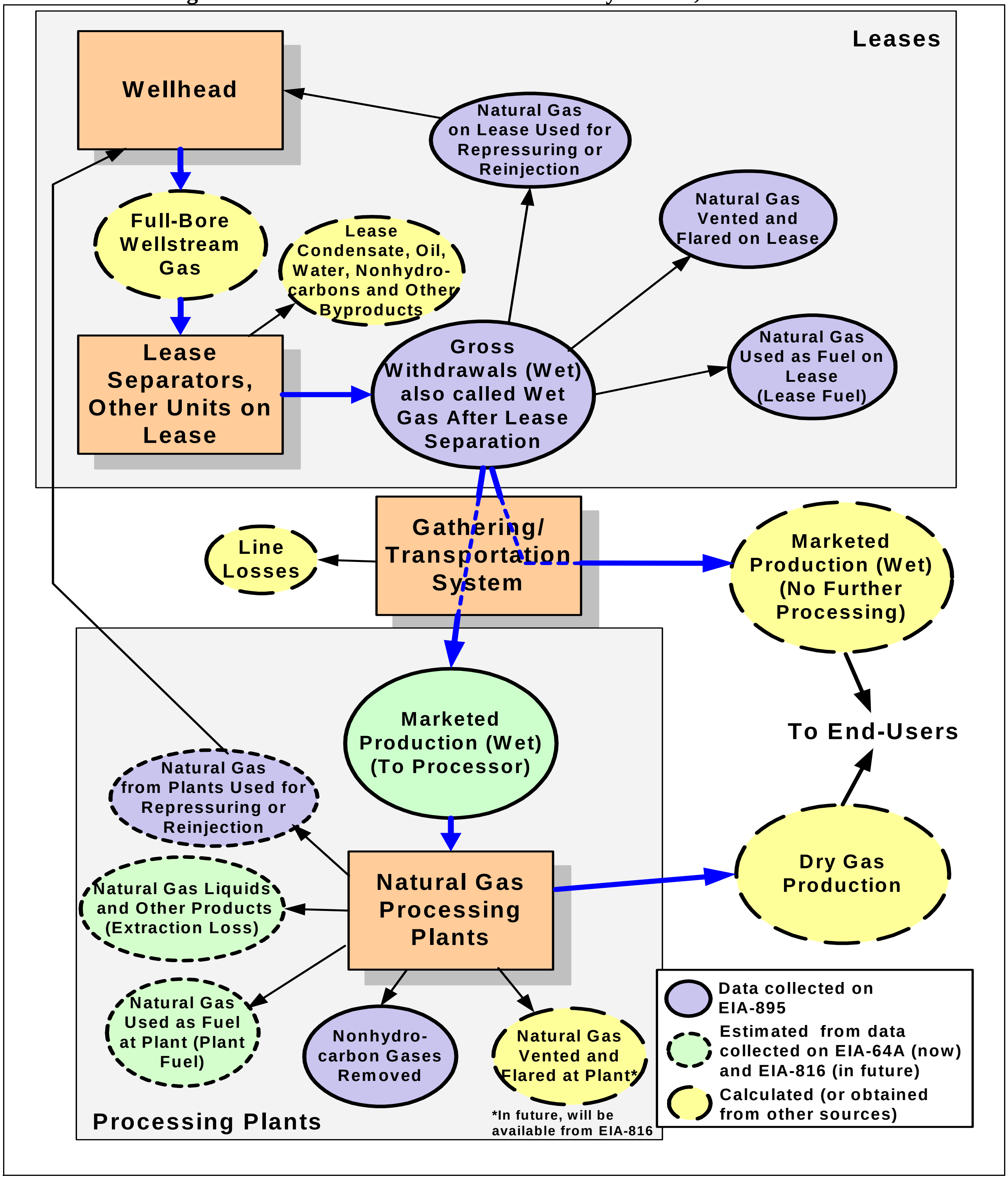


Table S-2. Definitions

Wellhead: The point at which the natural gas exits the ground.
Lease separation facility (lease separator): A facility installed at the surface for the purpose of (a) separating gases from produced crude oil and water at the temperature and pressure conditions set by the separator and/or (b) separating gases from that portion of the produced natural gas stream that liquefies at the temperature and pressure conditions set by the separator.
Processing plant: A surface installation designed to separate and recover natural gas liquids from a stream of produced natural gas through the processes of condensation, absorption, adsorption, refrigeration, or other methods and to control the quality of natural gas marketed and/or returned to oil or gas reservoirs for pressure maintenance, repressuring, or cycling.
Gross withdrawals: Full well stream volume, including all natural gas plant liquid and nonhydrocarbon gases, but excluding lease condensate. Also includes amounts delivered as royalty payments or consumed in field operations.
Wet natural gas: A mixture of hydrocarbon compounds and small quantities of various nonhydrocarbons existing in the gaseous phase or in solution with crude oil in porous rock formations at reservoir conditions. The principal hydrocarbons normally contained in the mixture are methane, ethane, propane, butane, and pentane. Typical nonhydrocarbon gases that may be present in reservoir natural gas are water vapor, carbon dioxide, hydrogen sulfide, nitrogen and trace amounts of helium. Under reservoir conditions, natural gas and its associated liquefiable portions occur either in a single gaseous phase in the reservoir or in solution with crude oil and are not distinguishable at the time as separate substances. *Note*: The Securities and Exchange Commission and the Financial Accounting Standards Board refer to this product as *natural gas*.
Dry natural gas: Natural gas which remains after: 1) the liquefiable hydrocarbon portion has been removed from the gas stream (i.e., gas after lease, field, and/or plant separation); and 2) any volumes of nonhydrocarbon gases have been removed where they occur in sufficient quantity to render the gas unmarketable. Note: Dry natural gas is also known as consumer-grade natural gas. The parameters for measurement are cubic feet at 60 degrees Fahrenheit and 14.73 pounds per square inch absolute.
Repressuring: The injection of gas into oil or gas formations to effect greater ultimate recovery.
Vented/flared: Gas that is disposed of by releasing (venting) or burning (flaring).
Extraction loss: The reduction in volume of natural gas due to the removal of natural gas liquid constituents such as ethane, propane, and butane at natural gas processing plants.
Marketed production: Gross withdrawals less gas used for repressuring, quantities vented and flared, and nonhydrocarbon gases removed in treating or processing operations. Includes all dry natural gas plus quantities of gas consumed in lease and processing plant operations.

EIA publishes information about these various parts of the natural gas production process in the *Natural Gas Monthly* (*NGM*) and the *Natural Gas Annual* (*NGA*). Gross production is the most accurate quantity because it is usually physically measured at the wellhead. Other measures of production (such as "dry production") are derived from gross production by subtracting gas removed in the production process; these quantities are more difficult to physically measure and track back to their source.

States vary as to which quantities in Table S-2 table they report and the timeliness in which they report them. For example, Oklahoma reports sales volumes, which closely approximate "marketed production," while most other States report quantities equivalent to "gross production." Information on gases vented and flared, repressuring, lease use and nonhydrocarbon gases removed are generally not directly available to EIA on a monthly basis. Details of the timing and completeness of data reporting are provided in the later sections of this report.

Part 1. Estimation Methodology for Natural Gas Production in the Gulf of Mexico

Introduction

The Gulf of Mexico (GOM) accounts for about 25% of the nation's total natural gas production and the only current source of production information in the GOM is the Minerals Management Service (MMS). However, the first release of Minerals Management Service (MMS) accepted preliminary monthly production data (95 percent of wells reported) lags by about 1.5 years. Reliable, final (essentially complete) production data from (MMS) are available 6 – 12 months after release of preliminary data, which is 24 – 36 months after the close of a report month. Because of these large lag times, the Reserves and Production Division (RPD) in EIA's Office of Oil and Gas (OOG) has developed and refined various methodologies to create GOM natural gas monthly production estimates based on data supplied by the MMS. This document focuses on the methodology currently being used.[3]

The MMS began releasing raw (unedited, uncertified) well-level production data in February 2003. After some editing, these data allow RPD to construct well production distributions of the early but incomplete well production data. The current method uses the early reported well production distribution as a sample of the final complete distribution to produce estimates of final monthly production data. The estimates for a given month become more reliable over time as the well data approach completeness. A monthly estimate can be verified at the well level by EIA when the well data are complete. As in any real-world estimation process, all of the methods may require expert judgment or analyst override, especially when unanticipated phenomena such as hurricanes occur in the GOM.

The production data provided by MMS represent gross withdrawals. Hence the methods described in this section refer to the estimation of monthly values for gross withdrawals in the Gulf of Mexico. Information on vented and flared, repressuring, lease use, and nonhydrocarbons gasses removed are not available monthly. Data are provided annually in request to special requests from EIA, though not in the EIA categories noted above. To estimate monthly values of these quantities, EIA assumes that they comprise the same percentage of production (gross withdrawals) during a given time period as they did during the same time period a year ago.

Data Description and Preparation

The process starts with receipt of the MMS data and is dependent on MMS's ability to collect and report these data. Starting in February 2003 the MMS made its new suspended well/completion production data available to EIA and the public on its website, http://www.gomr.mms.gov/homepg/pubinfo/freeasci/product/freeprod.html. In part, this was in response to EIA's request for these data. The MMS well/completion data are now available in three kinds of files:

1) Accepted or verified data that has 95% or more of well/completions reported (as of December 2003, the latest month meeting this condition is April 2002),
2) Accepted data that have less than 95% of wells reported (subsequent data), and

[3] Previous methodologies have been based on average month-to-month changes in historical GOM production data, average month-to-month changes in Texas production data, a simple linear model based on Texas production data, and the smoothed means of individual well production data. They are described in Appendix 2.

3) Suspended data (the newly available un-accepted, un-verified, un-edited data).

EIA combines and edits data from these three files to yield the total reported production. Having access to these data has permitted the development of the methodology to estimate final production described here. All these data are downloaded from the MMS website in 8 separate files.

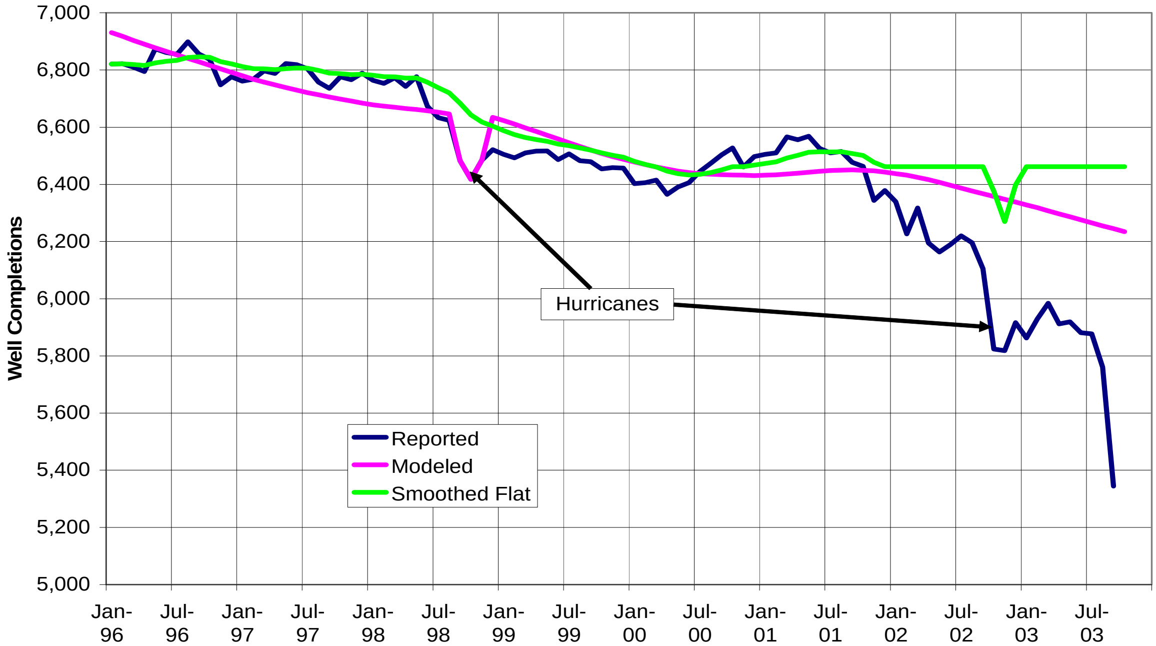
Historical data on the MMS website go back to 1996 and are updated occasionally (the last update occurred in October 2003). Current accepted data are updated monthly. Suspended data are currently updated approximately twice a month. Accepted data are in zipped delimited text files and the suspended data are in zipped Excel files. These detailed production data are by well completion by month.

SAS™ programs have been written to convert the downloaded delimited data into two SAS™ data sets of summarized monthly production. One data set contains all accepted data and the other contains the suspended data. Of the roughly 55,000 records in the suspended data (released date: 10-14-2003), about 3,700 were duplicates in the data processed in October 2003. Also, there are about 800 duplicates between the suspended and accepted data. SAS™ programs are used to identify and delete the duplicate records. There are approximately 6,500 records (completions) each month with gas production greater than zero (these include associated gas from oil completions).

Gas Production Estimation Methodology

As of December 2003, the MMS reported data are essentially complete for mid-2001 (and previous) and progressively less complete closer to current months. Figure 1-1 shows reported well completions dropping from about 6,500 in September 2001 to 5,900 in June 2003. A requirement of the methodology is to have an estimate for the number of expected well completions for each incomplete month. Then the expected completions and average production per completion can be used to estimate production in each month. Along with the reported well completions shown in Figure 1-1, two other estimates of expected well completions are shown. These will be discussed later.

Figure 1-1. Well Completions with Gas Production Greater than Zero



Simple Model

For any distribution the mean or average production per completion is:

$$M_i = \frac{P_i}{W_i}$$

Where:

M_i = Mean production per completion for month i
P_i = Total production for month i
W_i = Total producing completions for month i

The simplest model for the production is

$$P'_i = W'_i * M'_i$$

Where:

P'_i = Modeled production for month i
W'_i = Modeled or expected number of completions for month i
M'_i = Modeled mean production per completion for month i

However, the distribution of gas production per completion is skewed, so using the mean alone may not capture differences in production and reporting among large and small completions.

Figure 1-2. Gas Production per Completion (Month: May 2001)

Moments			
N	6574	Sum Weights	6574
Mean	2.16640601	Sum Observations	14241.9531
Std Deviation	5.91859516	Variance	35.0297687
Skewness	6.87008015	Kurtosis	63.735333
Uncorrected SS	261104.523	Corrected SS	230250.67
Coeff Variation	273.198797	Std Error Mean	0.07299679

Basic Statistical Measures			
Location		Variability	
Mean	2.166406	Std Deviation	5.91860
Median	0.318597	Variance	35.02977
Mode	0.000290	Range	91.78294

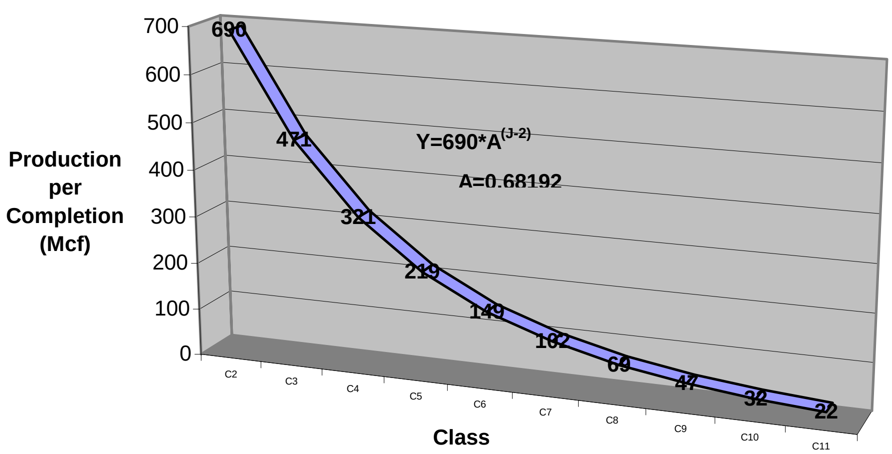
Tests for Normality				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.357172	Pr > D	<0.0100
Cramer-von Mises	W-Sq	261.0979	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	1284.144	Pr > A-Sq	<0.0050

Quantiles (Definition 5)	
Quantile	**Estimate**
100% Max	91.7830
99%	2.84755
95%	9.70106
90%	5.45597
75% Q3	1.71177
50% Median	0.318597
25% Q1	0.077419
10%	0.0234194
5%	0.0103548
1%	0.00135484
0% Min	0.000032258

12 Classes Model

The data are divided into 12 classes and the latest six months of complete production data (as of

Figure 1-3. Well Distribution by Class



December 2003, from May 2001 to October 2001) were used to create the expected distribution for each month (Table 1-1). If production per completion is less than 1.0 MMCF (most of these are oil well completions) or over 100 MMCF (rare but highly productive gas well completions), then they are defined as classes 1 and 12 respectively. As for other classes, the well completion is an exponential distribution (Figure 1-3.) with the formula of Wij=690(A)^(J-2). Excel solver is used to determine the coefficient A. Then based on the calculated well completions in the six-month calibration, the class boundaries are determined for each class (Table 1-1). These boundaries are used to divide all incomplete months into classes.

Table 1-1. Class Determination Method

Class(J)	Well Completions Number	Formula	Well Production Rate (MMCF/Day)
1	4402		0<P<1.0
2	690	Wij=690(0.68192)^(j-2)	1.0<=P<2.078
3	471	Wij=690(0.68192)^(j-2)	2.078<=P<3.601
4	321	Wij=690(0.68192)^(j-2)	3.601<=P<5.491
5	219	Wij=690(0.68192)^(j-2)	5.491<=P<7.951
6	149	Wij=690(0.68192)^(j-2)	7.951<=P<10.945
7	102	Wij=690(0.68192)^(j-2)	10.945<=P<15.15
8	69	Wij=690(0.68192)^(j-2)	15.15<=P<20.995
9	47	Wij=690(0.68192)^(j-2)	20.995<=P<31
10	32	Wij=690(0.68192)^(j-2)	31<=P<50
11	22	Wij=690(0.68192)^(j-2)	50<=P<100
12	Uncertain		100<=P

The basic concept assumes that incomplete data in recent months are a sample of what will ultimately be reported as the final distribution. The expected distribution determines the expected completions for each incomplete class. If the wells for a given month equal or exceed the expected number of wells, then that month is accepted as essentially complete. For months with fewer wells than expected, any class with completions numbering more than the expected distribution is considered complete and accepted as is. For all incomplete classes, the total number of missing completions is allocated to the incomplete classes proportional to the expected number of completions of all classes that are not full. Then, for each class, the product of the number of completions and the reported average production per completion is the estimate of production for that class.

Classes 11 and 12 (production rate > 50 MMcf/d) are treated differently. For class 11 we examine each well's historical production record and expert judgment is used to determine where monthly production is missing. The reported average production per completion for class 11 is used for the missing production value. For class 12, where each well can change the GOM monthly production by about 1 percent, we use class 12 well completions as reported.

The following charts show a comparison of two incomplete months to the expected distribution. Note that while not complete the January 2002 distribution is more complete than the January 2003 distribution. The third chart in the series shows the percentage of completions and production for May 2001 as an example of a complete distribution.

Figure 1-4. Distribution of the Expected Completions, January 2002 Reported Completions, and January 2003 Reported Completions

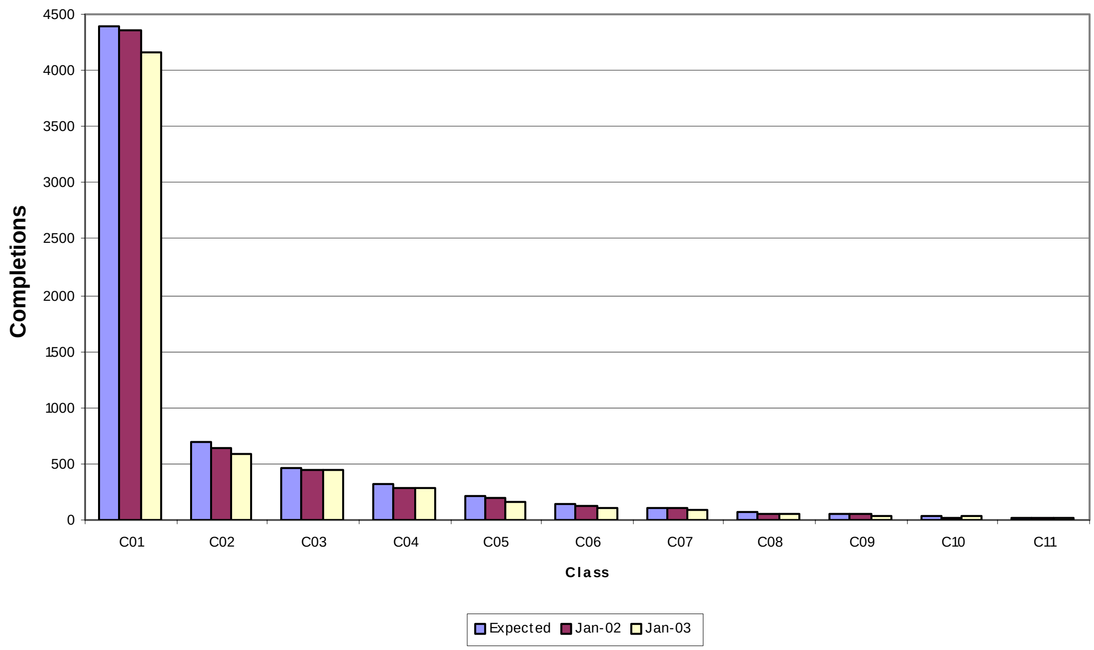
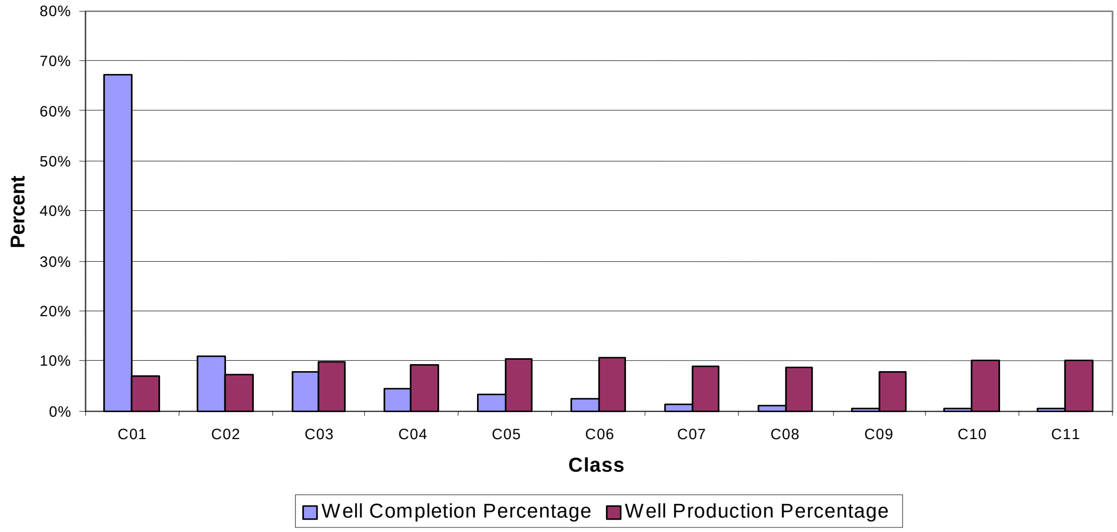


Figure 1-5. Distribution of Expected Production, January 2002 and January 2003 Reported Production

Figure 1-6. Distribution of Well Completion and Production Percentage in May 2001



Estimates for Class 11

As shown in Figure 1-7, class 11 (> 50 MMcf/day) shows significant growth from 1996 to 2002; hence it requires careful treatment. In 1996 this class had only 3 or 4 wells representing roughly 2 percent of GOM production. In the calibration period, May through October 2001, the class holds about 22 wells with about 10 percent of the production (Figure 1-7). Table 1-2 shows the historical production of some of these wells with several months of missing production. As an example, beginning with the August 2003 data, it is estimated that four more wells will likely be reported for Class 11 in June 2003, three in May 2003, one each in February through April 2003, two for January 2003, and one in March 2002 (cells highlighted in yellow). By the second update in September 2003, all but one well in January 2003 have been reported.

Figure 1-7. Class 11 Well Count Large Wells over 50 MMcf/day

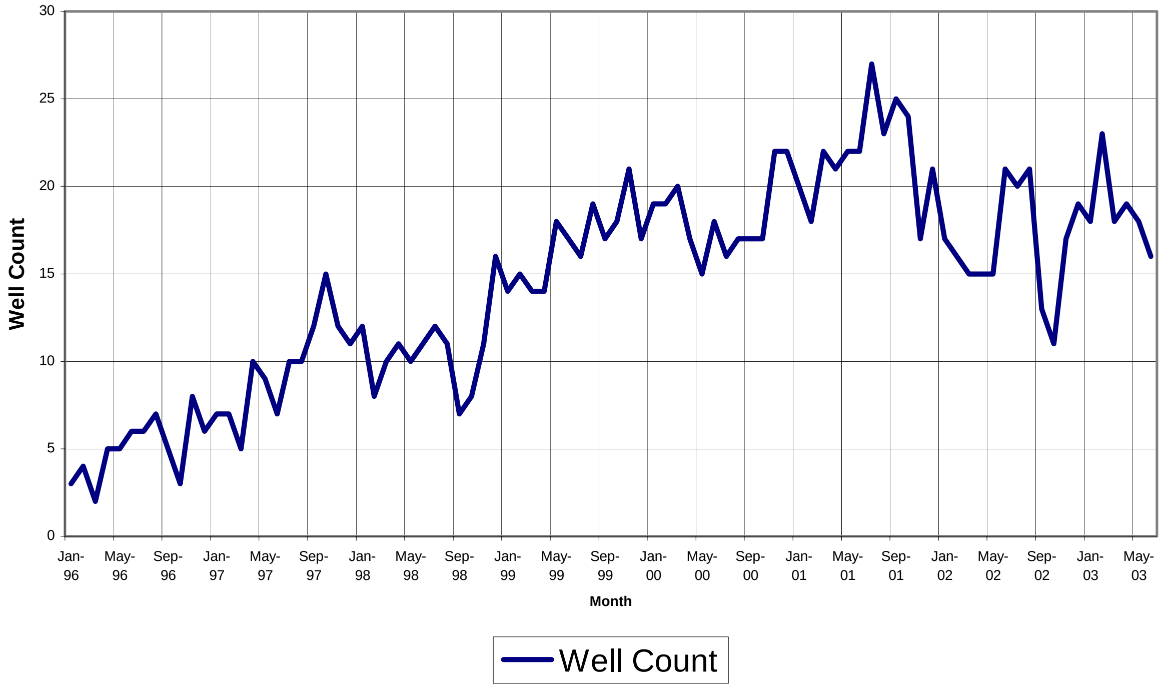


Table 1-2. Well Records Showing Missing Reported Values

Aug	Jun 03	May 03	Apr 03	Mar 03	Feb 03	Jan 03	Dec 02	Nov 02	Oct 02	Sep 02	Aug 02	Jul 02	Jun 02	May 02	Apr 02	Mar 02	Feb 02	Jan 02
177154108100	.	.	.	.	.	.	52.6	56.5	51.8	50.9	59.1	65.3	68.9	71.2	67.9	78.7	82.3	92.2
608044022101	.	52.2	52.1	55.2	54.2	.	60.9	60.1	56.9	56.6	59.7	57.6	.	.	.	.	.	.
608044022400	.	47.6	45.9	48.1	47.5	.	51.7	49.6	46.7	47.7	51.9	.	.	.	.	.	.	.
608044023400	98.1	99.0	96.7	52.1	.	.	.	.	.	.	.	.	.	.	.	.	.	.
608044023500	92.3	95.3	90.4	34.3	.	.	.	.	.	.	.	.	.	.	.	.	.	.
608164023900	.	.	45.3	45.4	37.4	48.1	50.9	53.7	45.2	43.6	52.0	57.5	57.6	59.8	59.6	56.4	57.8	61.1
608164024302	.	.	54.4	62.4	50.9	64.7	0.7	66.6	58.2	8.7	70.9	71.7	70.8	70.8	70.5	.	62.7	64.1
608164024700	.	.	78.6	73.0	60.2	77.5	79.2	80.7	63.9	69.0	72.2	80.2	79.0	76.9	51.3	30.9	35.7	38.6
608234000200	.	81.1	84.0	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Sep	Jun 03	May 03	Apr 03	Mar 03	Feb 03	Jan 03	Dec 02	Nov 02	Oct 02	Sep 02	Aug 02	Jul 02	Jun 02	May 02	Apr 02	Mar 02	Feb 02	Jan 02
177154108100							52.6	56.5	51.8	50.9	59.1	65.3	68.9	71.2	67.9	78.7	82.3	92.2
608044022101	50.2	52.2	52.1	55.2	54.2		60.9	60.1	56.9	56.6	59.7							
608044022400	46.4	47.6	45.9	48.1	47.5		51.7	49.6	46.7	47.7	51.9							
608044023400	98.1	99.0	96.7	52.1														
608044023500	92.3	95.3	90.4	34.3														
608164023900		51.2	45.3	45.4	37.4	48.1	50.9	53.7	45.2	43.6	52.0	57.5	57.6	59.8	59.6	56.4	57.8	61.1
608164024302		59.6	54.4	61.8	50.4	64.0	0.7	65.9	58.2	8.7	70.9	71.7	70.8	141.8			62.7	64.1
608164024700		75.4	78.6	73.0	60.2	77.5	79.2	80.7	63.9	69.0	72.2	80.2	79.0	76.9	51.3	30.9	35.7	38.6
608234000200	67.4	81.1	84.0	77.9	79.3	80.4	55.6	51.3	13.1									

Sep-Update	Jun 03	May 03	Apr 03	Mar 03	Feb 03	Jan 03	Dec 02	Nov 02	Oct 02	Sep 02	Aug 02	Jul 02	Jun 02	May 02	Apr 02	Mar 02	Feb 02	Jan 02
177154108100	53.2	55.1	50.7	50.9	53.8	56.1	52.6	56.5	51.8	50.9	59.1	65.3	68.9	71.2	67.9	78.7	82.3	92.2
608044022101	50.2	52.2	52.1	55.2	54.2	.	60.9	60.1	56.9	56.6	59.7	.	.	.	.	.	.	.
608044022400	46.4	47.6	45.9	48.1	47.5	.	51.7	49.6	46.7	47.7	51.9	.	.	.	.	.	.	.
608044023400	98.1	99.0	96.7	52.1	.	.	.	.	.	.	.	.	.	.	.	.	.	.
608044023500	92.3	95.3	90.4	34.3	.	.	.	.	.	.	.	.	.	.	.	.	.	.
608164023900	50.1	51.2	45.3	45.4	37.4	48.1	50.9	53.7	45.2	43.6	52.0	57.5	57.6	59.8	59.6	56.4	57.8	61.1
608164024302	58.7	59.6	57.7	61.8	50.4	64.7	65.7	65.9	58.2	8.7	70.9	71.7	70.8	70.9	79.6	.	62.7	64.1
608164024700	75.8	75.4	78.6	73.0	60.2	77.5	79.2	80.7	63.9	69.0	72.2	80.2	79.0	76.9	51.3	30.9	35.7	38.6
608234000200	67.4	81.1	84.0	77.9	79.3	80.4	55.6	51.3	13.1	.	.	.	.	.	.	.	.	.

The following are the formulas for the twelve-class model.

For a distribution with 12 classes:

$$P_i = \sum_{j=1}^{12} P_{i,j}$$

Where j is a class from 1 to 12.

The mean production per completion for any class is given by the following:

$$M_{i,j} = \frac{P_{i,j}}{W_{i,j}}$$

Where

$M_{i,j}$ = Mean production per completion of class j for month i
$W_{i,j}$ = Number of completions of class j for month i

Therefore, the production model for a class, and total are:

$$P'_{i,j} = W'_{i,j} * M_{i,j}$$

$$P'_i = \sum_{j=1}^{12} P'_{i,j}$$

Where $M_{i,j}$ is the actual reported mean production per completion of the sample distribution.

Production is then estimated by the following equation.

$$P'_i = \sum_{j=1}^{10} \left[M_{i,j} * WE'_{i,j} \right] + M_{i,11} * WE'_{i,11} + M_{i,12} * W_{i,12}$$

Where:

P'_i = Modeled production for month i
$M_{i,j}$ = Reported mean production per completion in class j for month i
$WE'_{i,j}$ = Modeled or expected number of completions in class j for month i
$WE'_{i,11}$ = Professional Expected number of completions in class 11 for month i
$W_{i,12}$ = Reported number of completions in class 12 for month i

We have the following options for estimating production:

1) Use the Expected/standard mean production per completion for each class.
2) Use the Reported mean production per completion for each class.
3) Use a Smoothed mean production per completion for each class.
4) Use a Flat Expected well completion count.
5) Use a Modeled Expected well completion count.

Modeled Well Completions from Rig Counts

For "normal" months a model based on the rig counts in the GOM can be used to estimate the expected number of wells. The model is benchmarked to the six-month reference period (as of December 2003, May to October 2001) and supplies an expected number of well completions for each month. The expected completions are computed via the recursive equation:

$$WE_i = \left[WE_{i-1} * e^{[A]} + B * SmRigs_i\right] * C$$

Where:

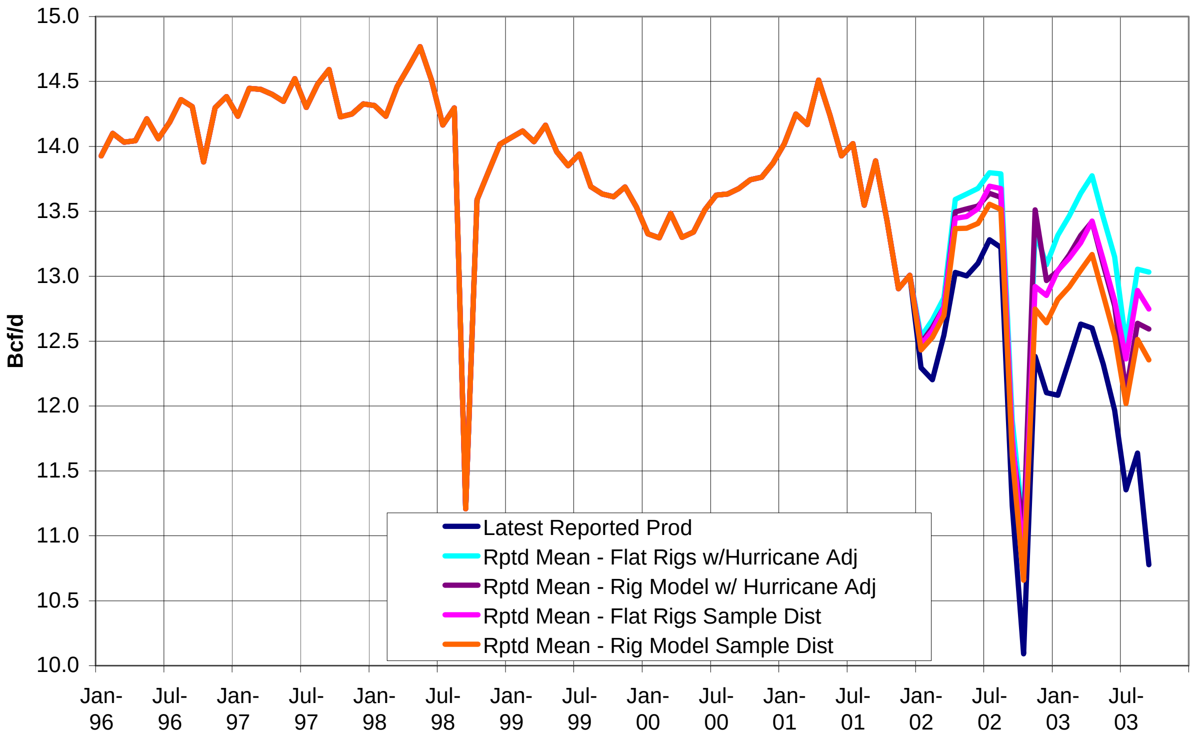
WE_i = Modeled or expected number of completions for month i
$SmRigs_i$ = Smooth GOM rig count (6 month exponentially smoothed (0.2857)) for month i
A = -0.00778 (fit parameter)
B = 0.379 (fit parameter)
C = 0.9998 (fit parameter used from December 1998 forward to reflect a change after a major storm in September 1998)
Initial WE_i = 6937 (fit parameter, starting point for the model)

The first half of the equation is a decline function that reduces the number of completions each month. The second half of the equation adds completions based on the smooth rig count. The resulting expected completions are shown in Figure 1-3.

Figure 1-8 is a plot of the reported production and estimated final production, as of December 2003. Reported production is essentially complete through October 2001 (accepted and edited data 99.5 percent complete plus RPD edited suspended data). Major hurricanes or storms occurred in the fall of 1998 and 2002. The estimated production shown here includes an empirical adjustment to the number of completions during the storms of 2002.

Figure 1-8 shows four estimates using the modeled well completions and flat well completions for both a twelve-class distribution and a single class distribution. All four cases include an empirical hurricane adjustment in 2002. As of December 2003, reported production is essentially complete through October 2001.

Figure 1-8. GOM Gas Production Estimates Compared to Reported Gas Production (as of December 2003)



Well Distribution Stability Test

A key part of the procedure is the determination of the expected/standard well distribution based on six months of essentially complete data. The stability of the well distributions was examined using a Chi-Square goodness of fit test. After applying this test to the distribution in the 12 months prior to the 6 standard months and later incomplete months, it was determined that all of the months have a similar distribution (an exception will be discussed later).

As an example, the following tables show the Chi-Square test for March 2001 and April 2003.

CHI-SQUARE Goodness of Fit Test

ProdDate=200103			
Frequency Count			
COUNT	Frequency	Percent	Test Percent
4429	4429	67.64	67.70
701	701	10.71	10.61
472	472	7.21	7.24
329	329	5.02	4.94
213	213	3.25	3.37
131	131	2.00	2.29
121	121	1.85	1.57
74	74	1.13	1.06
45	45	0.69	0.72
33	33	0.50	0.49

Chi-Square Test for Specified Proportions	
Chi-Square	6.5276
DF	9
Pr > ChiSq	0.6862
Sample Size = 6548	

ProdDate=200304			
Frequency Count			
COUNT	**Frequency**	**Percent**	**Test Percent**
4175	4175	69.53	67.70
631	631	10.51	10.61
405	405	6.74	7.24
273	273	4.55	4.94
170	170	2.83	3.37
129	129	2.15	2.29
90	90	1.50	1.57
57	57	0.95	1.06
48	48	0.80	0.72
27	27	0.45	0.49

Chi-Square Test for Specified Proportions	
Chi-Square	14.2235
DF	9
Pr > ChiSq	0.1146
Sample Size = 6005	

Hurricane Exceptions

As shown in the table below, for the months of September and October 2002 and January and July 2003 the Chi-Square test indicates that the monthly distributions are statistically significantly different from the standard 6-month distribution at the .05 percent level. November and December of 2002 are statistically significantly different at the .1 percent level. Tropical storm Isidore and Hurricane Lili in September and October affected production in these months and their effects lingered through January. When wells are shut in for a partial month or even several months, the distribution changes. An empirical downward adjustment to the number of expected completions is necessary for months with a major storm.

Chi Square Tests for stability

Month	Raw Count (10 Classes)	Raw Pr > Chisq	Rig Model Count	Rig Model Pr>ChiSq	Flat Model Count	Flat Model Pr>ChiSq
Feb-02	6218	0.3211	6380	0.4490	6393	0.4542
Mar-02	6287	0.9070	6370	0.9135	6393	0.9150
Apr-02	6190	0.2978	6362	0.5896	6393	0.6241
May-02	6157	0.5693	6354	0.7156	6393	0.7375
Jun-02	6164	0.6847	6337	0.7149	6393	0.7218
Jul-02	6152	0.3042	6327	0.3329	6393	0.3427
Aug-02	6127	0.5184	6319	0.5889	6393	0.6106
Sep-02	6020	0.0004	6319	0.0002	6393	0.0002
Oct-02	5628	<.0001	6311	<.0001	6393	<.0001
Nov-02	5734	0.0409	6295	0.0844	6393	0.0897
Dec-02	5798	0.0437	6285	0.0770	6393	0.0845
Jan-03	5884	0.0059	6276	0.0115	6393	0.0135
Feb-03	5902	0.2822	6265	0.3536	6393	0.3750
Mar-03	5897	0.6684	6252	0.7673	6393	0.7968
Apr-03	5884	0.4842	6241	0.5322	6393	0.5453
May-03	5883	0.7113	6232	0.7988	6393	0.8162
Jun-03	5911	0.1260	6226	0.1568	6393	0.1647
Jul-03	5915	0.0078	6216	0.0052	6393	0.0040
Aug-03	5801	0.0687	6203	0.0484	6393	0.0409
Sep-03	3742	0.5782	6199	0.1847	6393	0.1658

Part 2. Estimation Methodology for Natural Gas Production in Texas

Background

After a side-by-side evaluation, EIA recently replaced the model it used to estimate natural gas production in Texas with a new methodology, referred to as the "multinomial method." The report of the evaluation can be found at http://www.eia.doe.gov/pub/oil_gas/natural_gas/analysis_publications/ngprodtx/ngprodtx.pdf. The initial formulation of the multinomial methodology was presented in the master's thesis of Crystal Linkletter,[4] whose goal was to prepare timely estimates of natural gas production given the available data structure. The abstract is available at http://www.stat.sfu.ca/alumni/Theses/Linkletter.abs.shtml. The work was conducted under a research fellowship jointly sponsored by the American Statistical Association and the Energy Information Administration. The methodology had been used for product warranty estimation and in AIDS research.[5, 6] The model theory is based upon determining maximum likelihood estimates for probabilities associated with a multinomial distribution.

Data provided by Texas represent gross withdrawals. Preliminary monthly estimates for vented and flared; lease use; repressuring and nonhydrocarbon gasses removed are also available from the State of Texas.

Data Preparation

The Texas Railroad Commission posts natural gas production data on its website (http://www.rrc.state.tx.us/divisions/og/information-data/stats/ogismcon.html) and revises the data regularly over 24 months, with small revisions occurring sporadically over subsequent years. Data from the Texas Railroad Commission website are monthly updates of aggregate gross natural gas withdrawals for the most recent and all previous months. Data for the most recent month, denoted month t, are first available between 45 and 60 days after the close of the reference month.

These data are extracted and added to the historical data in a spreadsheet. The data are entered into a sheet in columns of monthly vintages. Each month a new column is entered with the first report for month t, denoted $P_{t,1}$, and revised reports for all previous months, denoted $P_{t-k,k+1}$ for k=1, … 96 (or the number of months from the first value included in the spreadsheet.) The data are then arranged into columns, one for each value of k, from k=1, …, 24. These data are the fundamental input for the model.

[4] Crystal Linkletter, "Predicting Natural Gas Production in the Presence of reporting Delays", Simon Fraser University, MSc Project, 2002.

[5] Brookmeyer, R. and Liao, J. (1990). "The Analysis of Delays in Disease Reporting: Methods and results for the Acquired Immunodeficiency Syndrome." *American Journal of Epidemiology*, **132**, 355-365.

[6] Kalbfleisch, J.D., Lawless, J.F. and Robinson, J.A. (1991). "Methods for the Analysis and Prediction of Warranty Claims." *Technometrics*, **33**, 273-285.

Basic Model

The theory for the multinomial model is based upon maximum likelihood estimates for certain probabilities associated with a multinomial distribution. Gas that is produced in month t will be included in either $P_{t,1}$ (the first report from the state of Texas), or $P_{t,2}$ (the second report from the state of Texas), or … $P_{t,24}$ (the 24th report from the State of Texas).[7] The partitioning of the gas produced into one of 24 reporting months can be viewed as defining a multinomial distribution with 24 possible report months for each tcf of gas produced. The basic probabilities in the multinomial distribution are the probabilities that a given tcf of gas will be reported in a given month, k. Based on the assumption that the multinomial distribution holds, a likelihood function can be written. At this step, the model is quite general, and the basic probabilities may change over time. However, to make it possible to compute a maximum likelihood estimate, the assumption is made that the probabilities remain constant (stationary) over the recent past (m reporting periods). With this assumption, maximizing the likelihood function with respect to the specific parameters needed to estimate the total production in a month at any point in the reporting process yields the expressions below. In particular, the model estimates $g_{t,k,}$ the conditional probability that gas produced in month t is reported in the k th report from the state of Texas given that it was reported on or before the kth report for k=1, …, 24

The stationarity assumption is that the reporting patterns have remained stable over the most recent m months, where m is a chosen time period (which can be specified parametrically). The model has been run with m=6 and with m=9. Larger values of m are preferred if the stationarity assumption holds because averaging more values results in a smaller variance. Smaller values of m are better if the assumption of stationarity does not hold. For the data currently available, results for m=6 appear to be somewhat better than for m=9 because there are increasing delays in company level reporting to the state.

The stationarity assumption is that $g_{t,k} = g_k$ over the most recently available m time periods. Under this assumption, maximum likelihood estimates for the conditional probabilities, $g_{t,k}$, are given by $g_{t,1} = 1$ and

$$g_{t,k} = \sum_{j=t+1-m-k}^{t+1-k} (P_{j,k} - P_{j,k-1}) \Big/ \sum_{j=t+1-m-k}^{t+1-k} P_{jk} \quad \text{for } k>1.$$

The $g_{t,k}$ are used to provide an estimate of the factor used to “weight up” a current report from the State of Texas, $P_{t,k}$ to prepare an estimate for the final reported production volume in month t.

The weight, which is used to adjust the k th estimate from the State of Texas for production at time t is the product of the conditional probabilities a unit of natural gas **not** being reported by time t+k

$$\hat{W}_{t,k} = \prod_{i=k+1}^{24} \left[1 - g_{t,i}\right]$$

---

[7] The number of months defining the multinomial distribution is a parameter of the model. Currently the value 24 is being used. In earlier years, 12 months might have been sufficient. However, delays in company level reporting to the State of Texas seem to be increasing.

Hence, the estimate for the final value of production for month t based on knowing the k^{th} preliminary value is obtained by dividing $P_{t,k}$ by $\hat{W}_{t,k}$, or

$$\tau_{t,k} = P_{t,k} / \hat{W}_{t,k}$$

For publication in the *Natural Gas Monthly* in its current production cycle, the third estimate for production in month t is used as the basis for estimation. Hence $\tau_{t,3}$ provides the estimate for publication. As the Natural Gas Monthly moves its production cycle forward $\tau_{t,2}$ or $\tau_{t,1}$ may be used to provide more timely estimates.

Prediction intervals

The variance of $\hat{W}_{t,k}$ is given by

$$V(\hat{W}_{t,k}) = \hat{W}_{t,k}{}^{2} \sum_{s=k+1}^{24} \frac{(\hat{g}_{t,s})}{(1-\hat{g}_{t,s})\sum_{j=t+1-m-k}^{t+1-k} P_{jk}} .$$

The approximate variance for the prediction interval $Y_{t,24} - \tau_{t,k}$ is given by

$$V(Y_{t,24} - \tau_{t,k}) = \frac{P_{t,k}{}^{2}}{\hat{W}_{t,k}{}^{4}} V(\hat{W}_{t,k}) .$$

The reader is referred to the references for more detailed information about the methodology and the derivation of the estimates and variances.

Part 3. Estimation Methodology for Natural Gas Production in Louisiana

Background

Louisiana reports production data similar to other States, in that there is a first preliminary value reported followed by a succession of revised reports for any given month. Relatively final production values are reached in 12 to 18 months. Vintage natural gas production[8] reports from Louisiana became available twice a year (every February and August) beginning in 1998, and have recently become available monthly.[9] These data are used to calculate a six-month median ratio of preliminary production values to final values; the six-month median ratio covers three years of data. The median preliminary to final ratios have an 18-month lag to assure that the ratio's final volume is very close to the ultimate final volume. The preliminary monthly reported value is multiplied by the appropriate six-month median ratio to yield an estimate of the final volume. This six-month median ratio is similar to the weights described in the Texas section of this document.

The data provided by Louisiana represent gross withdrawals. Louisiana provides estimates for aggregate volumes of vented and flared, repressuring, lease use and nonhydrocarbon gasses removed monthly, but does not provide the detail in terms of the EIA definitions. To estimate monthly values of these quantities, EIA assumes that they comprise the same percentage of production (gross withdrawals) during a given time period as they did during the same time period a year earlier.

Methodology

For a given month, the first preliminary estimate is labeled $P_{t,1}$, the second is labeled $P_{t,2}$, the third estimate, $P_{t,3}$, … the 24^{th} reported estimate is labeled $P_{t,24}$, and so on. As an example, the six-month median ratio described above for the $P_{t,3}$ preliminary value reported for June 2003 is the median of six ratios ($P_{t,24}/P_{t,3}$), for the six available months from June 1999 through December 2001. The $P_{t,24}$ reported estimate is assumed to be final. The reported value, $P_{t,3}$, is multiplied by the calculated six-month median ratio to determine the estimate for June 2003. The same process is used for all preliminary reported values from $P_{t,1}$ through $P_{t,23}$ and estimates are calculated for each.

$$\tau_{t,k} = \left(Median\ of\ \ 6_{k-17}^{k-54}\right) * \left(P_{t,k}\right)$$

EIA has asked Louisiana State officials to set up and make available on their web site a routine process to make production data available every month from their database. Currently, they run this retrieval only when asked. This custom retrieval built for EIA includes gross gas and all the reported dispositions.

[8] Vintage monthly data are years of historical monthly data for each report month, up to the month last reported.
[9] Louisiana provides these monthly data to EIA through a special database retrieval.

Prior to the implementation of this method, the method described below for Oklahoma was also used for Louisiana.

Part 4. Estimation Methodology for Natural Gas Production in Oklahoma

Background

Oklahoma and Texas serve some of the same natural gas markets. Oklahoma production data reports lag about four months, and their reporting patterns vary substantially and do not show the stable revision pattern that Texas data do. Relatively stable and close-to-final production data lag by at least a year. However, final production values, once they are reached, are relatively stable, although revisions are sometimes made after three years or more. EIA gets data on annual gross withdrawals from Oklahoma from a spreadsheet submitted by the State; it comes in time to be included in the *Natural Gas Annual (NGA)*.

Oklahoma provides monthly information on sales volumes, which most closely approximate marketed production. Oklahoma provides no information on vented and flared, repressuring, lease use or nonhydrocarbon uses.

Methodology

Using the similarities in production patterns for Texas and Oklahoma, (and the more reliable data from Texas), the current methodology for Oklahoma uses the ratio of month-to-month change in estimated Texas gas production (i.e., the Texas slope) to estimate production in Oklahoma.

In particular, let $\hat{V}tx_t$ represent the volumes estimated for Texas in month t using the methodology from Part 2 of this report. Then the "Texas slope" is estimated as

$$T\hat{S}_t = \frac{\hat{V}tx_t}{\hat{V}tx_{t-1}}$$

The Texas slope is applied to last month's estimate to obtain the current month's estimate of production. The estimated production in Oklahoma, $\hat{V}ok_t$, is given by:

$$\hat{V}ok_t = \hat{V}ok_{t-1} * T\hat{S}_t \quad .$$

This creates a series of monthly estimates based on the previous monthly estimate. Since final production is relatively stable, the results of this simple methodology have proven to be acceptable. The iterative process uses data from the State. Assuming that production data for Oklahoma are available for month t-k, Vok_{t-k} , the estimated production in month t, can be written:

$$\hat{V}ok_t = Vok_{t-k} * T\hat{S}_{t-k+1} * \ldots * T\hat{S}_t$$

Part 5. Estimation Methodology for Natural Gas Production in Other States

Background

EIA's primary source for natural gas production volume information is the Form EIA-895, "Monthly and Annual Quantity and Value of Natural Gas Production Report." The EIA-895 is completed by State agencies and the MMS. It is a voluntary survey designed to collect basic volumetric data associated with the natural gas product stream as it passes through several upstream activities including natural gas production, field gathering and processing. The goal of the data collection is to quantify the flow of natural gas from gross withdrawals through the volume of natural gas that is available for delivery to consumers. The natural gas product stream begins with gross withdrawals of gas from wells and is reduced along the way by various disposition and purification processes until the resulting product stream is total dry production ready for consumption. Appendix 1 provides a detailed description of the physical flow of the natural gas product stream and describes relationships between the form EIA-895 and that flow.

The EIA-895 is designed to collect most of the data components necessary to determine dry natural gas production, as presented on the *Natural Gas Annual* (*NGA*). Table 5-1 identifies data that are presented in the *NGM* and *NGA* and the associated data sources. The EIA-64A is the

Table 5-1. Natural Gas Production Data Concept: Sources of Data Presented in the *NGM* and *NGA*

Supply	Data Sources
A. Gross Withdrawals (excluding Lease Condensate)	Reported on EIA-895
Less	
B. Gas Used for Repressuring	Reported on EIA-895
C. Gas Vented and Flared (at Lease)	Reported on EIA-895
Equals	
D. Natural Gas Production (Wet) After Lease Separation	Calculated (D = A – B – C)
Less	
E. Nonhydrocarbons Removed	Reported on EIA-895
Equals	
F. Natural Gas Marketed Production (Wet)	Calculated (F = D – E)
Less	
G. Extraction Loss	Estimated from EIA-64 and EIA-816
Equals	
H. Dry Natural Gas Production (including lease fuel and plant fuel)	Calculated (H = F – G)
Disposition	
I. Lease Fuel	Estimated from EIA-895 and EIA-176
J. Plant Fuel	Estimated from EIA-64A and EIA-176

"Annual Report of the Origin of Natural Gas Liquids Production" survey and the EIA-816 is the "Monthly Natural Gas Liquids Report" survey. Both are required surveys completed by operators of natural gas processing plants.

States agencies that collect data on the volume of natural gas production are to submit the Form EIA-895 within 90 days of the end of the report month. However, incomplete, late or non-reporting by EIA-895 respondents has forced EIA to develop imputation methods for much of the data. Recently more States began making gas production information available on their websites.

The completeness and timeliness of reporting varies among the producing States. There are several categories of timing and completeness of reporting. Of the 32 natural gas producing States in the US, 10 States submit at least partial data for the EIA-895 in time for publication of the *Natural Gas Monthly*. The Natural Gas Division (NGD) extracts Wyoming data from its website, http://wogcc.state.wy.us/CntyTable.cfm, extracts Colorado data from different portions of its website, http://www.oil-gas.state.co.us/, and prepares estimates for the remaining 17 producing States that either do not respond to the EIA-895 or whose responses are not received by the *NGM* deadline. As described earlier, the Reserves and Production Division (RPD) bases its production estimates on data for Texas and MMS available on their respective websites, receives several special query reports from Louisiana, and receives an annual spreadsheet from Oklahoma.

More detailed information about the timeliness and completeness of the reporting of data on natural gas production is provided below and in Table 5.2

1. **Complete data and timely monthly reporting**. Only six States provide a value for all production and disposition items listed in the EIA-895 in a timely manner each month. These States, Alabama, Alaska, Arizona, Mississippi, Nevada and North Dakota, have relatively small production volumes, accounting for 5.6% of 2001 average monthly marketed production.

2. **Incomplete data but timely monthly reporting**. Four States, which represent 5.5% of the 2001 average monthly production, send partial data in response to the EIA-895. These states are California, Florida, Kansas, and Michigan. In addition, the timely but incomplete data that are collected by EIA staff from the Wyoming and Colorado websites accounted for another 9.8% of 2001 average monthly US production.

3. **Complete data but late monthly reporting**. Four States respond completely to the monthly form, but do not report in time to be included in that month's publication. The States in this category are Montana, New Mexico, Oregon, and Utah. They represented almost 10% of US production in 2001.

4. **No monthly data reported**. Fourteen States do not provide any monthly information on the EIA-895. These states are Arkansas, Illinois, Indiana, Kentucky, Maryland, Nebraska, New York, Ohio, Pennsylvania, South Dakota, Tennessee, Virginia, and West Virginia. These states accounted for 4.3% of U.S. production of natural gas in 2001.

5. **Estimates provided by the Reserves and Production Division (RPD).** Estimates for the States of Texas, Louisiana, and Oklahoma, as well as for the Federal Offshore in the Gulf of Mexico, are prepared by the RPD. Data for Louisiana, Oklahoma and the Federal Offshore are generally very late and incomplete. *{It is my understanding that they do not*

provide an EIA-895 form, instead we get the data from websites – but correct me if I am wrong.} Reliable information is available for the State of Texas on its website.

Table 5-2. Categories of Submission Completeness for Monthly Natural Gas Production Data[10]

Complete Data	Average Monthly Production (Mmcf)	Partial Data	Average Monthly Production (Mmcf)	Complete But Late Data	Average Monthly Production (Mmcf)	Does Not Submit Data	Average Monthly Production (Mmcf)
Alabama	43,483	California	31,485	Montana	6,783	Arkansas	13,900
Alaska	39,287	Florida	476	New Mexico	140,760	Illinois	15
Arizona	26	Kansas	40,012	Oregon	93	Indiana	89
Mississippi	8,962	Michigan	22,920	Utah	23,659	Kentucky	6,810
Nevada	1	Colorado*	58,020	Louisiana	437,449	Maryland	3
North Dakota	4,561	Wyoming**	113,657	Oklahoma	134,615	Nebraska	101
				Texas	543,325	New York	2,316
				GOM	(In 2001, was included in State totals)	Ohio	8,342
						Pennsylvania	13,083
						South Dakota	92
						Tennessee	167
						Virginia	5,962
						West Virginia	20,911
96,319		266,570		1,228,586		129,810	
5.6%		15.5%		71.4%		7.5%	

* Until recently Colorado submitted monthly data. Colorado website is accessed directly now.
** Similarly, Wyoming website is accessed directly now.

Because of the significant amount of missing EIA-895 data at the time of publication, NGD has developed a series of imputation and estimation procedures.

Data Imputation

If the completed Form EIA-895 from a State is not received by the monthly filing deadline, a staff member attempts to reach the contact person by phone to obtain the requested data. If attempts to receive data by phone are not successful, EIA prepares estimated (imputed) values for all missing items. When partial data are submitted or found of the State's website, values for the missing data elements are imputed.

Description of Algorithms

No data available for a State: The imputation method makes use of the seasonal month-to-month change patterns in past natural gas volumes of the same item to estimate missing values. In particular, if an estimate is needed for the volume in the current month, V_t, the month to

[10] Average monthly marketed production is shown, from *Natural Gas Annual* 2001.

month change factor is computed as the ratio of the average volume for the same month in the preceding 3 years to the average volume for the previous month in the preceding 3 years, or

$$F_t = \frac{\sum_{k=1}^{3} V_{t-12k}}{\sum_{k=1}^{3} V_{t-1-12k}}.$$

The imputed volume is given as the product of the month-to-month change factor and the estimate that was published in the previous month (the estimate in the previous month may have been an imputed value, or data provided by the State.)

$$\hat{V}_t = F_t x \hat{V}_{t-1}$$

This approach is used to estimate volumes of gross withdrawals, gas used for repressuring, gas vented and flared, nonhydrocarbon gases removed, and gas used as fuel on leases. Marketed production (excluding lease fuel) is equal to gross withdrawals minus gas used for repressuring, gas vented and flared, nonhydrocarbon gases removed, and gas used as fuel on the lease.

Partial data available for a State. Typically items missing on a partially completed form are natural gas used for repressuring, natural gas vented and flared, nonhydrocarbon gases removed, and natural gas used as fuel on leases. At least one of the natural gas production volumes (gross withdrawals, or marketed production) is provided. In this situation, the missing items are imputed based on their contribution to the available total production volume in the past three years as follows:

Let V_t represent the volume of the item to be imputed (repressuring, vented and flared, nonhydrocarbon gases removed or used as fuel on leases.) Let P_t represent the production volume reported on the form. The usual contribution factor for this item is computed as:

$$C_t = \frac{\sum_{k=1}^{3} V_{t-12k}}{\sum_{k=1}^{3} P_{t-12k}}$$

Hence, for each of the missing volumes on the form, the imputed volume for month t is given as the product of the usual contribution factor and the reported production volume for the current month.

$$\hat{V}_t = C_t x P_t$$

This approach is used to estimate volumes of gas used for repressuring, gas vented and flared, nonhydrocarbon gases removed, and gas used as fuel on leases. If gross withdrawals are reported on the form, marketed production (excluding lease fuel) is computed as gross withdrawals minus gas used for repressuring, gas vented and flared, nonhydrocarbon gases removed, and gas used as fuel on the lease. If marketed production is reported, gross withdrawals are computed from marketed production by addition.

If imputing an item and multiple totals exist, the order in which the programs perform their search is:

1st: Marketed Production excluding Lease Use
2nd: Marketed Production including Lease Use
3rd: Gross Withdrawals

Table 5-3. Categories of Submission Completeness for Partial Monthly Production Data

Monthly Data Collection for States that Submit Only Partial Data								
	Note: Source for each data item is either EIA-895, State website or imputed							
	Gross withdrawals from							
	Gas Wells	Oil Wells	Total	Repressuring	Vented and Flared	Non-HC gases removed	Lease Use	Marketed Production
California	895	895	895	895	Imputed	Imputed	Imputed	895
Florida	Imputed	Imputed	895	Imputed	Imputed	Imputed	Imputed	Imputed
Kansas	Imputed	Imputed	895	895	895	Imputed	Imputed	895
Michigan	895	895	895	Imputed	Imputed	895	Imputed	Imputed
Colorado	Imputed	Imputed	website	Imputed	Imputed	Imputed	Imputed	Imputed
Wyoming	Imputed	Imputed	website	Imputed	Imputed	Imputed	Imputed	Imputed

Annual

The annual EIA-895 is to be submitted along with the December monthly report – or 90 days after the close of the calendar year. However, the *Natural Gas Annual* is typically published in the fall of the year, which allows publication of data for many States whose EIA-895 data are received late. Table 4.x describes which States provide complete data in time for the Natural Gas Annual, which provide partial data, and which provide no data. When States do not provide any data and none can be obtained from websites or secondary sources, data are imputed as the average of the last 3 years of estimates.

Table 5-4. Categories of Submission Completeness for 2001 Annual Production Data

	Complete	Annual Production (MMcf)	RPD	Annual Production (MMcf)	Partial	Annual Production (MMcf)	Does not submit	Annual Production (MMcf)
1	Alabama	521,799	Texas	6,519,904	New Mexico	1,689,125	Colorado	696,237
2	Alaska	471,440	Louisiana	5,249,383	Oklahoma	1,615,384	West Virginia	250,932
3	California	377,824			Wyoming	1,363,879	Pennsylva nia	157,000
4	Mississippi	107,541			Kansas	480,145	Illinois	185
5	Kentucky	81,723			Michigan	275,036	Maryland	32
6	Virginia	71,543			Arkansas	166,804		
7	North Dakota	54,732			Ohio	100,107		
8	Tennessee	2,000			Montana	81,397		
9	South Dakota	1,100			New York	27,787		
10	Arizona	307			Florida	5,710		
11					Nebraska	1,208		
12					Oregon	1,110		
13					Indiana	1,064		
14					Utah	283,913		
15					Nevada	7		
Total Volume		1,690,009		11,769,287		6,092,676		1,104,386
% of Total		8.18%		56.98%		29.50%		5.35%

Note: Marketed production is from *Natural Gas Annual 2001*.

Appendix 1. The Natural Gas Product Stream and the EIA-895

The Form EIA-895, (shown in Figure A1-1) is a voluntary survey designed to collect basic volumetric and cost data associated with the natural gas product stream as it passes through several upstream activities including natural gas production, field gathering and processing. The data identified for collection on the EIA-895 begin with the produced raw natural gas and

Figure A1-1a. EIA-895 Form (Monthly Schedule)

www.eia.doe.gov
Energy Information Administration

Form Approved
OMB No. 1905-0175
Expiration Date: 11/30/05

MONTHLY AND ANNUAL QUANTITY AND VALUE OF NATURAL GAS PRODUCTION REPORT
FORM EIA-895

This report is voluntary under the Federal Energy Administration Act of 1974 (Public Law 93-275). For the provisions concerning the confidentiality of information and sanctions, see instructions.

Revised Report ☐

SECTION I. IDENTIFICATION DATA

EIA ID NUMBER: 8 9 5

State Name:
Name of Office/Agency:
Address 1:
Address 2:
City:

State: Zip Code: -
Contact Name:

Phone No.: - - Ext:
Fax No.: - -

Email address:

Report Period: **Mo** **Year** 2 0

Please submit this form by the 90th day after the report month to:

Energy Information Administration, EI-45
U. S. Department of Energy
P.O. Box 8279
Silver Spring, MD 20907
ATTN: Form EIA-895

Fax to: (202) 586-1076
OR
Email to: OOG.SURVEYS@eia.doe.gov

Questions? Call (202) 586-6119

SECTION II. MONTHLY NATURAL GAS VOLUMES (Report all volumes in thousands of cubic feet)

Item		**Volume (Thousand cubic feet)**
Gross Withdrawals (excluding lease condensate)		
Gas and Condensate Wells	**(1)**	
Oil Wells (Casinghead)	**(2)**	
Coalbed Methane Wells	**(3)**	
TOTAL (sum of 1+2+3)	**(4)**	0
Natural Gas Used for Repressuring or Reinjection	**(5)**	
Natural Gas Vented and Flared	**(6)**	
Nonhydrocarbon Gases Removed	**(7)**	
Natural Gas Used as Fuel on Leases (lease fuel)	**(8)**	
Marketed Production (excluding lease fuel) (equals 4 - 5 - 6 - 7 - 8)	**(9)**	0

(10) Pressure base at which all volumes are reported (psia at 60° Fahrenheit)

SECTION III. COMMENTS: *To enter a new line of comments, hold down the alt key and hit the enter key.*

Figure A1-1b. EIA-895 Form (Annual Schedule)

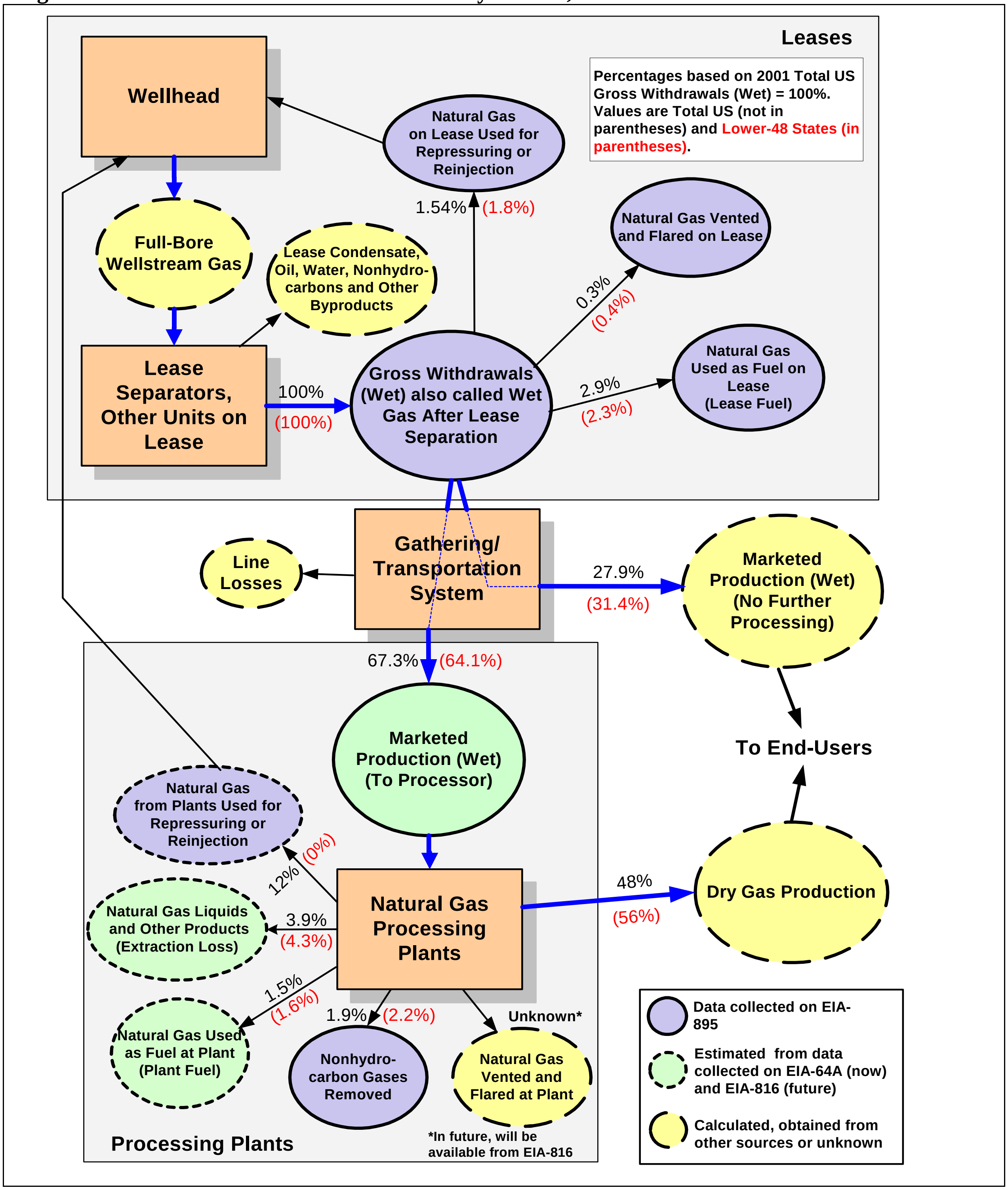
SECTION IV. ANNUAL SCHEDULE (To be completed when a calendar year of monthly reports has been completed									
Total number of producing gas wells in operation as of December 31 for the reporting year. Pressure base at which all volumes are reported (psia 60° Fahrenheit)									
	Gross Withdrawals (excluding lease condensate)								
Month	**Gas and Conden-sate Wells (1)**	**Oil Wells (Casing-head) (2)**	**Coalbed Methane Wells (3)**	**Total (4)**	**Used for Repressur-ing or Reinjection (5)**	**Natural Gas Vented and Flared (6)**	**Non-hydro-carbon Gases Removed (7)**	**Natural Gas Used as Fuel on Leases (Lease Fuel) (8)**	**Marketed Production (9)**
January									
February									
March									
April									
May									
June									
July									
August									
September									
October									
November									
December									
Total									

SECTION V. Comments

SECTION VI. ANNUAL SCHEDULE (To be completed when a calendar year of monthly reports has been completed)		
Month	**Value of Marketed Production**	**Quantity of Marketed Production (Value Based)**
January		
February		
March		
April		
May		
June		
July		
August		
September		
October		
November		
December		
Total		

SECTION V. Comments

Figure A1-2. Natural Gas Product and Activity Stream, Relative Volumes and EIA Data Sources



proceeds through various reductions in the flowing volume to result in the marketed natural gas, as depicted in Figure A1-2. The percentages shown (black is for total US; red is for total US less Alaska) represent the portion of Gross Withdrawals (Wet) of the item indicated. For example, Dry Gas Production is 48% of Gross Withdrawals (Wet) for total US. It is noteworthy that in

Alaska about 90% of natural gas produced comes from oil wells and that most of the dry natural gas coming out of processing plants is used for repressuring oil wells.

The natural gas production process begins with the lease, which contains wells from which natural gas is extracted. The lease boundary will hold one or more natural gas producing wells. Natural gas may be produced from three types of wells, natural gas and condensate wells that primarily produce raw natural gas; oil wells that produce oil as the primary product but also produce raw natural gas (casing head gas); and coal bed methane wells which are natural gas wells drilled into coal seams.

Raw natural gas contains water vapor, hydrogen sulfide (H_2S), carbon dioxide, helium, nitrogen, and other compounds. Once the raw gas is produced, it flows through field separators that remove free water, some hydrocarbons and other byproducts (chemicals, drilling fluids, etc.) from the gas stream. The separated gas is commonly referred to as unprocessed wet gas and is reported as the Gross Withdrawals on lines 1 through 4 of the EIA-895. At this point, some portion of the gas may be reinjected into the well. Reinjection will normally only take place in oil wells producing natural gas. The gas may be used to maintain pressure of the well or help lift the heavier hydrocarbons to the surface. The volume of gas reinjected, which is reported on Line 5 of the EIA-895, represents a reduction in the product stream before it leaves the lease.

One element of consumption of natural gas on lease sites is the volume of gas used for fuel on the lease for drilling operations, heaters, dehydrators, and field compressors. This volume of gas is recorded on line 8 of the EIA-895 as "Natural Gas Used as Fuel on Leases (lease fuel)" and on the Form EIA-176. Natural gas consumed as lease fuel is very common in production activities. Some of the lease fuel reported on the EIA-895 may be associated with activities that transcend lease boundaries. For example, fuel used for compressor station operations on gathering lines could be considered lease fuel although the activity takes place outside the lease boundary.

Another reduction in product stream volume at the lease is the volume of natural gas that is vented and flared. Venting is the controlled release of natural gas into the atmosphere while flaring is the controlled burning of natural gas, both for disposal purposes. Venting and flaring may be necessary for reasons of safety or operating practices. For example, venting and flaring may take place during well testing, purging of process systems with inert gas or disposal of unmarketable gas. In addition to disposal at the lease boundary, venting and flaring frequently occurs at the processing plant. However, the EIA-895 is not designed to collect processing plant venting and flaring although some States may be including venting and flaring volumes from natural gas plants on the EIA-895.[11]

The product stream remaining after reinjection, venting and flaring, and use as lease fuel flows into the gathering system where it is combined with natural gas from other leases and is transported to the processing plant. Some natural gas is able to bypass the processing plant if it already meets pipeline specifications. There may be additional minor reductions in product stream volume due to pressure differentials and leakage (line loss). The product stream on the inlet of the processing plant is referred to as 'wet marketed production excluding lease fuel'.

[11] At least two states, Utah and Wyoming, include natural gas processing plant vented and flared volumes in the vented and flared data for the EIA-895 submission, but NGD doesn't have information about which of the other states include natural gas processing plant vented and flared volumes with their EIA-895 submission. Without this information, it's not possible to know what vented and flared volumes are not accounted for.

The volume at the inlet side of natural gas processing plants is a physically measurable quantity, and is available annually on the EIA-64A. Beginning with data for the report month January 2003, it has also been collected on the EIA-816 but these data are not yet ready for use in accounting for production.

The product stream on the inlet side of the processing plant contains natural gas, natural gas liquids, nonhydrocarbon gases and other impurities. The natural gas processing activity separates the components of the product stream into individual marketable and non-marketable products. This process reduces the volume of the product stream by the removal of non-hydrocarbon gasses (not already removed before the plant), which are recorded on line 7 of the EIA-895.

The processing of natural gas also removes natural gas liquids. The volume of natural gas liquids removed from the product stream may vary depending on natural gas and natural gas liquids market conditions or other economic factors. There is a limit, imposed by transportation companies, on how much natural gas liquids may remain in the dry natural gas product stream. The reduction in product stream due to removal of natural gas liquids is based on data reported annually by natural gas plants on the EIA-64A, and data reported monthly by natural gas plants on the EIA-816.

EIA adds the volume of natural gas used as fuel on the lease to the product stream after the processing plant removal of natural gas liquids and other products to arrive at the dry production volume published in the *NGA.* Dry natural gas shipments have been collected as natural gas shipments from processing plants on the EIA-816 beginning in January 2003, but the data are not yet (August 2003) considered reliable. The following formula summarizes the various measurable quantities and how they are employed to arrive at the dry natural gas production for the *NGA.*

Marketed Production (wet) $MP_w = GW - R - VF_l - NHG - F_l$
Marketed Production (dry) $MP_d = MP_w - NGL + (F_l)$

MP_d = dry marketed production
MP_w = wet marketed production

GW = gross withdrawals (reported on the EIA-895)
R = natural gas used for repressuring or reinjection (reported on the EIA-895)
VF_l = natural gas vented and flared at the lease (reported on the EIA-895)
F_l = natural gas used as fuel on leases (includes gathering) (reported on the EIA-895 and EIA-176)
NGL = natural gas liquids (reported on the EIA-64A and on the EIA-816)
NHG = non-hydrocarbon gases (reported on the EIA-895)

The amount of both wet and dry gas production may be overstated by the amount of line losses, for which EIA does not have data. In addition, the amount of available dry gas may be overstated by the amount of venting and flaring at natural gas plants (VF_p) and the removal of nonhydrocarbon gases at processing plants to the extent they are not included in line 7 of the EIA -895. All of these factors would tend to inflate the amount of dry marketed natural gas.

In terms of consumption, EIA separately identifies the natural gas used as fuel on leases. The sources of data for this series are the EIA-176 and the incomplete reports of this series on the EIA-895. The information on fuel used at natural gas processing plants has been collected on the EIA-816 beginning with the report month of January 2003, but has not reached a satisfactory quality level. This monthly information will be examined to determine if the reported data can replace a portion of the estimate for lease and plant fuel now used in Table 3 of the *NGM*.

Appendix 2. Brief Description of Previous GOM Estimation Methodologies

The Reserves and Production Division (RPD) has used a variety of methods to estimate the Federal Offshore Gulf of Mexico monthly gas production. Since 1996 (about 5.5 years of "final" data) these methods have yielded an absolute average error of 2.5 percent and an average error of 0.43 percent. Below is a brief description of some of the previous methods employed.

Method A: Average month-to-month percent change (Used prior to October 2001)

- An average percent change for the same month for several selected past years is applied to the previous month's estimate.
- Specific years are selected for this, usually 3, 4, or 5, but not necessarily the most recent years.
- The analyst selects the specific years to use in calculating the average to drive the resulting estimate to a "reasonable" production volume.
- This method works well if production has been stable for a long time and does not change.
- Some expert judgment is required.
- A hurricane may require a complete analyst override of the process.
- Since the current estimate is based on the previous months estimate, that in turn may not be based on preliminary data, there is a potential for compounding errors especially given an infrequent revision policy

Method B: Applying the month-to-month percent change in Texas to the GOM (Used from October 2001 to December 2002 *NGM*)

- Method was used when GOM final production data were not available for 1.5 years or more.
- Month-to-month percent changes in the Texas production data or estimates were applied to the GOM beginning at the last final data month.
- The last final data month is the last month of 95 percent or more complete accepted GOM data.
- The last final data month and the previous 11 months were adjusted upwards by an average or a median of the preliminary-to-final data ratios for the GOM based on several years of revision history for the GOM.
- Like the GOM, Texas also supplies about 25 percent of the nation's total production.
- It is assumed that both the GOM and Texas supply most of the same market and therefore changes in market demand affect both areas about the same. While this appears to be true in a general sense, specific months may be quite different.
- Actual preliminary-to-final data ratios can vary from the average.
- Requires analyst override in the event of a hurricane or other extraordinary event that affects the GOM and Texas differently.

Method C: Texas Linear Model (Used for December 2002 through February 2003 *NGM*)

- This method uses smoothed Texas production data as input into a linear model to estimate GOM production.

- A least squares fit process determines model parameters.
- The particular time interval for fitting the model can be varied by the analyst to somewhat steer the resulting estimate; some expert judgment is required.
- This method was used when GOM final production data were not available for 1.5 years or more and applied after the last final data month.
- The last final data month is the last month of 95 percent or more complete accepted GOM data.
- The last final data month and the previous 11 months were adjusted upwards by an average or a median of the preliminary-to-final data ratios for the GOM based on several years of revision history for the GOM.
- Like the GOM, Texas also supplies about 25 percent of the nation's total production.
- Both the GOM and Texas supply most of the same market and therefore changes in market demand affect both areas about the same. While this appears to be true in a general sense, specific months may be quite different.
- Actual preliminary-to-final data ratios can vary from the average.
- Works well when both the GOM and Texas production are stable.
- Requires analyst override in the event of a hurricane or other extraordinary event that affects the GOM and Texas differently.

<u>Method D (Used for the March and April 2003 *NGM*)</u>:

- Product of the smoothed mean production per completion and the expected number of completions.
- Past models used only the 95 percent complete accepted GOM data.
- This method uses the 95 percent complete accepted data, the accepted subsequent data (less than 95 percent complete), and the newly available suspended data after some editing.
- The modeled or expected number of completions is determined by projecting forward an exponentially 12-month smoothed completion count of the latest approximately final data month.
- The modeled production per completion uses a modified exponential smoothing that adds a ratio of actual completions over the expected or modeled completion count to the smoothing term to estimate the production per completion.
- The modified smoothing requires the model to depend less on current data and more on history as the number of reported completions declines.
- When the completion ratio is applied it appears that very little if any actual exponential smoothing is required.
- Because this method uses all available data, the estimates are based on much more current data than the previously described methods.
- The suspended data are edited to remove duplicate records, about 9,000 for April 2003.
- As the MMS releases more suspended data, i.e., they catch up; estimates based on this methodology will be even better.
- This method requires analyst intervention only when an extraordinary event occurs like a hurricane.

<u>Method E (Used for the 2001 *NGA*, and for 2002-03 data in *NGM*)</u>

- Because MMS was not able to provide monthly data after July 2001, NGD used estimates provided by RPD for the remaining months of 2001 to prepare the 2001 estimate for 2001.
- For the months of January –October 2002,EIA/NGD published MMS natural gas preliminary production data taken from the MMS web page after they became available in winter 2003. Although these MMS data have been updated, the NGM estimates have remained the ones used for the March NGM.
- Estimates of GOM production for November 2002 through current months have been provided by RPD.

Appendix 3. Previous Estimation Method for Texas

Background

From about June 2001 through August 2003, EIA used the method described in this Appendix to estimate natural gas production in Texas. This previous method is referred to as the "parametric" method. It has been replaced by the multinomial method, described in Part 2 of this report. A discussion of the comparative evaluation that resulted in the replacement of the parametric method with the multinomial method can be found at http://www.eia.doe.gov/pub/oil_gas/natural_gas/analysis_publications/ngprodtx/ngprodtx.pdf.

The Texas Railroad Commission posts natural gas production data on its website (http://www.rrc.state.tx.us/divisions/og/information-data/stats/ogismcon.html) and revises the data regularly over 24 months, with small revisions occurring sporadically over subsequent years. EIA's previous estimation model is a spreadsheet model that uses this 24-month historical data revision pattern as a template to estimate final production values from preliminary data. Reported volume data generally start out low and approach their final reported values according to a relatively stable pattern (curve). The revision pattern is determined from history, for which 24 months of revisions are available for a given production month. The historical pattern is then applied to recent production months, for which reported production may have been revised up to 23 times (i.e., every month). The model also attempts to account for changes in the relationships between the preliminary data and final data over time.[12]

For production months with at least seven pieces of information (six revisions), the model works very well. For production months with fewer than seven pieces of information (the six most recent months with 0 - 5 revisions) some additional controls or parameters are used. An alternative method (multinomial model) using the Texas website data is described in Appendix 3.

Data Preparation

The data are entered into a sheet in columns of monthly vintages. Each month a new column is entered with the first report for the current month and revised reports for all previous months. The data are then organized into columns of first reported data, second reported, third reported . . . (i.e., first preliminary, second preliminary, third preliminary . . .columns of data) referred to as $P_{i,1}$, $P_{i,2}$, $P_{i,3}$, etc. These data, organized by "P's" are the fundamental input for the model.[13]

Basic Model

The fundamental model equation is below.

---

[12] Prior to the development of the model described in this section (in February 2003), an average historical month-to-month change was used to estimate Texas monthly gas production, as is now done for Louisiana and Oklahoma. The average was taken from 2 - 8 years of historical month-to-month changes for the particular month for which the estimate was being determined. The estimates were determined by adding an incremental volume to either a P_2 or P_3 reported volume. The incremental volume was a calendar year average of the difference between the reported P_2 or P_3 monthly volumes and the latest reported monthly volumes with a lag of at least one calendar year. The analyst selected the specific years used to calculate the average month-to-month change. The average was applied to last month's estimate or successive averages were applied beginning with the latest close-to-final reported production data.

[13] Found on the 1stSheet tab in the TexasModel29.xls workbook.

$$BF_{i,j} = \frac{P_{i,j}}{1 + MC_{i,j} + E_{i,j}}$$

where $MC_{i,j} = -\left(1 - \frac{P_{i,j}}{BF_{i,j}}\right)$

$BF_{i,j}$	= Best Final estimated production
$P_{i,j}$	= Preliminary reported production data
$MC_{i,j}$	= From the smoothed lagged 6 month median model of $MC_{i,j}$
$E_{i,j}$	= Error, amount not accounted for by $MC_{i,j}$
i	= Production month
j	= Number of the preliminary estimate for production month i

$MC_{i,j}$ Model

The $MC_{i,j}$ model is fit first. Since this term is based on a smoothed 6 month lagged, 6 month median, calculated value of $MC_{i,1}$, the $BF_{i,j}$ fit parameter can be used to calculate the $MC_{i,j}$ term which is then used to determine a later $BF_{i,j}$ parameter and so on. This "cascading" through the historical data carries the revision pattern forward through the current month's estimate, i.e., $BF_{i,1}$ from $P_{i,1}$.

The $MC_{i,j}$ model is based on the smoothed 6 month lagged, 6 month median, calculated value of $MC_{i,1}$ term as a starting point for the revision pattern. The Z term allows the revision pattern to change over time as the relationships between the preliminary values and the final values change.

$$MC_{i,j} = \frac{MC_{i,1}}{\left(1 + (j-1) * Z\right)^C} \qquad \text{For j = 2 to 24} \qquad \text{Where } Z_i = A * \left(1 + B * MC_{i,1}\right)$$

A	= 0.678
B	= 8.367
C	= 3

The $MC_{i,j}$ model is fit over about 6 years of historical data where the $P_{i,24}$ values are available. For this historical period $BF_{i,j}$ is equal to $P_{i,24}$. The fit parameters A, B, and C are determined by a least squares fit. The $MC_{i,j}$ model with its determined fit parameters A, B, and C is cascaded from the historical data fitting period down through the most recent months of reported data.

$BF_{i,j}$ Model

The $BF_{i,j}$ model determines the Best Final production value for up to 24 simultaneous equations for each production month. The Basic Model equation is rearranged as follows and a least squares fit is used to minimize the difference between modeled Ps and actual Ps. The $BF_{i,24}$ becomes a fit parameter. The additional control mentioned above for the first six months appears here as the error term $E_{i,j}$ (see 2ndSheet tab in TexasModel29.xls workbook).

$$P_{i,j} = BF_{i,24} * \left(1 + MC_{i,j} + \sum_{j}^{24} dE_{i,j} + E_{i,j=1\text{ to }6}\right)$$

$$dE_{i,j,j+1} = FC_{i,j} - FC_{i,j+1} + MC_{i,j+1} - MC_{i,j}$$

$$FC_{i,j} = \frac{P_{i,j}}{BF_{i-1,24}} - 1$$

This model is used without the error term from $P_{i,7}$ through $P_{i,24}$ and with the error term from $P_{i,1}$ through $P_{i,6}$. Where $P_{i,24}$ is available, $BF_{i,24}$ is set equal to $P_{i,24}$. Otherwise, $BF_{i,24}$ is a fit parameter.

$P_{i,1}$ through $P_{i,6}$ Models

For $P_{i,1}$ through $P_{i,6}$, a separate model is used for each $P_{i,j}$ to estimate the error term. For each estimate the current $P_{i,j}$ and all previous $P_{i,j}$'s and previous $BF_{i,j}$ estimates are used. Most of the fit parameters are in the error term (described below). $BF_{i,j}$ is a fit parameter and minimized in a least squares fit everywhere a $P_{i,24}$ is available in each $P_{i,j}$ model and all modeled $P_{i,j}$'s are minimized against actual $P_{i,j}$'s in the same least squares fit (see 3rdSheet tab in TexasModel29.xls workbook).

For example, the $BF_{i,3}$ estimate for the $P_{i,3}$ reported data uses $P_{i,1}$, $P_{i,2}$, and $P_{i,3}$ with an error term model to determine the $E_{i,3}$ term. The $P_{i,1}$, $P_{i,2}$, and $P_{i,3}$ model equations, listed below, and the error functions are simultaneously fit.

$$P_{i,1} = BF_{i,3} * \left(1 + MC_{i,1} + dE_{i,1,2} + dE_{i,2,3} + E_{i,3}\right)$$

$$P_{i,2} = BF_{i,3} * \left(1 + MC_{i,2} + dE_{i,2,3} + E_{i,3}\right)$$

$$P_{i,3} = BF_{i,3} * \left(1 + MC_{i,3} + E_{i,3}\right)$$

The error term is defined as follows:

$$E_{i,3} = FE_{i,3} * \left(FC_{i,3} - MC_{i,3} - PC_{i,3}\right)$$

Where

$$PC_{i,3} = \frac{BF_{i,2} - BF_{i-1,24}}{BF_{i-1,24}}$$

If

$$W_{i,3} = \left(FC_{i,3} - MC_{i,3} - PC_{i,3}\right)$$

For $W_{i,3} < 0$

$$FE_{i,3} = 1 - \frac{1 - A_{i,3}}{1 + \left(BN * |W|\right)^2}$$

For $W_{i,3} >= 0$

$$FE_{i,3} = \frac{A_{i,3}}{1 + \left(BP * |W|\right)^2}$$

Where

$$A_{i,3} = A * e^{\left[D * \left(1 - \frac{0.005 + |dE_{i,1,3}| + |W|}{0.005 + |dE_{i,1,3}|}\right)\right]}$$

A = 1.244
BN = 73.36
BP = 156.4
D = 0.395

The $P_{i,1}$ through $P_{i,6}$ Models are fit or minimized in sequence as part of an iterative process with the $BF_{i,j}$ Model. The $P_{i,1}$ through $P_{i,6}$ models are fit sequentially because each one depends on the $BF_{i,j}$ from the previous one. The last error term from each $P_{i,1}$ through $P_{i,6}$ model is used in the last 6 terms or months of the $BF_{i,j}$ model described above. The $BF_{i,j}$ Model is then fit using the supplied error terms ($E_{i,1}$ through $E_{i,6}$). Because the $P_{i,1}$ through $P_{i,6}$ Models are also dependent on the results of the $BF_{i,j}$ Model the $P_{i,1}$ through $P_{i,6}$ Models are fit again sequentially. The last error term from each $P_{i,1}$ through $P_{i,6}$ model is used again to revise the last 6 terms or months of the $BF_{i,j}$ model. Approximately five iterations of the $P_{i,1}$ through $P_{i,6}$ Models and the $BF_{i,j}$ Model are necessary to optimize the resulting monthly production rate estimates.
This appendix describes a multinomial model being tested by the Office of Oil and Gas to estimate natural gas production in Texas.[14] Results of evaluation tests are also described.

[14] This methodology was developed by C. Linkletter, R. Sitter and J. Vetter for the Office of Oil and Gas. It was presented to the American Statistical Association (ASA) Committee on Energy Statistics in 2003, which endorsed testing it as an alternative to the existing methodology.

Appendix 4. Percent Revision Between First Released and Current Best Estimate of Natural Gas Marketed Production

EIA tracks the accuracy of its data estimates by calculating the difference (i.e., revision) between its first released (i.e., published) estimate of data and its current best estimate of data. Table A4-1 shows the percent revisions for *Marketed Production of Natural Gas* for the five largest producing States and the Federal offshore Gulf of Mexico for 1999, 2000 and 2001. At the national level, the changes from the first published estimate to the current estimate have all averaged less than three percent each year. However, this would not reflect a data change made after the first release but which was later changed back, nor does it measure whether the month to month trend is consistent over time (i.e., whether production is increasing or decreasing).

Table A4-1. Percent Revision From First Released Estimate To Current Best Estimate for *Marketed Production of Natural Gas (1999-2001)*

1999	LA	NM	OK	TX	WY[15]	Federal Gulf	Total US
Max percent error	9.46	14.80	5.62	-.58	-14.16	4.52	.20
Min percent error	-.40	-.43	-7.92	-2.29	-26.40	.39	-2.70
Mean percent error	4.71	6.22	-2.26	-1.26	-19.33	2.22	-1.23
Mean absolute percent error	4.77	6.29	3.71	1.26	19.33	2.22	1.26
2000							
Max percent error	8.73	2.62	1.53	-.08	-4.95	6.20	.55
Min percent error	-2.26	-16.65	-7.30	-3.53	-34.97	.05	-3.06
Mean percent error	5.96	-8.60	-3.32	-2.25	-22.19	3.94	-2.06
Mean absolute percent error	6.38	9.04	3.58	2.25	22.19	3.94	2.15
2001							
Max percent error	4.56	-.30	4.55	-2.02	.46	5.15	.80
Min percent error	2.07	-17.11	-.80	-6.31	-24.89	-.84	-3.51
Mean percent error	3.09	-8.80	.98	-3.13	-13.43	1.72	-1.41
Mean absolute percent error	3.09	8.80	1.31	3.13	13.50	1.99	1.54
Three year mean absolute percent error	4.75	8.04	2.86	2.21	18.34	2.72	1.65

Sources: Table 7 of EIA *Natural Gas Monthly* for first release, internal EIA estimates for current best estimate, which may reflect data that has not yet been published.
Note: Percent revision is defined as: 100 * (first release – current) / current.

[15] The large revisions for Wyoming were due to corrections in the data supplied to EIA.