Accuracy of Petroleum Supply Data

by Tammy G. Heppner and Carol L. French

Overview

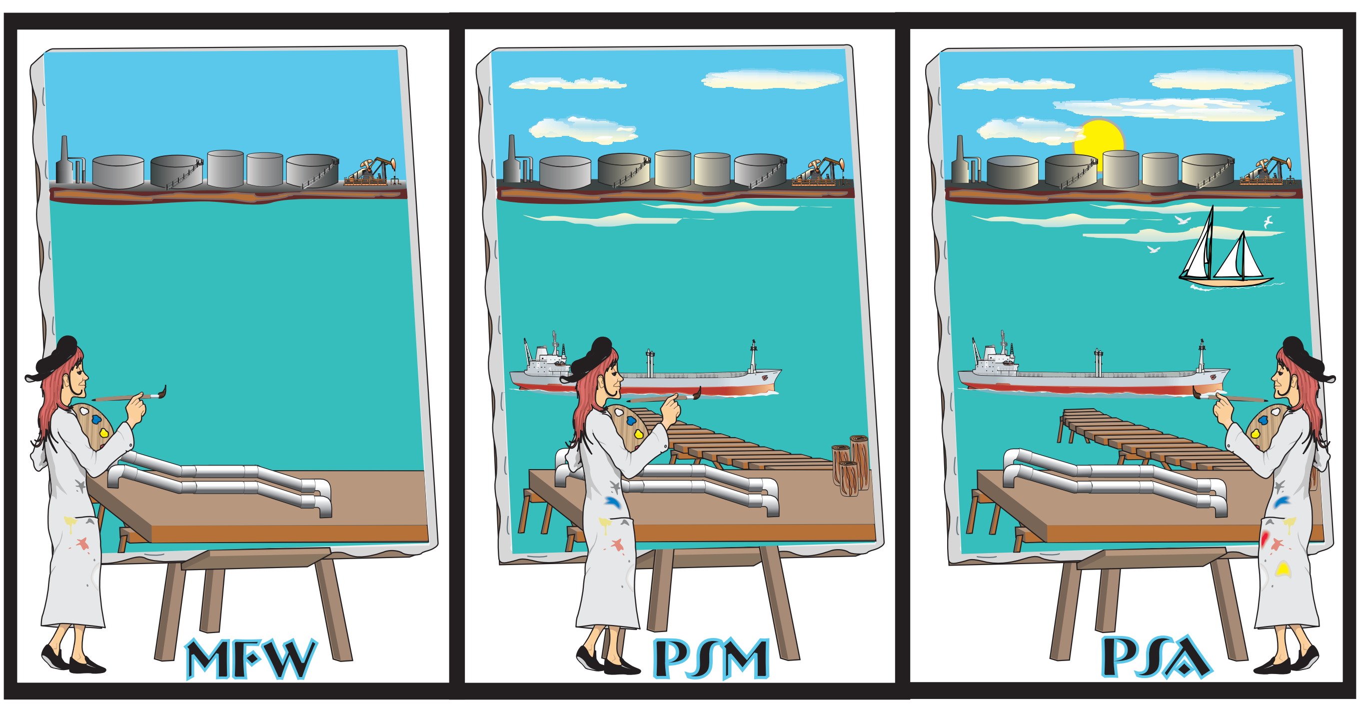
Petroleum supply data collected by the Petroleum Division (PD) in the Office of Oil and Gas (OOG) of the Energy Information Administration (EIA) showed an improvement in the accuracy of the 2004 data from initial estimates, to interim values, to final values. These data were presented in a series of PD products: the *Weekly Petroleum Status Report* (WPSR), *This Week in Petroleum* (TWIP), the *Petroleum Supply Monthly* (PSM), and the *Petroleum Supply Annual* (PSA). Weekly estimates in the *WPSR* and *TWIP* were the first values available.

Figure FE1 illustrates that as reporting and review time passes from the weekly estimates to the interim monthly values to the final petroleum supply values, the EIA is able to complete a more accurate supply picture. For the monthly-from-weekly (MFW) data, respondents have the shortest reporting time, analysts have the shortest review time, and the data are least accurate. For the *PSM* data, respondents have a longer reporting time than the weekly, analysts have a longer review time, and the data are more accurate. For the *PSA* data, respondents have the longest reporting time, analysts have the longest review time, and the data are the most accurate, completing the supply picture.

For 2004, 63 petroleum supply data series were analyzed to determine how close the *PSM* values were to the final *PSA* values. For these series, 34 out of the 63 were within 1 percent of the *PSA* values in terms of mean absolute percent error as compared to 46 out of 66 in 2003. Fifty-six petroleum supply data series were analyzed to see how close the MFW estimates were to the final *PSA* values. For these 56 series, 24 were within 2 percent of the *PSA* values in terms of mean absolute percent error and, of those, 13 were within 1 percent, compared to 27 and 11, respectively, out of 6l series for 2003.

Two major factors that contribute to the *PSM* values being more accurate than the MFW estimates are: (1) the greater length of time between the close of the reference period and the publication date of the *PSM*; and, (2) most MFW values (weekly data converted to a monthly value) are based on company's operational records whereas *PSM* values are generally extracted from company's accounting systems, the latter being more accurate. The greater length of time allows more in-depth review of the data by the respondents and EIA. Within 2 months of the close of a reference month, interim values are published in the *PSM*. The weekly data are more quickly available. The *WPSR* and *TWIP* are available electronically 5 days after the close of the reference week (excluding holiday weeks). About 6 months after the end of the reference year, final monthly values, reflecting resubmissions, are published in the *PSA*.

Figure FE1. With Time, the Supply Picture is Completed in the PSA



Historically, the weekly publication (*WPSR*) and the monthly publication (*PSM)* provided volumes of crude oil and petroleum products data at relatively increasing levels of accuracy. This article provides petroleum analysts with a measure of the degree to which, on average, estimates and interim values vary from their final values.

The Petroleum Supply Reporting System

The 16 surveys in the Petroleum Supply Reporting System (PSRS) track the supply and disposition of crude oil, petroleum products, and natural gas liquids in the United States. To maintain a database with historically accurate observations and current estimates from the petroleum industry, EIA administers three survey series: weekly, monthly, and annual.

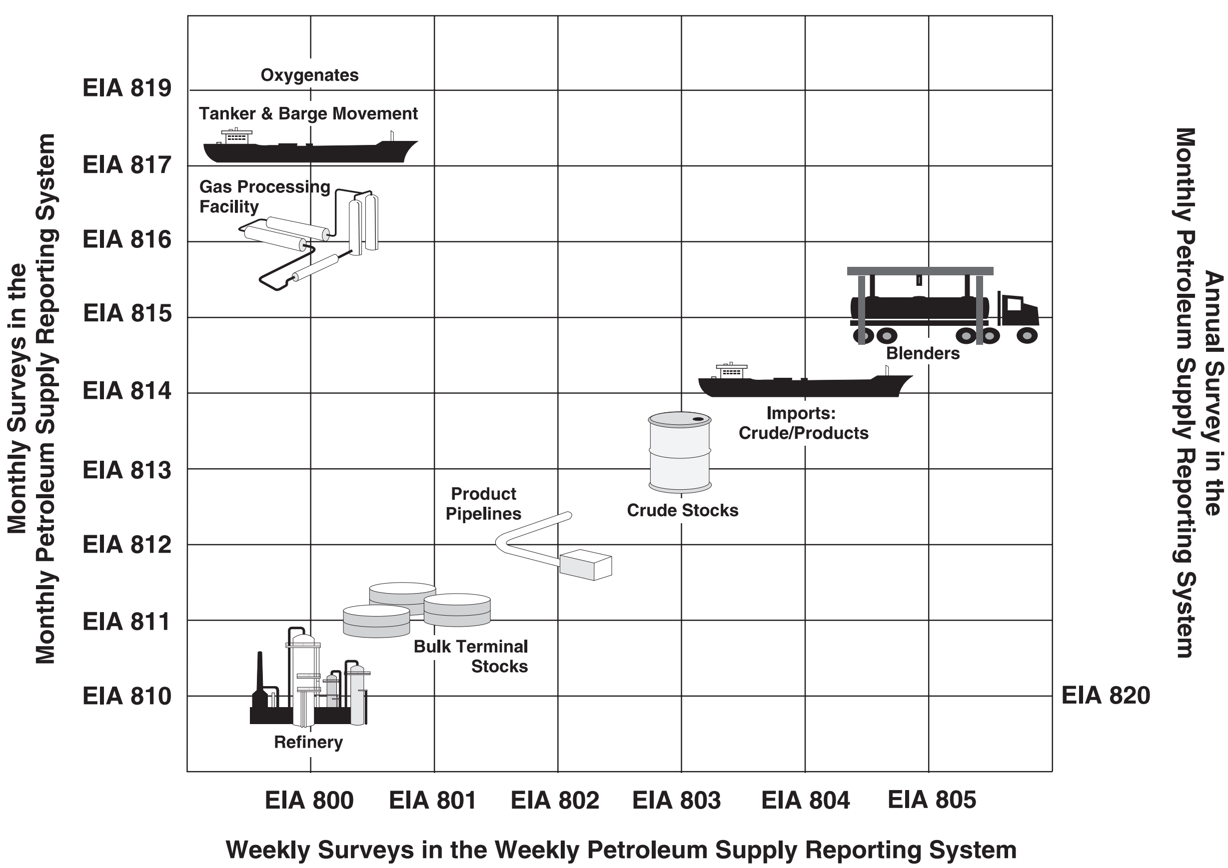
The PSRS is organized into two data collection subsystems, the Weekly Petroleum Supply Reporting System (WPSRS) and the Monthly Petroleum Supply Reporting System (MPSRS). The WPSRS processes data from the six weekly surveys. The MPSRS includes nine monthly surveys and one annual survey. Figure FE2 displays the petroleum supply and distribution system and indicates the points at which petroleum supply data are collected. Both weekly and monthly surveys are administered at six key points along the petroleum production and supply path: (1) refineries, (2) bulk terminals, (3) product pipelines, (4) crude oil stock holders, (5) importers, and (6) blenders.

Annual U.S. refinery capacity data are collected on the Form EIA-820, "Annual Refinery Report." These data were collected and published in Volumes 1 and 2 of the *PSA* for 2004, available only electronically.

The Weekly Petroleum Supply Reporting System

The WPSRS contains the data collected from the six weekly surveys. Each weekly survey is distributed to a sample of the corresponding monthly survey's universe. In Figure FE2, the icons represent the target population of the monthly and weekly surveys of the PSRS. For example, the target population for the survey Forms EIA-801 and EIA-811 is bulk terminals. Thus, the respondents to the Form EIA-801 are a sample of the respondents who report on Form EIA-811. For the weekly surveys, EIA aims for a minimum 90-percent multi-attribute-cutoff sample from the respondents to the corresponding monthly survey. In choosing the sample for each product, companies are ranked in descending order by

Figure FE2. Petroleum Supply Reporting System: Surveys and Subsystems



Source: Energy Information Administration, Petroleum Supply Reporting System.

volume. Respondents are chosen in order, down the list until the sample includes those companies contributing at least 90 percent of a variable's total volume. For example, for distillate fuel oil stocks, the weekly sample includes those respondents whose combined volumes of stocks for distillate fuel oil from refineries, bulk terminals, and pipelines constitute at least 90 percent of the total volume of distillate fuel oil stocks as reported in the corresponding monthly surveys.

These surveys enable EIA to provide timely, relatively accurate snapshots of the U.S. petroleum industry every week. The weekly surveys collect information on the supply and disposition of selected petroleum products and crude oil. The reference period for each weekly survey begins at 7:01 a.m. each Friday and ends at 7:00 a.m. the following Friday. Respondents report their data via telephone, facsimile, electronic spreadsheets, or EIA's electronic data collection software package, the Personal Computer Electronic Data Reporting Option (PEDRO). All respondents must submit their data by 5:00 p.m. on the Monday following the end of the reference period. During 2 working days, quality control procedures are executed. Cell values determined to be unusual or inconsistent with other cell values are flagged. The validity of the value of each flagged cell is investigated. Some flagged values are verified by the respondent to be correct; other flagged cells are corrected; and the remaining flagged values are referred to as unresolved. Nonrespondent and unresolved flagged data are imputed using an exponentially smoothed mean of the respondents' historical data.

As a new weekly web product in 2002, *This Week in Petroleum* (TWIP) provides analysis, data, and charts of the latest weekly petroleum supply and price data. Prior to October 11, 2002, weekly propane data were collected only during the heating season on Form EIA-807, "Propane Telephone Survey." Collection of weekly propylene (nonfuel use) inventory data began on January 10, 2003. In January 2004, the *WPSR* collection and processing system were rewritten using more advanced technology. Beginning with data for April 9, 2004, the weekly survey forms were modified to collect more detailed data on some products and incorporate propane data, previously collected on Form EIA-807, into the *WPSR*.

Within 5 days of the close of the reference week, weekly data are made available to the public on the EIA's internet web site (http://www.eia.doe.gov) through the *WPSR* and *TWIP*. Except when holidays delay data processing schedules, weekly data are available via the internet at 10:30 a.m. Eastern Time on the Wednesday following the close of the reference week. *TWIP* is generally available at 1:00 p.m. on Wednesdays at http://tonto.eia.doe.gov/oog/info/twip/twip.asp.

The Monthly Petroleum Supply Reporting System

The reference period for the monthly surveys starts on the first day of the month at 12:01 a.m. and ends on the last day of the month at midnight. The deadline for filing monthly surveys is the 20th calendar day following the end of the report month. Data are reported via mail, telephone, facsimile, electronic spreadsheets, or PEDRO. Beginning with the January 2004 EIA-819 data, the early collection and publication dates were changed to coincide with the other monthly surveys.

During the period of data editing, either the respondent or EIA staff may identify an error. If the respondent discovers an error, the EIA representative for a particular survey is notified and the value is corrected. If EIA's edits diagnose an unusual value, an EIA representative will determine if the value is correct or incorrect by calling the company and/or reviewing historical data.

Within 60 days of the close of the reference month, all of the interim monthly data are published in the *PSM* on the internet. Throughout the year, EIA accepts data revisions of monthly data. If a revision is made after the *PSM* has been published, it is referred to as a resubmission. Additionally, preliminary company-level imports data are released electronically between the 7th and 10th of each month.

Beginning with the March 2005 *PSM*, the formats of the tables were modified to accommodate the new product breakouts and the section on summary statistics was eliminated. In addition, Table H1, "Petroleum Supply Summary" which showed early estimates of monthly data based on weekly submissions (monthly-from-weekly) has been eliminated. These monthly-from-weekly estimates are published in the *WPSR* via the internet on the Wednesday following the first Friday of each month.

Within 6 months of the end of the calendar year, the final monthly values for the previous year are published in the *PSA*. These values reflect all *PSM* resubmissions and other data corrections. The values contained in the *PSA* are EIA's most accurate measures of petroleum supply activity.

Factors Affecting Data Accuracy

Maintaining an accurate database is a major goal of EIA. The quality of the data drives the quality of all qualitative and quantitative analyses conducted using these data. Accuracy and timeliness are primary attributes of high quality data. Accuracy of survey data is measured as the closeness of the published values to the true values (i.e., those values that would be obtained if the entire target population had been surveyed and all the data had been precisely recorded).

Table FE1. Average Coverage for Weekly Surveys, 2004 and 2003 (Percent of Final Monthly Volumes Included in Monthly-from-Weekly Sample)

Product	Stocks						Production		Imports	
	Refinery		Bulk Terminal		Pipeline					
	2004	2003	2004	2003	2004	2003	2004	2003	2004	2003
Total Motor Gasoline	98	98	94	93	96	97	98	98	95	97
Jet Fuel	98	98	94	92	98	99	98	99	93	91
Distillate Fuel Oil	97	97	93	87	97	98	97	97	95	95
Residual Fuel Oil	95	96	92	92	—	—	94	95	82	80
Crude Oil	97	97	—	—	—	—	—	—	97	97

— = Not Applicable.
Source: Energy Information Administration, Petroleum Supply Reporting System.

Respondents to the monthly surveys have more time to file than the weekly respondents, enabling them to collect, review, and revise their data more carefully than the weekly respondents. Additionally, EIA has more time to edit the monthly data. Also, some weekly respondents report estimates while many monthly respondents extract actual data from accounting systems. Thus, the monthly data are typically more accurate.

Some sources of error, such as nonresponse, are not totally preventable. Other errors, such as sampling errors, are unique to a particular type of survey. One situation where sampling error occurs is if the group of sampled respondents is dissimilar to the full population. Within the PSRS, only weekly surveys are at risk of having sampling errors. However, all surveys in the PSRS are at risk for nonsampling errors, such as: (1) insufficient coverage of respondents (the survey frame does not include all members of the target population); (2) nonresponse; (3) response error; and (4) errors due to lack of survey clarity. A detailed discussion of factors influencing data accuracy and how they are minimized in the PSRS follows.

Samples and Sampling Error

A sample is a subsection of a universe identifying members of a target population. The weekly surveys are administered to samples of the monthly populations to reduce respondent burden and to expedite the turnaround of data from survey respondents to the public. As with any sample, the values obtained are different from those obtained if the full universe had been surveyed. Sampling error is the difference between a sample estimate and a population value.

There are six samples, one for each weekly petroleum supply survey, in the WPSRS. For these surveys, the sampling error is minimized by using a minimum 90-percent multi-attribute-cutoff sample from the corresponding monthly survey's frame. At the end of each month, updates are made to the samples and survey frames if a 90-percent coverage was not obtained.

For the weekly surveys, better coverage will most likely reduce sampling error. As shown in Table FE1, 2004 coverage was comparable to 2003. Of the 21 product and supply type combinations, 20 had coverage of 90 percent or above in 2004. For 12 of the 21 combinations, 2004 coverage decreased from 2003. Bulk terminal distillate fuel oil stocks had the largest percentage increase from 2003 to 2004, increasing by 6.1 percent. The largest percentage decrease from 2003 to 2004 was for pipeline motor gasoline stocks, decreasing by 1.5 percent. Tabulations were done before rounding of the coverage values.

Nonsampling Error

Unlike sampling errors, all survey data, even those from a census survey, are at risk of incurring nonsampling errors. There are two categories of nonsampling errors, random and systematic. With random error, on average, and over time, values will be overestimated by the same amount they are underestimated. Therefore, over time, random errors do not bias the data, but they will give an inaccurate portrayal at any point in time. On the other hand, systematic error is a source of bias in the data, since these patterns of errors are made repeatedly. The following is a discussion of how the four most frequently occurring types of nonsampling error are minimized within the PSRS.

Frame Updates

The list of all companies identified as members of the target population is called a frame. If members of the target population are not included in the frame, there is an undercount of the aggregate data. To diminish the chance of undercounting, the PSRS frames are continually updated. New companies are identified through continual review of petroleum industry periodicals, newspaper articles, and correspondence from respondents.

Maintaining a Low Nonresponse

Survey respondents are required by law to report to EIA (see Explanatory Note 6 of the *PSM* for a description of action for chronic nonresponse). The 2004 response rates for the weekly

Table FE2. Average Response Rates for Monthly and Weekly Surveys, 2004

Survey Site	Respondents to Monthly Surveys			Respondents to Weekly Surveys		
	Average Universe Size	Average Number of Respondents	Percent [1]	Average Weekly Sample Size	Average Number of Respondents	Percent [2]
Refinery	157	156	99.4	131	126	96.3
Bulk Terminal	242	240	99.0	87	81	93.8
Pipeline	78	78	100.0	44	43	98.3
Crude Oil Stocks	147	141	95.6	58	57	98.0
Blenders	248	247	99.3	145	144	99.4

[1] The average response rates for monthly surveys are calculated by summing the individual monthly response rates and dividing by 12.
[2] The average response rates for weekly surveys are calculated by summing the individual weekly response rates and dividing by 52 (39 for refinery and blenders).
Note: Percents are calculated before rounding.
Source: Energy Information Administration, Petroleum Supply Reporting System.

surveys and their corresponding monthly surveys are enumerated in Table FE2. Compared to the 2003 response rates, one half of the 2004 response rates increased or stayed the same percent. The largest difference in response rate was for the monthly bulk terminal survey, increasing from 94.0 percent in 2003 to 99.0 percent in 2004. Beginning in 2004, response rates were collected for blenders.

To mitigate the effect of nonresponse, imputed values are calculated for all nonreported values except monthly imports. Weekly imputed values are the exponentially smoothed mean of that respondent's historical values for that variable. Monthly imputed values are the previous month's value for the particular respondent and variable. For imports, however, there is a great deal of fluctuation from one reference period to another, with respondents frequently having no imports of a particular product. As a result, the data for nonreported cells on the monthly imports survey are not imputed. In addition, the monthly imports are collected and published at a much greater level of detail than the weekly imports, which makes imputation impractical.

Reducing Response Error

Improvements to the PSRS system are continuously being made to reduce response error. To satisfy customer needs and meet the particular requirements of some respondents, computerized spreadsheets that resemble the actual survey forms have been developed, and are available for respondent reporting. Another improvement has been the increased participation in the PEDRO system, which permits all weekly and monthly survey data, except the Form EIA-819 to be submitted to EIA electronically. A respondent entering values via PEDRO may execute edit routines prior to transmission of the survey responses. These routines include consistency and outlier (extreme value) checks of the data. Unusual or nonreported cells are flagged and, prior to transmission of the data, a representative of the company is able to review and verify or correct data in the flagged cells.

Even with sophisticated edit checks, response error (the difference between the reported value and the actual value) remains the most likely cause of data inaccuracy. The weekly surveys are more susceptible to response error since some of their values are estimates or based on operational records. Many monthly respondents abstract their monthly data from accounting systems and thus are generally more accurate.

Maintaining accurate accounting records, however, does not ensure against response error. For example, numbers can be transposed within the correct cell; an otherwise correct value may be entered in the wrong cell; a respondent may misinterpret the intent of a question; or the wrong units may be used.

Survey Clarity

The terms, layout, and definitions on all survey forms are periodically reviewed for completeness, clarity, and consistency across surveys. At regular intervals, survey intent, as well as what data are collected, are subject to industry and government review. To the extent possible, industry changes in terminology and practice are incorporated into the PSRS on an ongoing basis.

Data Assessment

Each of the variables included in these analyses is of current and historical interest. Of the 63 variables for which both *PSM* and *PSA* values were published, only 56 of them were published weekly throughout 2004. For each variable, six measures of accuracy were calculated to compare the differences between the MFW and *PSM* values relative to the *PSA* values.

- **Error** is the difference between the estimate (MFW) or interim (*PSM*) value and the final (*PSA*) value for a given month. For inputs, production, stock change, imports, exports, and product supplied, values are expressed in units of thousands of barrels per day. For stocks, values are expressed in units of thousands of barrels.

MFW Error = MFW Volume - *PSA* Volume

PSM Error = *PSM* Volume - *PSA* Volume

Percent Error is the error for a given month divided by the final value for a given month, and multiplied by 100.

$$\text{MFW Percent Error} = \frac{\text{MFW Error}}{\textit{PSA}\text{ Volume}} \times 100$$

$$\textit{PSM}\text{ Percent Error} = \frac{\textit{PSM}\text{ Error}}{\textit{PSA}\text{ Volume}} \times 100$$

Mean absolute error is the weighted average over the 12 months of the year of the absolute values of the errors for each month. The mean absolute error measures the average magnitude of the revisions that took place over a year. Outliers increase the mean absolute error. The number of days in the month is used for weighting all product categories except stocks. Stocks are weighted equally for each of the 12 months.

Mean absolute percent error is the weighted average over the 12 months of the year of the absolute values of the percent errors. It provides a measure of the average magnitude of the revisions relative to final values. The mean absolute percent error has an inverse relationship with data accuracy; i.e., the smaller the mean absolute error, the closer the interim data are to the final data; conversely, the larger the mean absolute percent error, the greater the difference in the interim value and the final value. Outliers inflate the mean absolute percent error.

Range is the difference between the smallest and largest percent errors. The range shows the dispersion of the percent differences between interim and final values.

Median of the percent errors is the point at which half the values are higher and half are lower. Unlike the mean, the median is not affected by an outlier. In these analyses, each distribution has 12 observations. The median is the average of the sixth and seventh ordered observation.

The average final absolute volumes and the mean absolute percent error for MFW estimates and *PSM* interim values for 2004 and 2003 are presented in Table FE3. The average final absolute volumes are presented to give the reader an idea of the magnitude of these volumes. Variables with very small volumes are prone to larger percent changes because a modest volume change is being compared to a small final volume. The mean absolute error and the size of the volumes involved must both be included in the interpretation of data accuracy.

The 2004 MFW mean absolute percent errors which were within 2 percent of their respective *PSA* values (24 of the 56 MFW series), and the 2004 *PSM* mean absolute percent errors which were within 1 percent of their *PSA* values (34 of the 63 *PSM* series), are distinguished by a single asterisk. Mean absolute percent errors that were greater than 10 percent are marked by a double asterisk. There were 14 such MFW series and 4 *PSM* series, compared to 12 and 3, respectively, for 2003.

For 2004, 8 of the 10 weekly production series increased in mean absolute percent error from 2003. Nine of the 13 production series have a single asterisk in the *PSM* column, indicating a mean absolute percent error of less than 1 percent from the *PSA*. Additionally, 11 of the 13 *PSM* production series in 2004 increased in mean absolute percent error from 2003. Weekly propane production data was not available for all of 2004.

The single asterisks in Table FE3 by the stock series show that, as in prior years, the stock values for both MFW estimates and *PSM* interim values are very close to the final *PSA* values. Fuel ethanol and methyl tertiary butyl ether stocks are not collected weekly, but are collected on the Form EIA-819, "Monthly Oxygenate Telephone Report." Prior to 2004, these data were collected earlier than other monthly data and served as a preliminary estimate for stocks. These data now follow the same collection cycle as other monthly data. Oxygenate data are displayed in Appendix D of the *PSM*. Ten of the 16 weekly stock series and 14 of the 18 monthly stock series for 2004 increased in mean absolute percent error from 2003.

Stock change is the difference between stocks at the beginning of the month and stocks at the end of the month. Since the monthly change in stock levels is small compared to the stock levels themselves, a large percent error in stock change can occur even when the percent errors in stock levels are small.

Crude oil stock change is one of the components in the calculation of unaccounted for crude oil (calculated disposition minus calculated supply of crude oil). For both the MFW and the *PSM* numbers, the volume of the unaccounted for crude oil may be increased by a combination of factors including an understatement of imports, an overstatement of exports, an understatement of crude oil production, an understatement of stock withdrawals, and an overstatement of crude oil inputs. The overstatement of crude oil inputs can be caused by injections along crude oil pipelines of natural gas liquids. When refiners receive this mixture, they process it as crude oil. As seen in Table FE3, the production, imports, and refinery inputs of crude oil have a small mean absolute percent error relative to crude oil stock change.

For petroleum products, stock change is a component in the calculation of product supplied (representing the consumption of petroleum products). Unlike the other variables, stock change values can be negative. Stock change thus has an added dimension by which to evaluate accuracy; this is the

Table FE3. Summary Statistics for Differences Between Interim and Final Data, 2004 and 2003

Variable	PSA Average Absolute Volumes		Monthly-from-Weekly Mean Absolute Percent Error		PSM Mean Absolute Percent Error	
	2004	2003	2004	2003	2004	2003
Crude Oil Production (thousand barrels/day)	5,419	5,681	* 0.90	1.64	* 0.68	1.07
Refinery Operations						
Refinery Crude Oil Inputs (thousand barrels/day)	15,475	15,304	* 0.77	0.48	* 0.06	0.02
Operating Utilization Rate (percent)	93	93	* 0.72	0.65	* 0.29	0.07
Production (thousand barrels/day)						
Total Production	20,044	19,630	—	—	* 0.20	0.09
Refinery Production	17,814	17,487	* 1.69	1.20	* 0.23	0.11
Finished Motor Gasoline	8,723	8,501	* 0.83	1.02	* 0.72	0.29
Reformulated Motor Gasoline	2,844	2,715	2.72	1.95	* 0.81	0.44
Conventional Motor Gasoline	5,593	4,787	7.85	1.54	1.41	0.31
Jet Fuel	1,547	1,488	* 1.17	0.67	* 0.06	0.00
Distillate Fuel Oil	3,814	3,707	* 0.93	0.77	* 0.12	0.18
Low Sulfur Distillate Fuel Oil	2,847	2,719	2.17	1.22	1.08	0.08
High Sulfur Distillate Fuel Oil	967	988	8.02	2.10	2.49	0.63
Residual Fuel Oil	656	660	2.94	3.40	* 0.86	0.43
Other Products	5,304	5,273	—	—	1.10	0.50
Propane	1,110	1,075	—	—	* 0.19	0.30
Other Products Refinery Production	3,531	3,438	** 17.59	8.97	* 0.41	0.30
Stocks (thousand barrels)						
Total Stocks	1,615,468	1,544,719	* 0.33	0.91	* 0.25	0.13
Total Stocks, excl. SPR	953,326	930,810	* 0.54	1.49	* 0.42	0.21
Total Crude Stocks	951,525	895,912	* 0.39	0.26	* 0.25	0.15
Crude Oil Stocks, excl. SPR	289,382	282,002	* 1.33	0.74	* 0.83	0.47
SPR Stocks	662,143	613,909	* 0.07	0.05	* 0.00	0.00
Refined Products Stocks	663,943	648,808	* 0.95	2.09	* 0.31	0.18
Total Motor Gasoline Stocks	207,330	202,766	* 1.35	0.61	* 0.65	0.28
Reformulated Motor Gasoline Stocks	23,261	32,832	9.71	2.29	3.86	1.24
Conventional Motor Gasoline Stocks	114,410	116,130	2.29	1.12	* 0.37	0.23
Jet Fuel Stocks	39,079	38,723	* 1.75	1.32	* 0.18	0.46
Distillate Fuel Oil Stocks	117,181	117,130	* 1.88	1.14	* 0.32	0.34
Low Sulfur Distillate Fuel Oil Stocks	71,283	72,088	2.26	1.99	* 0.64	0.16
High Sulfur Distillate Fuel Oil Stocks	45,898	45,041	3.05	2.48	1.51	0.73
Residual Fuel Oil Stocks	37,758	33,077	2.45	2.46	* 0.24	0.81
Other Products Stocks	262,596	257,111	* 0.67	5.15	* 0.65	0.26
Propane Stocks	47,405	44,768	2.20	3.48	2.07	0.65
Fuel Ethanol Stocks	5,959	6,653	—	5.47	6.31	2.03
Methyl Tertiary Butyl Ether Stocks	4,311	6,079	—	15.36	2.78	1.44
Stock Change (thousand barrels/day)						
Total Stock Change	429	724	** 72.77	83.12	** 78.61	29.17
Crude Stock Change	280	231	** 72.43	109.61	** 98.44	11.57
Refined Products Stock Change	499	603	** 60.52	95.34	** 17.87	17.29
Imports (thousand barrels/day)						
Total Imports	13,145	12,264	2.31	2.12	1.88	0.97
Total Crude Imports	10,088	9,665	* 1.29	1.81	* 0.49	1.06
Crude Oil Imports, excl. SPR	10,088	9,665	* 1.29	1.81	* 0.49	1.06
SPR Imports	0	0	* 0.00	0.00	* 0.00	0.00
Refined Products Imports	3,058	2,599	7.91	3.58	6.53	0.94
Finished Motor Gasoline Imports	496	518	6.82	4.51	3.66	1.58
Reformulated Motor Gasoline Imports	212	249	** 12.20	7.39	1.76	0.44
Conventional Motor Gasoline Imports	284	269	** 10.39	11.66	4.96	2.68
Jet Fuel Imports	127	109	** 14.34	21.07	9.22	3.50

See footnotes at end of table.

Table FE3. Summary Statistics for Differences Between Interim and Final Data, 2004 and 2003 (Continued)

Variable	PSA Average Absolute Volumes		Monthly-from-Weekly Mean Absolute Percent Error		PSM Mean Absolute Percent Error	
	2004	2003	2004	2003	2004	2003
Distillate Fuel Oil Imports	326	333	** 10.83	7.23	2.56	0.98
Low Sulfur Distillate Fuel Oil Imports	148	135	** 14.85	12.85	1.84	0.91
High Sulfur Distillate Fuel Oil Imports	177	198	** 17.68	11.53	3.18	1.32
Residual Fuel Oil Imports	426	327	** 24.14	13.80	** 13.18	3.02
Other Products Imports	1,682	1,312	8.59	5.85	6.47	1.82
Propane Imports	209	168	—	—	1.19	4.05
Exports (thousand barrels/day)						
Total Exports	1,048	1,027	** 11.29	6.67	* 0.00	1.00
Crude Oil Exports	27	12	** 53.44	87.27	* 0.00	0.00
Refined Products Exports	1,022	1,014	** 10.58	6.35	* 0.00	1.00
Total Net Imports (thousand barrels/day)	12,097	11,238	* 1.92	1.73	2.04	1.14
Products Supplied (thousand barrels/day)						
Total Products Supplied	20,731	20,034	* 1.88	1.81	1.03	0.31
Finished Motor Gasoline Supplied	9,105	8,935	* 0.84	1.02	* 0.56	0.32
Jet Fuel Supplied	1,630	1,578	* 1.82	2.19	* 0.85	0.63
Distillate Fuel Oil Supplied	4,058	3,927	2.78	2.69	* 0.41	0.48
Residual Fuel Oil Supplied	865	772	9.58	6.28	7.01	1.93
Other Products Supplied	5,073	4,822	5.31	6.51	1.96	1.38
Propane Supplied	1,276	1,215	—	—	2.25	0.99

— = Not Applicable.
* = For MFW values, mean absolute percent error less than or equal to 2; for PSM values, mean absolute percent error less than or equal to 1.
** = Mean absolute percent error greater than or equal to 10.
SPR = Strategic Petroleum Reserve
Notes: Error is the difference between Monthly-from-Weekly estimates or interim monthly data published in the Petroleum Supply Monthly and the final value as published in the Petroleum Supply Annual. Percent error is the error multiplied by 100 and divided by the final published value. Mean absolute error is the weighted average of the absolute errors. Mean absolute percent error is the weighted average of the absolute percent errors. The number of days in the month is used for weighting all product categories except stocks. Stocks are weighted equally for each of the 12 months. •Totals may not equal sum of components due to independent rounding. •For 2003 production and stocks, oxgenated and other motor gasoline were combined into conventional.
Source: Energy Information Administration, Petroleum Supply Reporting System.

correctness of the direction of the change. Table FE4 provides a measure of accuracy of the direction of MFW and *PSM* stock change values for 2004 and 2003. Three of the six stock change values for 2004 increased the number of months that differed from the direction of the *PSA* values compared to 2003.

For imports, one reason for the large mean absolute percent errors in the MFW values is that shipments do not always arrive during the week in which they were expected. This has a greater impact when the end of the month occurs in the middle of the week. Nine of the 14 MFW import series in Table FE3 showed an increase in mean absolute percent error from 2003 to 2004. For the *PSM*, 11 of the 15 import series increased in mean absolute percent error from 2003 to 2004.

With the exception of refinery receipts in the U.S. Territories, EIA does not collect export data. They are gathered by the U.S. Bureau of the Census on a monthly basis. They are received by EIA on a monthly basis approximately 7 weeks after the close of the

Table FE4. Number of Months In Which the Direction of NonFinal Stock Change Values Differed From PSA

	Number of Months	
	2004	2003
Total Stock Change		
MFW and PSA Values	2	1
PSM and PSA Values	1	0
Crude Stock Change		
MFW and PSA Values	1	2
PSM and PSA Values	1	0
Refined Products Stock Change		
MFW and PSA Values	1	2
PSM and PSA Values	0	0

Source: Energy Information Administration, Petroleum Supply Reporting System.

reporting month. The weekly estimates for exports are projections based on past monthly data. Because the export data are highly variable, it is difficult to obtain estimates of comparable quality to domestic estimates.

Products supplied is the calculation of field production, plus refinery production, plus imports, plus unaccounted for crude oil, minus stock change, minus crude oil losses, minus refinery inputs, minus exports. Therefore, the accuracy of products supplied is affected by the individual components.

Box and Whisker Plots

Example 1 in the shaded box titled "Structure of Box and Whisker Plots," is a simplified illustration of the box and whisker plots that follow. The box and whisker plots map the 5-year trends in historical accuracy of weekly estimates and monthly interim values. The details provided by the box and whisker plots include: historical trends, the range of monthly percent errors, direction of the error (i.e., overestimation or underestimation), and the identification of unusual values.

Each box and whisker plot is placed on a graph, where the horizontal axis represents the year and the vertical axis represents the percent error. The center horizontal axis for all the box and whisker plots is zero percent error. For each variable studied, a pair of charts, each containing five box and whisker plots (one for each year, from 2000 through 2004), are presented side-by-side; the chart on the left contains the percent errors for the MFW estimates, and the chart on the right contains the percent errors for the *PSM* values. To facilitate the comparison of MFW percent errors and the *PSM* percent errors, the plots have the same scale.

The position of the box along the y-axis denotes whether the MFW or *PSM* values are predominantly overestimates or underestimates of the *PSA* values. For example, if the majority of the MFW values were overestimates, more than half of the box would be above the zero percent error line.

The outliers, represented by an asterisk, are usually the result of resubmissions sent in throughout the year by respondents due to misreporting or reporting problems. Some of these problems were due to the respondent's unfamiliarity with the new survey forms.

Crude Oil Production and Crude Oil Inputs

Crude oil production data are not collected through any of EIA's surveys. EIA's Dallas Field Office assembles data collected from State agencies responsible for measuring crude oil production. Based on historical trends and data reported on Form EIA-182, "Domestic Crude Oil First Purchase Report," EIA estimates weekly and monthly production. Final estimates based on revised Form EIA-182 data, State government agencies, and the U.S. Department of Interior's Minerals Management Service data are published in the *PSA*. Figure FE3 presents errors of MFW and *PSM* values relative to *PSA* values for crude oil production and crude oil inputs. More than half of the 2004 MFW estimates for crude oil production overestimated the final *PSA* values. All but three of the 2004 *PSM* interim values overestimated the final *PSA* values.

Similar to 2003, most of the 2004 MFW estimates for refinery crude oil inputs underestimated the final *PSA* values. The range (2.77) of the 2004 MFW percent errors was the largest range over the 5-year period but was the smallest range of all other MFW plots analyzed for 2004, ranging from -2.59 to 0.18 percent. The one outlier in April (-2.59) had the largest absolute percent error over the past 60 months studied. As in prior years, the 2004 *PSM* refinery crude oil inputs were extremely close to the final *PSA* values, with percent errors within 0.23 percent. There were two outliers in January (0.23) and May (-0.19).

Product Production

As expected, *PSM* interim values for production of each of the four major petroleum products were superior to their comparable MFW estimates. Figures FE4 and FE5 contain the box and whisker plots for motor gasoline and distillate fuel oil production, and residual fuel oil and jet fuel production, respectively.

The range (4.71) of the 2004 MFW motor gasoline production percent errors, displayed in Figure FE4, was the largest range over the past 5 years, ranging from -3.13 to 1.58 percent. In addition to the outlier in November (1.58), the outlier in June (-3.13) had the largest absolute percent error over the past 60 months. Similarly, the range (3.55) of the 2004 *PSM* percent errors for motor gasoline production was the largest range over the 5-year period, ranging from -1.73 to 1.82 percent. The outlier in January (1.82) had the largest percent error over the 60 months studied.

Unlike 2003, more of the 2004 MFW estimates for distillate fuel oil production overestimated the final *PSA* values, ranging from -1.45 to 2.28 percent. The percent errors were evenly distributed around the median of 0.38 percent. All but one of the 2004 *PSM* interim values overestimated the final *PSA* values. The percent errors for 2004 were tightly distributed around the median of 0.03 percent. There was one outlier in February (0.61).

The box and whisker plots for residual fuel oil production and jet fuel production are shown in Figure FE5. All but three of the 2004 MFW estimates for residual fuel oil production underestimated the final *PSA* values, ranging from -8.56 to 1.97 percent. In contrast to prior years, all but one of the 2004 *PSM* interim values underestimated the final *PSA* values. The 2004 median of -1.01 percent was the smallest median over the 5-year period.

Structure of Box and Whisker Plots

All box and whisker plots discussed in this article are the visual presentation of a variable's distribution of 12 values of percent errors for either MFW or *PSM* values relative to *PSA* values for a given year. In general, box and whisker plots group data, ordered from smallest to largest, into four areas of equal frequency, quartiles, and show the range and dispersion of data within the quartiles. Sometimes the values of quartiles must be interpolated, i.e., if there are two values that meet the criteria of a quartile, then the average of the two must be taken. Presented below is a discussion of components of box and whisker plots and how they apply to the 12-value distribution illustrated in Example 1: -35, -20, -11, -9, 0, 0, 0, 0, 4.5, 5.5, 15, and 20.

- **First Quartile**

Twenty-five percent of the values are equal to or below the first quartile. In Example 1, the first quartile is the average of the third and fourth ordered observations, i.e., (-11+(-9))/2=-10. The first quartile demarcates the lower boundary of the box.

- **Second Quartile**

The second quartile is the median, and it intersects the box. Fifty percent of the observations are equal to or below the median; in our example, the values of these six observations are: 0, 0, -9, -11, -20, and -35. Also, for this example, the median is the average of the sixth and seventh value, 0, i.e., (0+0)/2. The plot provides the value of the median (the second quartile) as well as information on how the median compares in magnitude to the rest of the observations. Outliers distort the magnitude of the mean, whereas a median is not distorted since it is the actual value that falls in the middle of the distribution. Since outliers have occurred in the distributions of values of PSRS variables, a median is preferred to a mean when assessing accuracy.

- **Third Quartile**

Seventy-five percent of the observations (9 in this case) have values equal to or below the third quartile. In Example 1, the third quartile is 5, i.e., (4.5+5.5)/2. The third quartile demarcates the upper boundary of the box.

- **Box**

The box contains half of all the values. In Example 1, as well as in each box found in Figures FE3-FE11, a minimum of six values are contained within the box. The interquartile range is the length of the box, the difference between the first and third quartiles. The interquartile range for Example 1 is 15, i.e., 5-(-10).

- **Whiskers**

Each whisker extends out from the box, one from the first quartile and the other from the third quartile, to the most extreme value that still falls within 1.5 times the interquartile range. In Example 1, a whisker extends from the third quartile, 5, to 20, which is the maximum value and is within 1.5 interquartile ranges of 5 (as it is less than 5+(1.5*15)=27.5). Also in Example 1, the lower whisker extends from the first quartile -10, to -20, which is the lowest value of the distribution within 1.5 interquartile ranges of the first quartile.

- **Fourth Quartile**

The fourth quartile is the maximum value of the distribution. In Example 1, the fourth quartile, 20, also demarcates the upper value of the top whisker as it is within 1.5 interquartile ranges of the third quartile.

- **Outlier**

An outlier, identified as an asterisk, is an observation that is more than 1.5 interquartile ranges greater than the third quartile, or more than 1.5 interquartile ranges less than the first quartile. In Example 1, there is one outlier, -35. It is less than the lower whisker's threshold value, which is -32.5 (-10-(1.5*15)). The importance of the occurrence of an outlier depends on the distribution of the variable. If the interquartile range is very tight and the outlier is in close proximity, then there is little concern about the occurrence of that outlier. (See Figure FE3, MFW vs *PSA* of Crude Oil Production for 2000.)

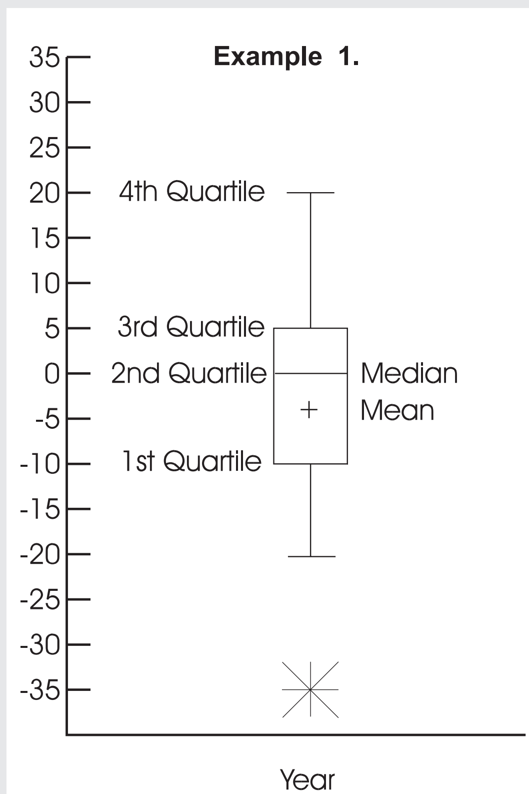
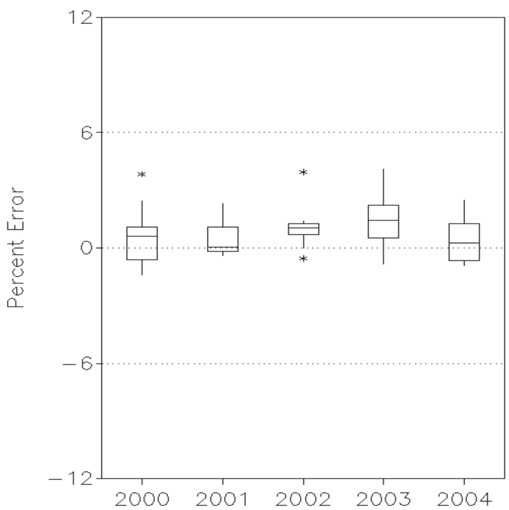


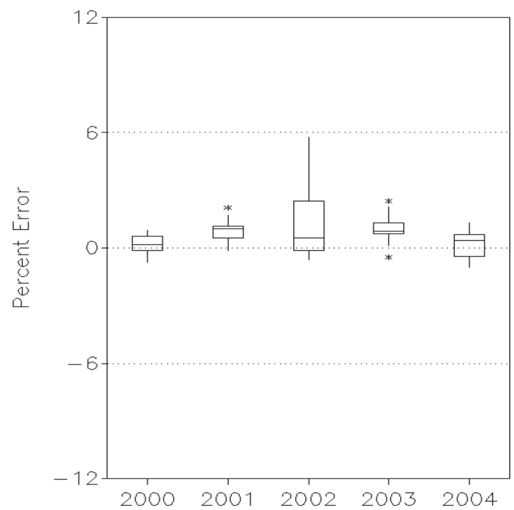
Figure FE3. Range of Percent Errors for MFW and *PSM* Crude Oil Production and Refinery Crude Oil Inputs Data, 2000 - 2004

Crude Oil Production

MFW vs.*PSA*

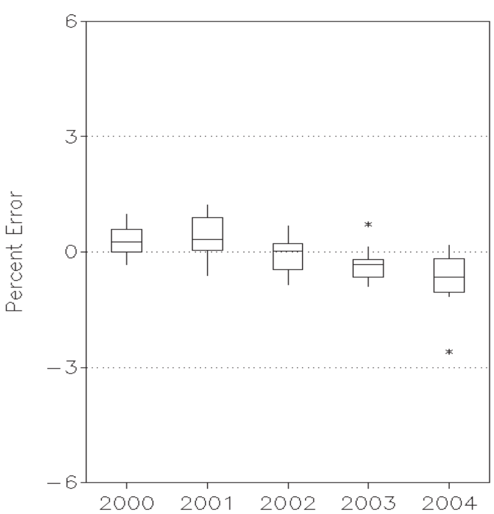


PSM* vs. *PSA

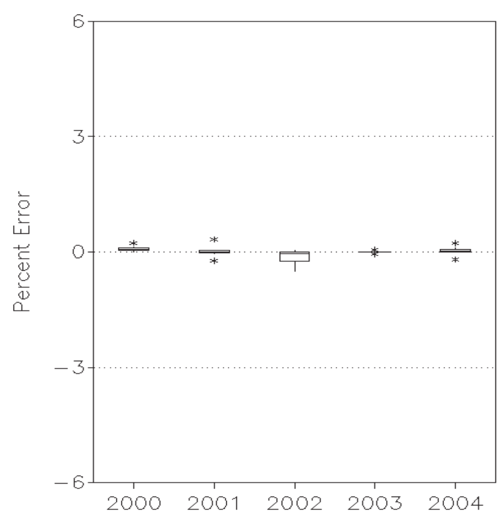


Refinery Crude Oil Inputs

MFW vs.*PSA*



PSM* vs. *PSA

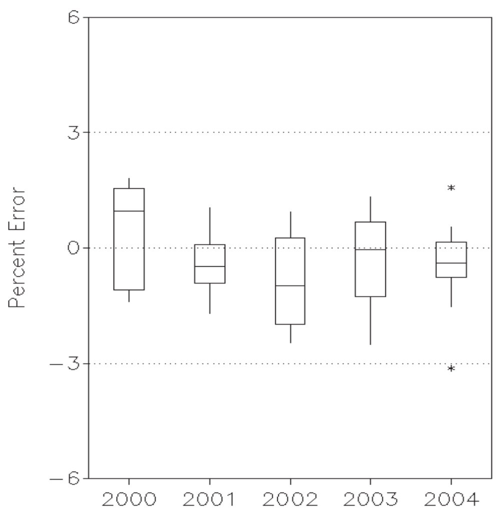


Source: Energy Information Administration, Petroleum Supply Reporting System.

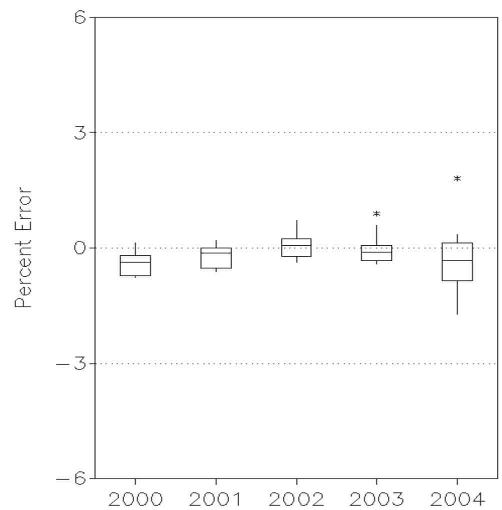
Figure FE4. Range of Percent Errors for MFW and *PSM* Motor Gasoline and Distillate Fuel Oil Production Data, 2000 - 2004

Motor Gasoline Production

MFW vs.*PSA*

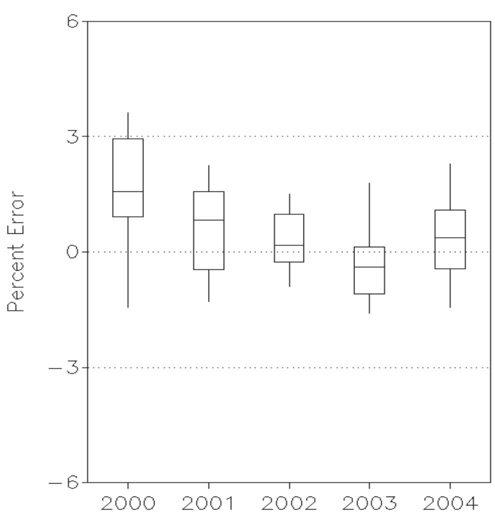


PSM* vs. *PSA

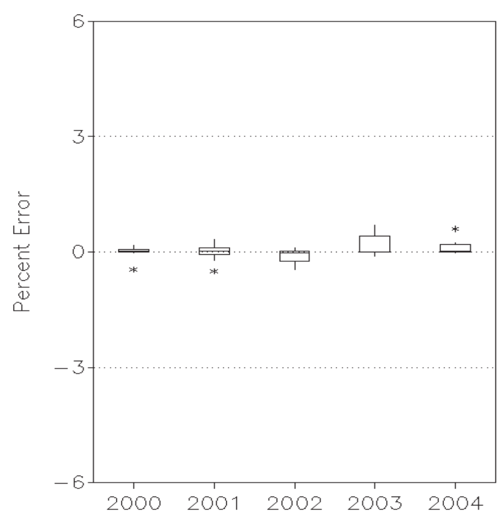


Distillate Fuel Oil Production

MFW vs.*PSA*



PSM* vs. *PSA

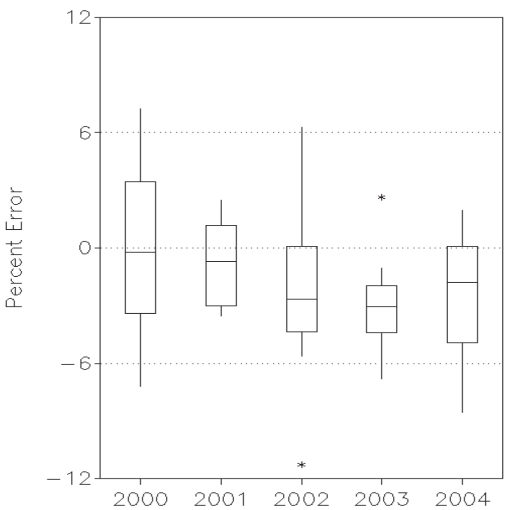


Source: Energy Information Administration, Petroleum Supply Reporting System.

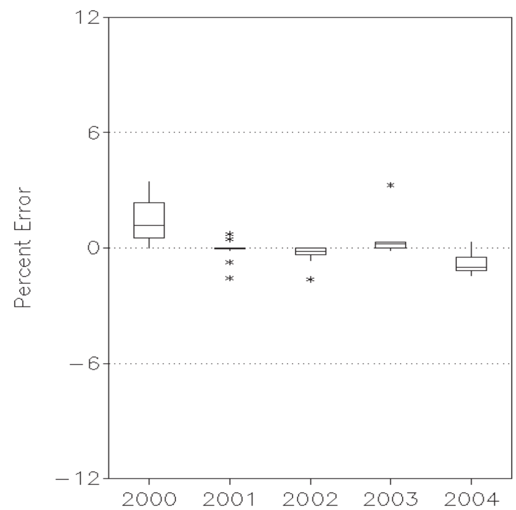
Figure FE5. Range of Percent Errors for MFW and *PSM* Residual Fuel Oil and Jet Fuel Production Data, 2000 - 2004

Residual Fuel Oil Production

MFW vs.*PSA*

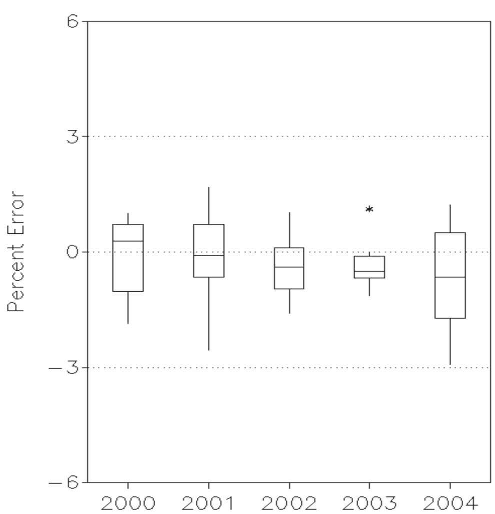


PSM* vs. *PSA

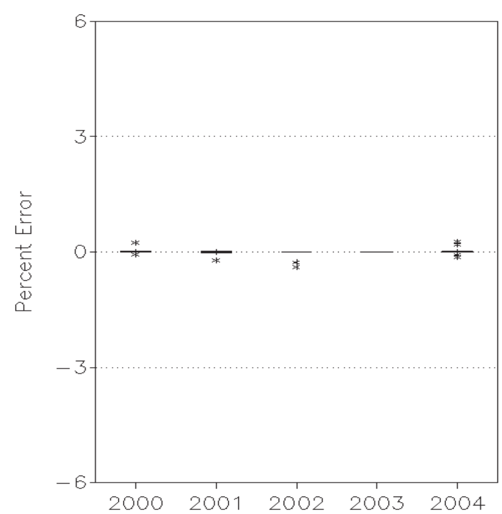


Jet Fuel Production

MFW vs.*PSA*



PSM* vs. *PSA



Source: Energy Information Administration, Petroleum Supply Reporting System.

For jet fuel production, the 2004 median (-0.66) of MFW percent errors, ranging from -2.93 to 1.22 percent, was the smallest median over the 5 years studied. March 2004 (-2.93) had the largest absolute percent error over the past 60 months. The range (0.40) of the 2004 *PSM* percent errors was the largest range over the 5-year period but was the smallest range of all other *PSM* plots analyzed for 2004. The two outliers occurred in March (0.27) and April (-0.13).

Stocks

Figures FE6, FE7, and FE8 show the yearly distribution of percent errors for stocks of crude oil, motor gasoline, distillate fuel oil, residual fuel oil, jet fuel, and propane. Figure FE6 shows the box and whisker plots for crude oil stocks and motor gasoline stocks. The 2004 range (5.11) of MFW percent errors for crude oil stocks was the largest range over the 5-year period, ranging from -2.44 to 2.67 percent. Similarly, the range (3.84) of the 2004 *PSM* percent errors for crude oil stocks was the largest range over the 5 years, ranging from -2.69 to 1.15 percent. February 2004 (-2.69) had the largest absolute percent error over the 60 months studied.

As in prior years, most of the 2004 MFW estimates for motor gasoline stocks underestimated the final *PSA* values. The 2004 median (-1.32) was the smallest median over the 5-year period. Unlike 2003, most of the 2004 *PSM* interim values for motor gasoline stocks were overestimates. The 2004 range (2.39) of *PSM* percent errors was the largest range over the past 5 years. December 2004 (-1.20) had the largest absolute percent error over the 60 months studied.

Figure FE7 shows box and whisker plots for distillate and residual fuel oil stocks. The range (9.55) of the 2004 MFW percent errors for distillate fuel oil stocks was the largest range over the 5-year period, ranging from -3.88 to 5.67 percent. April 2004 (5.67) was the largest percent error over the 60 months studied. Unlike prior years, all of the 2004 *PSM* interim values for distillate fuel oil stocks underestimated the final *PSA* values. The percent errors were tightly distributed around the median of -0.28 percent except for one outlier in February (-0.81). The 2004 range (0.70) of *PSM* percent errors was the smallest range over the past 5 years.

Residual fuel oil stocks typically have larger percent errors than other stock series. Most of the 2004 MFW estimates underestimated the final *PSA* values. The median (-1.80) of the 2004 MFW percent errors was the smallest median for the 5 years analyzed. There was one outlier in August 2004 (-6.82). The 2004 *PSM* percent errors for residual fuel oil stocks were tightly distributed around the median of zero percent. There were three outliers in March (0.50), January (-0.31), and April (-1.81).

The box and whisker plots for jet fuel stocks and propane stocks are shown in Figure FE8. The range (8.46) of the 2004 MFW percent errors for jet fuel stocks was the largest range over the 5-year period, ranging from -5.31 to 3.15 percent. The outlier in August 2004 (-5.31) had the largest absolute percent error over the 60 months studied. The other outlier was in December (3.15). The range (0.67) of the 2004 *PSM* percent errors for jet fuel stocks was the smallest range over the 5-year period. The 2004 *PSM* percent errors were within 0.38 percent.

The range (6.38) of the 2004 MFW percent errors for propane stocks was the smallest range over the 5-year period, ranging from -4.03 to 2.35 percent. Conversely, the 2004 range (10.42) of the *PSM* percent errors was the largest range over the 5-year period. All of the 2004 *PSM* interim values for propane stocks underestimated the final *PSA* values. The outlier in May (-10.49) had the largest absolute percent error over the past 60 months.

Imports

Figures FE9, FE10, and FE11 show the yearly distributions of percent errors for the imports of crude oil and four products: motor gasoline, distillate fuel oil, residual fuel oil, and jet fuel. Because of the irregularity of imports for crude oil and petroleum products, the magnitude and range of percent errors for both the MFW and the *PSM* imports numbers can be expected to be much larger and wider than for production and stocks.

Figure FE9 shows that compared to the prior 4 years, more of the 2004 MFW estimates of crude oil imports overestimated the final *PSA* values. The 2004 median of -0.39 percent was the closest to zero over the 5-year period. In addition, the 2004 range (4.27) of MFW percent errors was the smallest range over the 5-year period, ranging from -2.63 to 1.64 percent. Unlike 2003, all but one of the 2004 *PSM* interim values for crude oil imports underestimated the final *PSA* values. The 2004 median of -0.31 percent was the closest to zero over the 5-year period. The 2004 range (1.31) of *PSM* percent errors was the smallest range over the 5-year period and of all other *PSM* import series analyzed for 2004.

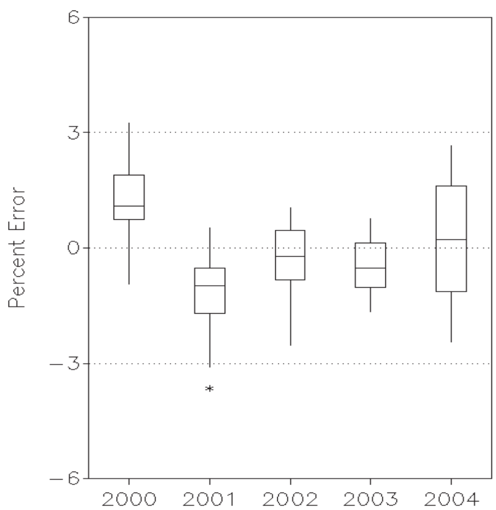
The distributions of percent errors of the MFW estimates and *PSM* interim values for 2000 through 2004 of motor gasoline and distillate fuel oil imports are shown in Figure FE10. The 2004 MFW percent errors for motor gasoline imports were evenly distributed about the median of -0.20 percent, meaning that one half of the estimates were overestimates and one half were underestimates. All but one of the 2004 *PSM* percent errors for motor gasoline imports underestimated the final *PSA* values.

The 2004 median of -1.56 percent for the MFW percent errors for distillate fuel oil imports was the closest to zero for the 5-year period. October 2004 (30.04) had the largest percent error over the 60 months studied. All but one of the 2004 *PSM* interim values for distillate fuel oil imports underestimated the

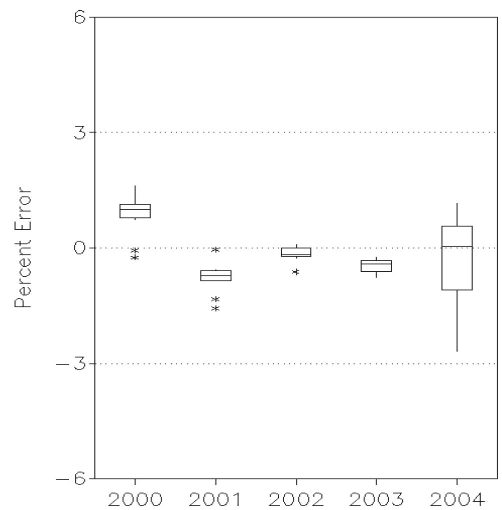
Figure FE6. Range of Percent Errors for MFW and *PSM* Crude Oil Stocks Excluding Strategic Petroleum Reserve (SPR) and Motor Gasoline Stocks Data, 2000 -2004

Crude Oil Stocks Excluding SPR

MFW vs. *PSA*

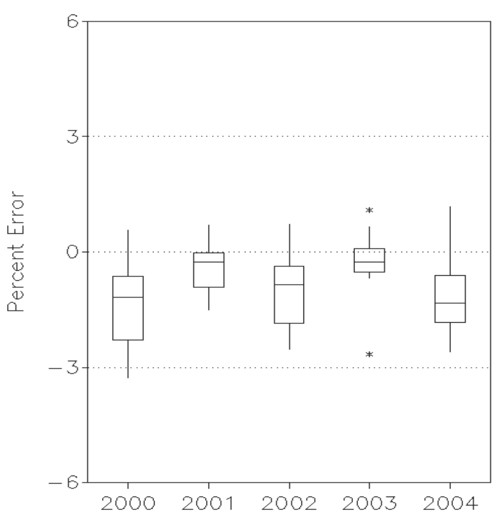


PSM vs. *PSA*

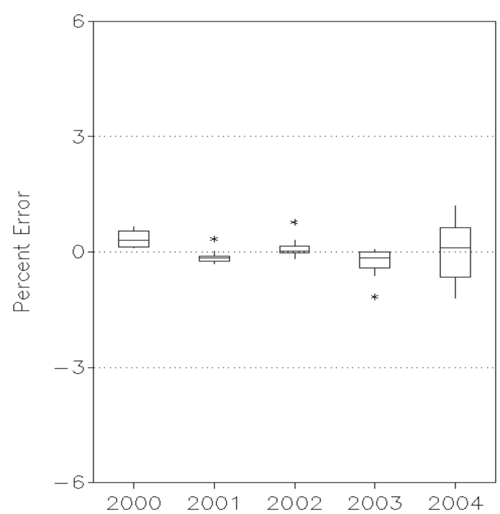


Motor Gasoline Stocks

MFW vs. *PSA*

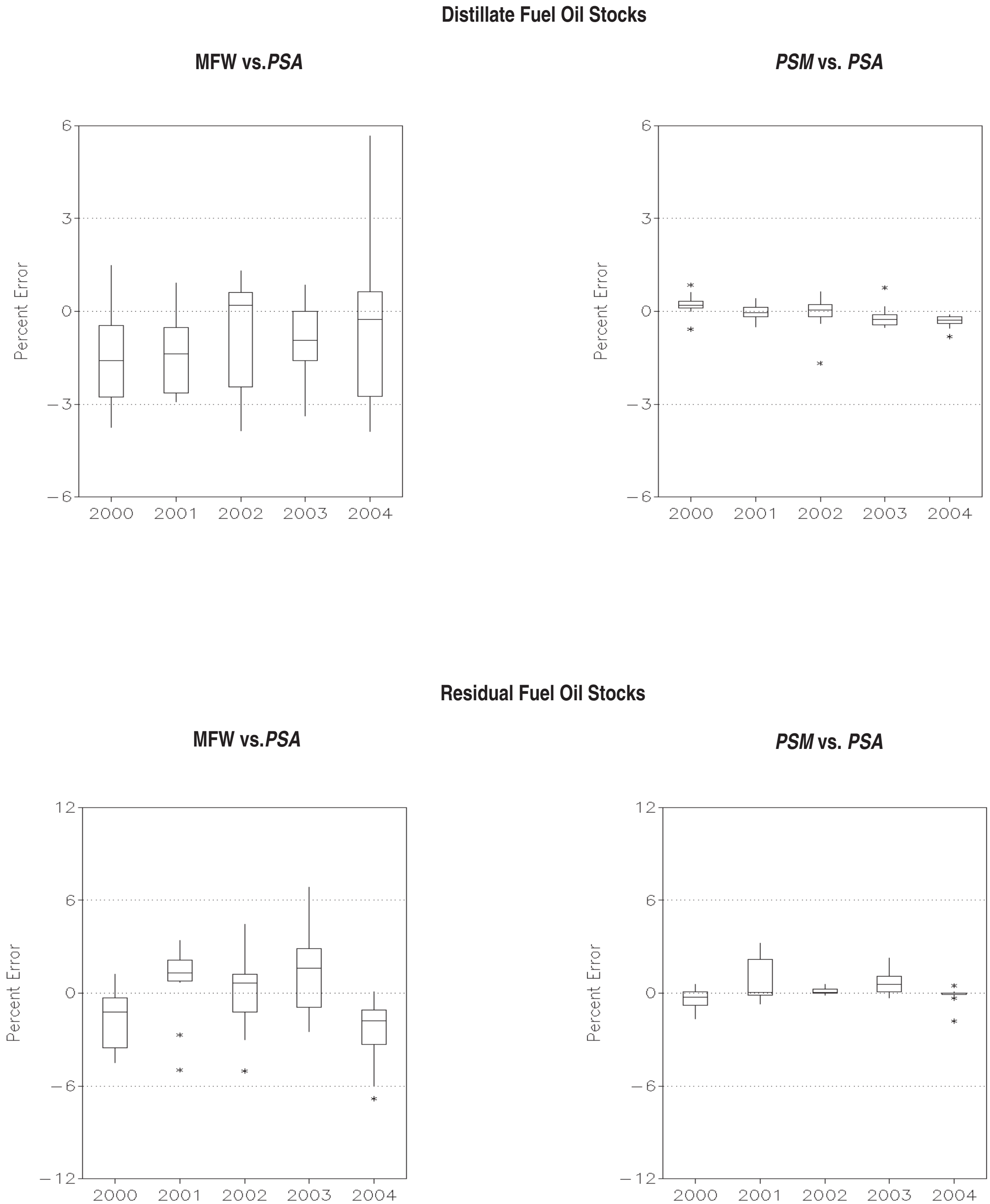


PSM* vs. *PSA



Source: Energy Information Administration, Petroleum Supply Reporting System.

Figure FE7. Range of Percent Errors for MFW and *PSM* Distillate Fuel Oil and Residual Fuel Oil Stocks Data, 2000 - 2004

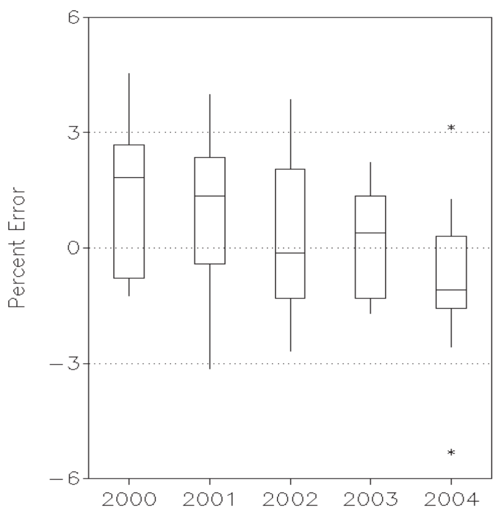


Source: Energy Information Administration, Petroleum Supply Reporting System.

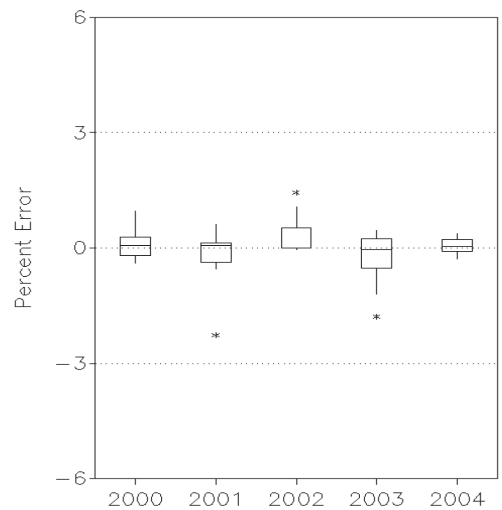
Figure FE8. Range of Percent Errors for MFW and *PSM* Jet Fuel Stocks and Propane Stocks Data, 2000 - 2004

Jet Fuel Stocks

MFW vs. *PSA*

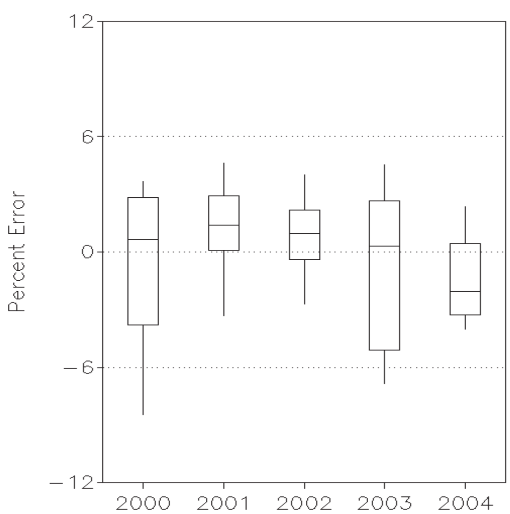


PSM vs. *PSA*

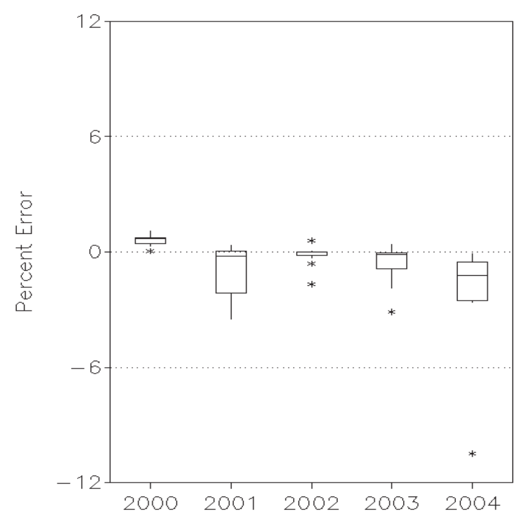


Propane Stocks

MFW vs. *PSA*

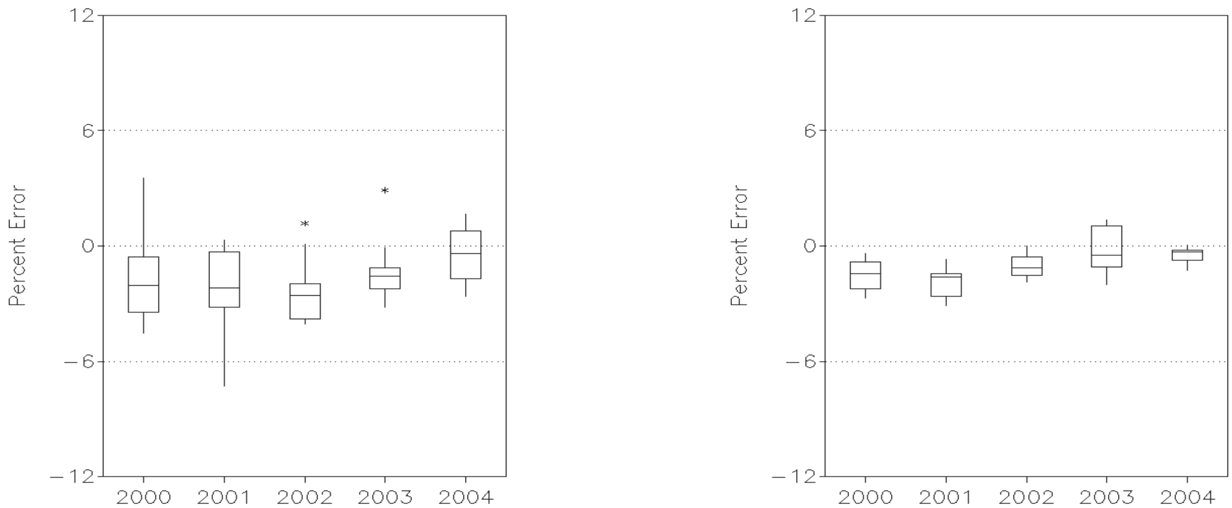


PSM vs. *PSA*



Source: Energy Information Administration, Petroleum Supply Reporting System.

Figure FE9. Range of Percent Errors for MFW and *PSM* Crude Oil Imports Excluding SPR Data, 2000 - 2004



Source: Energy Information Administration, Petroleum Supply Reporting System.

final *PSA* values. Two outliers occurred in April (-8.61) and July (6.01).

Figure FE11 shows the box and whisker plots for residual fuel oil imports and jet fuel imports. All but one of the 2004 MFW estimates for residual fuel oil imports underestimated the final *PSA* values. The range of 62.33 percent was the largest range of all other MFW plots analyzed for 2004, ranging from -51.22 to 11.11 percent. All but one of the 2004 *PSM* interim values for residual fuel oil imports underestimated the final *PSA* values.

Unlike 2003, more of the 2004 MFW estimates for jet fuel imports were underestimates. The range (30.00) of the 2004 *PSM* percent errors for jet fuel imports was the largest range over the 5-year period and of all other *PSM* plots analyzed for 2004, ranging from -30.00 to 0 percent. September 2004 (-30.00) had the largest absolute percent error over the past 60 months. There was one outlier in February (-15.45).

Conclusion

In summary, similar to previous years, the interim *PSM* data were closer in value to the final *PSA* volumes than the MFW estimates. This is largely a result of the longer time period provided to process the monthly data and monthly respondents' accounting systems.

In 2004, 34 of 63 *PSM* interim values were within 1 percent (mean absolute percent error) of the final values; 24 of 56 MFW estimates were within 2 percent (mean absolute percent error) of the final values; and 13 of those 24 were within 1 percent. As in previous years, the accuracy of 2004 preliminary and interim values varied by product and by petroleum supply type. As a group, stocks continued to have the most accurate MFW estimates and *PSM* interim values.

The good coverage for weekly surveys across petroleum supply type and product combinations has contributed to the accuracy of weekly estimates. In 2004, for 20 of the 21 categories, coverage was 90 percent or above. One half of the 2004 response rates for the weekly and monthly surveys increased or stayed the same percent compared to the 2003 response rates.

To successfully maintain and improve the accuracy of these data, the Petroleum Division (PD) is participating in several Office of Oil and Gas initiatives in the areas of data collection, survey processing, and data dissemination.

As a result of a comprehensive review of current petroleum industry operations and an analysis of pending product changes resulting from the Clean Air Act, the EIA made significant changes in the survey data collected starting with January 2004. These changes included the initiation of two new surveys, the EIA-805, "Weekly Terminal Blenders Report," and the EIA-815, "Monthly Terminal Blenders Report." Propane weekly data, that had formerly been collected through a separate EIA-807 survey and processing system, was eliminated and the collection of propane data was included as a major product on the primary weekly petroleum surveys (EIA-800-804). While there were numerous small changes to many product categories, such as the inclusion of a new ultra-low sulfur level diesel category and new categories for oxygenate production, the most significant product category changes occurred in motor gasoline. To better track the increasing volumes of special reformulated fuels meeting new Federal and State regulations, petroleum weekly and monthly surveys now track six separate categories of blending components and five categories of finished gasoline. All these changes will provide our Federal, State, and private customers with valuable new data from which to analyze and assess the U.S. petroleum market.

In addition, in January 2004, the PD implemented a new Weekly Petroleum Supply Status Report System. The previous system was written in Clipper, used the DOS Operating System, and was on a Local Area Network. It was rewritten to run in Access using a Windows Operating System and reside on a SQL Server. Enhancements to the system included more reports for assessing quality. The publication system was upgraded to a web-based system.

Some other areas of improvement during 2004 included enhancements to the Data Collection Module (DCM) which allows data to be collected in a common system, enhancements to the Standard Energy Processing System (STEPS) which is designed to handle different surveys with different needs using generalized programs and data structures to process survey data, the continuation of nonresponse follow-up and customer outreach, the expansion and improvement of electronic data dissemination on the EIA web site, including many new user-friendly information retrieval options; and the continuation of efforts to insure compliance with reporting requirements.

In 2004, the PD continued to expand the Survey Information System (SIS) which contains information needed for data validation and ad hoc queries. The system is a valuable link between the output from STEPS and data repository systems which produce the publications. The system was modified to handle the new surveys and product categories, and it was upgraded to the newest software versions of Oracle and PowerBuilder. The historical data were partitioned by year to improve performance. The capability to edit saved queries was enhanced.

In 2005, a new web product, Petroleum Navigator, was introduced. Petroleum Navigator provides an integrated and consistent interface for accessing a comprehensive set of EIA's petroleum data. Features include: downloadable spreadsheets containing complete data history, data tables which "pivot" to present different perspectives, and selection boxes to easily change the product, area, process, period, and unit of measure. Petroleum Navigator can be accessed at the following website: http://tonto.eia.doe.gov/dnav/pet/pet_sum_top.asp

In 2005, the Electronic Data Extraction System (EDES) was implemented. This system automatically extracts data from Excel spreadsheets submitted by some survey respondents through Secure File Transfer or email, and transforms the data into a format that can be sent to the DCM and then to STEPS.

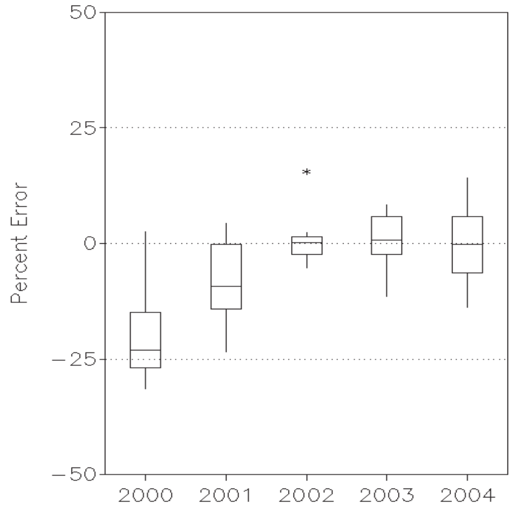
Also in 2005, a new survey processing system was developed in Access to process data on exports, thus eliminating multiple steps on multiple applications. The new system will allow better country detail information and be easier to execute and maintain.

Results of these efforts should enable the PD to continue to provide accurate weekly and monthly data estimates.

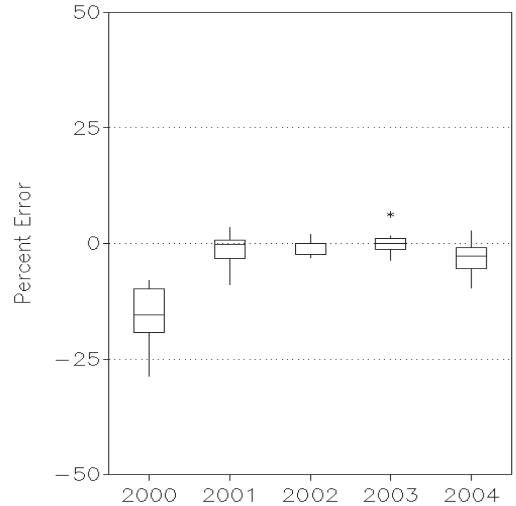
Figure FE10. Range of Percent Errors for MFW and *PSM* Motor Gasoline and Distillate Fuel Oil Imports Data, 2000 - 2004

Motor Gasoline Imports

MFW vs.*PSA*

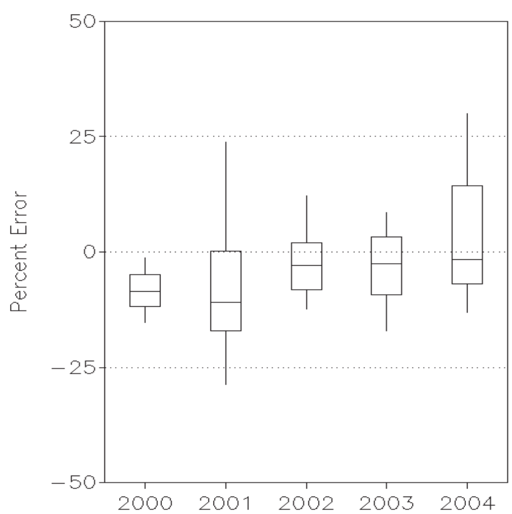


PSM* vs. *PSA

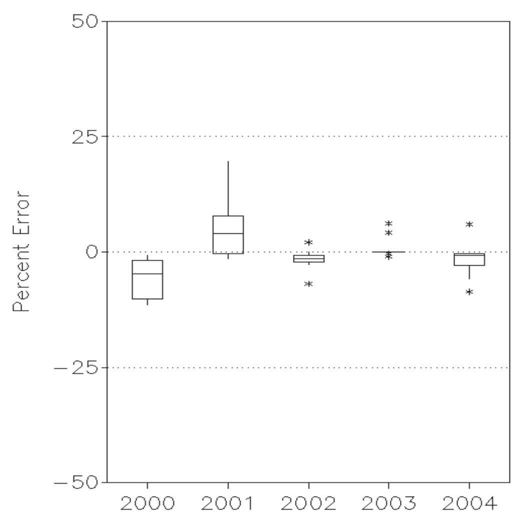


Distillate Fuel Oil Imports

MFW vs.*PSA*



PSM* vs. *PSA

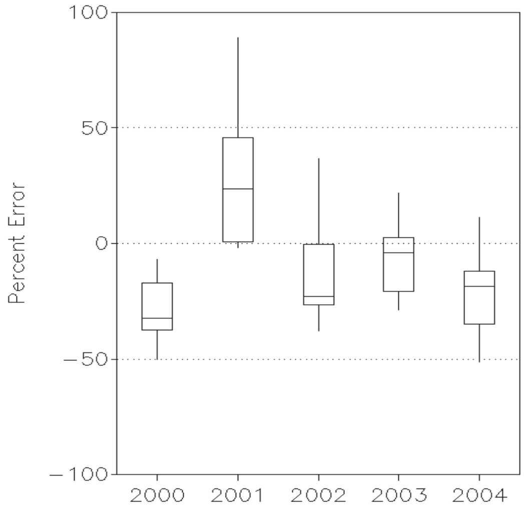


Source: Energy Information Administration, Petroleum Supply Reporting System.

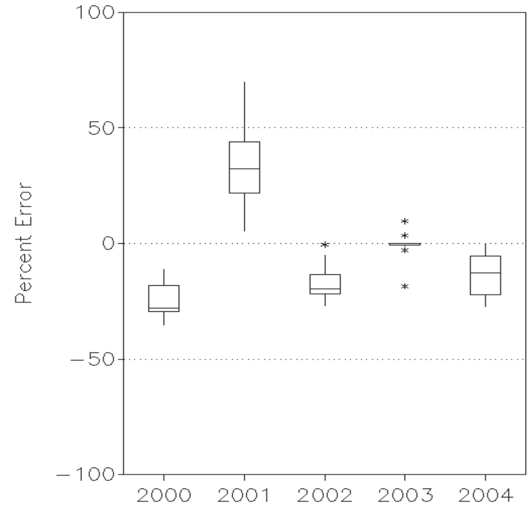
Figure FE11. Range of Percent Errors for MFW and *PSM* Residual Fuel Oil and Jet Fuel Imports Data, 2000 - 2004

Residual Fuel Oil Imports

MFW vs.*PSA*

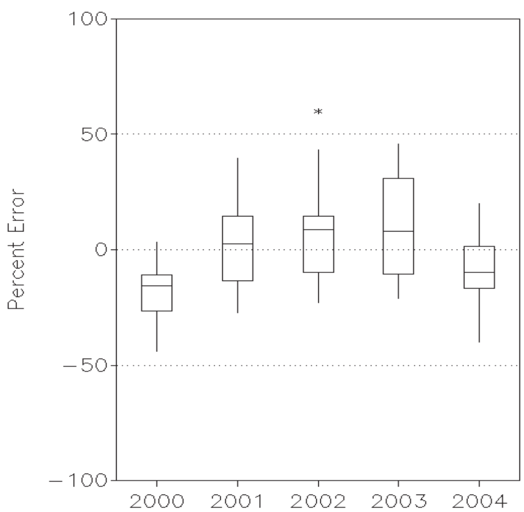


PSM* vs. *PSA

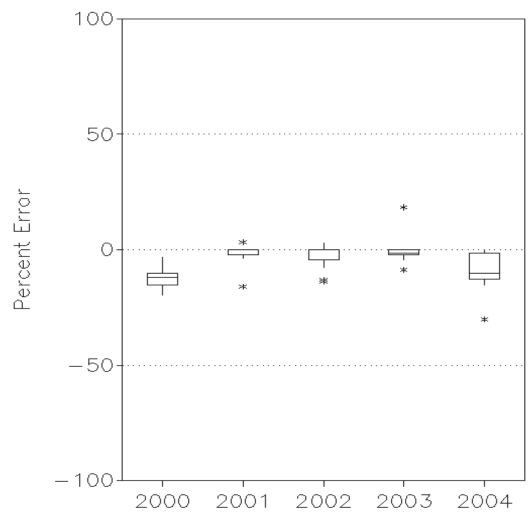


Jet Fuel Imports

MFW vs.*PSA*



PSM* vs. *PSA



Source: Energy Information Administration, Petroleum Supply Reporting System.