



Independent Statistics & Analysis
U.S. Energy Information
Administration

Appendix D

Statistical Methodology of Estimating Petroleum Exports Using Data from U.S. Customs and Border Protection

Updated July 8, 2020



This report was prepared by the U.S. Energy Information Administration (EIA), the statistical and analytical agency within the U.S. Department of Energy. By law, EIA's data, analyses, and forecasts are independent of approval by any other officer or employee of the United States Government. The views in this report therefore should not be construed as representing those of the U.S. Department of Energy or other federal agencies.

Introduction

On August 31, 2016, the U.S. Energy Information Administration (EIA) started incorporating export data collected by the U.S. Customs and Border Protection (CBP) through the Automated Commercial Environment (ACE) system to help estimate U.S. petroleum exports reported in the *Weekly Petroleum Status Report* (WPSR). The updated WPSR methodology provided a more accurate estimate of U.S. petroleum exports to better inform EIA's weekly estimates of domestic quantity demanded for crude oil, petroleum products, and biofuels.

Exports are an important component for calculating EIA weekly estimates of quantity demanded for petroleum products. EIA calculates domestic quantity demanded (measured as product supplied) for petroleum products as:

$$\text{Domestic quantity demanded} = \text{Production} + \text{Imports} - \text{Stock change} - \text{Exports.}$$

EIA conducts weekly surveys to collect the data that are used to estimate production, imports, and stocks but not exports. Export data are published monthly by the U.S. Census Bureau (USCB). To estimate weekly exports, EIA previously used forecasts based on an autoregressive integrated moving-average (ARIMA) procedure to model weekly export statistics using the latest USCB monthly data. EIA obtained official USCB export data for crude oil, petroleum products, and biofuels on a monthly basis approximately six weeks after the close of the reporting month and published them in its *Petroleum Supply Monthly* (PSM). These data provided the inputs to the weekly export ARIMA model. However, as the volume and variability of U.S. exports increased, significant discrepancies were sometimes observed between modeled weekly estimates and final monthly exports data provided by the USCB. Because of the growing impact of exports on EIA's weekly estimates of U.S. petroleum quantity demanded, EIA recognized the need to improve the weekly export estimates used in the WPSR.

Beginning in 2013, EIA investigated data sources and methodologies to improve its weekly estimates of petroleum and biofuels exports. EIA found that administrative data collected by CBP through the ACE system, which cover U.S. exports except for most exports to Canada, could be used to improve EIA's weekly export estimates. To obtain these data, EIA established a Memorandum of Understanding with CBP in September 2014 through the International Trade Data System (ITDS) process. In August 2016, the Office of Management and Budget (OMB) granted EIA an exception to the provisions of OMB's *Statistical Policy Directive No. 3 on the Compilation, Release, and Evaluation of Principal Federal Economic Indicators*. This exception permits EIA to publish weekly estimates of export volumes of crude oil and six refined petroleum products and biofuel categories in advance of the release of official U.S. trade data from USCB.

Since August 31, 2016, EIA has continuously monitored and analyzed the estimates derived from the recently acquired near real-time CBP export data. The remainder of this document describes the method deployed for the WPSR release on July 8, 2020, onward. You can find a description of changes to EIA's estimation methodology over time in Attachment B.

Trade Data Collection and Dissemination

To derive weekly estimates of domestic petroleum product supplies, the WPSR relies on estimates of weekly U.S. exports of

- Crude oil
- Finished motor gasoline
- Kerosene-type jet fuel
- Distillate fuel oil
- Residual fuel oil
- Propane
- Other oils, including
 - Biofuels
 - Asphalt
 - Aviation gasoline
 - Butanes
 - Ethane
 - Lubricants
 - Gasoline blending components
 - Miscellaneous products
 - Non-biofuel oxygenates
 - Petroleum coke
 - Unfinished oils
 - Wax

CBP collects the export data that EIA uses to develop these weekly estimates.

In general, entities exporting products from the United States are required to file export transaction data with CBP. This information is included in CBP's ACE system. As a result of an agreement with CBP, EIA now receives these unedited transactional ACE data files from CBP and uses them to develop aggregate estimates of weekly exports for the WPSR. The data from CBP include an export date that makes it possible to group transactions into WPSR weekly reporting periods. EIA has found that data received from CBP require editing and imputation to avoid using questionable data values for analysis, modeling, and aggregation. This document later describes the edit and imputation methods EIA uses. EIA is unable to follow up with exporters to resolve questions about the CBP data. USCB may follow up with exporters to address questionable data and, if necessary, obtain corrections during its preparation of official U.S. trade statistics.

Per U.S.-Canadian agreements, entities exporting to Canada are not required to report to CBP, so these files generally exclude Canadian transactions. Products that require an export license are an exception, and export transactions involving licensed products must be reported to CBP regardless of the destination country. Crude oil was a licensed export product until December 2015. USCB has an agreement with Statistics Canada to obtain monthly Canadian imports data from the United States, which USCB uses to produce monthly estimates of U.S. exports in official U.S. trade statistics. However, EIA does not have early access to Canadian import data. To address these issues, EIA developed methodologies to edit the raw CBP data and to estimate weekly exports to Canada.

The CBP data EIA uses to estimate weekly exports begin with the same data that the USCB uses as the initial input to prepare the monthly *International Trade in Goods and Services* report (FT900), a principal economic indicator. The FT900 has a six-week publication lag, while WPSR data are released each Wednesday for the weekly report period ending the previous Friday. Because the monthly publication cycles are more closely aligned, EIA uses the FT900 data to report monthly export volumes that are included in EIA's *Petroleum Supply Monthly* (PSM) and *Petroleum Supply Annual* (PSA).

Estimation Methodology

Before producing the published estimates for the WPSR, EIA performs the following edit steps after it receives CBP data for all products:

- Remove duplicate records
- Exclude records with export dates beyond the reporting period
- Exclude products that are out of the WPSR's scope
- Convert all quantities to barrels exported
- Calculate the dollar value per barrel and compare against daily spot prices from Refinitiv (formerly Thomson-Reuters) to identify extreme values. Use a procedure based on shipment or commodity weight to replace these values, if necessary. Shipment weight is the combined weight of the commodity being shipped and its packaging. If replacements based on commodity value or shipment weight are unacceptable, replace based on historical CBP data.

See Attachment A for a more complete list of edit rules that EIA applies to the CBP data.

Following editing, EIA's estimation method consists of two parts:

1. Estimates based on aggregations of edited CBP data

For those products that have adequate coverage in the CBP data, EIA reports the sum of the quantities exported using the edited data. These products include distillate fuel, propane, and finished motor gasoline. EIA will continue to monitor weekly export estimates for these products and may use some form of a post-editing statistical model if the weekly estimates consistently diverge from USCB's monthly exports data.

2. Estimates based on post-editing statistical models

For those products that appear to have coverage differences in the CBP data based on EIA's comparisons to the USCB monthly volumes that are published in the PSM, EIA uses a post-editing statistical model to correct for the coverage differences. These products include crude oil, kerosene-type jet fuel, residual fuel oil, and other oils. The post-editing statistical model is described below:

EIA uses a two-component approach to generate estimates. The first component is a linear regression model to estimate the exports to destinations other than Canada. The formula for the updated regression component is:

$$Y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$$

where t denotes a month and:

- Y_t is the non-Canadian (from PSM) export of crude oil, kerosene-type jet fuel, residual fuel oil, or other oils in barrels per day as published in the PSM for month t
- x_t is non-Canadian edited CBP export data in barrels per day aggregated for month t to be used as the regressor
- β_0 is the intercept
- β_1 is the regression coefficient for the regressor x_t
- ε_t is the error term

The estimated regression coefficients are applied to the non-Canadian edited CBP export data on a weekly basis to produce the published weekly exports data. This method assumes a similar linear relationship between the non-Canadian PSM and CBP data on a weekly basis for the weeks in a given month. EIA will continue to assess the results of this approach and identify alternative methods to develop weekly export volume estimates, if needed.

The second component is a three-month moving average of Canadian exports in thousand barrels per day as reported in the latest series of the *Petroleum Supply Monthly*.

Summary

Using near real-time petroleum export data from CBP, published monthly trade data from the USCB, and revised estimation methodologies, has improved the accuracy of EIA's WPSR estimates of U.S. weekly exports. EIA will continue to monitor the performance of these methods and will modify them as needed to further improve their accuracy.

Attachment A: Edit Rules

The following are the calculations and edit rules in place to prepare the CBP data for estimates:

1. Identify and remove records that are outside the reporting period.
2. Identify and remove duplicate records.
3. Exclude exports originating from U.S. territories (for example, Puerto Rico and U.S. Virgin Islands).
4. Map Schedule B/Harmonized Tariff Schedule codes to WPSR products.
5. Keep only products that are within scope of WPSR products.
6. Convert units of quantity not already in barrels (for example, kilograms, liters, and tons) to barrels.
7. Calculate a dollar value per barrel using spot prices from Refinitiv (formerly Thomson-Reuters).
8. Based on value (in dollars) divided by quantity of commodity exported, calculate a unit price (dollars per barrel).
9. Compare the unit price with spot prices to determine outliers based on an acceptable range around spot prices.
10. Flag extreme values. Then use the following procedure:
 - a. Recalculate weight of the commodity exported based on quantity (in barrels) and a conversion factor.
 - b. Compare the recalculated weight against a proxy quantity such as shipment or commodity weight based on an acceptable range around shipment or commodity weight.
 - c. If the recalculated weight is within an acceptable range around shipment or commodity weight, then use the original reported quantity in barrels.
 - d. If the recalculated weight is not within an acceptable range around the shipment or commodity weight, then use the gross weight to impute reported quantity.
 - i. Recalculate a unit price (dollars per barrel), compare with spot prices, and flag those outside of an acceptable range.
 - e. For flagged values, compare the original and recalculated weights with maximum capacity. If both weight values exceed an acceptable range around maximum capacity, use a specified upper percentile based on an analysis of the upper tail of the distribution of weight from historical CBP data relative to maximum capacity.

Attachment B: Updates to WPSR's Methodology for Estimating Exports of Crude Oil, Finished Motor Gasoline, Residual Fuel Oil, and Kerosene-Type Jet Fuel

Phase 1: Implemented starting with the WPSR published on March 8, 2017

- Updated the regression model (Model 1), which was used to estimate total exports from edited CBP data. The regression model used CBP data as an input and accounted for U.S. exports to Canada that are missing from CBP data. The updated model (Model 2), included an intercept term and used daily CBP data that EIA edited to identify and impute questionable quantities and then aggregated to monthly totals as input. Model 1 used monthly USCB data that were interpolated to a weekly series.
- Began using Model 2 to estimate finished motor gasoline exports. The previous method had been used to report finished motor gasoline exports equal to the aggregation of CBP data after editing and imputation.

Phase 2: Implemented starting with the WPSR published on June 7, 2017

- Model 2 was used to estimate crude oil exports, replacing the Unobserved Components Model (UCM).

Phase 3: Implemented starting with the WPSR published on March 14, 2018

- Continued same methodology using edited CBP data for estimating distillate fuel oil and propane (unchanged since August 2016).
- Updated the methodology to estimate exports of crude oil, kerosene-type jet fuel, residual fuel oil, finished motor gasoline, and other oils.
 - For residual fuel oil and finished motor gasoline, EIA reports edited CBP data without further modeling.
 - For crude oil, kerosene-type jet fuel, and other oils, EIA uses an updated methodology (called Model 3) as described in the estimation methodology section above.

Phase 4: Implemented starting with the WPSR published on July 8, 2020

- Updated the methodology to estimate exports of residual fuel oil using Model 3.
- Continued using the same methodology for all other products (unchanged since March 2018).