



*Independent Statistics & Analysis*

U.S. Energy Information  
Administration

---

# U.S. Energy-by-Rail Data Methodology

June 2016



---

*Independent Statistics & Analysis*

[www.eia.gov](http://www.eia.gov)

U.S. Department of Energy

Washington, DC 20585

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.

## Table of Contents

---

Introduction .....	1
Overview of Methodology .....	2
Carload waybill sample data .....	3
Converting to barrels .....	5
Crude oil .....	5
Ethanol and biodiesel.....	6
Near-month estimates.....	6
PSM revision schedule .....	7
Illustrative Energy-by-Rail Calculations.....	8

## Figures

---

Figure 1. Petroleum Administration for Defense District (PADD) map .....	1
Figure 2. EIA's energy-by-rail data process .....	2

## Tables

---

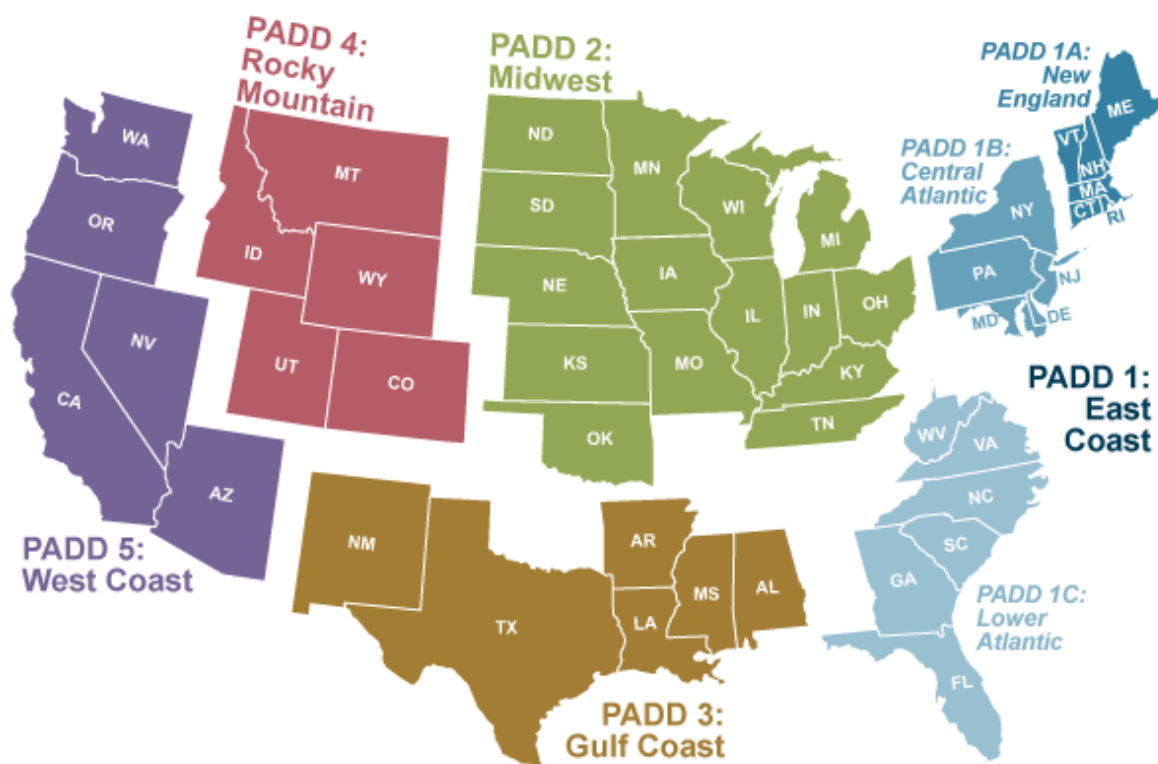
Table 1. Calculation of the EIA reporting period.....	4
Table 2. Standard Transportation Commodity Codes (STCC) used in EIA rail data .....	5
Table 3. Illustrative calculation from a waybill .....	8
Table 4. Converting carloads to barrels of crude oil in any quarter .....	9
Table 5. Choosing near-month estimate models.....	9

## Introduction

In March 2015, the U.S. Energy Information Administration (EIA) began to provide monthly data on rail movements of crude oil, which have significantly increased over the past six years. Monthly ethanol and biodiesel-by-rail data were first introduced in April 2016. EIA reports energy-by-rail data back to January 2010 and integrates the rail data with EIA's existing monthly petroleum supply statistics, which already include movements by pipeline, tanker, and barge. The rail data improve the overall quality of EIA's petroleum statistics by providing additional information and transparency into the movement of crude oil and selected petroleum products and biofuels.

EIA uses Petroleum Administration for Defense Districts (PADDs, see Figure 1) to track energy-by-rail movements between pairs of PADD regions (inter-PADD), within each region (intra-PADD), and across the U.S.-Canada border. EIA developed the new series using information provided by the U.S. Surface Transportation Board (STB) and EIA survey data. Canada's National Energy Board (NEB) provides select information used to calculate cross-border crude-by-rail data.

**Figure 1. Petroleum Administration for Defense District (PADD) map**

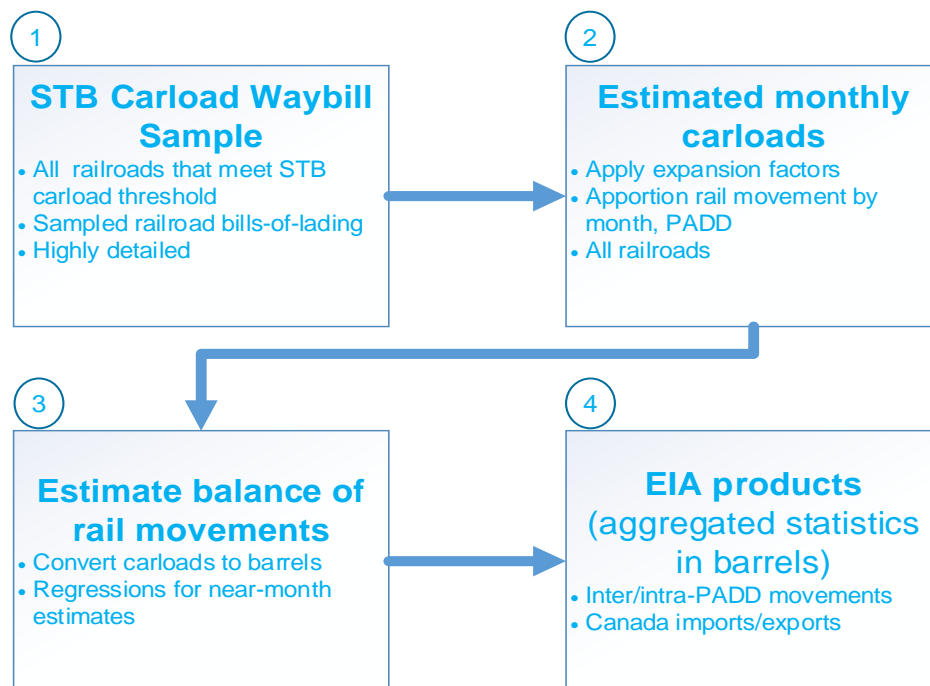


Source: U.S. Energy Information Administration

## Overview of Methodology

EIA energy-by-rail data are calculated based on data from the STB Carload Waybill Sample, which is discussed more fully in the next section. Railroads that terminate, i.e., deliver to a final destination, more than 4,500 carloads per year in the United States must submit a waybill sample quarterly to the STB.<sup>1</sup> The sampling rate varies with the number of carloads in the waybill.<sup>2</sup> Figure 2 shows how the carload waybill sample data are used to calculate energy-by-rail data. The carload waybill sample provides the origin and destination rail station of each sampled waybill, which EIA aggregates by PADD and month.

**Figure 2. EIA's energy-by-rail data process**



Source: U.S Energy Information Administration

To estimate the number of carloads shipped, EIA divides the number of carloads reported in the waybill sample by the sampling rate. EIA calculates a quarterly barrel-per-carload factor for each origination region from carload weights and either densities (for ethanol or biodiesel) or API gravities of crudes<sup>3</sup> produced in the United States and Canada. EIA uses the barrel-per-carload factor to convert the number of expanded carloads to barrels for integration with published EIA petroleum and biofuel data series. For near-month estimates that are issued before STB data are available, EIA models energy-by-

<sup>1</sup> 49 CFR. § 1244.2, *Title 49 – Transportation*, “Waybill analysis of transportation of property – railroads,” <http://www.gpo.gov/fdsys/pkg/CFR-2012-title49-vol9/pdf/CFR-2012-title49-vol9-part1244.pdf>.

U.S.-based Class I railroads also submit monthly reports.

<sup>2</sup> STB carloads per waybill sampling rates: 1–2 carloads → 1:40; 3–15 carloads → 1:12; 16–60 carloads → 1:4; 61–100 carloads → 1:3; and > 100 carloads → 1:2

<sup>3</sup> U.S. Energy Information Administration, “Crude oils have different quality characteristics,” *Today in Energy*, July 16, 2012, <https://www.eia.gov/todayinenergy/detail.cfm?id=7110>.

rail data using other available EIA survey data like field production, refinery and blender inputs, imports, and exports, as well as lagged energy-by-rail data. Like all of EIA's monthly petroleum and biofuel data, the new energy-by-rail data are subject to revision as new information becomes available. Near-month estimates are updated upon receipt of information from STB.

## Carload waybill sample data

STB provides a stratified carload waybill sample on a monthly, quarterly, and annual basis to EIA under a special confidential waybill agreement for specific Standard Transportation Commodity Codes (STCC). The carload waybill sample consists of waybills submitted by railroads annually terminating 4,500 or more revenue carloads passing through the United States.<sup>4</sup>

The carload waybill sample includes waybills for both hazardous material (hazmat) and non-hazmat STCCs at the 7-digit level. Release of the annual waybill sample can be delayed up to one year beyond the reporting period, but the quarterly and monthly waybill samples are both available with an approximate two-month lag after the end of the reporting period. The monthly waybill sample is not complete for the first two months in a given calendar quarter; the first two months' sample does not include waybills from some Class II and Class III railroads.<sup>5</sup>

Each railroad selects its own systematic sample of waybills, and waybills for larger shipments are sampled with higher probabilities. Sampling rates ranging from 1:2 for 101+ carload unit train shipments to 1:40 for single-carload manifest shipments. Railroads sample their own waybills and provide 192 fields of data<sup>6</sup> to the STB in the annual CWS. Of interest to EIA, regardless of the time frequency of the waybill, are the waybill date, the number of carloads, the sampling rate, the carload weight, the origin and destination of the shipment, and any interchange of the shipment between railroads.

EIA does not use the sample accounting periods identified by the railroads, because they do not always align with the waybill dates, e.g., some Class I railroads have bundled all quarterly transactions in the third month of a quarterly accounting period, and other railroads have sampled waybills that appear in accounting periods two-to-five months following the waybill date. EIA calculates an "EIA reporting period" from the carload waybill sample waybill date that varies by the type of train,<sup>7</sup> commodity carried, and number of railroads involved in the movement, as shown in Table 1. Energy-by-rail movements can involve multiple waybills and railroads, such as Rule 11 transactions involving the transfer of ethanol cars in the Chicago area.<sup>8</sup> The time to complete a transaction increases with the

---

<sup>4</sup> 49 CFR. § 1244.2, *Title 49 – Transportation*, "Waybill analysis of transportation of property – railroads," <http://www.gpo.gov/fdsys/pkg/CFR-2012-title49-vol9/pdf/CFR-2012-title49-vol9-part1244.pdf>.

<sup>5</sup> The seven Class I railroads are Burlington Northern Santa Fe (BNSF), CSX, Kansas City Southern (KCS), Norfolk Southern (NS), Union Pacific (UP), Canadian National (CN), and Canadian Pacific (CP). All other railroads are either Class II or Class III.

<sup>6</sup> Surface Transportation Board, *2014 Surface Transportation Board Carload Waybill Sample Reference Guide*, November 1, 2015, <http://www.stb.dot.gov/stb/docs/Waybill/2014%20STB%20Waybill%20Reference%20Guide.pdf>.

<sup>7</sup> Unit trains are composed of 80 or more like cars hauling a single bulk commodity, while manifest trains are composed of mixed rail cars (i.e., tank cars, boxcars) carrying multiple commodities.

<sup>8</sup> Rule 11 is an accounting rule governing shipments involving two or more railroads in which each railroad submits its own waybill for the same shipment. See Association of American Railroads, *Railway Accounting Rules*, [https://www.railinc.com/rportal/alf\\_docs/RAR.pdf](https://www.railinc.com/rportal/alf_docs/RAR.pdf) for background.

number of railroads involved in a single movement. The “EIA reporting period” month is used to aggregate the energy-by-rail data.

**Table 1. Calculation of the EIA reporting period**

Commodity	Number of railroads involved in waybill	Number of days added to the waybill date by train type	
		Unit	Manifest
Crude oil	1	4	12
	2+	6	16
Ethanol/Biodiesel	1	6	14
	2+	8	18

Source: U.S. Energy Information Administration

Table 2 shows the STCCs that EIA currently uses to calculate movements of petroleum and biofuels. Because STCCs used by the railroads can change without notice, the table is subject to revision. Bitumen and condensate are included in the crude-by-rail data.

**Table 2. Standard Transportation Commodity Codes (STCC) used in EIA rail data**

Commodity	STCC	Hazmat STCC	Commodity	STCC	Hazmat STCC	
<b>Crude oil</b>	1311110	4910165	<b>Ethanol</b>	2991240	4908179	
		4910191				4910236
		4910599				4912336
		4915165			2085120	4907603
<b>Bitumen</b>	2899885	4914109				4910102
<b>Condensate</b>	2911976	4910164				4910240
<b>Biodiesel</b>	2899416	None				4914105
					2818414	4912269
					2818419	4909117
						4909363
					2818445	4909104
						4909105
						4909159
						4914146
				2818446	4908180	
					4909146	
		4909152				
		2818447	4908170			
			4909109			

Notes: Other hazmat STCCs corresponding to 2899885 or 2911976 are excluded from the crude oil category based on EIA research. Waybills that have origin stations near facilities that EIA has determined to produce ethanol for non-fuel use are omitted.

Source: U.S. Energy Information Administration

## Converting to barrels

### *Crude oil*

EIA's National Energy Modeling System/Oil and Gas Supply Module (OGSM) computes unpublished estimates of crude oil production volumes by crude type and region. EIA used crude oil production estimates by OGSM regions as defined in the *Annual Energy Outlook*, Appendix F,<sup>9</sup> combined with a representative API gravity for each crude type, to calculate a volume-weighted API, which is a measure of how heavy or light a petroleum liquid is compared to water. If its API gravity is greater than 10, a liquid is lighter than and floats on water; if its API gravity is less than 10, it is heavier than and sinks in water. Even though API gravity is an inverse measure of the relative density of a petroleum liquid with respect to the density of water, it is used to compare the relative densities of petroleum liquids. For 2010-14 crude-by-rail movements, EIA used the volume-weighted API calculated from OGSM and the car weights found in the carload waybill sample data to convert carloads to barrels for U.S.-originated crude-by-rail movements. EIA chose representative API gravities for the primary crude oil in each

<sup>9</sup> U.S. Energy Information Administration, *Annual Energy Outlook 2015*, April 2015, <http://www.eia.gov/forecasts/aeo/pdf/0383%282015%29.pdf>.



originating region to calculate barrel-per-car averages by quarter and originating PADD. The barrel-per-car averages by quarter are used to convert U.S.-originating carloads to barrels. For 2015 onwards, the midpoint API gravity from the EIA-914 is used for the carload-to-volume conversion.<sup>10</sup>

EIA's approach is different for Canadian-originating carloads. The National Energy Board (NEB) of Canada regulates the export of crude oil and petroleum products and requires export orders or licenses for these commodities leaving Canada. Holders of export authorizations report monthly on their export activities to NEB, indicating the mode of transport: rail or pipeline. The aggregated export data are publicly available on a monthly basis from 2012, but by special agreement between NEB and EIA, unpublished details such as API gravity, province of origin, and destination are shared. EIA calculates barrel-per-car averages by quarter for Canadian exports, which are used to convert Canadian-originating carloads to barrels.

### *Ethanol and biodiesel*

For ethanol and biodiesel, the density of the respective liquids in combination with the weights found in the carload waybill sample are used to convert carloads to barrels on a quarterly basis. Biodiesel carload weights are averaged on an originating-PADD basis. Ethanol carload weights are averaged on origin-destination PADD basis.

## Near-month estimates

EIA produces energy-by-rail movement data timed with the release of the *Petroleum Supply Monthly (PSM)*, EIA's monthly publication of petroleum and biofuels data. *PSM* published data include supply field production, refinery and blender net production, imports, receipts, adjustments and disposition, stock change, refinery and blender net inputs, shipments, exports, and product supplied. EIA includes energy-by-rail movements along with pipeline and waterborne movements in the calculation of PADD-to-PADD shipments and receipts. The *PSM* is released on the last business day of the month and contains data for the penultimate month prior to publication. Energy-by-rail estimates are published in *PSM* Tables 57-62.<sup>11</sup> However, the most current monthly STB waybill sample is not received with sufficient lead time to be processed and included in the *PSM* publication. In addition, STB aggregates and provides waybills based on the accounting period, while EIA is interested in the waybill date, which can differ by several months from the accounting period. As a result, EIA must estimate anywhere from one to three months of rail movements, relying on near-month model estimates for those months to be published in the *PSM* for which incomplete or no STB data are available.

To develop near-month estimates, EIA runs four different models:

- Ordinary least squared model (OLS)
- Auto-regressive model (AR)
- Average of the previous six months of data

---

<sup>10</sup> API gravity table for data entry on new EIA-914 form, *Monthly Crude Oil, Lease Condensate, and Natural Gas Production Report*, section 4.1, page 9, [http://www.eia.gov/survey/form/eia\\_914/form.pdf](http://www.eia.gov/survey/form/eia_914/form.pdf), is reported monthly on <http://www.eia.gov/petroleum/production/>.

<sup>11</sup> U.S. Energy Information Administration, *Petroleum Supply Monthly*, Tables 57-62, <http://www.eia.gov/petroleum/supply/monthly/>.

- Repeat of the last available calculated regional STB data

EIA also calculates the following hybrid models with equal weighting for each component:

- OLS/AR
- AR/6-month average
- AR/OLS/6-month average

Each month, EIA prepares one to three months of near-month estimates and evaluates the accuracy of the estimates using three-month, out-of-sample forecasts (OSF). EIA uses a pair of three-month OSFs (six monthly data points) to calculate a root-mean squared error (RMSE). For each of the region-to-region rail movements, EIA uses the RMSE estimates, in combination with analyst judgment, to select model-based near-month estimates. The procedure is repeated each time STB releases a new set of quarterly data.

### PSM revision schedule

EIA may revise energy-by-rail data back to January of the penultimate year during any *PSM* monthly publication cycle. This means that for any *PSM* published through July 2016, EIA can revise energy-by-rail data back to January 2014. When the *Petroleum Supply Annual* is published, typically in August, EIA may revise energy-by-rail data back 10 years.

EIA revises energy-by-rail data monthly, revising model-based near-month estimates with sample-based data from the STB waybill sample and revising other historical data as more accurate information becomes available.

## Illustrative Energy-by-Rail Calculations

The first step in calculating energy-by-rail movements is to examine carload waybill sample data and concatenate any Rule 11 (i.e., multiple railroad) waybills that exist. EIA assigns a train type either manifest or unit as well as both an origin and destination PADD, based on the origin and destination stations, respectively. EIA calculates the “EIA Reporting Period” for each waybill and aggregates expanded carloads by pairs of origin and destination PADDs (and Canada). The last step is to convert the carload units to barrels. Table 3 illustrates how train type, origin, destination, and EIA reporting period are derived from a crude oil waybill.

**Table 3. Illustrative calculation from a waybill**

STB Waybill*		Derived by EIA	
Data Item Description	Data Value	Data Value	Data Item Description
STCC	1311110	Crude oil	Commodity
Hazmat STCC	4910191	BNSF	Origin railroad
Origin railroad	777	CSX	Destination railroad
Destination railroad	712	2	Number of railroads
Number of carloads	104	Unit train	Type of movement
Expanded carloads	208	6	Number of days added to waybill date
Waybill date	3/28/2016	16-Apr	EIA reporting period
Accounting period	16-Mar	Trenton, ND	Origin station
Origin FSAC**	BNSF-1049	Philadelphia, PA	Destination station
Destination FSAC**	CSX-70005	2	Origin PADD
		1	Destination PADD

\* For illustration purposes only

\*\*FSAC = Freight Station Accounting Code

Source: U.S. Energy Information Administration

Note that in this example in Table 3, the transaction would be reported as an April 2016 crude-by-rail movement from PADD 2 to PADD 1 even though the waybill date is in March. Six days are added to the waybill date, because it is a unit-train move involving two railroads (see Table 1). The aggregated monthly-regional carloads are converted to barrels by multiplying the number of carloads by a barrel-per-carload factor that is calculated quarterly based on API gravity, carload weight, and carload origination. Table 4 provides an example carload-to-barrel calculation for crude-by-rail originating in PADD 2 in a quarter. Changing the PADD of origin or quarter will result in different results.

**Table 4. Converting carloads to barrels of crude oil in any quarter**

Origin PADD	2
Average carload weight (lbs)	175,000
Volume Weighted API	40
	Carload Weight
Formula	$\left(\frac{141.5}{\text{API} + 131.5}\right) * \text{Density of Water}$
<b>Barrels per car</b>	<b>607</b>

Source: U.S. Energy Information Administration

For months without STB data, EIA develops near-month estimates using regression models and other types of models. Table 5 provides the different models EIA estimates, illustrative forecast errors, regression samples for those models, and RMSEs. EIA uses the model with the lowest RMSE, shaded in gray for this example in Table 5, to calculate near-month estimates for a particular region-to-region movement for the next three months. EIA runs regressions quarterly when new STB data are available to re-evaluate the models.

**Table 5. Choosing near-month estimate models**

Model	Last month with STB data: June 2015			Forecast Error			6-month RMSE
	15-Jan	15-Feb	15-Mar	15-Apr	15-May	15-Jun	
OLS	90	-44	-94	93	8	14	68
AR	19	-49	30	1	25	-13	27
6-month average	87	58	-43	-72	49	20	59
Repeat	77	81	79	50	55	45	50
OLS/AR	66	-29	76	57	82	18	60
AR/6-month average	-51	26	-13	-79	93	-77	64
OLS/AR/6-month average	-58	55	-56	-95	29	80	66

**Regression Sample**      January 2010 - December 2014      January 2010 - March 2015

Notes: The text shaded in gray indicates that the AR model was chosen for the next three months of model estimates. Model types include: Ordinary least squared model (OLS), Auto-regressive model (AR), 6-month average from the previous six months of data, Repeat of the last available calculated regional STB data, and hybrid models with equal weighting for each component.

Source: U.S. Energy Information Administration