Short-Term Energy Outlook: Petroleum Refining Forecasts

Concepts, Data Sources, and Techniques
Handbook of Energy Modeling Methods

www.eia.gov
U.S. Department of Energy
Washington, DC 20585
Last updated: October 2025

The U.S. Energy Information Administration (EIA), the statistical and analytical agency within the U.S. Department of Energy (DOE), prepared this report. By law, our data, analyses, and forecasts are independent of approval by any other officer or employee of the U.S. Government. The views in this report do not represent those of DOE or any other federal agencies.

Table of Contents

1. Introduction

Crude oil refining is an essential step in the petroleum value chain that links crude oil production to the consumption of finished products and liquid fuels by consumers. In 2022, the United States refining fleet averaged 18.0 million barrels per day of crude oil processing capacity (measured as atmospheric distillate capacity). U.S. refineries are some of the most sophisticated in the world, particularly along the U.S. Gulf Coast where significant conversion capacity offered by secondary units (such as catalytic crackers, hydrocrackers, and cokers) allow refiners to produce high yields of liquid fuels such as gasoline, distillate fuel oil, and jet fuel. The timing of changes to refinery capacity, logistical complications and bottlenecks, changes in refinery margins based on crude oil and liquid fuels prices, and technical limitations based on unplanned outages and maintenance schedules are among the largest contributors to uncertainty in our refining outlook.

The *Short-Term Energy Outlook* (STEO) refining module provides forecasts for refinery inputs of several products, most notably crude oil inputs (or *refinery runs*). It also produces forecasts for refinery outputs of major liquid fuels, including motor gasoline, distillate fuel oil, and jet fuel that together account for more than 75% of total refinery output in most years. The module also includes an outlook for total atmospheric crude oil distillation capacity in the United States, based on announced changes to capacity such as expansions and closures. Our refining capacity outlook is not an attempt to predict decisions by refining firms or produce a forecast of changes in refinery capacity based on market pressure.

The refining module contains 11 estimated regression models and four identities (accounting equations). The majority of the input variables for the regression equations rely on data series produced internally within the *Short-Term Integrated Forecasting System* (STIFS), either within other modules or within the refining module itself.

2. Module Outputs

The petroleum products supply module of the STIFS provides forecasts of petroleum refinery inputs and refinery outputs (or production). The module forecasts for refinery inputs are:

- Crude oil inputs
- Unfinished oils inputs
- Pentanes plus inputs
- Liquified petroleum gas (LPG) inputs
- Motor gasoline blending component inputs
- Aviation gasoline blending component inputs
- Total refinery inputs

The module forecasts for refinery outputs are:

- Finished motor gasoline production
- Distillate fuel oil production
- Jet fuel production
- Residual fuel oil production
- Liquefied petroleum gas production

- Other petroleum products
- Total refinery outputs

Refinery yield of each output (expressed as a percentage) represents the percentage, by volume, of each refinery output from inputs of crude oil and net inputs of unfinished oils. Although we do not produce a yield forecast, an approximation of the yield may be calculated for each month by computing the ratio of net production of a product to the sum of crude oil inputs and unfinished oils inputs.

To calculate the yield for finished motor gasoline, subtract natural gas liquids inputs, other hydrocarbons and oxygenates inputs, and net inputs of motor gasoline blending components from the net production of finished motor gasoline.

3. Data Sources

The sources for monthly U.S. refinery inputs and outputs are:

- EIA Weekly Petroleum Status Report (WPSR)
- EIA *Petroleum Supply Monthly* (PSM) for preliminary monthly data
- EIA *Petroleum Supply Annual* (PSA) for revised final monthly data
- Monthly average U.S. spot market prices of motor gasoline and No. 2 diesel fuel, as well as spot market prices of kerosene-type jet fuel

We publish PSM data with a two-month lag. For example, the PSM released at the end of March contains data through January. The module uses monthly estimates based on weekly data from the WPSR to fill in data for the two most recent months. Data from the PSM series are considered final data. In contrast, we sometimes adjust STEO values from WPSR estimates based on analyst judgment. The module previously used refiner prices to resellers and end users provided by the *Petroleum Marketing Monthly* as spot market prices, but these series have been discontinued. We currently estimate equivalent prices using data provided in Tables 11, 12, and 13 of the WPSR. We expect to adjust this spot price estimation method in the future.

4. Linear Regression Models

4.1 Refinery input equations

The refinery inputs section of the petroleum products supply module contains five regression models and one identity. The regression models for refinery inputs are for crude oil, unfinished oils, pentanes plus, motor gasoline blending components, aviation gasoline blending components, and other petroleum products.

The refinery input section of the module produces forecasts of preliminary values for refinery inputs of crude oil and unfinished oils. The crude oil and unfinished oils refinery inputs are then adjusted upward or downward so that total refinery inputs plus refinery processing gain equal total refinery outputs (see the *Balancing refinery items* section). The module is not intended to capture internal refinery processes but rather to reflect the total volumes of net input and output products. All inputs and outputs reflect net values given that certain refinery processes may produce products such as unfinished oils or motor gasoline blend components that are then consumed by another process within the refinery.

4.1.1 Crude oil inputs

U.S. refinery inputs of crude oil are a positive function of current and prior-month consumption of the primary petroleum products. Crude oil inputs are also a negative function of fuel imports and other refinery inputs that may be considered substitutes in production such as unfinished oils and liquefied petroleum gas. The primary independent variables in our regression equation for crude oil inputs are:

- Domestic consumption of petroleum gasoline (defined as finished gasoline consumption minus ethanol consumption and inputs of motor gasoline blend components)
- Domestic consumption of petroleum distillate fuel (excluding biodiesel)
- Domestic consumption of jet fuel
- Net imports of total petroleum products
- Monthly dummy variables to capture seasonal effects

Refinery profitability and the incentive to process crude oil is often measured by refining margins, or the difference between product wholesale, or spot, price and the cost of crude oil. A commonly used measure of refining margins that is included in the crude oil refinery inputs equation is the 3-2-1 crack spread, which assumes 3 gallons of crude oil is processed to make 2 gallons of gasoline and 1 gallon of distillate fuel. The 3-2-1 crack spread in this model is calculated as two-thirds times the refiner price of gasoline for resale plus one-third times the refiner price of diesel fuel for resale minus the price of the refiner average acquisition cost of crude oil, all expressed in cents per gallon. That is,

$$\text{Crack Spread} = \frac{2}{3}\text{Refiner Price of Gasoline} + \frac{1}{3}\text{Refiner Price of Diesel} - \text{Cost of Crude Oil.}$$

Crack spreads derived from current prices are likely to reflect expectations for both current and future fuels production and crude oil consumption by refineries. Therefore, we do not include the 3-2-1 crack spread as an independent variable in the regression formula because it can introduce a greater risk of endogeneity than the use of fuel consumption, which we primarily determine using exogenous variables. However, consideration of current price levels, refinery margins, and the 3-2-1 crack spread are still used as analytical guidelines in assessing the outlook for crude oil inputs.

4.1.2 Unfinished oils inputs

Unfinished oils are all oils requiring further processing after undergoing all or part of the refining process. Unfinished oils include naphthas and lighter oils, kerosene and light gas oils, heavy gas oils, and residuum. Heavy gas oils and residuum represent most of the refinery inputs of unfinished oils. Refinery inputs of unfinished oils are seasonal (highest during the summer and winter months) but are also highly variable because refiners may purchase and process unfinished oils as they become available. However, unfinished oils inputs most often reflect oils produced by lower conversion refineries that are sold to and processed by more complex, higher conversion refineries that rely on available streams of unfinished oils to supplement their crude oil inputs. In recent years, the supply of unfinished oils has fallen as simpler refining capacity has closed, and less unfinished oil are available to U.S. refineries. To reflect this declining availability of unfinished oil our regression includes a trend variable that reduces the forecasted inputs of unfinished oils over time. It also includes seasonal variables to reflect the typical seasonality observed in unfinished oil inputs. The other inputs for unfinished oils are:

- Monthly dummy variables to capture seasonal effects

4.1.3 Pentanes plus

Pentanes plus is a mixture of hydrocarbons, mostly pentanes and heavier hydrocarbons (such as isopentane, natural gasoline, and plant condensate), extracted from natural gas. Pentanes plus may be blended directly into gasoline (reported in the PSM as “refinery and blender net input”) or may be blended into fuel ethanol as a denaturant (reported in the PSM as “renewable fuel and oxygenate plant net production”). Pentanes plus blended into ethanol as a denaturant is reported in the PSM as a negative value for renewable fuel and oxygenate net production. Consequently, we calculate the total volume of pentanes plus blended into gasoline as refinery and blender net input minus renewable fuel and oxygenate plant net production. We estimate pentanes plus renewable plant production as a fixed 2% of ethanol production.

We assume pentanes plus gasoline blending to be inversely related to gasoline yields. Increasing the share of gasoline produced from crude oil generally requires operating downstream processing units such as catalytic crackers at higher rates and severities. This process increases production of lighter by-products such as pentanes plus, which reduces the demand for refinery inputs of pentanes plus from other sources. Pentanes plus blended into gasoline is estimated as a function of:

- Refinery production of motor gasoline
- A lagged dependent variable
- Monthly dummy variables to capture seasonal effects

4.1.4 Motor gasoline blending components

Motor gasoline blending components (for example, straight-run gasoline, alkylate, reformate, benzene, toluene, and xylene) are used to blend or combine together to produce or augment finished motor gasoline. These components include reformulated gasoline blend stock for oxygenate blending (RBOB) but exclude oxygenates (alcohols, ethers), butane, and pentanes plus. The motor gasoline blending component *adjustment* represents the unaccounted for supply volume in the balance of blending component total supply (inputs) and disposition (refinery inputs, stock build, and exports). Motor gasoline blending component refinery inputs are a positive function of unaccounted for motor gasoline blending components, that is, additional unaccounted for supply. The independent variables in the regression model for inputs of motor gasoline blending components are:

- A lagged dependent variable
- Motor gasoline blend component adjustment
- Monthly dummy variables to capture seasonal effects

4.1.5 Aviation gasoline blending components

Aviation gasoline blending components are naphthas that will be used for blending into finished aviation gasoline (for example, straight run gasoline, alkylate, reformate, benzene, toluene, and xylene). Net refinery inputs of aviation gasoline components are generally close to zero. The independent variables for inputs of aviation gasoline blending components are:

- A lagged dependent variable
- Monthly dummy variables to capture seasonal effects

4.1.6 Total refinery input

The module calculates total refinery input as the sum of:

- Aviation gasoline blending component refinery inputs
- Crude oil refinery inputs
- Motor gasoline blending component refinery inputs
- Total hydrocarbon gas liquids (HGL) refinery inputs
- Other hydrocarbons and oxygenates refinery inputs
- Unfinished oils refinery inputs

4.2 Refinery output equations

The refinery outputs section of the module contains five regression models and one identity. The regression models are for motor gasoline, distillate, jet fuel, residual fuel, and other petroleum products. Refinery outputs are generally estimated as functions of refinery inputs and relative product prices.

4.2.1 Finished motor gasoline

Finished motor gasoline is a complex mixture of relatively volatile hydrocarbons with or without small quantities of additives, blended to form a fuel suitable for use in spark-ignition engines. Motor gasoline is characterized as having a boiling range of 122°F to 158°F at the 10% recovery point to 365°F to 374°F at the 90% recovery point. Motor gasoline includes conventional gasoline, all types of oxygenated gasoline (including gasohol), and reformulated gasoline but excludes aviation gasoline.

Refinery output of finished motor gasoline does not include some gasoline blending components and oxygenates that are reported as an “adjustment” to finished motor gasoline supply. Motor gasoline output is estimated as a function of:

- Crude oil refinery inputs
- Unfinished oils refinery inputs
- Motor gasoline blend component refinery inputs
- Other hydrocarbons refinery inputs
- The average price spread, in real dollars, between motor gasoline and distillate
- The previous month’s total gasoline stock relative to the four-year average stock level
- Monthly dummy variables to capture seasonal effects

4.2.2 Distillate fuel oil

Distillate fuel is a general classification for one of the petroleum fractions produced in conventional distillation operations. It includes diesel fuels and fuel oils. Products known as No. 1, No. 2, and No. 4 diesel fuel are used in on-highway diesel engines, such as those in trucks and automobiles, as well as off-highway engines, such as those in railroad locomotives and agricultural machinery. Products known as No. 1, No. 2, and No. 4 fuel oils are used primarily for space heating and electric power generation. Refinery output of distillate fuel is estimated as a function of:

- Crude oil refinery inputs
- Unfinished oils refinery inputs
- The average price spread, in real dollars, between motor gasoline and distillate

- The previous month's distillate fuel stock relative to the four-year average stock level
- The deviation of heating degree days in the U.S. Northeast from the 30-year monthly average
- Monthly dummy variables to capture seasonal effects

4.2.3 Jet fuel

Jet fuel is a refined petroleum product used in jet aircraft engines. It includes kerosene-type jet fuel and naphtha-type jet fuel. Jet fuel production in the United States had been rising relatively steadily from 1.4 million barrels per day in 2009 to about 1.8 million barrels per day in 2019. In 2020, jet fuel consumption declined significantly due to COVID-19-related travel restrictions. Consumption slowly rose in 2022 and 2023 but still remained below 2019 demand levels. The regression model for jet fuel refinery output is similar to the models for the other major products. Jet fuel refinery output is a positive function of refinery inputs and a negative function of the spread between gasoline or diesel fuel wholesale prices and jet fuel wholesale price. Refinery output of jet fuel is estimated as a function of:

- Crude oil refinery inputs
- Unfinished oils refinery inputs
- The average price spread, in real dollars, between motor gasoline and jet fuel
- The average price spread, in real dollars, between distillate fuel and jet fuel
- The previous month's jet fuel stock relative to the four-year average stock level
- Monthly dummy variables to capture seasonal effects

4.2.4 Residual fuel

Residual fuel oil is a general classification for the heavier oils, known as No. 5 and No. 6 fuel oils, that remain after the distillate fuel oils and lighter hydrocarbons are distilled away in refinery operations. No. 5, a residual fuel oil of medium viscosity, is also known as Navy Special and is defined in Military Specification MIL-F-859E, including Amendment 2 (NATO Symbol F-770). It is used in steam-powered vessels in government service and inshore power plants. No. 6 fuel oil includes Bunker C fuel oil and is used for the production of electric power, space heating, vessel bunkering, and various industrial purposes. The residual fuel refinery output equation is a simple function of refinery inputs. Refinery output of residual fuel is estimated as function of:

- Crude oil refinery inputs
- The previous month's residual fuel stock relative to the four-year average stock level
- Monthly dummy variables to capture seasonal effects

4.2.5 Other petroleum products

Refineries also output a range of other products, including aviation gasoline, kerosene, petrochemical feedstocks, lubricants, waxes, special naphthas, petroleum coke, asphalt, and still gas. Refinery output of other products is estimated as a function of:

- Crude oil refinery inputs
- Unfinished oils refinery inputs
- Monthly dummy variables to capture seasonal effects

4.2.6 Total refinery outputs

The module calculates total refinery output as the sum of:

- Finished motor gasoline refinery output
- Distillate fuel oil refinery output
- Jet fuel refinery output
- Total HGL refinery output
- Residual fuel refinery output
- Other petroleum products refinery output

5. Refinery Balancing Items

5.1 Refinery processing gain

Refinery processing gain is the volumetric increase that occurs when crude oil and heavy unfinished oils are cracked into lighter products in refinery secondary processing units such as cokers, catalytic crackers, and hydrocrackers. The overall mass of material that enters and leaves the refinery remains the same, however, because long-chain hydrocarbon molecules are *cracked* into smaller molecules. The density of the products decreases, and the total volume produced increases.

Refinery processing gain is estimated as a function of:

- Total distillation column inputs
- Crude oil inputs
- Monthly dummy variables to capture seasonal effects

5.2 Balancing refinery items

The separate forecasts of refinery inputs, refinery outputs, and refinery processing gain must be adjusted to maintain the volume balance that outputs equal inputs plus processing gain. If estimated refinery outputs are greater or less than the initial value of refinery inputs plus refinery processing gain, the initial values for crude oil and unfinished oils refinery inputs are proportionally adjusted upward or downward to create a balance.

5.3 Refinery capacity and utilization rate

We report refinery utilization rates as total inputs to crude oil atmospheric distillation units divided by total operable atmospheric distillation capacity, expressed as a percentage. Total inputs to crude oil distillation units are less than total refinery inputs. Some refinery inputs such as unfinished oils may be fed to processing units downstream of atmospheric distillation such as vacuum distillation units and catalytic crackers. Total distillation inputs are estimated independently using crude oil inputs, unfinished oils inputs and seasonality as independent variables. The difference between total refinery inputs (crude oil inputs plus unfinished oil inputs) and total distillation inputs is assumed to be unfinished oil inputs to other secondary refining units.

The STEO does not attempt to predict any unannounced changes in operable refining capacity outside of increases publicly announced by refiners, which we assume are reasonably likely to come online at the announced time. Refining capacity is otherwise set equal to operable capacity published in the most recent WPSR. In cases of discrepancy between figures from the WPSR and the PSM, we default to the capacity indicated in PSM. The average refinery utilization rate is equal to total inputs to crude oil atmospheric distillation units divided by total capacity.