



*Concepts, Data Sources, and Techniques*

**Handbook of Energy  
Modeling Methods**

---

# Short-Term Energy Forecasting Overview



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.

## 1. Introduction

---

The *Short-Term Energy Outlook* (STEO) is a monthly release of EIA's forecasts of energy supply, consumption, and prices through the end of the next calendar year. At the end of each year, the forecast period is extended twelve months, so the forecasts cover a 13-to-24 month horizon. STEO includes forecasts for petroleum, natural gas, coal, renewables, electricity, biofuels, and carbon dioxide emissions. Most of the forecasts in STEO are based on an underlying historical data series reported in EIA statistical data release publications. STEO also includes forecasts from non-EIA sources for macroeconomic, weather, and other variables that are important drivers of energy markets. EIA forecasts data in STEO in monthly intervals and presents the data in monthly, quarterly, and annual interval. The forecasts cover a 13-to-24 month horizon.

EIA uses the Short-Term Integrated Forecasting System (STIFS) to create the forecasts published each month in STEO. STIFS comprises two main models, programmed in the Eviews software package. However, STIFS uses the output of several other models and forecasts, the results of which are fed exogenously to STIFS. The two main models of STIFS are

- *U.S. short-term energy model (USSTEM)*
- *Global short-term oil model (GSTOM)*

USSTEM, which forecasts many more series than GSTOM, is the main engine of STIFS. Although USSTEM and GSTOM can be run independently, each includes information from the other. This document will mainly focus on the USSTEM part of STIFS, while referencing the global oil market model periodically when discussing the connections between the two models. EIA will release documentation for GSTOM in a separate report.

## 2. Historical Data Inputs

---

With few exceptions, STEO publishes forecasts of data series that are reported in other EIA historical data publications. USSTEM generally accesses historical EIA data from

- 1) Internal Oracle databases
- 2) Publicly available Microsoft Excel files on the EIA website

These data are then stored on a network drive for USSTEM to access. These historical data come mainly from 10 EIA publications:

- *Petroleum Supply Monthly*
- *Petroleum Marketing Monthly*
- *EIA-914, Monthly Production Report*
- *Weekly Petroleum Status Report*
- *Natural Gas Monthly*
- *Weekly Natural Gas Storage Report*
- *Coal Quarterly*
- *Electric Power Monthly*

- *Monthly Energy Review*
- *Hourly Grid Monitor*

In addition, USSTEM uses non-EIA data historical data from a number of sources:

- Refinitiv
- S&P Global Market Intelligence: SNL Energy Data
- IHS Markit
- Oxford Economics
- U.S. Census Bureau
- U.S. Federal Highway Administration
- U.S. Federal Aviation Administration
- U.S. Bureau of Labor Statistics
- U.S. Bureau of Economic Analysis
- National Oceanic and Atmospheric Administration (NOAA)
- North American Electric Reliability Corporation (NERC)
- U.S. Regional Transmission Organizations (RTO) and Independent System Operators (ISO)

### 3. Exogenous Models and Forecasts Used by USSTEM

---

Although most of the forecasts EIA publishes in STEO are produced endogenously within USSTEM, the model also relies on several forecasts produced by outside sources:

- U.S. macroeconomics – IHS Markit
  - For the U.S. macroeconomic forecast, EIA runs the IHS Markit model of the U.S. economy. For STEO, the only input EIA adjusts when running the IHS model is the oil prices, which are set to match EIA’s internally generated oil price forecast.
- Global macroeconomics – Oxford Economics
  - This model provides macroeconomic data used mostly by GSTOM but also read by USTEMM.
- U.S. weather – NOAA
  - EIA gets state-level heating degree day (HDD) and cooling degree day (CDD) history and forecasts from NOAA. EIA uses contemporaneous population weights to aggregate the state-level data to the regional and national levels published in STEO. For example, historical HDD for January 2010 are weighted using historical state populations from January 2010, and forecast HDD for January 2021 are weighted using forecast state populations for January 2021. So, changes in U.S. and regional HDD and CDD data represented in STEO over time reflect not only changes in temperature but also shifts in population.

### 4. USSTEM module components

---

Although EIA STEO analysts execute USSTEM by running a single Eviews program file, USSTEM comprises several sub-modules, which are automatically executed each time the main USSTEM model is run. These sub-modules generally represent the various energy sources. Each module consists of econometric

linear regression and/or identity equations that are solved simultaneously within the Eviews software. Identity equations calculate values for some series that result from simple arithmetic relationships between other variables.

Seven module components are endogenous to USSTEM:

- Electricity demand models
  - Electricity retail prices
  - Electricity retail sales
  - Electricity load
- Electricity supply models
  - Electricity generation
  - Fuel consumption by generators
  - Fuel costs for power generators
  - Wholesale electricity prices
- Renewable energy models
  - Renewable energy capacity
  - Renewable energy consumption
  - Distributed solar generation
- Natural gas models
  - Natural gas supply
  - Natural gas demand and inventories
  - Natural gas prices
- Petroleum models
  - Jet fuel
  - Residual fuel
  - Distillate fuel
  - Motor gasoline
  - Other petroleum products
  - Hydrocarbon gas liquids
  - Petroleum refinery supply
  - Petroleum product prices
- Coal model
  - Coal production, consumption, stocks, and trade
- CO2 emissions model
  - Emissions by fuel source

EIA maintains the model objects and associated equations in an Eviews database, which are read in when USSTEM is executed. Once USSTEM is executed, it solves all of the equations in the model objects simultaneously to produce the forecasts.

EIA also produces several key forecasts using models that are outside of the USSTEM framework. These forecasts are done outside of USSTEM because they require modeling techniques or data compiling complexities that do not fit neatly into the econometric structure of USSTEM. Although they are produced outside of USSTEM, these forecasts include many of the foundational variables used in USSTEM that are both key energy market variables and also important explanatory variables in regression models.

- Global oil production and consumption
  - These forecasts come from the GSTOM component of STIFS.
- Brent and WTI crude oil spot prices
  - EIA uses two models as bases for oil price forecasts. The first is a pooling model that provides a single Brent price path that is an average of five separate models that use both EIA and non-EIA oil market data. The second is a linear regression model based on inputs from the STEO global oil market forecast. Ultimately, EIA considers the output of the two models as guides and constructs a price forecast using analyst judgment.
- U.S. crude oil and natural gas production
  - Using the crude oil and Henry Hub natural gas spot price forecasts as inputs, EIA generates crude oil and natural gas production forecasts using a linear regression model in Eviews. This model forecasts regional-level production using historical well-level production based on decline rate data.
- U.S. liquefied natural gas (LNG) imports and exports
  - As the United States steadily increases its LNG export capacity, LNG exports will be a function of this capacity, which is difficult for a model to forecast. EIA analysts track export project construction and determine utilization rates for new export capacity.
- U.S. electricity generation and wholesale electricity prices
  - EIA uses LCG Consulting's UPLAN software, which is an integrated modeling tool for simulating electricity industry operations and power markets on an hourly basis. Analysts run UPLAN scenarios using data from USDEMM and data provided by LCG. USSTEM uses the model solution output from UPLAN to produce the monthly STEO electricity supply forecasts.

## 5. Solving the Model

---

When the Eviews software executes USSTEM, it first reads the historical data, exogenous forecasts, and EIA forecasts produced out of USSTEM, and then it solves the model for the endogenous forecast variables. EIEWS uses the Gauss-Seidel iterative method to solve USSTEM.

USSTEM contains nearly 600 equations to forecast variables representing different energy sources and market activities. Most of the equations are linear regression models involving historical EIA data series. Many of these series are published with a two-month lag (for example, data for January are published at the end of March). For some data series, EIA publishes weekly data, in which case USSTEM uses this weekly data to fill in the intervening months between the last monthly data point and the forecast. Often STEO analysts adjust the weekly data point using analyst judgment because these series are generally considered preliminary, but the monthly data are considered final and are never adjusted. The system aggregates data from EIA's hourly grid monitor to provide recent monthly estimates for some electricity data series.

Whether or not preliminary data are available to provide estimates of recent historical data, USSTEM will prepare modeled data estimates for any historical month for which no data exist and then execute the forecasting equations for any future month. The following are two examples of the progression from actual data to forecast in a hypothetical STEO published in April.

	Historical data availability		STEO data series			
	Last monthly data	Weekly data (yes/no)	January	February	March	April
<b>Example 1</b>	January	yes	historical monthly	weekly	weekly	forecast
<b>Example 2</b>	January	no	historical monthly	forecast	forecast	forecast

The accuracy of a forecast created using coefficients estimated from historical data depends on the relationships among the historical data series, as represented by the coefficients, holding in the future. In reality, energy markets are constantly evolving, and prior estimated relationships often lose their forecasting significance. As a result, STEO analysts regularly monitor market developments to account for events not reflected in the historical data. In addition, exogenous factors outside of available data sets (for example, regulatory changes) are often known and can affect forecasts. Many equations in USSTEM have associated adjustment factors (*add factors*) that STEO analysts use to reflect expected energy market changes that may not be adequately reflected by the current structure of the model equations. Add factors are stored in an Eviews database file and read in when USSTEM is executed.

As a general practice, EIA analysts who work on STEO reestimate the forecasting equations once a year. However, some equations are reestimated more frequently and some less frequently.<sup>1</sup> In the reestimation process, analysts consider whether the set of independent variables is still appropriate from a standpoint of both description and statistical significance. Analysts also consider the estimation period to ensure it captures a period for which the historical data relationships observed are relevant for the current forecast period.

<sup>1</sup> For details on recommended updating frequencies for STEO regression models, see “Estimating Re-fitting Frequencies for Short-term Energy Models” by Lent, J. and George, R. (2019), available from EIA on request.

# Model diagram

