Concepts, Data Sources, and Techniques

Handbook of Energy Modeling Methods

Short-Term Energy Outlook: Natural Gas Module

Table of Contents

Table of Figures

Table of Tables

1. Introduction

From 2000 to 2021, U.S. natural gas consumption grew by more than 19 billion cubic feet per day (Bcf/d) to 83 Bcf/d (Figure 1). The largest consumers of natural gas in the United States are the electric power and industrial sectors. Consumption of natural gas in the electric power sector has grown the most during this time period, increasing from 22% of total natural gas consumption in 2000 to 37% in 2021. As of 2021, the industrial sector accounted for 27% of total natural gas consumption, and the residential and commercial sectors combined accounted for 26%. The remaining 10% of natural gas was consumed as vehicle fuel, as lease and plant fuel, and in pipelines and distribution.

Figure 1. Natural gas delivered to consumers in the United States (2000–2021)
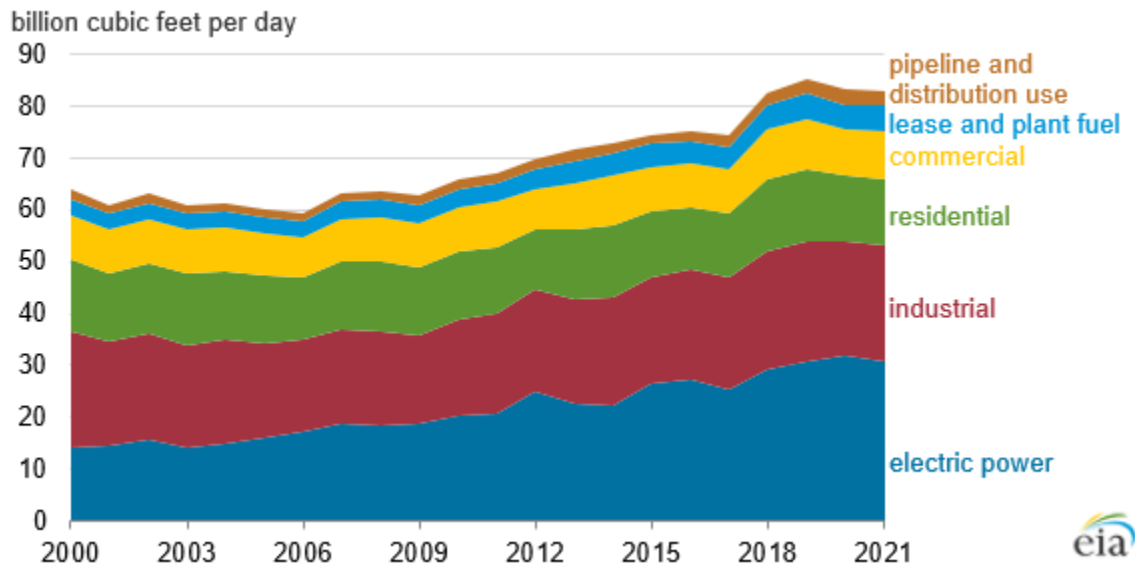
Data source: U.S. Energy Information Administration, *Natural Gas Annual*, Table 15, August 2022

Exports of natural gas are also growing in the United States. In 2021, the United States exported 18.2 Bcf/d of natural gas: 46% as pipeline exports and 54% as liquefied natural gas (LNG) exports. U.S. exports of natural gas have more than quadrupled since 2010, driven by increases in pipeline exports to Mexico and the construction of LNG export facilities. Of the natural gas that the United States exported by pipeline in 2021, 70% went to Mexico, and 30% went to Canada. U.S. LNG exports went largely to Asia in 2021. Exports to Europe, however, have increased significantly in 2022.

Figure 2. U.S. natural gas exports and imports (January 2018–May 2022)
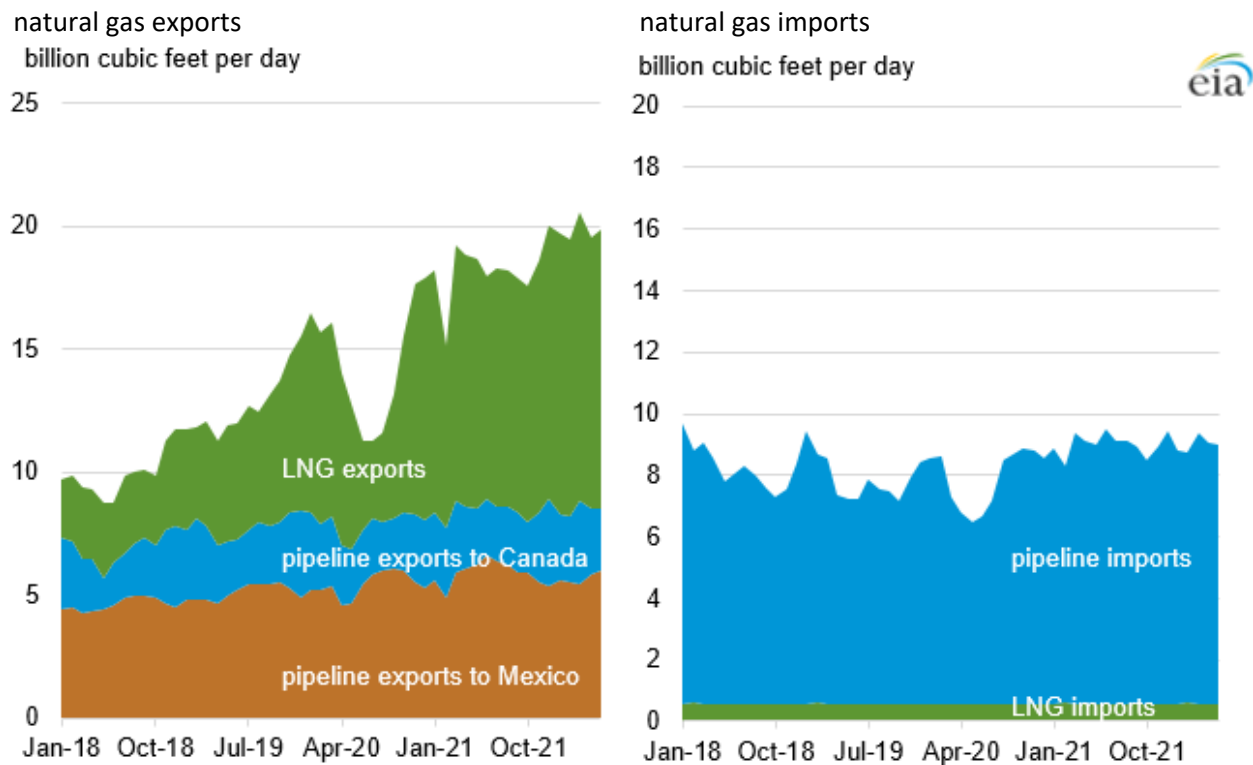
Data source: U.S. Department of Energy, Office of Fossil Energy, *Natural Gas Imports and Exports,* August 2021

Although most of the natural gas consumed in the United States is produced domestically, the United States also imports natural gas. In 2021, the United States imported 7.7 Bcf/d of natural gas, 99% of which was imported by pipeline (Figure 2). Almost all pipeline imports of natural gas come from Canada, and a small amount comes from Mexico. Overall, the United States exports more natural gas than it imports.

Natural gas inventory withdrawals provide the second-largest source of supply (after field production) during the winter heating months (October through March). During the peak heating-demand months (December, January, and February), natural gas inventory withdrawals account for about 20% of total supply. Because of the high demand for natural gas during the winter heating season and relatively lower demand in April through September, working natural gas storage injections and withdrawals follow a seasonal pattern (Figure 3).

Inventory levels reported in our *Short-Term Energy Outlook* (STEO) represent end-of-month inventories. We refer to April through October as the *injection season* and November through March as the *withdrawal season*. Working natural gas in storage typically reaches its peak at the end of the injection season (October), although depending on weather and market conditions, injections often continue on a weekly basis into early November. Working natural gas in storage typically falls to its low point at the end of the withdrawal season (March).

Figure 3. U.S. underground storage net injections and withdrawals (January 2018–May 2022)
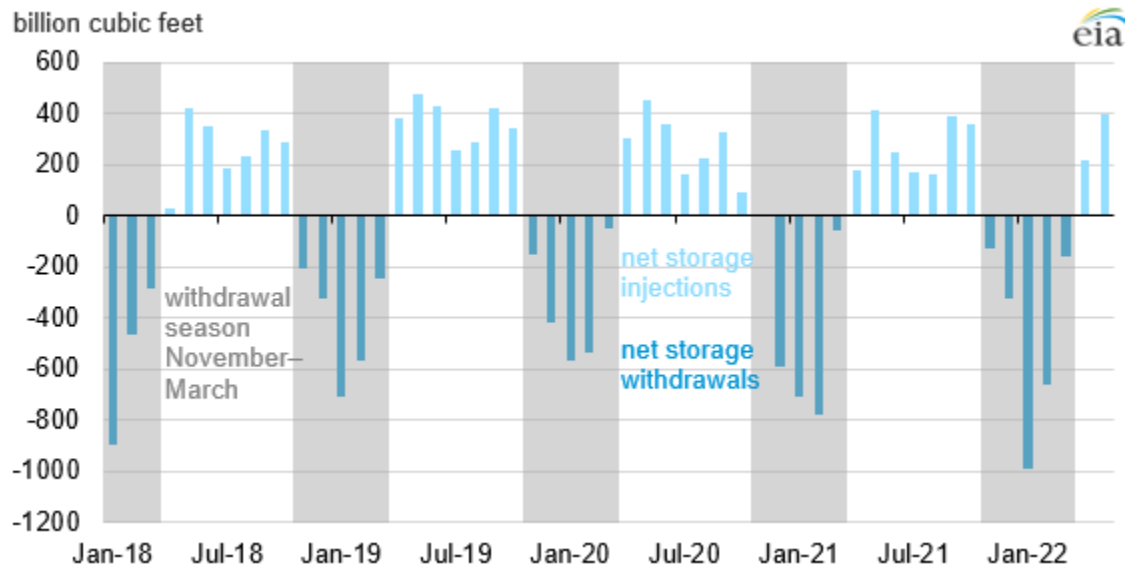
Data source: U.S. Energy Information Administration, *Monthly Underground Gas Storage Report*, August 2022

The Henry Hub is a pipeline hub in Louisiana that serves as the delivery point for the natural gas futures contracts on the New York Mercantile Exchange (NYMEX) and is used as the U.S. benchmark natural gas price. Excluding taxes and fees, the price that consumers pay for natural gas is made up of the commodity cost and the transmission and distribution costs. The commodity cost is the cost of the natural gas itself, and the transmission and distribution costs include the cost of moving the natural gas from where it was produced to storage and the cost to deliver it to customers.

Figure 4. Natural gas prices by sector (January 2018–May 2022)
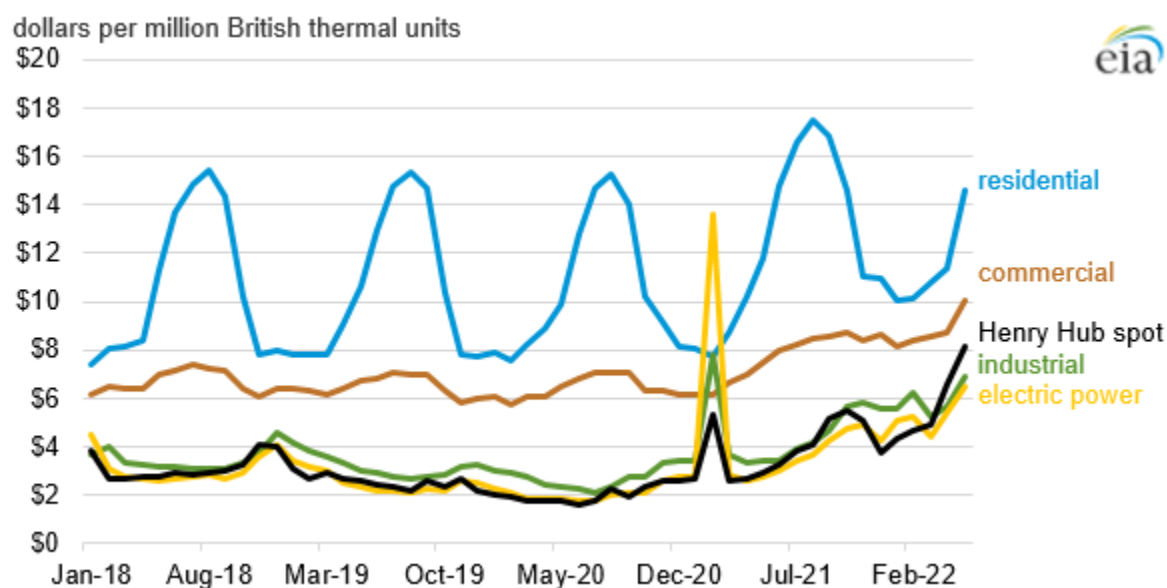
Data source: U.S. Energy Information Administration, *Natural Gas Monthly,* August 2022

The STEO Natural Gas Module forecasts:

- Natural gas consumption of by sector
- End-of-period natural gas inventories by storage region
- Natural gas trade (imports and exports)
- Natural gas prices by sector and census region

The Natural Gas Module contains 173 equations, and 66 of those equations are linear regression models. The remaining 107 equations are balancing equations, used to ensure that forecasts add up; for example, the sum of natural gas consumption across sectors must equal total natural gas consumption. Independent variables in the linear regression models typically include items such as heating and cooling degree days, employment, and number of households that use natural gas as a primary source for space heating. Some input variables to the Natural Gas Module come from other STEO modules, such as natural gas consumption by the electric power sector, which comes from the Electric Supply Module. Other independent variables come from forecasts produced by other organizations, such as weather forecasts from the National Oceanic and Atmospheric Administration (NOAA).

2. Natural Gas Module Outputs

The STEO Natural Gas Module forecasts data series for total United States and for U.S. regions, including:

- Natural gas consumption by end-use sector
    - Residential natural gas consumption
    - Commercial natural gas consumption
    - Industrial natural gas consumption

- o Other natural gas consumption
    - Lease and fuel plant use (of natural gas)
    - Pipeline and distribution use (of natural gas)
- End-of-period working natural gas inventories by storage region (Figure 5)
    - o Net natural gas inventory withdrawals
    - o Working gas inventory
- Natural gas pipeline gross imports and exports
- Retail prices of natural gas by end-use sector and census division (Figure 6)
    - o Residential sector prices
    - o Commercial sector prices
    - o Industrial sector prices

Figure 5. EIA natural gas storage regions
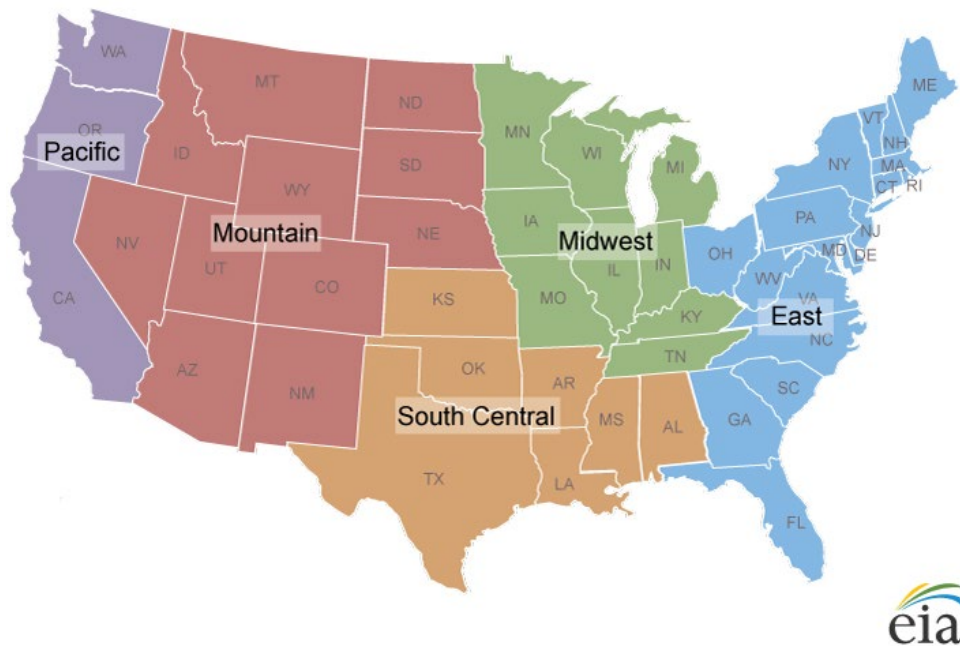
Data source: U.S. Energy Information Administration
Note: Alaska is considered its own storage region

Figure 6. U.S. census regions and divisions
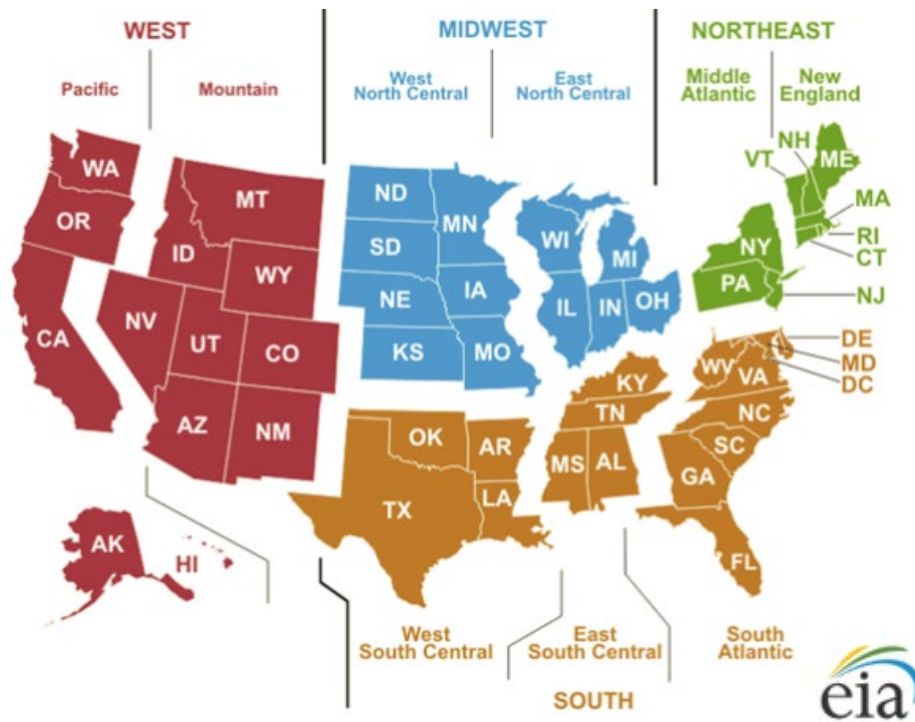
Data source: U.S. Energy Information Administration, <u>Regional Dashboards and Tables</u>

3. Data Sources and Aggregations

Data in the Natural Gas Module primarily comes from our *Natural Gas Monthly* (NGM). The NGM provides monthly preliminary historical data for U.S. natural gas consumption, end-use prices, and inventories, which we later update to final monthly data in our *Natural Gas Annual* (NGA).

Each month, we obtain model input data from several EIA publications and external resources (Table 1).

Table 1. Sources of model input data in the Natural Gas Module

Input data	Source
Natural gas consumption in the residential, commercial, and industrial sectors and for vehicle use	*Natural Gas Monthly* (NGM) table, "Natural Gas Consumption by End Use"
Number of households that use natural gas as their primary space heating fuel	American Community Survey table, "Physical Housing Characteristics for Occupied Housing Units" (S2504)
Commercial sector employment	S&P Global Macroeconomic model
Industrial subsector shares of natural gas consumption	*Manufacturing Energy Consumption Survey* (MECS), Table 5.1
Natural gas use by industrial consumers (as a feedstock and for heating)	MECS, Table 2.1
Combined-heat-and-power use	Electric Power Monthly (EPM), Table 2.4.A

Natural gas-weighted industrial production index	Growth rates of individual manufacturing subsectors (sourced from S&P Global) and their respective shares of the total natural gas consumed by all manufacturing subsectors
Heating and cooling degree day histories	National Oceanic and Atmospheric Administration (NOAA), via the STEO Weather model
Natural gas inventories	*Weekly Natural Gas Storage Report* (WNGSR) and the NGM, "Underground Natural Gas Storage by All Operators"
Pipeline and distribution use and lease and plant fuel use	NGM table, "Natural Gas Consumption by End Use"
Pipeline and liquefied natural gas (LNG) trade	NGM tables, "U.S. Natural Gas Imports by Country" and "U.S. Natural Gas Exports and Re-Exports by Country"
Natural gas end-use prices	NGM table, "Natural Gas Prices"
Henry Hub spot price history	Refinitiv via EIA

Data source: U.S. Energy Information Administration

In addition, several key variables are exogenous inputs into the Natural Gas Module, including:

- U.S. natural gas production in the Lower 48 states (from the Crude Oil and Natural Gas Production Module)
- U.S. natural gas consumption in the electric power sector (from the Electricity Supply Module)
- U.S. exports and imports of LNG (based on analyst judgment)
- Henry Hub spot price forecast
- S&P Global U.S. economic forecast
- NOAA weather forecast

We aggregate state-level end-of-period working natural gas inventory data to the *storage region* level for the inventory equations. The storage regions aggregate to the U.S. total including Alaska (Figure 5 and Table 2).

Table 2. EIA natural gas storage regions

Region	State or district	
East region	Connecticut	New York
	Delaware	North Carolina
	District of Columbia	Ohio
	Florida	Pennsylvania
	Georgia	Rhode Island
	Maine	South Carolina
	Maryland	Vermont
	Massachusetts	Virginia
	New Hampshire	West Virginia
	New Jersey	
Midwest region	Illinois	Minnesota
	Indiana	Missouri

	Iowa	Tennessee
	Kentucky	Wisconsin
	Michigan	
South Central region	Alabama	Mississippi
	Arkansas	Oklahoma
	Kansas	Texas
	Louisiana	
Mountain region	Arizona	New Mexico
	Colorado	North Dakota
	Idaho	South Dakota
	Montana	Utah
	Nevada	Wyoming
	Nebraska	
Pacific region	California	Washington
	Oregon	

Data source: U.S. Energy Information Administration
Note: Alaska is consider its own storage region

We aggregate retail end-use natural gas price data for the end-use sectoral pricing equations. The regions sum to the U.S. total and correspond to the U.S. census regions (Figure 6 and Table 3).

Table 3. U.S. census regions

U.S. census region	State or district	
Midwest	Illinois	Missouri
	Indiana	Nebraska
	Iowa	North Dakota
	Kansas	Ohio
	Michigan	South Dakota
	Minnesota	Wisconsin
Northeast	Connecticut	New York
	Maine	Pennsylvania
	Massachusetts	Rhode Island
	New Hampshire	Vermont
	New Jersey	
South	Alabama	Mississippi
	Arkansas	North Carolina
	Delaware	Oklahoma
	District of Columbia	South Carolina
	Florida	Tennessee
	Georgia	Texas
	Kentucky	Virginia

	Louisiana	West Virginia
	Maryland	
West	Alaska	Nevada
	Arizona	New Mexico
	California	Oregon
	Colorado	Utah
	Hawaii	Washington
	Idaho	Wyoming
	Montana	

Data source: U.S. Census Bureau

4. Linear Regression Models

The STEO Natural Gas Module uses linear regressions to estimate coefficients that represent relationships between important variables, for example, supply, demand, and prices. To achieve satisfactory model performance many of the regression models described in this chapter also include four types of independent variables:

- Monthly dummy variables to capture seasonal effects
- Dummy variables for months before or after a certain date to pick up significant shifts in market conditions
- A one-month lagged dependent variable to capture the effect of the previous month's dependent variable on the current month's dependent variable and to reduce autocorrelation in the regression model residuals
- Month-year specific dummy variables (sometimes called *intervention effects*) to reduce the influence of anomalous monthly data observations caused, for example, by survey errors or extreme weather

4.1. Natural gas consumption

Total U.S. natural gas consumption in the STEO model is the sum of end-use sector consumption estimates. The end-use sectors included in total U.S. natural gas consumption are residential, commercial, industrial, lease and plant fuel, pipeline and distribution use, and vehicle use. End-use natural gas consumption is driven primarily by the deviation of monthly heating degree days (HDDs) from normal HDDs; *normal HDDs* are the average monthly HDDs over the past 30 years. HDDs measure how cold the temperature was relative to a reference temperature (65°F), and cooling degree days (CDDs) measure how warm the temperature was relative to the reference temperature. Monthly HDD and CDD variables are sums of the daily measures.

If HDDs increase, then consumption of natural gas increases, and vice versa, because colder weather in the winter leads to an increase in heating demand for natural gas. Residential natural gas consumption reflects the share of households that use natural gas as a primary source of space heating, and commercial and industrial end use reflect economic activity. If the natural gas price increases, the

industrial activity associated with converting natural gas into other products, such as ammonia, will decrease, and vice versa.

To capture the primary effects driving natural gas consumption, the linear regression models for end-use consumption include combinations of six monthly independent variables:

- U.S. HDD deviation from normal (degree days)
- Measures of natural gas as a primary source of space heating (number of households in thousands)
- Commercial sector employment as a proxy for commercial floor space (thousand employees)
- Combined-heat-and-power use, which represents onsite electricity generation at industrial facilities (Bcf/d)
- Natural gas-weighted industrial production index (index; year 2017=100)
- Price of natural gas to the industrial sector (dollars per million British thermal units ($/MMBtu))

We calculate a region's average deviation of HDDs or CDDs from the seasonal norm for a given month by taking the difference between the monthly value for regional HDDs or CDDs and the corresponding 30-year average and dividing this difference by the number of days in the month.

Residential sector

We model residential end-use consumption of natural gas as a function of these variables:

- U.S. HDD deviation from normal, weighted by the number of residential natural gas customers
- Number of households that use natural gas as a primary source of space heating
- Monthly seasonal dummy variables
- A lagged dependent variable

About half of the homes in the United States use natural gas for space and water heating. As a result, weather (measured by HDDs) is a main driver of residential consumption.

Commercial sector

We model commercial end-use consumption of natural gas as a function of:

- U.S. HDD deviation from normal weighted by commercial natural gas customers
- Commercial sector employment (proxy for commercial floor space)
- Monthly seasonal dummy variables
- A lagged dependent variable

In the commercial sector, natural gas is used for space and water heating, refrigeration and cooling, cooking, drying clothes, and lighting. Similar to the residential sector, weather is a main driver of commercial consumption.

Industrial sector

We model industrial end-use consumption of natural gas as a function of:

- Natural gas consumption by industrial consumers as a feedstock and for heating

- Natural gas-weighted industrial production index (based on manufacturing indexes from S&P Global) and our own weighting methodology (based on our *Manufacturing Energy Consumption Survey*)
- Price of natural gas to the industrial sector
- U.S. HDD deviations from normal
- Combined heat- and- power use
- Monthly seasonal dummy variables
- A dummy variable for each month from January 2016 through December 2021

The industrial sector uses natural gas as a fuel for process heating; a fuel for combined-heat-and-power systems (represents onsite electricity generation at industrial facilities); and a feedstock to produce chemicals, fertilizer, and hydrogen. The forecast of the industrial sector's natural gas consumption for combined heat and power is determined in the STEO Electricity Supply Module.

Lease and plant fuel
We model lease and plant fuel end-use consumption of natural gas based on the assumption that the last published monthly estimate of lease and plant fuel use as a share of marketed natural gas production continues over the forecast period.

Pipeline and distribution consumption
We model pipeline and distribution end-use consumption of natural gas based on the assumption that the last published monthly estimate of pipeline and distribution use as a share of natural gas deliveries, including LNG feed gas, continues over the forecast period.

Vehicle consumption
We assume vehicle end-use consumption of natural gas is equal to the last published monthly estimate in the NGM, "Natural Gas Consumption by End Use."

4.2. Natural gas inventories
We forecast monthly working natural gas inventory changes by region, which are summed to forecast U.S. net inventory withdrawals and total U.S. inventory level. The change in working natural gas inventories is mostly driven by weather, both hot and cold. When temperatures are high during summer months, the electric power sector uses natural gas to meet air-conditioning demand, potentially offsetting natural gas storage injections or causing withdrawals, in some instances. When temperatures are low during the winter, the residential and commercial sectors use natural gas for space heating, potentially accelerating withdrawals.

We model storage withdrawals and builds by storage region as a function of:

- Regional HDD deviation from normal
- Regional CDD deviation from normal
- Deviation in the region's end-of-month inventory level from the prior seven-year average for the prior month (for example, in August 2022, this variable would measure the deviation in inventory level from July 2022 compared with July 2015–2021)

- Intervention effect dummy variables, specific to each region

Working gas inventory is the quantity of natural gas in the reservoir that exceeds the base gas volume. Base gas inventory is the quantity of natural gas needed to maintain adequate reservoir pressures and deliverability rates throughout the withdrawal season. Base gas typically is not withdrawn and remains in the reservoir.

4.3. Natural gas trade

Imports and exports by pipeline vary with the level of heating demand, based on HDDs, as well as recent trends in volumes shipped. Imports by pipeline enter the United States primarily from Canada, and imports increase if HDDs in the Midwest (the highest natural gas-largest consuming region from Canada) increase, and vice versa. Exports by pipeline also increase or decrease based on the deviation of U.S. HDDs from normal because we assume these deviations impact natural gas consumption in Mexico and Canada. We also assume imports and exports change over time based on energy policies and trends in Mexico and Canada. The forecast for U.S. LNG exports is driven by:

- Current and new LNG export facility capacity, including assumptions about ramp-up periods for new capacity
- Assumptions about utilization of LNG export capacity in the context of global natural gas market conditions
- Seasonal maintenance periods

We model pipeline imports as a function of:

- HDD deviations from normal in the Midwest
- A dummy variable for each month from September 2014 through December 2021
- A dummy variable for each month from January 2016 through December 2021
- A dummy variable for each month from January 2018 through December 2021
- A lagged dependent variable

We model pipeline exports as a function of:

- U.S. HDD deviation from normal for the winter heating months
- A linear trend variable
- Intervention effect dummy variables for August 2014, March 2018, and February 2020
- A dummy variable for each month from January 2016 through December 2021
- A lagged dependent variable

4.4. Natural gas prices

Generally, prices and supply have a negative relationship, meaning an increase in supply will result in a decrease in prices, and vice versa. Prices and demand generally have a direct relationship, meaning that an increase in demand will result in an increase in prices. The degree of shortfall or surplus of working natural gas storage compared with average storage levels for a particular time of year affects end users' and commodity traders' perceptions of how constrained the market may be during the withdrawal and refill seasons.

To capture the primary effects driving price formation, the linear regression models for national and regional prices include combinations of two monthly independent variables:

- U.S. HDD and CDD deviation from normal (degree days)
- Deviation in working gas inventory from the previous three-year average (Bcf)

Henry Hub spot price

We develop our Henry Hub spot price forecast each month by an analyst-driven iterative process that takes into account

- The NYMEX futures price
- The previous month's forecast
- Recent market developments
- Several exogenous price models

Each month, we run a number of price forecasts through our models, iterating until we find a balance between supply and demand. The two modules most sensitive to the Henry Hub price forecast are the Crude Oil and Natural Gas Production Module and the Electricity Supply Module. Because these two modules do not interact with each other, we use the iterative process for forecasting price to balance supply and demand.

End-use sector prices

End-use sector prices include all fees and taxes and represent the per-unit cost to the consumer. Rather than forecast absolute prices, the Natural Gas Module forecasts the premium or discount of the regional prices to the Henry Hub spot price. We forecast prices at the U.S. census division level for each sector. The U.S. average natural gas price for each sector is a volume-weighted average of the regional prices.

For the residential and commercial sectors, we model the regional price differentials to the Henry Hub spot price as a function of:

- Change in Henry Hub spot price from the previous month
- Deviation in working gas inventory from the previous three-year average
- Regional HDD deviation from normal in the current month and in the previous month
- Monthly seasonal dummy variables
- A lagged dependent variable

For the industrial sector, we model regional price differentials to the Henry Hub spot price as a function of:

- Change in Henry Hub spot price from the previous month
- Deviation in working gas inventory from the previous three-year average
- Regional HDD deviation from normal
- Monthly seasonal dummy variables
- A lagged dependent variable

Industrial prices have lower associated distribution costs, are much lower than residential and commercial prices, and more closely follow the Henry Hub spot price (Figure 4). Although the same factors that drive residential and commercial prices drive the industrial price, the spot price plays a larger role, and the spread between the industrial end-use price and the Henry Hub price is less influenced by the weather.

5. Natural Gas Balance

The amount of natural gas supplied and the sum of the components of natural gas disposition are typically not exactly equal. The difference between the two may be due to quantities lost during transmission, distribution, or processing or to the effects of data reporting problems. We measure this difference by the natural gas balancing item in the STEO. We calculate this discrepancy as:

Balancing item = Total consumption – Dry natural gas production – Net imports
– Inventory withdrawals – Production of supplemental gaseous fuels

In the forecast period, we generally assume natural gas supply and demand to be in equilibrium and the balancing item to be zero. In practice, however, markets do not operate at peak efficiency, and therefore, a non-zero balancing item is common. Typically, the balancing item is less than 2% of total natural gas consumption.

We adjust inventory withdrawals to resolve the difference between supply and demand as much as reasonably possible, based on analyst judgment. We allocate these adjustments across the regions with the largest amount of storage inventories: the East, Midwest, and South Central storage regions.