Independent Statistics & Analysis
U.S. Energy Information Administration

Natural Gas Market Module of the National Energy Modeling System: Model Documentation 2020

May 2020

Independent Statistics & Analysis
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 Department of Energy or other Federal agencies.

Table of Contents

1. Introduction 1
Documentation purpose and scope 1
Model Summary 1
Documentation organization 3
Model archival citation 3
2. Model purpose 4
Model objectives 4
Reflect current and future natural gas market 4
Representation of the natural gas market in the NEMS 8
Relation to other modules 11
Data transfer 11
3. Natural Gas Market Module design and structure 17
Model design 17
Relationship between calendar dates and model periods: Time horizon 20
Mapping parameters 21
Solver 21
Model structure 22
Identifiers in QP 22
Transfer data into QP parameters 25
Piecewise linear curve definitions 26
Objective function 30
Constraints 31
Additional forms of mathematical program 34
Capacity expansion 34
STEO benchmarking 36
4. Pre-processing routines 40
First year initialization 40
Supply 41
Nonassociated and associated-dissolved natural gas production 41

Supplemental Supplies ... 43
LNG imports ... 44
Demand ... 44
End-use consumption ... 44
Lease and plant fuel ... 45
Pipeline fuel ... 45
Balancing item ... 48
Storage ... 49
Canada ... 51
Mexico ... 52
Mexican supply ... 52
Mexican demand ... 54
LNG exports ... 56
World LNG prices ... 56
Price of U.S. LNG ... 58
Net present value of LNG export capacity ... 61
Alaska ... 62
5. Post-processing routines ... 65
Production and supply prices ... 65
LNG exports ... 65
Imports and exports ... 66
Pipeline import and export volumes ... 67
Import and export prices ... 68
Delivered end-use prices ... 68
Spot prices ... 68
Citygate prices ... 68
End-use natural gas prices ... 70
Reporting to the NEMS ... 78
Appendix A. Model abstract ... 79
Appendix B. References ... 84
Appendix C. Documentation variables mapped to model identifiers ... 85
Appendix D. Documentation equations mapped to procedures in the NGMM code ... 90

Appendix E. Model variables mapped to input files 91
Appendix F. Global data transferred between the NEMS and the NGMM 100
Appendix G. Documentation of estimations 105
G-1. Mexico crude oil production 106
G-2. Mexico associated-dissolved natural gas production 108
G-3. Mexico non-associated natural gas production 110
G-4. Mexico industrial natural gas consumption 113
G-5. Alaska residential and commercial sector consumption and Alaska citygate price 116
G-6. Alaska lease and plant fuel consumption 120
G-7. Lower 48 states natural gas citygate prices by month 122
G-8. Distributor tariff markup for delivered prices to the residential sector by Census division 129
G-9. Distributor tariff markup for delivered prices to the commercial sector by Census division 138
G-10. Historical delivered end use prices to the industrial sector 147

Tables

Table 3.1 Time and horizon parameters set in the NGMM 21
Table 3.2 Set names and indexes representing the dimensions of the NGMM quadratic program 22
Table 3.3 Parameter names, abbreviations used in NGMM documentation, dimensionality, and descriptions 23
Table 3.4 Decision variable names, abbreviations used in NGMM documentation, dimensionality, and descriptions 25
Table 3.5 NGMM quadratic program parameters and corresponding pre-processing parameters 25
Table 3.6 STEO factors calculated in the NGMM during STEO benchmarking 37
Table 4.1 Mapping offshore regions to Lower 48 states 42

Figures

Figure 1.1 NGMM natural gas pipeline network representation 2
Figure 2.1 Schematic of upstream, midstream and downstream industries in the natural gas market 5
Figure 2.2 Schematic of the NEMS and flow of information between modules 9
Figure 2.3 Inputs and outputs of the NGMM, including relationships between other NEMS modules 13
Figure 2.4 NEMS Oil and Gas Supply Regions and corresponding tight oil and shale gas regions 14
Figure 2.5 Natural Gas (NG) regions used to report regional flows and capacity 15
Figure 2.6 NGMM demand regions (Census) 16
Figure 2.7 Natural Gas-Electricity Market Module (NGEMM) regions 16
Figure 3.1 Schematic price-quantity graph illustrating maximized producer-consumer surplus 17
Figure 3.2 Process flow diagram representing the NGMM and its operation within NEMS in a given year 18
Figure 3.3 Relationship between time in the NEMS, time in the NGMM, and the NGMM time horizon 20
Figure 3.4 Schematic representation of short-term supply curve used in NGMM 27
Figure 3.5 Schematic representation of tariff curve structure used in NGMM 29
Figure 3.6 Schematic representation of LNG export demand curve used in NGMM 30
Figure 3.7 Flow diagram of capacity assignments and capacity expansion quadratic program 35
Figure 3.8 Flow diagram describing the assignment of STEO benchmarking factors in the NGMM 39
Figure 4.1 Flow chart of first year initialization 40
Figure 4.2 Mexico market regions in NGMM 52

Abbreviations

AD	Associated-dissolved natural gas production
AIMMS	Advanced Integrated Multidimensional Modeling Software
AEO	Annual Energy Outlook
Bcf	Billion cubic feet
Bcf/d	Billion cubic feet per day
Btu	British thermal unit
CDM	Commercial Demand Module
CNG	Compressed natural gas
EIA	Energy Information Administration
EMM	Electricity Market Module
IDM	Industrial Demand Module
IEM	International Energy Module
IEO	International Energy Outlook
LDC	Local distribution company
LFMM	Liquid Fuels Market Module
LNG	Liquefied natural gas
MAM	Macroeconomic Activity Module
Mcf	Thousand cubic feet
MMBtu	Million British thermal units
MMcf	Million cubic feet
MMcf/d	Million cubic feet per day
NA	Nonassociated natural gas production
NEB	National Energy Board (Canada)
NEMS	National Energy Modeling System
NG	Natural Gas (regions)
NGEMM	Natural Gas-Electricity Market Module (regions)
NGMM	Natural Gas Market Module
NGTDM	Natural Gas Transmission and Distribution Module
OGSM	Oil and Gas Supply Module
QP	Quadratic program
RDM	Residential Demand Module
SENER	Secretaría de Energía de México
SNG	Synthetic natural gas
STEO	Short Term Energy Outlook
Tcf	Trillion cubic feet
TDM	Transportation Demand Module

1. Introduction

The Natural Gas Market Module (NGMM) is the component of the National Energy Modeling System (NEMS) that is used to represent the North American natural gas transmission and distribution system. The NEMS was developed by the U.S. Energy Information Administration (EIA) and is the third in a series of computer-based energy-economy modeling systems used since 1974 by EIA and its predecessor, the Federal Energy Administration, to analyze and project U.S. domestic markets. The NEMS was designed to provide 25-30 year projections and permit the analysis of a broad range of energy issues at both national and regional levels. While the NEMS was first used in 1992, the model is updated each year; updates in individual modules range from simple historical data updates to complete replacements of submodules. The NGMM is an entirely new model incorporated into the NEMS for the *Annual Energy Outlook 2018*, replacing the Natural Gas Transmission and Distribution Module (NGTDM).

Documentation purpose and scope

The purpose of this report is to provide a reference document for model analysts, users, and the public that defines the objectives of the Natural Gas Market Module (NGMM) in the NEMS. It is intended to fulfill the legal obligation of EIA to provide adequate documentation in support of its models under Public Law 93-275, Federal Energy Administration Act of 1974, Section 57(B)(1) (as amended by Public Law 94-385, Energy Conservation and Production Act).

The report describes NGMM's basic design, provides detail on the methodology employed, and details the model inputs, outputs, and key assumptions. Since the NGMM was first incorporated into the NEMS for the *Annual Energy Outlook 2018*, the documentation also describes the decision to build a new model in the NEMS to represent natural gas markets and the differences between the NGMM and its predecessor.

In addition, this report also serves as a reference document for how the NGMM utilizes Advanced Integrated Multidimensional Modeling Software (AIMMS),[1] as well as AIMMS best practices for use in the NEMS. The NGMM is the second module (after the Coal Market Module) developed and implemented in the NEMS using the AIMMS modeling language and user interface, and EIA expects to develop all future optimization models in AIMMS.[2] Therefore, the documentation report uses AIMMS terminology to provide detailed descriptions of the most efficient, flexible, and transparent techniques and methods employed in the NGMM. This is particularly important for ensuring the reproducibility of results given the complexity of NEMS runs and the exchange of data between the NEMS Fortran code, the NGMM AIMMS code, and various external files.

Model Summary

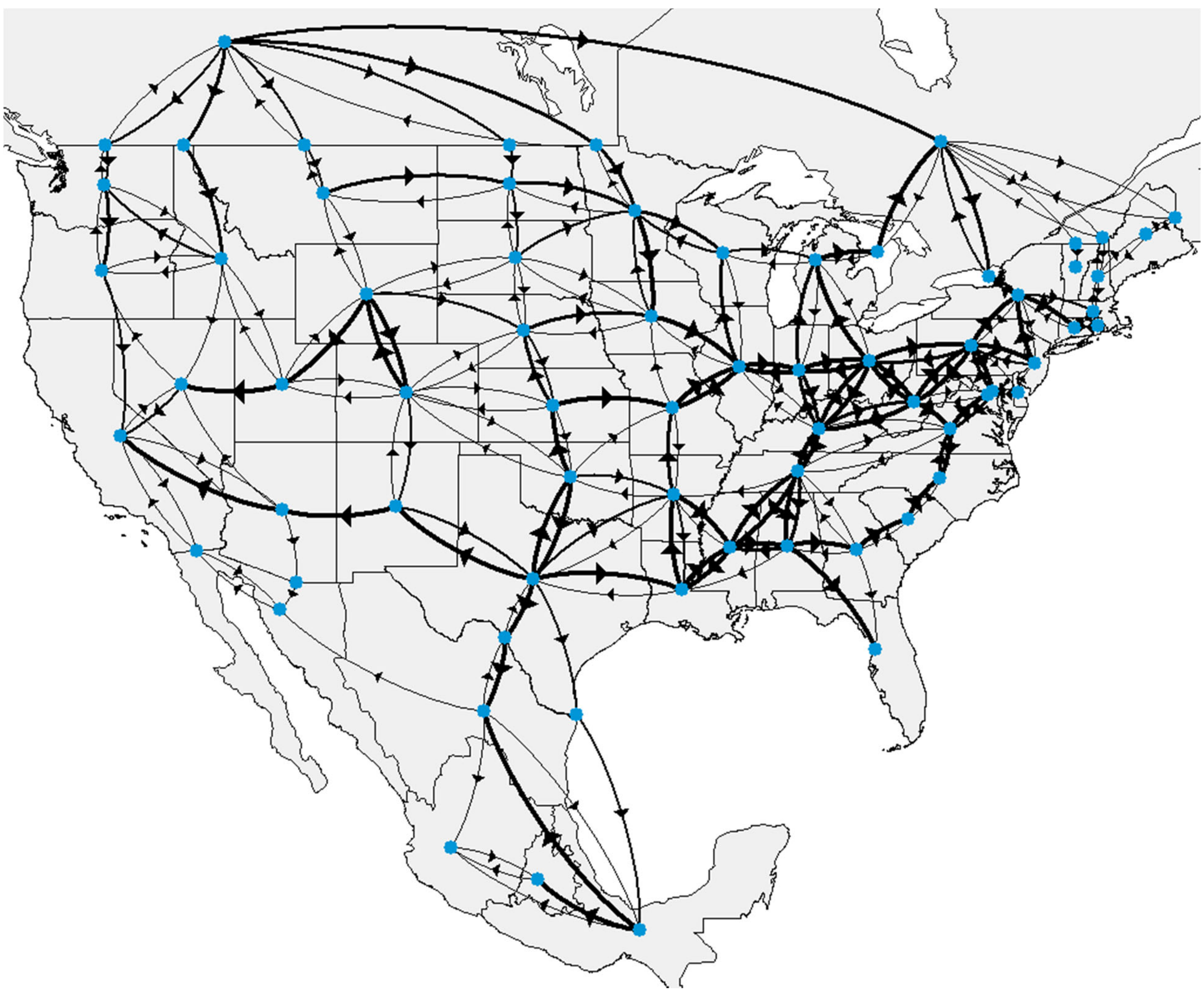
The NGMM models the transmission, distribution, and pricing of natural gas in the NEMS. The model code is written in AIMMS and is a quadratic program that maximizes consumer plus producer surplus

[1] *AIMMS Development Environment* is software that integrates the AIMMS mathematical modeling language, a graphical user interface, and numerical solvers. It is used to design and build optimization models and includes diagnostic tools as well as the ability to construct graphical reports of model results. Available AIMMS documentation includes *AIMMS—The Language Reference* and *AIMMS—The User's Guide*.

[2] AIMMS, *The U.S. Department of Energy expands its use of AIMMS for its NEMS Electricity Market Module*

minus transportation costs, subject to linear mass balance and capacity constraints. For all months in a year, the NGMM determines the production, flows, and prices of natural gas within a state-level representation of the U.S. pipeline network[3] and a regional representation of the Canadian and Mexican pipeline network (Figure 1.1[4]), connecting domestic and foreign supply regions with demand regions. End-use natural gas consumption by sector, storage, and liquefied natural gas (LNG) export terminals are all integrated into the network by demand region. The NGMM projects lease fuel, plant fuel, pipeline fuel, fuel used for liquefaction, LNG export capacity builds, and pipeline capacity expansions. Distributor tariffs are also projected to arrive at the delivered price of natural gas to domestic consumers. Since most other NEMS modules operate on an annual basis, NGMM results are generally passed to other NEMS modules as annual totals or quantity-weighted annual average prices.

Figure 1.1 NGMM natural gas pipeline network representation



[3] The Alaskan natural gas market is modeled in the NGMM independent of the integrated network.

[4] Blue circles represent transshipment nodes. Arcs represent pipeline capacity existing between nodes in 2020.

Documentation organization

The document is intended to provide a framework for understanding how the natural gas market is represented in EIA's long-term U.S. energy market projections. Subsequent chapters of this report provide:

- *Overview of natural gas market representation in the NEMS (Chapter 2)*
- *NGMM model structure, design, and mathematical formulation (Chapter 3)*
- *NGMM input data preoprocessing routines, including model initialization in the first year (Chapter 4)*
- *NGMM output data post-processing routines and reporting to other NEMS modules (Chapter 5)*
- *NGMM assumptions, inputs, and outputs (Chapter 6)*

It includes a number of appendices to support the material presented in the main body of the report:

- *Appendix A: Model abstract*
- *Appendix B: References*
- *Appendix C: Table relating the variable names used in the documentation to the specific variable, or identifier, used in the model code*
- *Appendix D: Table relating the equations presented in the documentation to the relevant procedure in the code*
- *Appendix E: Table relating the input data parameters in the model code and the data input files from which they are read and where detailed descriptions of the input data, including variable names, definitions, sources, units and derivations can be found*[5]
- *Appendix F: Table that Identifies all global data passed between other NEMS modules and the NGMM, as well as a brief description of the variable and the related module, where applicable*
- *Appendix G: Documents the derivation of all empirical estimations used in the NGMM*

Model archival citation

This documentation refers to the NEMS Natural Gas Market Module as archived[6] for the *Annual Energy Outlook 2020* (AEO2020). The model contact is

Kathryn (Katie) Dyl
Energy Information Administration
U.S. Department of Energy
1000 Independence Avenue, SW
Washington, D.C. 20585
(202) 287-5862

[5] The NGMM data files are available upon request from the model contact. Alternatively, an archived version of the NEMS model (source code and data files) can be downloaded here.

[6] U.S. Energy Information Administration, *Availability of the National Energy Modeling System (NEMS) Archive*

2. Model purpose

The purpose of the Natural Gas Market Module (NGMM) is to represent the U.S. natural gas market in the National Energy Modeling System (NEMS), both as it operates today and how it may evolve in the future. The NGMM balances natural gas supply and demand estimates in North America, projecting the volume and price of natural gas supply, its transmission through the pipeline network, and its distribution to end-use consumers. This chapter will give a brief overview of the U.S. natural gas market from wellhead to end user, discusses how the recent evolution of the natural gas market motivated the decision to build the NGMM, and how the NGMM interacts with other NEMS modules.

Model objectives

Reflect current and future natural gas market

Natural gas market overview

The natural gas market refers to the transportation of natural gas from the source of supply (e.g. gas processing plants) and its distribution to the end-use consumer. As of 2018, natural gas accounted for 31% of the primary energy consumed in the United States.[7] And unlike other energy sources like petroleum, which is primarily consumed in the transportation sector, or coal and renewables, which are primarily used to generate electricity, natural gas is widely consumed across many demand sectors. Natural gas is used in the residential and commercial sectors for heating, in the industrial sector for heating, power, and in the petrochemical industry as feedstock; natural gas has also become an important fuel in electric power generation.

An illustration of the entire natural gas market is shown in Figure 2.1 and can be classified by three industries: the upstream industry, encompassing activities relating to the exploration and production of natural gas; the midstream industry, encompassing the transmission and distribution of the natural gas; and the downstream industry, encompassing companies or facilities that deliver natural gas to consumers or transform it into other energy products. Generally speaking, the natural gas market, as it is represented in the NGMM, refers to the midstream portion of the natural gas industry as a whole; however, there are components of the upstream and downstream sectors that are also represented in the model.

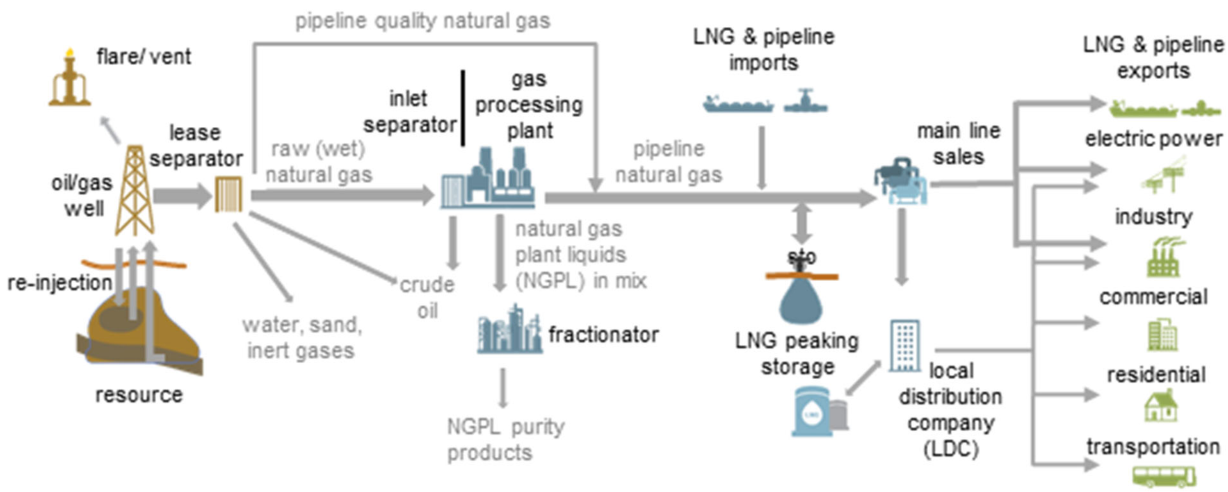
The **upstream** natural gas industry refers to the sector that identifies, characterizes, and produces natural gas resources. This includes investigating the potential of a resource with geological and geophysical (e.g. seismic) surveys, the development of a formation or basin, drilling wells, and operating producing wells. From the wellhead, the production stream will usually enter a lease separator where it is separated into three parts: liquids (either crude oil or lease condensate), wet natural gas,[8] and water.

---

[7] U.S. Energy Information Administration, *U.S. primary energy consumption by source and sector, 2018* , *Monthly Energy Review* (April 2019)

[8] EIA's definition of wet natural gas is a mixture of hydrocarbon compounds and small quantities of various non hydrocarbons existing in the gaseous phase or in solution with crude oil in porous rock formations at reservoir conditions. The principal hydrocarbons normally contained in the mixture are methane, ethane, propane, butane, and pentane. Typical nonhydrocarbon gases that may be present in reservoir natural gas are water vapor, carbon dioxide, hydrogen sulfide, nitrogen and trace amounts of helium.

Figure 2.1 Schematic of upstream, midstream and downstream industries in the natural gas market



In oil-directed wells, natural gas may be flared or reinjected if there is no natural gas pipeline infrastructure in place (or available pipeline capacity). In the United States, however, the vast majority of natural gas is commercialized. Once it leaves the lease separator, the marketed natural gas production[9] is sent to the processing plant. In some cases, the natural gas will be of sufficient quality to bypass the processing plant and directly enter the pipeline network.

The **midstream** natural gas industry encompasses the wide range of infrastructure required to process the natural gas produced from wells, transport it through the pipeline network, and distribute it to end users. The processing plant is the nexus of the upstream and midstream industries where marketed natural gas production is separated into natural gas plant liquids (NGPL) and dry natural gas.[10] The quality, or heat content, of this gas can vary considerably depending on the processes used, which can include condensation, absorption, adsorption, or refrigeration. Most of the heat content variability results from how much ethane is removed from the natural gas stream as it is the most similar in chemical properties to natural gas. From the processing plant, the NGPL will be sent to a fractionator to be separated into their individual compounds while the natural gas will enter the pipeline network.

The vast majority of natural gas pipeline capacity in the United States is on interstate transmission pipelines: a network of large-diameter pipes, often operating at high pressures, that can transport natural gas hundreds of miles from supply basins to demand markets. It is this pipeline network that the

[9] EIA's definition of marketed natural gas production is gross withdrawals of natural gas from production reservoirs, less gas used for reservoir repressuring, nonhydrocarbon gases removed in treating and processing operations, and quantities vented and flared.

[10] EIA's definition of dry natural gas is natural gas which remains after: 1) the liquefiable hydrocarbon portion has been removed from the gas stream (i.e., gas after lease, field, and/or plant separation); and 2) any volumes of nonhydrocarbon gases have been removed where they occur in sufficient quantity to render the gas unmarketable. Note: dry natural gas is also known as consumer-grade natural gas. The parameters for measurement are cubic feet at 60 degrees Fahrenheit and 14.73 pounds per square inch absolute.

NGMM represents in the NEMS. All interstate pipelines are regulated by the Federal Energy Regulatory Commission (FERC), including both their construction and tariff structure. There are also intrastate pipelines—most commonly in Texas, Oklahoma, and California—that are regulated by individual states and, despite not crossing state lines, serve an important role in regional natural gas transmission.

The midstream industry also includes assets that aid in maintaining adequate line pressures on the pipeline network and supplementing natural gas production during periods of high demand. These functions are primarily accomplished by storage operators. During the injection season, which is defined from April 1 to October 31, natural gas is typically injected into underground storage facilities from the interstate pipeline system; these facilities can be old natural gas wells or reservoirs no longer producing, salt caverns, or aquifers. Natural gas is then withdrawn from storage and delivered back into the pipeline network during the withdrawal season—November 1 to March 31—as needed to meet customer demand during the winter season.

In regions that lack underground storage facilities or have insufficient pipeline capacity to meet peak demand periods, small-scale liquefied natural gas (LNG) peak-shaving facilities may exist. Liquefying natural gas efficiently stores and transports large quantities since the volume of natural gas in its liquid state is about 600 times smaller than its volume in a gaseous state. These facilities, many of which produce LNG during periods of low demand, will store LNG until it is needed, regasify it, and send natural gas out into the market during periods of peak demand. In addition, the Lower 48 states also has 11 LNG import terminals which can receive, store, and regasify large cargoes from overseas via marine vessels.[11] However, only one— Everett, Massachusetts—still regularly receives LNG cargos.

At the end of the supply chain, the **downstream** natural gas industry comprises the end-use sectors that receive natural gas deliveries from the pipeline network and distribute it to customers. For some types of consumers and end-use sectors, such as LNG export facilities, industrial facilities, and electric generators, individual facilities have direct access to the interstate pipeline network. However, all residential and commercial consumers, the transportation sector, and a portion of the industrial sector receive natural gas from a local distribution company (LDC). This is a retailer that procures natural gas from the transmission system and distributes, or sells, it to end users through its own distribution pipeline system. In general, energy-intensive industries and facilities that regularly consume large volumes of natural gas (e.g. LNG export facilities, electric generators) will purchase natural gas directly from the interstate pipeline system as it is cheaper for these entities to directly purchase natural gas from marketers or suppliers. Obtaining natural gas from an LDC, while more expensive, ensures the delivery of natural gas and has regulatory requirements in place to guarantee that it will be available during periods of peak demand.

Shale gas production and the transformation of the U.S. natural gas market

In the United States, the natural gas market underwent a fundamental shift from 2005 to 2015 as horizontal drilling and hydraulic fracturing of shale formations transformed how oil and natural gas are produced. Prior to this timeframe, U.S. natural gas production had peaked in 1973 at 60 billion cubic feet per day (Bcf/d).[12] While natural gas production averaged 51 Bcf/d from 1990 to 2005, it had begun

[11] Federal Energy Regulatory Commission, North American LNG Import/Export Terminals: Existing.

[12] U.S. Energy Information Administration, U.S. Dry Natural Gas Production data from the *Natural Gas Annual*

to decline, and the South Central region of the country—Texas, Oklahoma, Louisiana, and Arkansas—and the Gulf of Mexico accounted for 60% of natural gas production in the United States. The interstate pipeline network was designed to transport natural gas from the Gulf Coast, western Canada, and the Rocky Mountains to demand centers in the Northeast and Midwest. While Canada's pipeline network was integrated with that of the United States, only 3.6 Bcf/d of capacity existed between the United States and Mexico.[13] The expectation was that domestic production would not be able to meet demand in the future; as a result, LNG import terminals were being constructed and proposed.

Today, the U.S. natural gas market resembles the opposite of what characterized it only a decade or so ago. Marketed natural gas production exceeded 99 Bcf/d in 2019, an estimated 30% of which was produced in the Northeast.[14] The interstate pipeline system has been transformed as pipeline reversals and bi-directional capabilities have resulted in natural gas flowing out of the Northeast toward demand centers on the Gulf Coast. Natural gas pipeline capacity between the United States and Mexico has tripled; as of the end of 2019, 13.7 Bcf/d of cross-border capacity into Mexico exists, with more currently under construction. And in February 2016, Sabine Pass became the first LNG export facility in the Lower 48 states to export LNG to global markets. Several more LNG export facilities are currently under construction, and by the end of 2020, the United States is expected to have 9.6 Bcf/d of LNG export capacity.[15]

Decision to replace the Natural Gas Transmission and Distribution Module (NGTDM) with the NGMM

The NGMM, which was first implemented in the Annual Energy Outlook 2018, replaces the Natural Gas Transmission and Distribution Module (NGTDM). The NGTDM was initially developed in 1991 as a linear program (LP), but was revised significantly in 1994, becoming a model that utilized a heuristic algorithm to balance flows in the natural gas market based on historical trends. While numerous modifications have been made since then, fundamental changes have occurred in the U.S. natural gas market that were unanticipated when the NGTDM was incorporated into the NEMS in 1994. The unprecedented growth in natural gas production in the Northeast, enabled by hydraulic fracturing and horizontal drilling of shale gas and tight oil formations, has resulted in rapid changes to natural gas pipeline flows, regional price differentials, and trade patterns. As a result, the decision was made to redesign the natural gas representation in the NEMS, allowing it to better capture dramatic changes to the market.

EIA published its requirements for a new model in the August 2014 document, *Requirements for a Redesigned Natural Gas Transmission and Distribution Model in the National Energy Modeling System*. The primary requirements of the redesigned NGTDM were to 1) project delivered end-use prices, wellhead prices, and import and export prices given delivered volumes and a set of regional supply curves, and 2) produce results that balance the natural gas market, projecting volumes of production, imports, and exports, as well as lease, plant, and pipeline fuel, and supplemental supplies. A secondary requirement was to project interregional flows and pipeline capacity. Finally, the model has to align well

[13] U.S. Energy Information Administration, *U.S. State-to-State Capacity* (Excel file, updated quarterly).
[14] This includes dry natural gas production from Pennsylvania, Ohio, and West Virginia. Regional fraction estimated from gross withdrawals of natural gas.
[15] U.S. Energy Information Administration, *U.S. Liquefaction Capacity* (Excel file, updated as information becomes available).

with history, capture likely future market behavior, and be relatively easy to maintain, update, and modify.

Several potential modelling approaches for representing the natural gas market in the NEMS were reviewed in the September 2014 Leidos report, *Review of Natural Gas Models In Support of U.S. Energy Information Administration Natural Gas Transmission and Distribution (NGTDM) Redesign Effort*. The approaches were as follows: a linear (or nonlinear) program that maximized social welfare, a mixed-complementarity formulation, and an agent-based approach. While EIA concluded that mixed-complementarity and agent-based models are useful when modeling markets without perfect competition, in the case of perfect competition, these formulations yield the same solution as linear (or nonlinear) program. Given that the U.S. natural gas market is a competitive market,[16] EIA concluded that a nonlinear program could effectively model these dynamics; furthermore, a nonlinear program would be easier to develop and maintain.

While the methodology proposed for the redesigned NGTDM is the same as that by EIA in 1991, several different approaches were adopted to address the issues that arose. These modifications included the following: represent pricing at a more disaggregate level where the marginal price for the region/period is more likely to align closely with the historical average price for the region/period; set pipeline rates based on historical price differentials (i.e., state-to-state differences in spot and citygate prices) rather than on regulated rates; set flows based on variable charges, accounting for reservation fees separately; and allow pipeline capacity to increase in the current solution year if volumes and prices warrant, rather than in a planning model for a future year, as was done in the 1991 version.

A complete discussion of the natural gas model redesign is available in the *Natural Gas Transmission and Distribution Module Component Design Report*, published by EIA in August 2015.

Representation of the natural gas market in the NEMS

NEMS Overview

The NEMS is structured as a modular system. The modules include the Integrating Module and a series of relatively independent modules that represent the domestic energy system, the international energy market, and the economy. The domestic energy system is decomposed into fuel supply markets, conversion activities (e.g., refineries and power generation), and end-use consumption sectors.[17] The projections in the NEMS are developed assuming that energy markets are in equilibrium[18] using a recursive price adjustment mechanism.[19] For each fuel and consuming sector, the NEMS balances

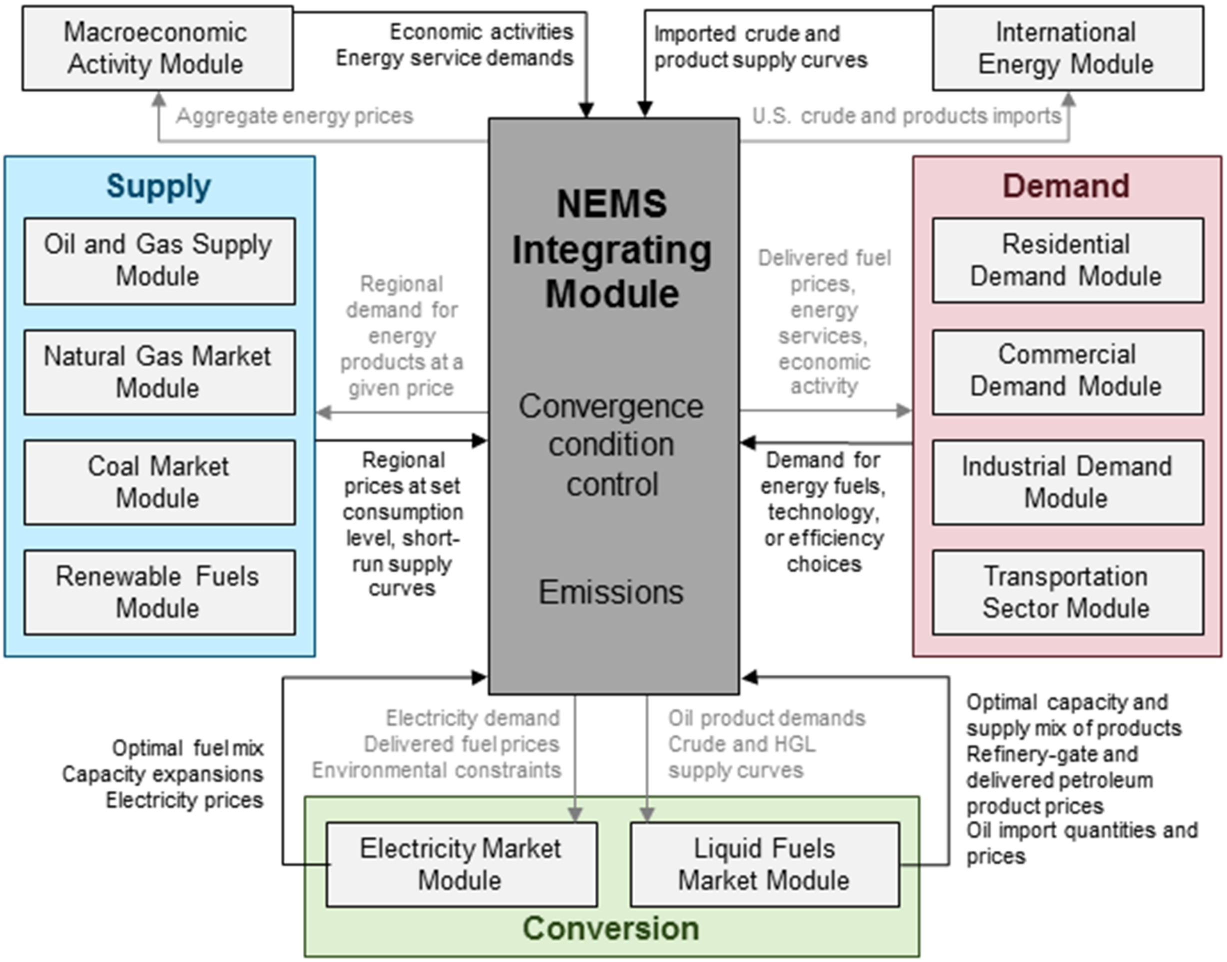
[16] In some cases the existence of long-term contracts could result in markets operating in a less than optimal manner, which could require some special handling (e.g., minimum flows) to properly reflect market dynamics in the first years of the projection period. Most notably long-term contracts are resulting in gas flows into the Northeast when market prices seem to indicate that gas should be flowing in the opposite direction.

[17] U.S. Energy Information Administration, *Integrating Module of the National Energy Modeling System: Model Documentation 2018*.

[18] Markets are said to be in equilibrium when the quantities demanded equal the quantities supplied at the same price; that is, at a price that sellers are willing to provide the commodity and consumers are willing to purchase the commodity.

[19] The central theme of the approach used is that supply and demand imbalances will eventually be rectified through an adjustment in prices that eliminates excess supply or demand.

Figure 2.2 Schematic of the NEMS and flow of information between modules



energy supply and demand, accounting for the economic competition between the various fuels and sources. The system includes a routine that can simulate a carbon emissions cap-and-trade system with annual fees to limit carbon emissions from energy-related fuel combustion. The primary flows of information between each of these modules are the delivered prices of energy to the end user and the quantities consumed by product, Census division, and end-use sector. The information flows also include other data such as economic activity, domestic production activity, and international petroleum supply availability. Figure 2.2 illustrates the modules that comprise the NEMS as well as the flow of information in the system.

For each projection year, the NEMS solves by iteratively calling each module in sequence (once in each NEMS iteration) until the delivered prices and quantities of each fuel in each region have converged within tolerance between the various modules, thus achieving an economic equilibrium of supply and demand in the consuming sectors. For some applications the model is also run in multiple cycles, generally to converge on a solution that involves the need to look ahead at other projected values for future years when solving the current projection year. Module solutions are reported for each projection year through the midterm horizon. While each module can operate at the level of detail—both regionally and temporally—most appropriate for its particular sector, they all aggregate (or disaggregate) their solutions to the Census-division structure on an annual basis to transfer information within the NEMS.

Natural Gas Market Module (NGMM) overview

Within the NEMS, the NGMM represents the transmission, distribution, and pricing of natural gas. Based on information received from other NEMS modules, the NGMM also includes representations of the end-use demand for natural gas, the production of domestic natural gas, and the availability of natural gas traded on the international market. The NGMM links natural gas suppliers (including importers) and consumers (including liquefied natural gas (LNG) export terminals) in the Lower 48 states and across the Mexican and Canadian borders through transmission between market hubs. For all months in a year, the NGMM determines the production, flows, and market clearing prices of natural gas within a state-level representation of the U.S. pipeline network and a regional representation of the Canadian and Mexican pipeline network.

Natural gas pricing and flow patterns are derived by obtaining a market equilibrium across the three main elements of the natural gas market: the supply element, the demand element, and the transmission and distribution network that links them. This equilibrium is obtained by optimizing for producer plus consumer surplus minus transportation costs and takes the form of a quadratic program (QP). The methodology employed allows for the analysis of impacts of regional capacity constraints in the interstate natural gas pipeline network and the identification of primary pipeline capacity expansion requirements. Distributor tariffs are also projected in order to arrive at the delivered price of natural gas to domestic consumers.

The Lower 48 states' demand regions are represented at the state level. Canada is represented as an eastern and western region,[20] while Mexico is represented as five regions.[21] For all regions, consumption is represented for five end-use sectors: residential, commercial, industrial, electric power, and transportation (or natural gas vehicles). The U.S. transportation sector is separated into compressed and liquefied natural gas for use in vehicles (retail and fleet), ships, and trains. In addition, the NGMM is responsible for projecting natural gas consumed in lease and plant operations, consumed or lost during interstate transport of natural gas via pipeline, and used for liquefaction at LNG export facilities. Canadian and Mexican demand projections are not provided by other NEMS modules but by EIA's *International Energy Outlook* (IEO); however, there are exceptions where either external sources are used, or the NGMM projects this demand endogenously.[22]

One or more domestic supply regions are represented in each NGMM region. Both the Canadian and Mexican supply regions match the demand regions. While the Oil and Gas Supply Module (OGSM)

---

[20] The eastern Canadian region includes the provinces of Ontario, Quebec, Newfoundland and Labrador, Nova Scotia, New Brunswick, and Prince Edward Island. The western Canadian region includes the provinces of Manitoba, Saskatchewan, Alberta, and British Columbia, as well as the three territories.

[21] The Mexican demand regions are consistent with the regionality used by the Secretaría de Energía de México (SENER) in reporting natural gas market statistics and modeling natural gas markets.

[22] Details may vary by AEO. Refer to Assumptions to the Annual Energy Outlook (Natural Gas Market Module) for specifics.

projects U.S. and Canadian expected production of both associated-dissolved (AD) and nonassociated (NA) gas, the NGMM determines the realized, or actual, NA natural gas production required to meet demand at a given price. Mexican natural gas production is represented within the NGMM and is a function of both world oil price (for associated-dissolved gas) and the Henry Hub price (for nonassociated gas).

To determine import and export volumes, border crossing hubs are represented for each of the Lower 48 states where pipeline capacity to a Canadian or Mexican region exists. Imports of LNG into North America are set to historical levels in the United States and set exogenously for Canada and Mexico according to IEO results. U.S. LNG exports are modeled within NGMM for each state where it is assumed future liquefaction facilities will be allowed to be built. Any LNG facilities in existence or under construction are included in the model.

To summarize, the following volumes and prices are projected by the NGMM:

- *Realized nonassociated natural gas production and supply prices by oil and gas district (84), annual*
- *Total dry gas production and supply prices by oil and gas region (13), annual*
- *Realized nonassociated natural gas production and supply prices by Canada region, annual*
- *Henry Hub spot price, annual*
- *Delivered end use prices by sector and Census division, annual*
- *Delivered end use prices to the transportation sector by transportation mode and Census division, annual*
- *Delivered end use prices to the electric power sector by Natural Gas-Electricity Market Module (NGEMM) region (17) and season (3)*
- *Lease, plant, pipeline, and liquefaction fuel use by Census division, annual*
- *Natural gas pipeline import and export volumes for Canada, Mexico, and LNG, annual*
- *LNG export capacity and volumes by Census division (plus western Canada and Alaska), annual*
- *Natural gas pipeline flows and capacities by Natural Gas market region (11) or Canada/Mexico region, annual*

Relation to other modules

Data transfer

The NGMM both requires and provides input to other NEMS modules. Data in the global data structure that are required by the NGMM to project the natural gas market include the following:

- *Gross domestic product (GDP) inflation adjustment factors and unemployment rates by year from the Macroeconomic Activity Module (MAM)*
- *Brent crude oil price and non-U.S. crude oil demand by type from the International Energy Module (IEM)*
- *Expected NA gas production and AD gas production by oil and gas district (84) and Canadian region (2) from the OGSM*
- *Alaska crude oil production by Alaska region from the OGSM*

- *Natural gas consumed during gas-to-liquids (GTL) and hydrogen fuel production and U.S. demand for crude oil by type from the Liquid Fuels Market Module (LFMM)*
- *Annual consumption by Census division from the Residential, Commercial, Industrial, and Transportation Demand and Electricity Market Modules (RDM, CDM, IDM, TDM, EMM)*
- *Seasonal (winter, summer,spring/fall)*[23] *consumption by Natural Gas-EMM (NGEMM) region from the EMM*
- *Annual consumption by transportation mode (personal, fleet, rail, and marine vehicles) and by Census division from the TDM*
- *Number of residential customers by Census division from the RDM*
- *Commercial floor space by Census division from the CDM*
- *Heating degree days by Census division from the CDM*

The NGMM also sends data to the NEMS global data structure for use by other modules. These include data used by the Integration Module to calculate the total natural gas supply-demand balance and data for publication in the *Annual Energy Outlook*. NGMM outputs, and the other modules that use them, include the following:

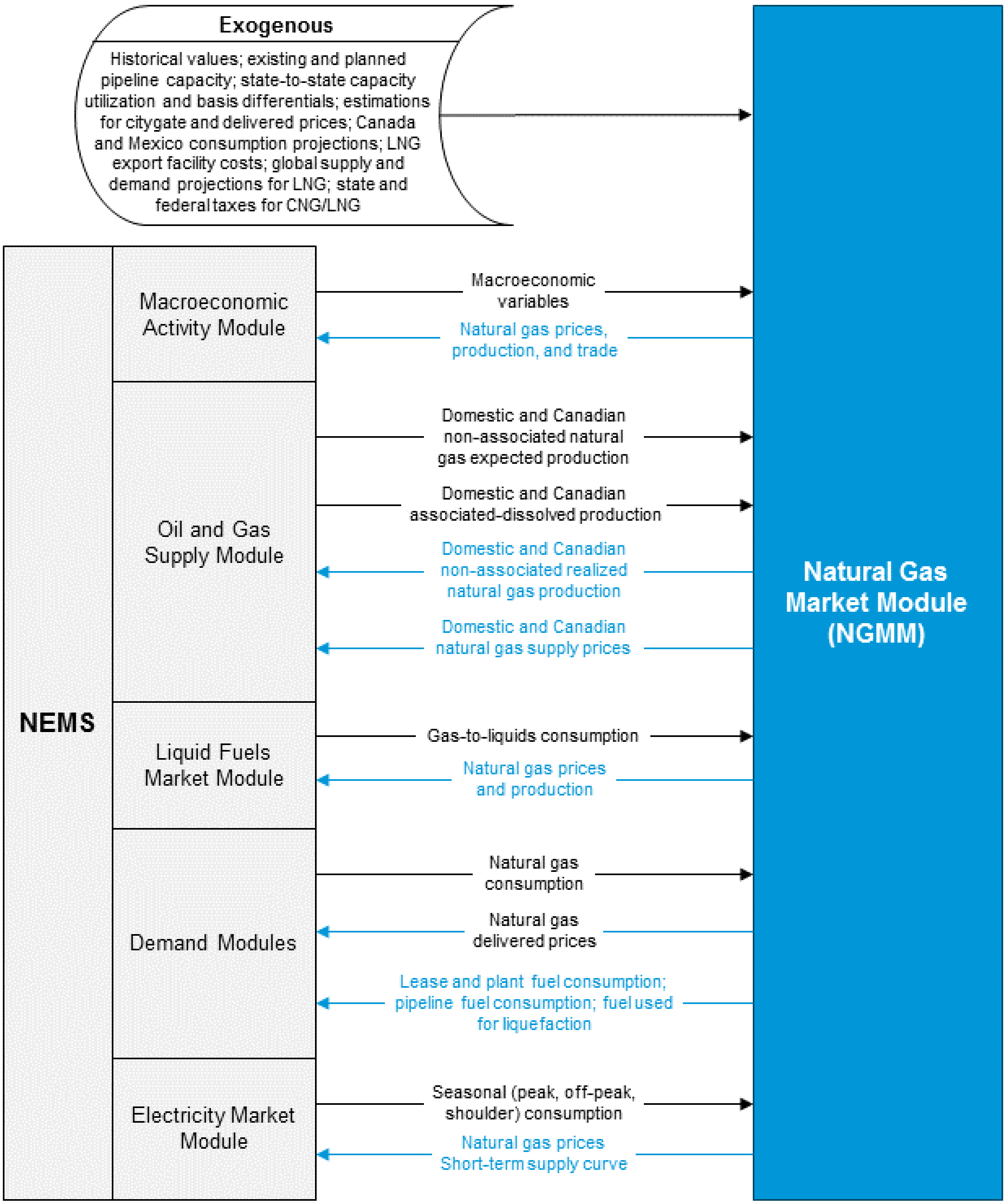
- *Realized total annual NA production by oil and gas district (84) and Canadian region (2) to OGSM*
- *Annual annual gas supply prices by oil and gas district and Canadian region to the OGSM*
- *Total annual dry natural gas production and supply prices by oil and gas region (13) to the NEMS*
- *Total annual natural gas supplemental supply volumes by oil and gas region (13) to the NEMS*
- *Total annual natural gas balancing item by Census division to the NEMS*
- *Total annual natural gas consumption used in lease and plant operations (lease and plant fuel) by Census division to the IDM*
- *Total annual natural gas consumed for liquefaction at LNG export facilities by Census division to the IDM*
- *Total annual natural gas consumed by pipelines (pipeline fuel) by Census division to the TDM*
- *Annual delivered prices for natural gas by Census division to the RDM, CDM, IDM, TDM, EMM*
- *Seasonal delivered prices to the electric power sector by NGEMM region to the EMM*
- *Natural gas supply curve parameters to the EMM*[24]
- *Annual delivered prices by transportation mode (personal, fleet, rail, and marine vehicles) and by Census division to the TDM*
- *Annual volumes and prices of U.S. natural gas imports and exports to Canada and Mexico (by pipeline), and as LNG by vessel, for use by the MAM*
- *Henry Hub spot price to the NEMS*

A complete representation of the exchange of data between the NGMM and other NEMS modules is illustrated in Figure 2.3.

[23] Winter months are defined as December through March. Summer months are defined as June through September. Spring/fall, or shoulder season, is defined as April, May, October, and November.

[24] The Electricity Capacity Planning Submodule of the EMM uses a reduced form, national natural gas supply curve representation in order to improve NEMS convergence and help determine the future utilization of natural gas generators.

Figure 2.3 Inputs and outputs of the NGMM, including relationships between other NEMS modules

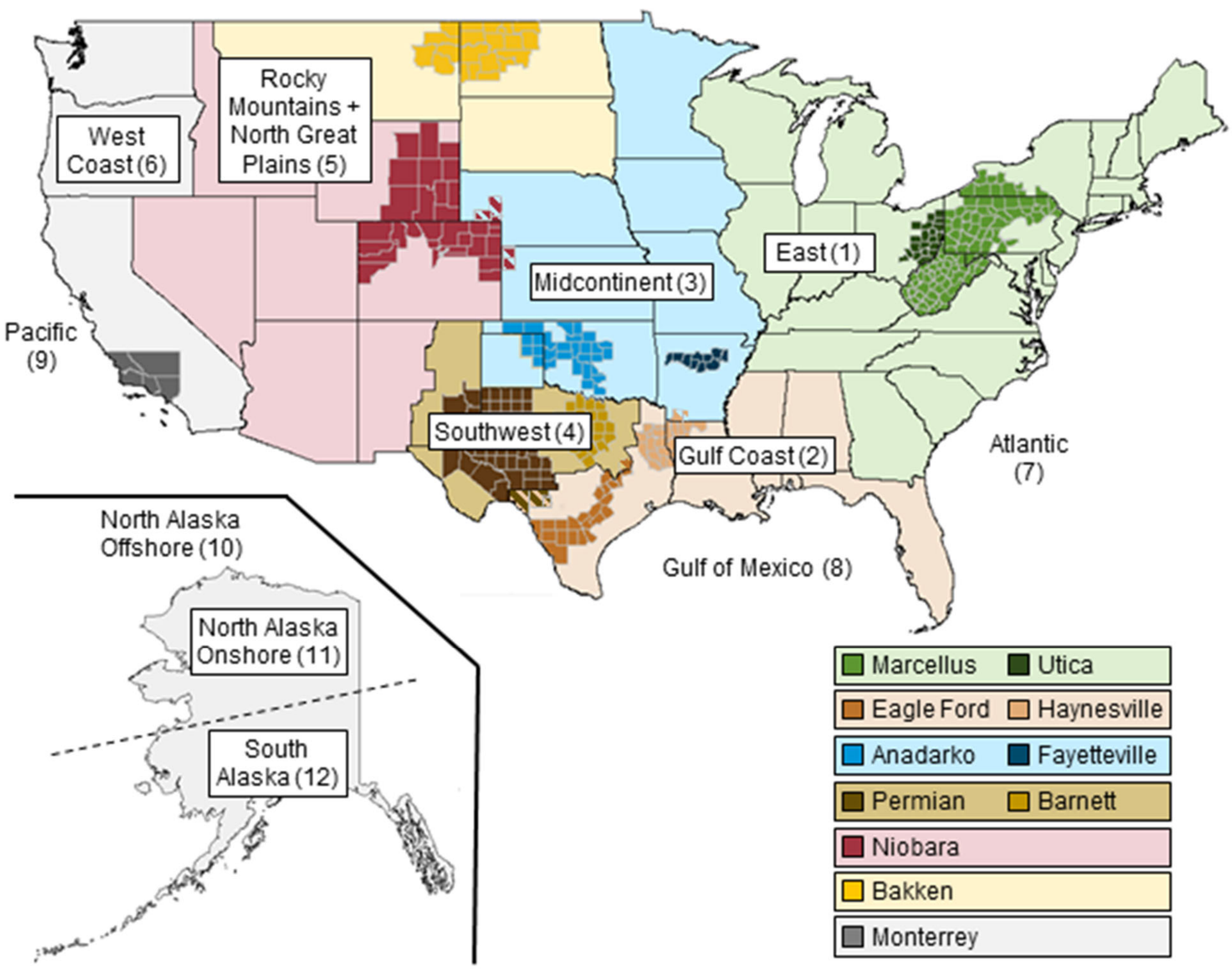


Regionality

Since the NEMS operates on an annual basis, NGMM results are generally passed to other NEMS modules as annual totals or quantity-weighted annual averages. While the NGMM and the OGSM pass expected/realized NA production, AD production, and supply prices to each other by OGSM district (of which there are 84), these results are ultimately passed to the NEMS by "oil and gas supply region." There are 14 of these regions in *Annual Energy Outlook 2020*: 7 onshore regions, 3 offshore regions, 3 Alaska regions, and a U.S. total. These regions, as well as their relationship to state boundaries and county-level tight oil and shale gas regions,[25] are shown in Figure 2.4.

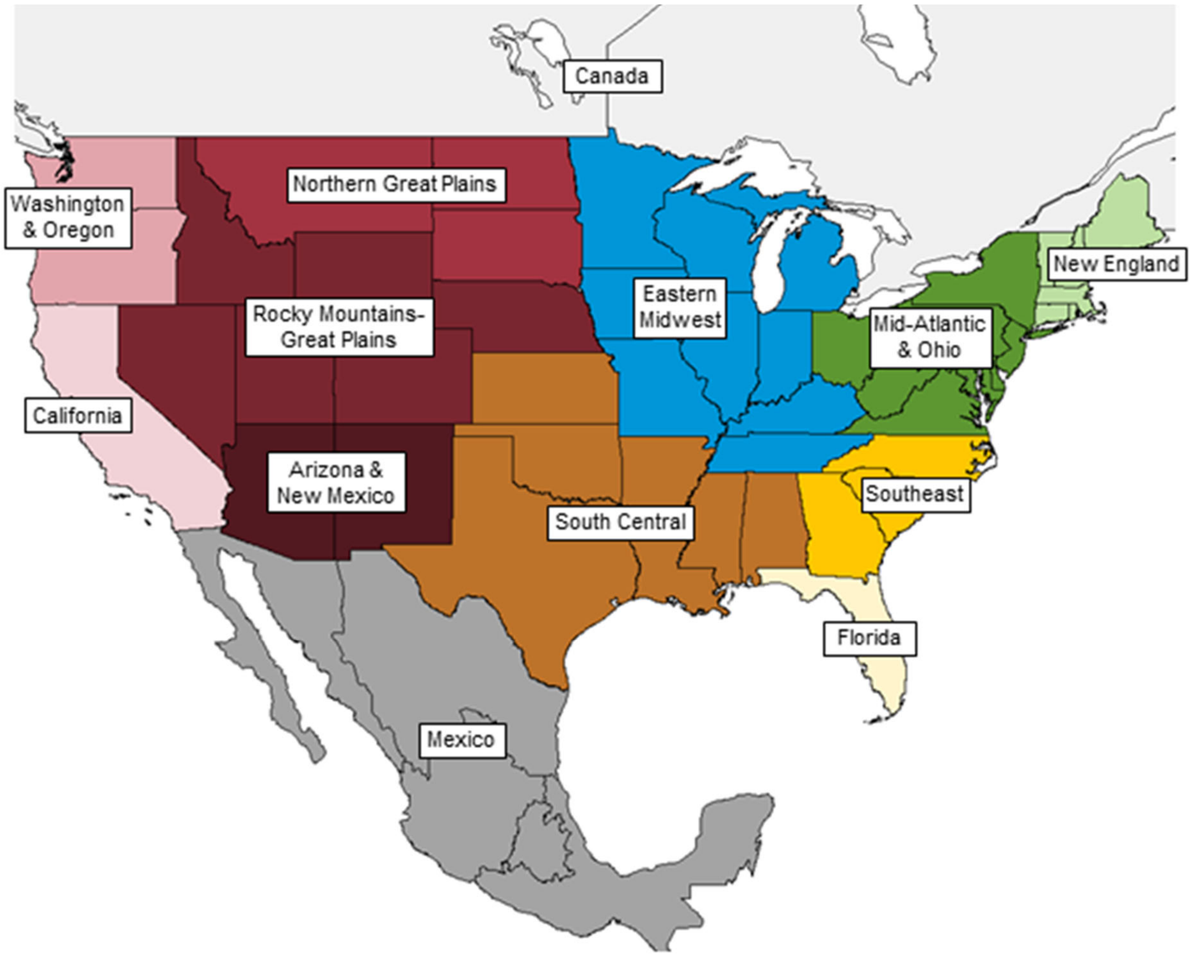
For reporting natural gas regional flows and pipeline capacities, the NGMM uses Natural Gas (NG) regions (Figure 2.5). This regionality is consistent with the natural gas storage regions used by EIA in its *Weekly Natural Gas Storage Report* and other natural gas data publications; however, NG regions are

Figure 2.4 NEMS Oil and Gas Supply Regions and corresponding tight oil and shale gas regions



[25] Tight oil and shale gas regions are used by EIA's *Drilling Productivity Report* (DPR) to estimate changes in oil and natural gas production in selected key basins. Regions for additional plays are included in order to interpret which shale gas basins contribute to the total production for a given region. The DPR combines the Marcellus and Utica regions, projected Appalachia basin production.

Figure 2.5 Natural Gas (NG) regions used to report regional flows and capacity



further disaggregated to provide greater insight into projected flows from and to specific supply and demand regions in the United States.

For all end-use demand modules and sectors except electric power generation, consumption volumes and delivered end-use prices are passed within the NEMS at the Census division level (Figure 2.6). In the case of the electric power sector, natural gas consumption and prices are transferred by Natural Gas-EMM (NGEMM) regions (Figure 2.7). These regions approximate the relationship between the North American Electric Reliability Corporation (NERC) regions at which the EMM operates and the demand regions used in the the NGMM.

Figure 2.6 NGMM demand regions (Census)

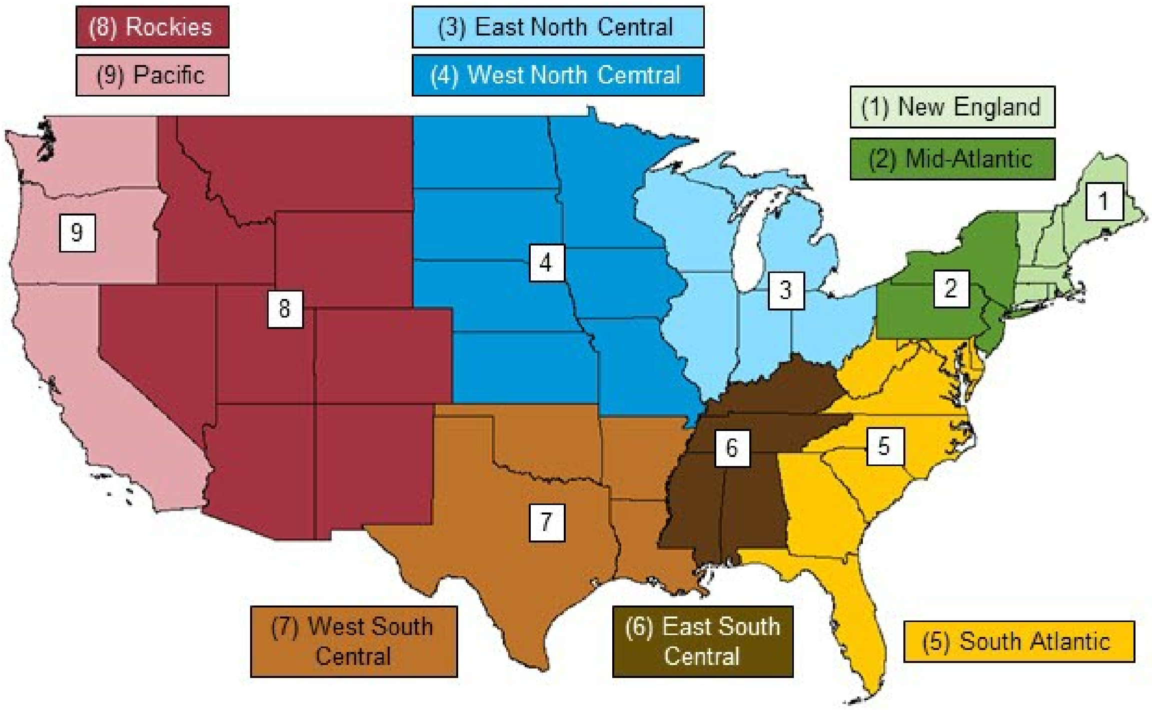
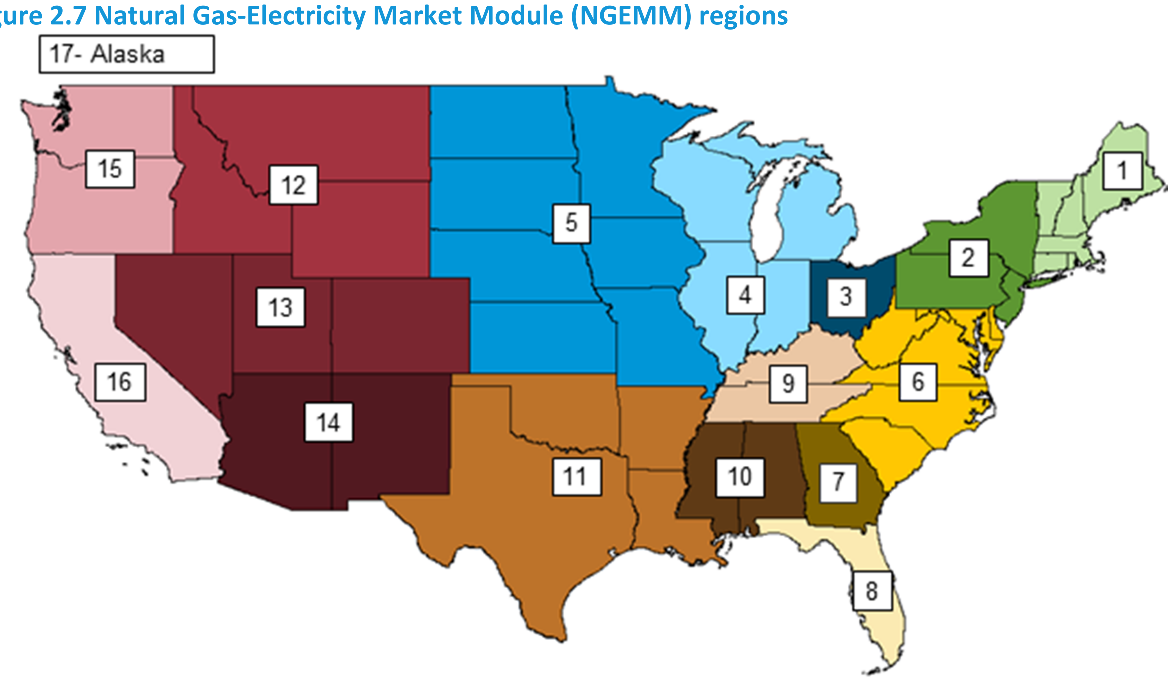


Figure 2.7 Natural Gas-Electricity Market Module (NGEMM) regions



3. Natural Gas Market Module design and structure

Model design

The Natural Gas Market Module, or NGMM, is a quadratic optimization model that balances natural gas supply and demand by maximizing consumer plus producer surplus minus variable transport costs, ensuring mass balance at each node. While the model is specified by a quadratic objective function, it is subject to linear constraints. Supply and demand elements are represented by either price-responsive curves or as fixed volumes, with the model code accommodating user selection of one or the other. The representative network contains a market hub in each state, as well as international and border hubs, and solves each month in a given year independently of all other months.

The objective function in the NGMM is an application of economic surplus, or the maximum economic benefit that an economy can obtain. The consumer surplus represents the amount of money saved by consumers who would buy natural gas at a given price, but are able to obtain it at a lower one. The producer surplus represents the added revenue of suppliers who could sell natural gas at a lower price, but are able to charge a higher one. Therefore, by maximizing this combined surplus, and subtracting transportation costs, the model arrives at an equilibrium price for the market, as seen in Figure 3.1.

A flow diagram of how NGMM operates in a given National Energy Modeling System (NEMS) iteration is illustrated in Figure 3.2. After reading in global data from other NEMS modules, which is transferred via

Figure 3.1 Schematic price-quantity graph illustrating maximized producer-consumer surplus

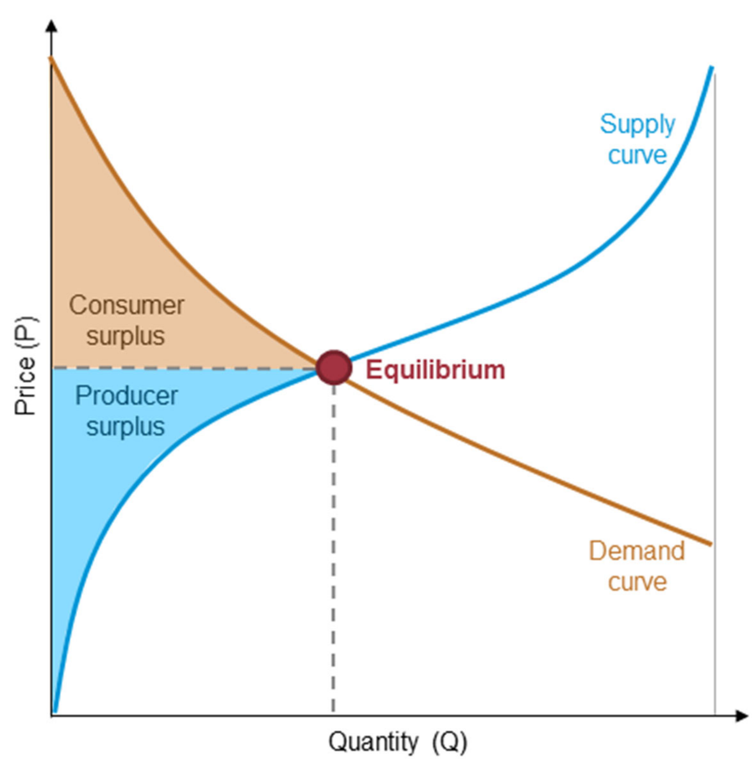
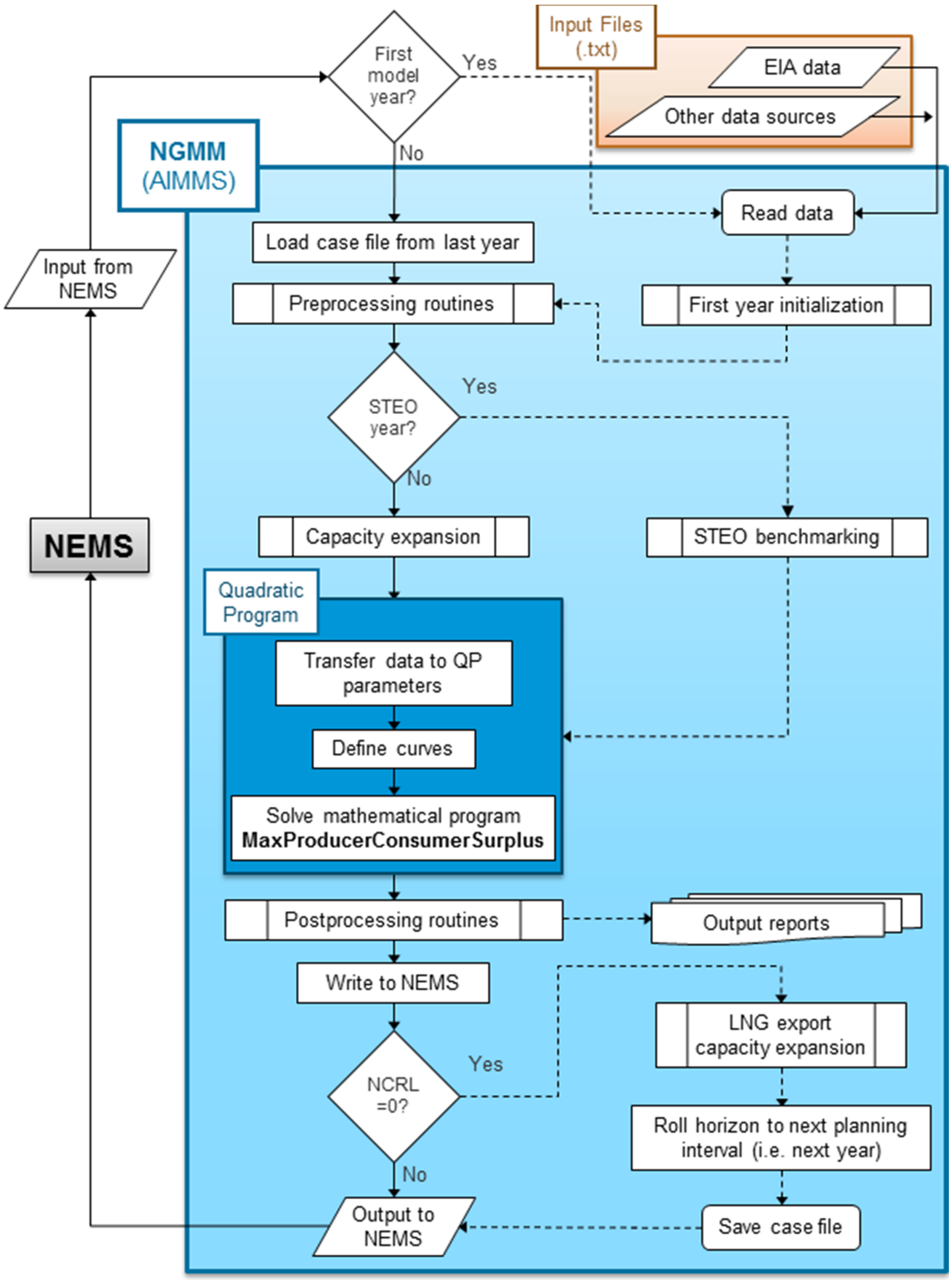


Figure 3.2 Process flow diagram representing the NGMM and its operation within NEMS in a given year



text file, a (binary) case file is loaded into AIMMS, a sparse-execution programming language. This case file contains the data saved during the NEMS report loop from the prior year (except during the first model year, where data are read in from input text files in order to initialize and process historical data). Data are disaggregated from the NEMS level, which is annual and by Census division (or, in the case of electricity demand, by season and Natural Gas-Electric Market Module (NGEMM) region), to the state and monthly level required by the NGMM. Next, if the model is solving over a year designated for benchmarking to the *Short Term Energy Outlook* (STEO), the quadratic program (QP) is first run iteratively to calculate benchmark factors. Otherwise, a second capacity expansion QP (with the same form) is run to determine if any capacity builds are required. After the QP is solved, data are then converted into report variables to return back to the NEMS.

Because of this design, the NGMM model code can be broadly divided into 6 sub-modules:

- ***First year initialization****: Done during the first NGMM model year (FirstModelYear), this is the section of code where all input files are read in, the time horizon is set and synchronized with the calendar, and all historical data are processed in order to calculate market shares or averages that are used throughout the projections.*
- ***Pre-processing routines****: Performed for all projection years (including the first model year), this is where all data (both historical and input data from other NEMS modules) are disaggregated and assigned to parameters that will be used in the QP. It is also where the parameters for both Mexican and Canadian production and consumption are calculated. The liquefied natural gas (LNG) export capacity expansion pre-processing routine is run only during the NEMS report loop.*
- ***Capacity expansion****: Run for all projection years beginning with the first year where new builds are allowed (NoBuildYear),*[26] *the capacity expansion QP is solved in order to determine if additional pipeline capacity is needed along any arcs in the model where capacity currently exists. Capacities are then assigned according to last year's capacity, including any planned capacity expansions and the additions built in the NGMM capacity expansion QP. This QP has the same structure as the main QP described below.*
- ***STEO benchmarking****: Only performed in years where STEO benchmarking needs to be applied, this is where the STEO factors are either set or determined by iteratively running the QP until all STEO factors are converged (convergence is achieved when all model results that are benchmarked to STEO are less than or equal to 2% of the STEO target value).*
- ***Quadratic program (QP)****: This is the section where parameters are transferred from their monthly/yearly assignment to their respective QP parameter (indexed by period in the planning horizon, as described below). The piecewise linear curves are assigned, and the mathematical program (objective function subject to constraints) is solved.*
- ***Post-processing routines****: In this section, the QP results are converted and aggregated into report variables to be transferred to other NEMS modules. All production and prices, including delivered end-use prices, are assigned here.*

[26] All years before *NoBuildYear* do not run the capacity expansion QP. This year is typically two to three years after the first model year, a typical amount of time allowing for regulatory approval and construction or upgrading of facilities.

The mathematical definition of the QP, as well as sub-modules where the QP is run — capacity expansion and STEO benchmarking — are described in the model structure section below. Pre-processing (including first year initialization and LNG capacity expansion) and post-processing routines are discussed in subsequent chapters.

The design of the NGMM also incorporates several features and techniques in AIMMS to minimize computational runtime, debug results, and visualize output.

Relationship between calendar dates and model periods: Time horizon

The NGMM handles time within the model code using a horizon – a set of planning periods. These periods are divided into three groups for a given model year, as see in Figure 3.3:

- *Planning interval: the main group of periods*
- *Past: all periods prior to the planning interval*
- *Beyond: all periods after the planning interval*

The time horizon is initialized in the first model year using a timetable, which maps the horizon to a calendar. The identifiers used to define both the calendar and horizon in the NGMM are in Table 3.1. Each planning interval represents the current model year and contains 12 periods, representing months. During the NEMS report loop, after the system has converged for a given year, the time horizon is rolled over to the next planning interval. This ensures that the calendar and horizon stay linked to time in the NEMS system despite the fact that any number of iterations may be performed for a given projection year in the NEMS.

By indexing the variables and constraints over a horizon, the model tells AIMMS to restrict the generation of constraints and variables to the periods within the current planning interval. This has the benefit of limiting the number of parameters and variables that must be assigned and fixing all variables in these periods, significantly decreasing the size of the mathematical program and the time required to solve it. However, the time horizon dimension also allows for past data to be stored within all parameters and variables, making debugging easier and allowing for more intuitive data visualization. Furthermore, as the NGMM evolves, inclusion of lagged parameters and variables in the QP is possible

Figure 3.3 Relationship between time in the NEMS, time in the NGMM, and the NGMM time horizon

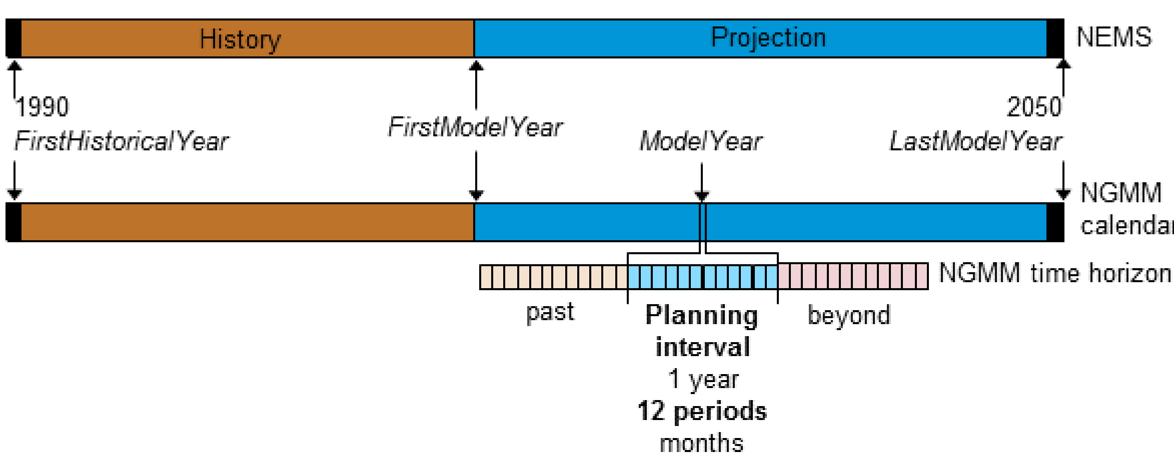


Table 3.1 Time and horizon parameters set in the NGMM

Parameter	Value	Notes
Calendar		
BeginDateOfCalendar	Jan, 1990	
EndDateOfCalendar	Dec, 2050	
FirstModelYear	2018	
LastModelYear	2050	
FirstHistoricalYear	1990	
LastFutureYear	2080	
Horizon		
NumberOfPeriodsInPlanningInterval	12	
NumberOfPeriods	396	(LastModelYear – LastHistoricalYear) * 12
CapExpNumberOfPeriodsInPlanningInterval	2	January + August
CapExpNumberOfPeriods	68	(LastModelYear – LastHistoricalYear) * 2

Note: The parameters *FirstModelYear* and *NumberofPeriods* will change each year depending upon the last complete year of historical data available. The parameter *LastModelYear* will only change when NEMS extends its projection (every 5-10 years).

(e.g. to calibrate to historical data or reflect the influence of prior years on the solution without requiring the model to resolve for all years).

Mapping parameters

Within the NGMM, there are many examples of mapping input and output data from one set to another. The most common example is assigning values to different regionalities. While the NGMM operates at the U.S. state level with (two) Canadian and (five) Mexican regions, most NEMS parameters are defined by U.S. Census division. Electric power sector demand and prices are defined by the 17 NGEMM regions, and expected dry production volumes from Oil and Gas Supply Module (OGSM) are defined by the 84 oil and gas districts. Examples are also found in converting between calendars and from a calendar to a horizon, assigning monthly and seasonal parameters, and defining hub-to-hub arcs where pipeline capacity exists (thereby allowing flow).

Binary mapping parameters are used to easily aggregate and disaggregate data as well as to apply domain conditions to the assignment of identifiers. In AIMMS, a binary parameter is the most computationally-efficient means to perform these calculations or restrict allowable indexes. A full list of these mapping parameters is in Appendix E.

Solver

The NGMM uses the CPLEX solver to optimize the quadratic objective function. The barrier method, the AIMMS default for QPs, is used; however, the dual crossover option must be selected in order for the NGMM to provide useful results.[27] Since the barrier method is an interior point method, its solutions are not basic, resulting in meaningless reduced costs and dual values (i.e. shadow prices). A non-basic solution would make the shadow prices of the constraints, which are used to assign the spot price at a

[27] IBM CPLEX Optimizer for z/OS 12.7.0 : *Solving problems with a Quadratic Objective.*

given hub, erroneous. Furthermore, the barrier method is designed to avoid placing the decision variables at upper or lower bounds; however, the curves in the NGMM are piecewise linear, which means that a decision variable indexed by step should reach its upper bound before the next step's value is non-zero. Because of these reasons, the dual crossover step is used to produce a basis and provide the nearest basic feasible solution to the QP.

Model structure

This section details both the mathematical formulation of the NGMM and the process by which the QP is set up in a given model year. It includes the following:

- *Identifiers in the QP*
  - *Sets, or dimensions, by which the parameters and variables in the QP are indexed*
  - *Parameters that are included in either the objective function or the constraints*
  - *Decision variables*
- *Transference of data into the QP parameters: the process by which parameters indexed by month and year in the pre-processing routines are transferred into the parameters, dimensioned by planning period, listed above*
- *Piecewise linear curve definitions*
  - *Supply curve*
  - *Tariff curve*
  - *LNG export demand curve*
- *Objective function*
- *Constraints*

Identifiers in QP

Sets

Table 3.2 Set names and indexes representing the dimensions of the NGMM quadratic program

NGMM	Index(es)	Notes
SupplyType_	**suptype**	
NA_AD_	naadgas	only associated-dissolved and nonassociated supply types
Regions_	**reg,reg1**	**root set**
QPSupplyNode_	qps	
DemandNode_	d,storage	demand nodes include consumption and storage
Hubs_	h,h1	
BorderCrossings_	bx	
InternationalRegions_	r_int	
L_48_	l48	lower 48 states (+ DC)
LNGTerminals_QP_	lngexp_qp	
Supply_Curve_Step_	step	also used to set tariff and LNG export curves
Years_	year	
Months_	mon	
MonthlyHorizon_	tmon	this index is suppressed in tables and equations below

QP parameters

Table 3.3 Parameter names, abbreviations used in NGMM documentation, dimensionality, and descriptions

NGMM	Abbrev.	Index	Definition
QP_Consumption	CONS	d	Consumption in region d
QP_Storage_Withdrawals	WTH	(storage,h)	Net storage withdrawals from storage at hub h
QP_Storage_Injections	INJ	(h,storage)	Net storage injections at hub h
QP_PlantFuel	PLT	d	Consumption of fuel during lease and plant operations in region d
QP_StorageLoss	Q^{store}	storage	Percentage of volumes in storage lost during injection/withdrawal
QP_DistributionLoss	Q^{dist}	d	Percentage of residential and commercial volumes at region d lost during distribution
QP_IntrastatePipeFuelLoss	Q^{intra}	d	Percentage production in region d lost during transport on intrastate pipelines
QP_GatheringCharge	P^{gath}	qps	Fixed charge to transport supply from region qps to market
QP_StorageFee	P^{store}	(storage,h)	Fixed fee per unit volume charged when withdrawing from storage at hub h
QP_Discrepancy	DISC	d	Average historical discrepancy between supply and demand
PipeFuelLossFactorIN	f^{pip}	h	Factor describing percentage of fuel lost from flow along arc (h1, h) assigned to hub h
PipeFuelLossFactorOUT	f^{pip}	h	Factor describing percentage of fuel lost from flow along arc (h, h1) assigned to hub h
Supply Curve			
Pbase	PBASE	(suptype, qps,step)	Base supply price for a supply type on a supply curve step
Qbase	QBASE	(suptype, qps,step)	Base production of a supply type on a supply curve step
QP_Supply	Q0	(suptype, qps)	Production of a supply type in region qps including lease and plant fuel
QP_SupplyPrice	P0	(suptype, qps)	Supply price for a given supply type in region qps
ParameterSupCrv	CRV	(suptype,qps,step)	Percentage of base production of a supply price in region qps on a supply curve step
ParameterSupElas	ELAS	(suptype,qps,step)	elasticity (percent change in quantity over percent change in price) of supply type in region qps on a supply curve step
QbaseMin	QMIN	(suptype, qps)	Minimum production of a supply type in region qps
QbaseMax	QMAX	(suptype, qps)	Maximum production of a supply type in region qps

NGMM	Abbrev.	Index	Definition
Q_UpperBound	MAXQ	(suptype,qps,step)	Maximum production of a supply type in region qps on a supply curve step
Q_LowerBound	MINQ	(suptype,qps,step)	Minimum production of a supply type in region qps on a supply curve step
SupCrv_MaxStep	SSMAX		Maximum allowable step on the supply curve
Tariff Curve			
PipeTariffCurveQty	QTAR	(h, h1,step)	Base flow volume along arc (h, h1) on a tariff curve step
PipelineTariff	PTAR	(h, h1,step)	Base cost to transport along arc (h, h1) on a tariff curve step
Parameter_CapacityUtilization	UTIL	(h, h1,step)	Percentage of capacity along arc (h, h1) on a tariff curve step
QP_Capacity	CAP	(h, h1)	Maximum capacity along arc (h, h1)
PipeTarCrvQty_UpBound	MAXQT	(h, h1,step)	Maximum capacity along arc (h, h1) on a tariff curve step
CapacityMaxBuild	MAXADD	(h, h1)	Maximum additional capacity allowed to be built along arc (h, h1) during capacity expansion
PTCrv_MaxStep	PTSMAX		Maximum allowable step on the supply curve
LNG Export Demand Curve			
QP_LNGExportCapacity	LNGCAP	(lngexp_qp)	Maximum capacity at region lngexp_qp to export LNG
QP_LNGExportPrice	PEXP	(lngexp_qp)	Price where it is economical in region lngexp_qp to operate an LNG export facility excluding any sunk costs (i.e. capital expenses) or shipping costs to destination
LNGExportQty	QLNG	(lngexp_qp, step)	Base volume of LNG exports from region lngexp_qp on a demand curve step
LNGExportPrc	PLNG	(lngexp_qp, step)	Base price of LNG exports from region lngexp_qp on a demand curve step
LNGExport_UpBound	MAXLNG	(lngexp_qp, step)	Maximum volume of LNG exports from region lngexp_qp on a demand curve step
LNGExport_LoBound	MINLNG	(lngexp_qp, step)	Minimum volume of LNG exports from region lngexp_qp on a demand curve step
LNGExpCrv_MaxStep	LSMAX		Maximum allowable step on LNG demand curve

Decision variables

Table 3.4 Decision variable names, abbreviations used in NGMM documentation, dimensionality, and descriptions

NGMM	Abbrev.	Index	Definition
QSupplyStep	SSTEP	(suptype,qps,step)	Supply type suptype taken from a given step on the supply curve for supply region qps
QProduction	PROD	(suptype,qps)	Total supply of supply type suptype taken from supply region qps
QLNGexp	LNG	(lngexp_qp,step)	LNG export demand from a given step at LNG export region lngexp_qp; includes fuel used for liquefaction
QTariffCurve	TAR	(h,h1,step)	Total volume along arc (h, h1) under a given step in the tariff curve
FlowHubToHub	FLOWH2H	(h,h1)	Flow from hub h to hub h1
FlowSupplyToHub	FLOWS2H	(qps,h)	Flow from supply region qps to hub h
FlowHubToDemand	FLOWH2D	(h,d)	Flow from hub h to demand region d
FlowHubToLNGExport	FLOWH2L	(h,lngexp_qp)	From from hub h to LNG export region lngexp_qp
FlowStorageToHub	FLOWT2H	(storage,h)	Flow from storage region to hub h
FlowHubToStorage	FLOWH2T	(h,storage)	Flow from hub h to storage

Transfer data into QP parameters

Table 3.5 NGMM quadratic program parameters and corresponding pre-processing parameters

QP parameter	Index	Pre-processing parameter	Index
QP_Consumption[a]	(tmon,d)	TotalConsumption	(mon,d)
QP_Storage_Withdrawals	(tmon,storage,h)	StorageWithdrawals	(mon,storage)
QP_Storage_Injections	(tmon,h,storage)	StorageInjections	(mon,storage)
QP_PlantFuel	(tmon,d)	PlantFuel	(mon,l48)
QP_StorageLoss	(tmon,storage)	StorageLosses	(mon,l48)
QP_DistributionLoss	(tmon,d)	DistributionLosses	(mon,l48)
QP_IntrastatePipeFuelLoss	(tmon,d)	IntrastatePipeFuelLosses	(mon,l48)
QP_GatheringCharge	(tmon,qps)	GatheringCharge + GatherChargeAdd	qps
QP_Discrepancy	(tmon,d)	BalanceItem	(mon,d)
Supply Curve			
QP_Supply[b,c]	(tmon,suptype, qps)	Supply	(mon,suptype,qps)

QP parameter	Index	Pre-processing parameter	Index
	(tmon,naadgas, qps)	Supply * (1–LeaseFuelFactor)	qps
QP_SupplyPrice[d]	(tmon,suptype, qps)	WellhdPrice	(year-1, suptype, qps)
Tariff Curve			
QP_Capacity	(tmon,h, h1)	CurrentPipeCapacity	(mon,h,h1)
LNG Export Demand Curve			
QP_LNGExportCapacity	(tmon, lngexp_qp)	LNGExports * (1+Pct_Liquefaction_Fuel)	(mon,lngexp_qp)
QP_LNGExportPrice	(tmon, lngexp_qp)	USLNGExportPrice	(year,lngexp_qp)

[a] Adjust Northeast Mexico (MX_NE) for STEO factor adjusting exports
[b] Adjust nonassociated natural gas supply in west Canada (CN_W) for STEO factor adjusting imports
[c] Adjust lease fuel factor by STEO factor
[d] Adjust wellhead price by STEO wellhead price factor

Piecewise linear curve definitions

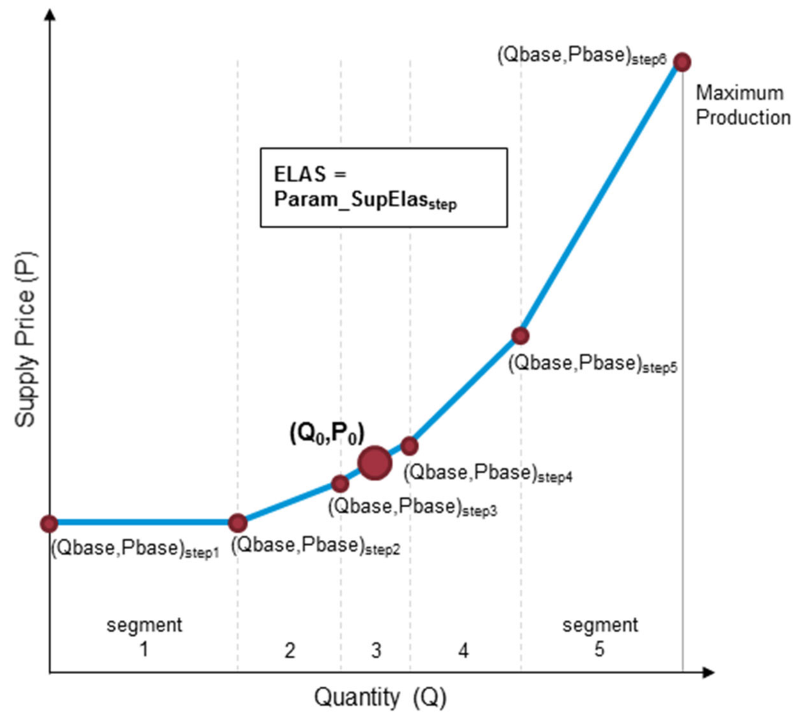
Supply curve definition

Variable, or price-responsive, supply is represented as producer surplus, or the area under a short-term supply curve. This short-term supply curve takes the form of a piecewise linear function based off of a price/quantity pair that represents the expected or baseline level of production or supply with an associated price. In the NGMM, the price/quantity pair is represented by last year's supply price and the expected production from the OGSM. Segments are then built off of that point by assuming a price elasticity of supply ($ELAS_{qps,step}$) for a given percent change from the expected production ($CRV_{qps,step}$).

The expected production represents an economically viable level and mix of production that producers are planning to make available to the market without either stressing the system or needing to cut back because of over-supply. As such, the supply curves are built around this expected production point with a shape that drives the solution towards that point while allowing some adjustment to balance the market. This is done by assuming a change in production values will be less responsive to a change in price at volumes less the expected production (i.e., once drilled, wells will produce regardless of price) and more responsive at volumes greater than the expected production (i.e., it will be more costly to speed up the drilling of new wells). The general form of this supply curve is shown in Figure 3.4, where the supply price $P_{qps,step}$ is a function of quantity of variable supply $Q_{qps,step}$ (i.e. nonassociated gas), its step's base quantity ($QBASE_{qps,step}$) and price ($PBASE_{qps,step}$), and an assumed elasticity:

$$P_{qps,step} = PBASE_{qps,step} * \left\{ \left[\left(\frac{1}{ELAS_{qps,step}} \right) * \left(\frac{Q_{qps,step} - QBASE_{qps,step}}{QBASE_{qps,step}} \right) \right] + 1 \right\} \qquad \forall (qps, step) \quad (1)$$

Figure 3.4 Schematic representation of short-term supply curve used in NGMM



Each supply region is assigned one of four options for the form of the short term supply curve, allowing for different levels of price responsiveness depending upon the region and analyst judgment. The values of the base supply and base price are calculated using the input parameters $CRV_{qps,step}$, the adjustment for each segment from expected production, and $ELAS_{qps,step}$, the price elasticity for each segment:[28]

For steps 1-3 below (Q_0,P_0):

$$QBASE_{qps,step} = (Q_0)_{qps} * \prod_{step}^{step=3} (1 - CRV_{qps,step}) \qquad \forall(qps, step) \tag{2}$$

$$PBASE_{qps,step} = (P_0)_{qps} * \prod_{step}^{step=3} \frac{(1 - CRV_{qps,step})}{ELAS_{qps,step}} \qquad \forall(qps, step) \tag{3}$$

For steps 4-6 above (Q_0,P_0):

[28] For AEO 2018, the elasticities defining each segment (1-5) are 0.8 (segment 1), 0.7, 0.5, 0.3, and 0.2 (segment 5).

$$QBASE_{qps,step} = (Q_0)_{qps} * \prod_{step=4}^{step} \left(1 + CRV_{qps,step}\right) \qquad \forall(qps, step)$$

(4)

$$PBASE_{qps,step} = (P_0)_{qps} * \prod_{step=4}^{step} \frac{\left(1 + CRV_{qps,step}\right)}{ELAS_{qps,step}} \qquad \forall(qps, step)$$

(5)

Pipeline tariff curve definition

For all arcs between two different hubs, a tariff curve was created to represent the variable cost of transportation per unit of flow as a function of capacity utilization (minus cost due to pipeline fuel used during transport). These curves are based on historical basis differentials between the spot prices at the two hubs.[29] They are specified so that the tariff increases rapidly as the flow approaches the pipeline capacity, or nears complete utilization; the final step is extended by a set percentage above the existing capacity.[30] This difference between the last two steps represents the maximum capacity build in a given year for most arcs. Exceptions for larger capacity builds in a given year are allowed on arcs in two cases:

- *For arcs where capacity is greater in the opposite direction, the model can simulate a decision to build additional capacity up to that level, representing pipeline reversals on large pipelines.*
- *For arcs with capacity below a user-defined level,[31] projected capacity is allowed to double in a given year, allowing small markets to grow at a faster rate.*

The formulation of the tariff curve is designed to act as a hurdle rate by which the model decides to add pipeline capacity when representative peak day consumption levels are flowed through the network. If pipeline capacity is added in one projection year, it indicates that either consumption has exceeded existing pipeline capacity or the cost of adding capacity along the arc is less than the cost of transporting natural gas via another existing route. The additional cost is assumed to be recovered by charging the same variable tariff rate to larger volumes of flow over time.

A representative tariff curve is depicted in Figure 3.5. The points ($QTAR_{h,h1,step}$, $PTAR_{h,h1,step}$) are largely defined from exogenous input parameters derived from historical data for monthly flows and spot prices; however, to allow for capacity builds, the quantity QTAR is calculated each projection year:

$$QTAR_{h,h1,step} = CAP_{h,h1} * UTIL_{h,h1,step} \qquad \forall(h, h1, step)$$

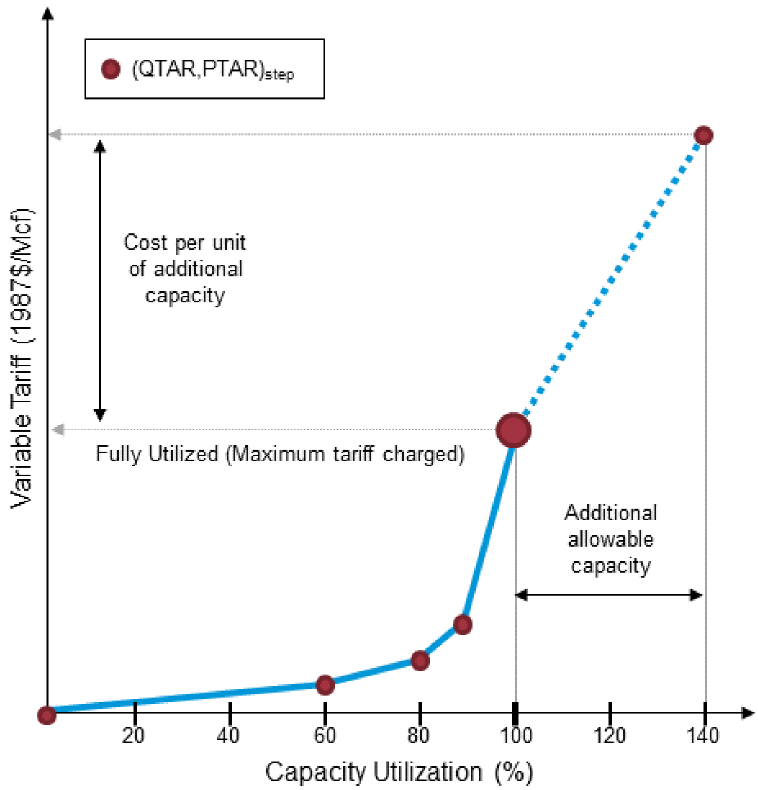
(6)

[29] Monthly average spot price history begins in 2009; data are used through latest available month and provided by *Natural Gas Intelligence*.

[30] For AEO 2018, additional capacity of up to 40% of the existing capacity can be added in a given year.

[31] For AEO 2018, this capacity was defined as 30 Bcf, or approximately 1 Bcf/d.

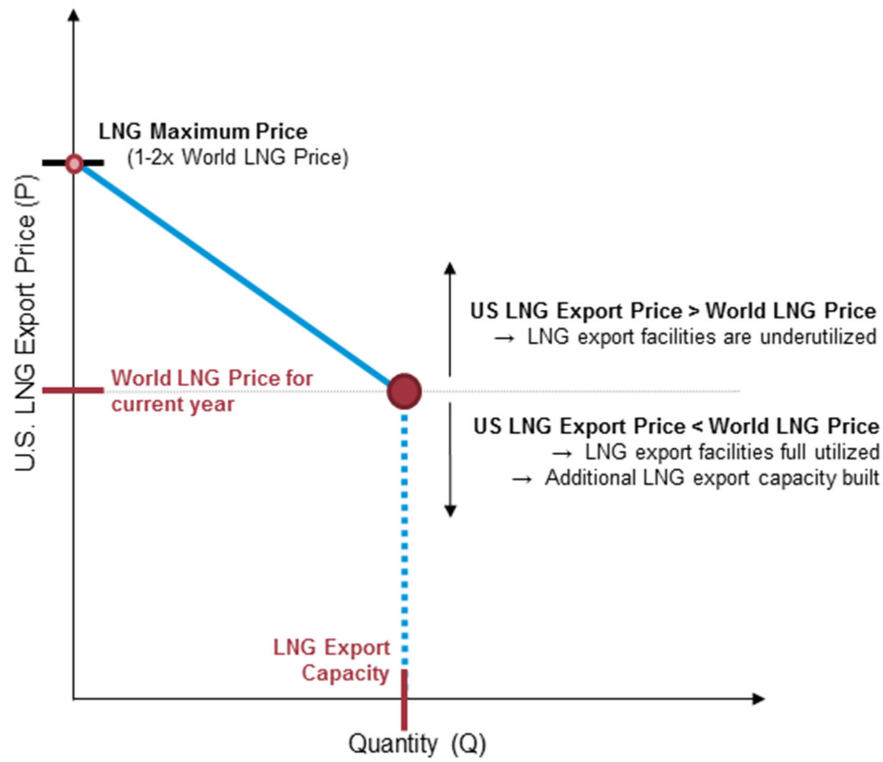
Figure 3.5 Schematic representation of tariff curve structure used in NGMM



LNG export demand curve definition

The decision to build additional LNG export capacity in a given demand region is determined outside of the quadratic program. However, the utilization of existing LNG export capacity in a projection year is endogenously determined using a linear demand curve as shown in Figure 3.6. The world LNG price is determined for the current year using equation (49) (discussed in Preprocessing: LNG Exports). If the U.S. LNG export price in that region, minus sunk costs, is at or below this value, then the LNG export capacity there will be fully utilized. If the U.S. LNG export price in the demand region is greater than the world price, however, then capacity is underutilized. At or above a certain price, which we assume is ~150% of the world LNG price, LNG exports reach zero (a consequence of the demand curve structure).

Figure 3.6 Schematic representation of LNG export demand curve used in NGMM



Objective function

The QP's solution maximizes consumer plus producer surplus, minus variable transport costs. The objective function to be maximized represents consumer plus producer surplus as the the area below the LNG export demand curves for all LNG export regions minus the area below the supply curves for all supply types and supply regions. It subtracts the variable transport costs, which include the gathering charges applied to all flows from a supply region to a hub and the area beneath the pipeline tariff curve for all flows between hubs (i.e. along transportation links).

All volumes are in billion cubic feet (Bcf), and all prices are in 1987 real dollars.

The mathematical description of the objective function is as follows with decision variables in bold font (see QP Parameters and QP Decision Variables for a description of variable names):

$$
\begin{aligned}
max.\\
& - \sum_{suptype} \sum_{qps} \sum_{step}^{SSMAX} \Bigg[\left(PBASE_{suptype,qps,step} * \boldsymbol{SSTEP_{suptype,qps,step}} \right) \\
& + \left(\frac{1}{2} * \sqrt{\boldsymbol{SSTEP_{suptype,qps,step}}} * \frac{PBASE_{suptype,qps,step+1} - PBASE_{suptype,qps,step}}{QBASE_{suptype,qps,step+1} - QBASE_{suptype,qps,step}} \right) \Bigg] \\
& - \sum_{h} \sum_{h1} \sum_{step}^{PTSMAX} \Bigg[\left(PTAR_{h,h1,step} * \boldsymbol{TAR_{h,h1,step}} \right) \\
& + \left(\frac{1}{2} * \sqrt{\boldsymbol{TAR_{h,h1,step}}} * \frac{PTAR_{h,h1,step+1} - PTAR_{h,h1,step}}{QTAR_{h,h1,step+1} - QTAR_{h,h1,step}} \right) \Bigg] \\
& - \sum_{qps} \sum_{h} P_{qps}^{gath} * \boldsymbol{FLOWS2H_{qps,h}} \\
& + \sum_{lngexp\_qp} \sum_{step}^{LSMAX} \Bigg[\left(PLNG_{lngexp\_qp,step} * \boldsymbol{LNG_{lngexp\_qp,step}} \right) \\
& + \left(\frac{1}{2} * \sqrt{\boldsymbol{LNG_{lngexp\_qp,step}}} * \frac{PLNG_{lngexp\_qp,step+1} - PLNG_{lngexp\_qp,step}}{QLNG_{lngexp\_qp,step+1} - QLNG_{lngexp\_qp,step}} \right) \Bigg]
\end{aligned}
\tag{7}
$$

Constraints

The QP constraints are described below. Decision variables are in bold font. For a further decription of individual decision variables and parameters, see QP Parameters and QP Decision Variables.

Supply Accounting (SupplyAccounting)

For all supply types and all supply regions, in a given month, the total production of a given supply type suptype in supply region qps equals the sum of the supply type under all supply steps (including the minimum production allowed).

$$
\boldsymbol{PROD_{suptype,qps}} = \sum_{step}^{SSMAX} \boldsymbol{SSTEP_{suptype,qps,step}} + QMIN_{suptype,qps} \qquad \forall (suptype, qps)
\tag{8}
$$

Supply Mass Balance (SupplyMassBalance)

For all supply types and all supply regions, in a given month, total production (or supply) of all supply types suptype must equal the flow from supply region qps to hub h.

$$
\sum_{suptype} \boldsymbol{PROD_{suptype,qps}} = \sum_{h} \boldsymbol{FLOWS2H_{qps,h}} \qquad \forall\ qps
\tag{9}
$$

Demand Mass Balance (DemandMassBalance)

For all demand regions, in a given month, the flow from hub h to demand region d must equal the sum of all sources of demand.

$$\sum_{h} \boldsymbol{FLOWH2D_{h,d}} = CONS_d + Q_d^{dist} + Q_d^{store} + Q_d^{intra} + PLT_d + DISC_d \qquad \forall\, d \tag{10}$$

Flow Balance at Hubs (HubBalance)

For all hubs, in a given month, total flow into hub h is equal to total flow out of hub h.

$$\begin{aligned}\sum_{h1}\left[\boldsymbol{FLOWH2H_{h1,h}} * \left(1 - \left(f_{h1}^{pip} + f_{h}^{pip} - f_{h1}^{pip} * f_{h}^{pip}\right) * STEO_{pip}\right)\right] + \sum_{storage} \boldsymbol{FLOWT2H_{storage,h}} \\ + \sum_{qps} \boldsymbol{FLOWS2H_{qps,h}} \\ = \sum_{h2} \boldsymbol{FLOWH2H_{h,h2}} + \sum_{d} \boldsymbol{FLOWH2D_{h,d}} + \sum_{lngexp\_qp} \boldsymbol{FLOWH2L_{h,lngexp\_qp}} \\ + \sum_{storage} \boldsymbol{FLOWH2T_{h,storage}} \qquad \forall\, h\end{aligned} \tag{11}$$

Flow Balance at Border Crossings for Pipeline Imports into the United States (HubBalance_BXtoUS)

For all border crossings, in a given month, total flow into the United States at border crossing bx is equal to the total flow out of international regions r_int into border crossing bx.

$$\begin{aligned}\sum_{r\_int}\left[\boldsymbol{FLOWH2H_{r\_int,bx}} * \left(1 - \left(f_{r\_int}^{pip} + f_{bx}^{pip} - f_{r\_int}^{pip} * f_{bx}^{pip}\right) * STEO_{pip}\right)\right] \\ = \sum_{l48} \boldsymbol{FLOWH2H_{bx,l48}} \qquad \forall\, bx\end{aligned} \tag{12}$$

Flow Balance at Border Crossings for Pipeline Exports out of the United States (HubBalance_UStoBX)

For all border crossings, in a given month, total flow out of the United States to border crossing bx is equal to the total flow from bx into international region r_int.

$$\begin{aligned}\sum_{l48}\left[\boldsymbol{FLOWH2H_{l48,bx}} * \left(1 - \left(f_{l48}^{pip} + f_{bx}^{pip} - f_{l48}^{pip} * f_{bx}^{pip}\right) * STEO_{pip}\right)\right] \\ = \sum_{r\_int} \boldsymbol{FLOWH2H_{bx,r\_int}} \qquad \forall\, bx\end{aligned} \tag{13}$$

LNG Export Demand Mass Balance (LNGExportBalance)

For all regions that have LNG export capacity, in a given month, the total demand for LNG exports, including the fuel used for liquefaction, equals the sum of flows from all hubs h to their corresponding LNG export regions lngexp_qp.

$$\sum_{step}^{LSMAX} \boldsymbol{LNG_{lngexp\_qp,step}} = \sum_{h} \boldsymbol{FLOWH2L_{h,lngexp\_qp}} \qquad \forall\ lngexp\_qp$$

(14)

Tariff Curve Quantity Balance (TariffCurveQtyBalance)

For all arcs, the flow along arc (h,h1) equals the total volume of natural gas under the tariff curve defining arc (h, h1).

$$\boldsymbol{FLOWH2H_{h,h1}} = \sum_{step}^{PTSMAX} \boldsymbol{TAR_{h,h1,step}} \qquad \forall (h, h1)$$

(15)

Storage Withdrawal Balance (StorageWthBalance)

For all storage regions, in a given month, the flow out storage equals the total withdrawals from storage.

$$\sum_{h} \boldsymbol{FLOWT2H_{storage,h}} = WTH_{storage} \qquad \forall\ storage$$

(16)

Storage Injection Balance (StorageInjBalance)

For all storage regions, in a given month, the flow into storage equals the total injections into storage.

$$\sum_{h} \boldsymbol{FLOWH2T_{h,storage}} = INJ_{storage} \qquad \forall\ storage$$

(17)

Supply Curve Range

For all supply curve steps for all supply types in all regions qps, in a given month, the quantity under the step must be between its defined minimum and maximum volume.

$$MINQ_{suptype,qps,step} \le \boldsymbol{SSTEP_{suptype,qps,step}} \le MAXQ_{suptype,qps,step} \qquad \forall (suptype, qps, step)$$

(18)

Tariff Curve Range

For all tariff curve steps for a given arc (h, h1), in a given month, the quantity under the step must be less than or equal to its maximum volume.

$$\boldsymbol{TAR_{h,h1,step}} \le MAXQT_{h,h1,step} \qquad \forall (h, h1, step)$$

(19)

LNG Export Demand Curve Range

For all LNG export demand curve steps in all regions lngexp_qp, in a given month, the quantity under the step must be between its defined minimum and maximum volume.

$$MINLNG_{lngexp\_qp,step} \leq LNG_{lngexp\_qp,step} \leq MAXLNG_{lngexp\_qp,step} \qquad \forall(lngexp\_qp, step) \tag{20}$$

Flow capacity

For all flows along arc (h, h1), for a given month, flow along (h, h1) cannot exceed its capacity.

$$\boldsymbol{FLOWH2H_{h,h1}} \leq CAP_{h,h1} \qquad \forall(h, h1) \tag{21}$$

Non-negativity

$$\boldsymbol{PROD_{suptype,qps}} \geq 0 \qquad \forall(suptype, qps) \tag{22}$$

$$\boldsymbol{FLOWH2H_{h,h1}} \geq 0 \qquad \forall(h, h1) \tag{23}$$

$$\boldsymbol{FLOWH2D_{h,d}} \geq 0 \qquad \forall(h, d) \tag{24}$$

$$\boldsymbol{FLOWH2L_{h,lngexp\_qp}} \geq 0 \qquad \forall(h, lngexp\_qp) \tag{25}$$

$$\boldsymbol{FLOWH2T_{h,storage}} \geq 0 \qquad \forall(h, storage) \tag{26}$$

$$\boldsymbol{FLOWT2H_{storage,h}} \geq 0 \qquad \forall(storage, h) \tag{27}$$

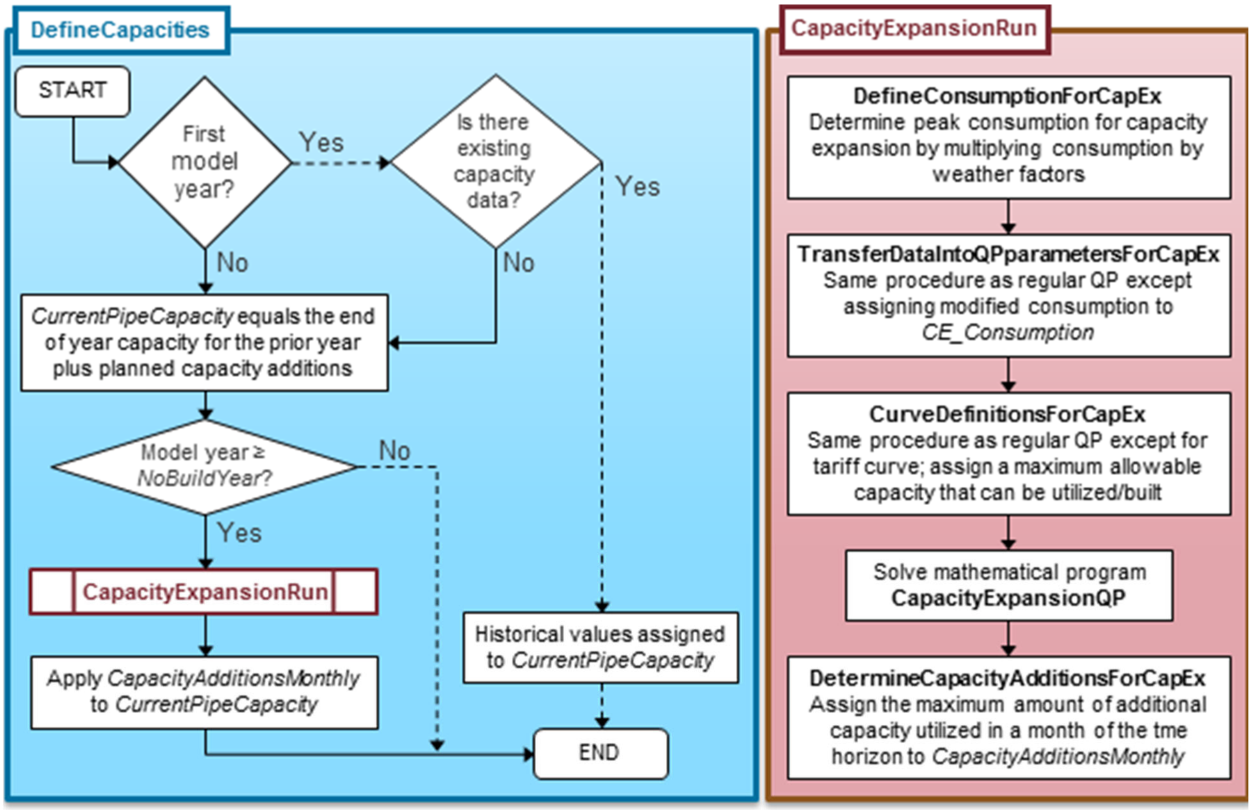
$$\boldsymbol{TAR_{h,h1,step}} \geq 0 \qquad \forall(h, h1, step) \tag{28}$$

Additional forms of mathematical program

Capacity expansion

For the first years of the projection (all years prior to *NoBuildYear*), pipeline expansion will be set to historical levels plus planned pipeline capacity expansions. These are defined as pipelines either under construction, approved by the Federal Energy Regulatory Commission (FERC), or those deemed likely to move forward. For subsequent years, before running the QP to balance supply and demand by month, the QP will be run for a peak summer and winter month to determine if the market needs and supports additional pipeline capacity. Figure 3.7 details the procedures used to assess this need.

Figure 3.7 Flow diagram of capacity assignments and capacity expansion quadratic program



The NGMM assumes a colder-than-normal winter and warmer-than-normal summer. The anticipated increase in demand is applied through two weather factors that are exogenous inputs, one by sector and one by state:

$$Q\_CAPEX_{mon,sec,st} = Q\_MONTH_{mon,sec,st} * f^1{}_{mon,sec,st} * f^2{}_{mon,sec} \qquad \forall\ (mon, sec, st) \tag{29}$$

where

$Q\_CAPEX_{mon,sec,st}$ = augmented consumption for sector sec, month mon, and state st, after accounting for extreme weather(Bcf)

$Q\_MONTH_{mon,sec,st}$ = consumption for sector sec, month mon, and state st (Bcf)

$f^1{}_{mon,sec,st}$ = weather factor applied to sector sec, month mon, and state st, to reflect increases in consumption due to extreme weather

$f^2{}_{mon,sec}$ = weather factor applied to sector sec and month mon, to reflect increases in consumption due to extreme weather

mon = projection month

sec = end use sector

st = lower 48 state (+ DC)

The assumption is that while the geographic location of a given state determines how much consumption changes (e.g., for residential heating needs), there also may be a sector-wide impact that is equally felt amongst all states (i.e. warmer weather increases electric power sector consumption across all states). All other QP parameters are the same and transferred using the planning horizon assumed for capacity expansion except for the maximum allowable step on the tariff curve (*PTCrv_MaxStep*), which is increased by one to allow for additional capacity to be built.

This approach will allow pipeline capacity to be added incrementally as needed in each projection year. While this approach only loosely represents how expansion projects are built, it is a reasonable approximation. Since pipeline capacity is being added to satisfy current year needs, the assumption is that the actual projects were planned ahead of time, as necessary, in anticipation of the future need.

STEO benchmarking

A basic requirement of the NGMM is that it produce model results for the STEO years[32] which are within 2% of the STEO results from a selected publication. Most of the STEO projections are national numbers, except for regional delivered prices to residential, commercial, and industrial customers. While many of the STEO values can be benchmarked with a straightforward additive or multiplicative factor to model input or output, there are several exceptions due to the interdependence of the NGMM model results:

- *Henry Hub spot price, which is assigned the shadow price of the hub mass balance constraint*
- *Pipeline fuel consumption, which depends on decision variable FlowHubToHub*
- *Lease fuel consumption, which depends on the decision variables QSupplyStep and FlowSupplyToHub*
- *Gross pipeline imports and exports, which are assigned using the decision variable FlowHubToHub to and from border crossing nodes and depend on several other model outputs*
    - *Canadian nonassociated production, specifically from western Canada*
    - *Mexican nonassociated production in the Northeast region*
    - *The decision to flow gas via the TransCanada pipeline from western to eastern Canada versus through the Midwestern United States*
    - *The competitiveness in eastern Canada of natural gas from the Appalachia basin with western Canadian natural gas*

To benchmark these results to STEO, the QP is executed over multiple iterations, adjusting the STEO correction factors based on the difference between the target STEO value and the model solution, until convergence is achieved for all interdependent values. These factors are then applied in the NGMM. This may be done while

- *Implementing the pre-processing routines, if the adjustment is made to NGMM model inputs*
- *Solving the QP, if the adjustment is made directly to a quantity within the QP formulation*

[32] STEO years are defined as all years for which the most recently released STEO available during AEO modeling efforts publishes its forecast (*NumberofSTEOYears*). In generally, this usually corresponds to the publication released 3-4 months prior to the AEO public release.

- *Completing the post-processing routines, if the adjustment is made to a report variable for other NEMS modules*

After the last STEO year, the STEO benchmark factors will be phased out over a given number of years (*NumberOfSTEOPhaseOutYears*). Table 3.6 lists the various STEO factors used in the NGMM. The flow process diagram demonstrating this algorithm is shown in Figure 3.8.

Table 3.6 STEO factors calculated in the NGMM during STEO benchmarking

STEO Factor[a]	NGMM Parameter	Applied to indices	Type
Applied during data preprocessing			
STEOStorageWithdrawalFactor	StorageWithdrawals	(mon,l48)	*
STEOStorageInjectionFactor	StorageInjections	(mon,l48)	*
STEOSupplementalSupplyFactor	Supply	(mon, sng, l48)	*
STEOLNGImportsFactor	Supply	(mon, 'LNG', l48)	*
STEOLNGExportsFactor	LNGExports	(mon,lngexp_l48)	±
Applied in the QP			
STEOMXExportFactor	TotalConsumption	(mon, 'Electric', 'MX_NE')	±
STEOLeaseFuelFactor	LeaseFuelFactor	qps	*
STEOPipelineImportFactor	Supply	(mon, 'NA', 'CN_W')	*
STEOWellhdPriceFactor	WellhdPrice	(year,suptype,qps)	*
STEOPipeFuelFactor	TranFuelLosses	(mon,h)	*
Applied during data postprocessing			
STEOCNExportFactor	Exports_Canada	year	*
STEOMXExportFactor	Exports_Mexico	year	±
STEOEndUsePriceFactor(sec,r_cen)	Price_Enduse	(year, sec, r_cen)	*
STEOElectricPriceFactor	AveragePrice_EnduseElectric	year	*

[a] All STEO factors have a year dimension. After last STEO phaseout year, all additive factors are 0 and multiplicative factors are 1.

Exceptions to STEO Benchmarking

There are several cases where the NGMM does not benchmark directly to a given STEO value due to definitional differences. In some of these cases, additional assumptions are required while in others no benchmarking is applied.

In STEO, as well as in EIA historical data, pipeline fuel use includes not only natural gas consumed during transport, but also natural gas used during the liquefaction of LNG exports. Within the NGMM, however, pipeline fuel consumption refers to natural gas specifically used or lost during transmission and distribution throughout the pipeline network. The natural gas used during the liquefaction of natural gas for exports is accounted for separately and assigned to the industrial sector. In order to benchmark to the STEO results, the NGMM assumes a specific percentage of natural gas is consumed during liquefaction (*Pct_Liquefaction_Fuel*). This volume is subtracted from the STEO target value prior to calculating the benchmark factor for pipeline fuel use.

Additional assumptions are also required for benchmarking U.S. natural gas imports and exports by pipeline. Since the STEO does not distinguish between U.S. pipeline trade for Canada and Mexico, only reporting gross volumes, the NGMM makes assumptions about the volume shares for the two countries. For the case of gross U.S. natural gas pipeline exports, a user-defined input parameter (*STEOCNExportPercent*) sets the STEO target export volumes to Canada and Mexico. The NGMM assumes all U.S. pipeline imports are from Canada during STEO years.

The NGMM does not apply any benchmark factor to the delivered price to industrial consumers. The STEO forecast for this price relies on historical data from EIA's *Natural Gas Annual*, which only surveys 15% of the market.[33] Furthermore, it is likely not surveying the largest energy-intensive consumers buying natural gas from the spot market, but smaller customers behind local distribution companies (LDCs). Benchmarking to the STEO value would therefore likely over-estimate industrial prices.

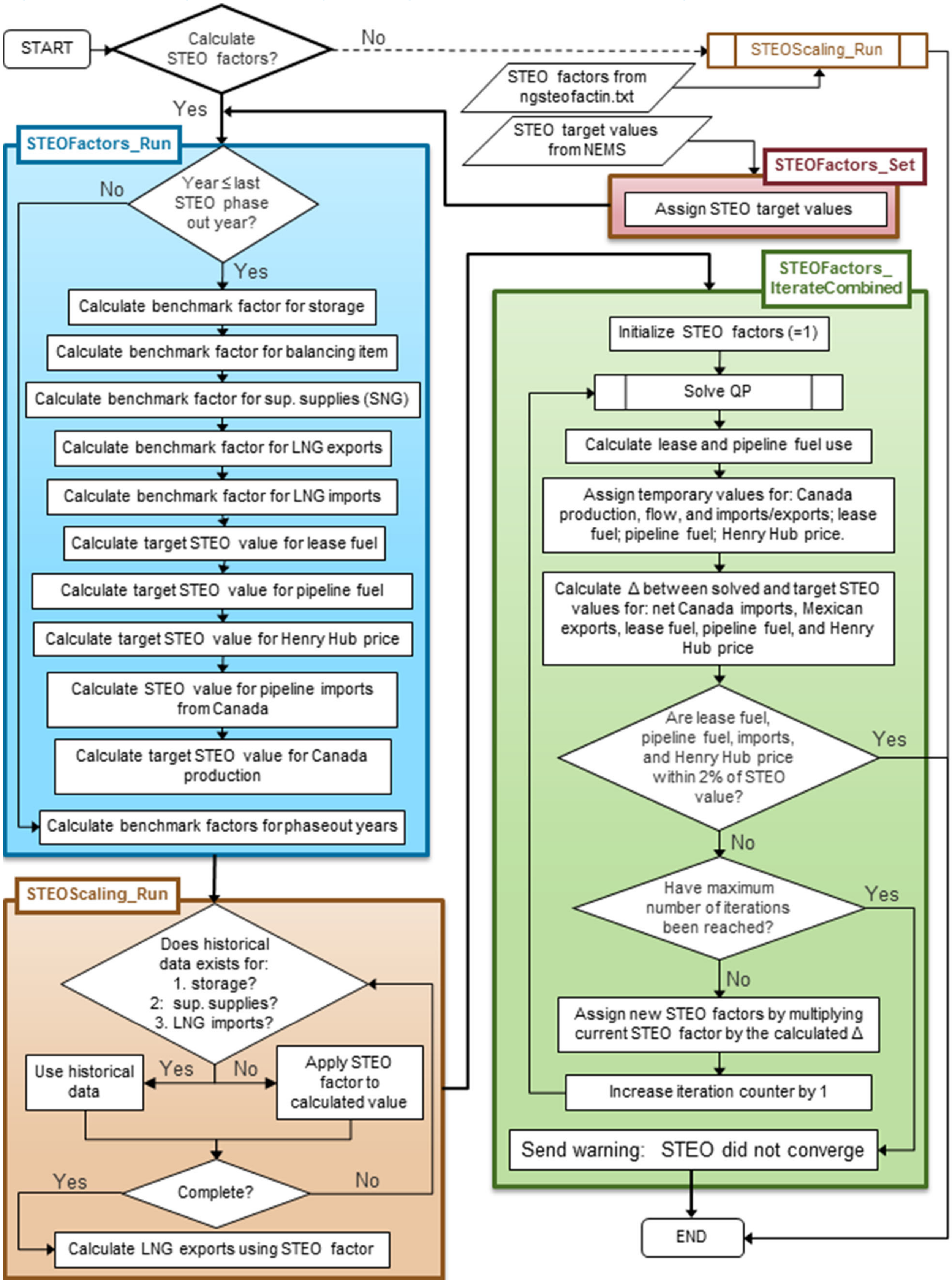
There are other examples where STEO benchmarking is not performed; however, in these cases it is because of the small volumes of natural gas concerned (i.e. the transportation sector).

Reading in STEO factors

When running side cases, it is not desirable to benchmark to the STEO values; rather, we want to see how the model results change during this time period due to the alternative assumptions. To allow this feature, the STEO benchmarking factors are written out to a text file (ngsteofactin.txt) during the NEMS reporting loop in the last STEO phase out year. Using the runtime option STSCALNG, the user can determine whether the STEO factors are calculated during a given run (STSCALNG=0) or read in (STSCALNG=1). The factors calculated and used during an AEO Reference case run can be applied to all side case runs.

[33] U.S. Energy Information Administration, *Natural Gas Annual*, Natural Gas Prices

Figure 3.8 Flow diagram describing the assignment of STEO benchmarking factors in the NGMM



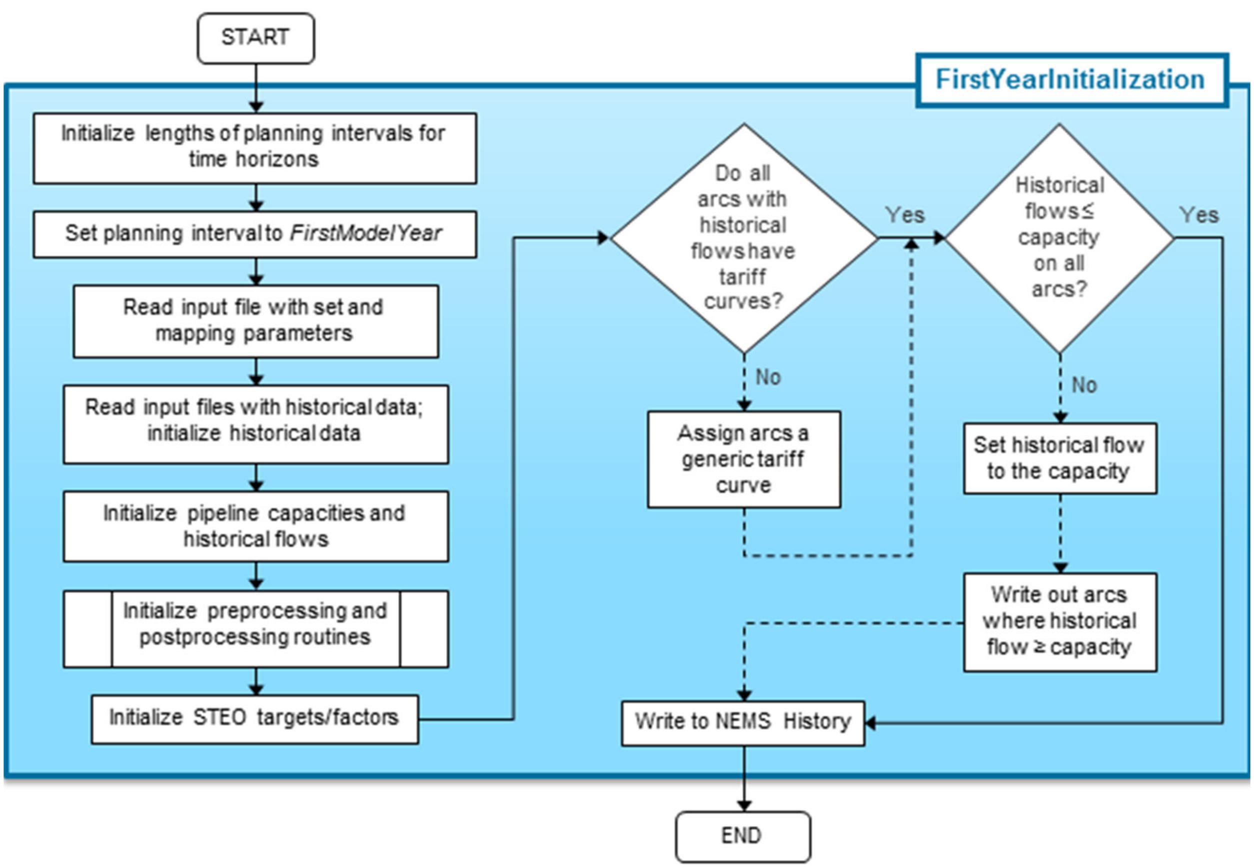
4. Pre-processing routines

This chapter describes all of the pre-processing routines required to send parameters to the quadratic program (QP). This includes the routines to convert data from other National Energy Modeling System (NEMS) modules into the state and monthly levels at which the Natural Gas Market Model (NGMM) represents supply and demand, as well as the routines which represent international markets for both pipeline and liquefied natural gas (LNG) trade. For all variables defined in this chapter, Appendix C provides a reference to the full identifier name used in the NGMM, Appendix D identifies where to find specific equations within the model code, and Appendix E indicates which input files contain input assumptions or parameters.

First year initialization

When called by the NEMS in the first model year, the NGMM needs to initialize model parameters and its structure. This is done in the procedure *FirstYearInitialization*, which is shown in Figure 4.1. After AIMMS is opened by the NEMS and global data is read in from a text file for all historical years, the NGMM first structures and initializes the time horizon according to the timescales and planning intervals defined in hard-coded parameters (i.e. these are defined in the AIMMS interface directly). Set definitions—the dimensions used for all parameters, decision variables, and constraints— and mapping parameters that define relationships between set elements are read in from an input text file.

Figure 4.1 Flow chart of first year initialization



After the model structure is in place, historical data, model parameters, and assumptions are read in from input text files. Historical pipeline capacities, as well as planned capacity additions, are fixed, which define the allowable transportation links along which natural gas can flow. In addition, historical data are processed in order to calculate historical averages, historical shares, and other historical trends that are used during data pre- and post-processing. This includes assigning factors and target values that are used to benchmark NGMM projections to the forecast published in *the Short Term Energy Outlook* (STEO).

As a quality check on the input data, the NGMM verifies that all historical pipeline flows are along arcs that have pipeline capacity, as well as ensuring that total annual flow volumes do not exceed that capacity. For flows that violate these conditions, the NGMM adds sufficient capacity to the historical data and records any changes made in an output text file. Finally, the NGMM writes out historical values for any global variables the model is responsible for assigning to a text file, which is read in by NEMS.

A further description of all associated input text files and their contents is in Appendix E.

Supply

There are 6 types of supply represented in the NGMM: nonassociated (NA) gas, associated-dissolved (AD) gas, liquefied natural gas (LNG) imports, synthetic natural gas (SNG) from coal (SNG_coal), synthetic natural gas from liquids (SNG_liq), and other synthetic natural gas (SNG_oth). All supply regions, including those in Canada and Mexico, can have any number of these supply types. Only NA gas is considered a variable supply (i.e., it is solved for in the mathematical program and allowed to change dynamically in response to the supply price in a given region). The structure of the supply curves is discussed in relation to the setup of the QP. The supply levels for the remaining categories are fixed at the beginning of each projection year (i.e., before market clearing prices are determined). Alaskan natural gas supply is not represented in the QP; rather, it is set to equal the projected consumption as determined in the Alaska pre-processing routine.

Nonassociated and associated-dissolved natural gas production

Nonassociated (NA) natural gas is largely defined as gas that is produced from gas wells and is assumed to vary in response to a change in the natural gas price. Associated-dissolved (AD) gas is defined as gas that is produced from oil wells and can be considered a byproduct of the oil production process.[34] For the United States and Canada, expected NA and AD natural gas production volumes are provided to the NGMM from the Oil and Gas Supply Module (OGSM) in the NEMS. Canada is represented as 2 regions: western Canada (comprising largely the Western Canadian Sedimentary Basin) and eastern Canada. The algorithm for setting annual production in Mexico is handled in the NGMM, and monthly levels are set by assuming no variation throughout the year.

While natural gas supply activities within OGSM are modeled at a highly disaggregate level (generally by county), U.S. dry gas production levels are provided to the NGMM at the oil and gas district level, or 84 regions. All 66 onshore regions represent a single state or are contained within a single state. In such cases, gas is assumed to initially flow to the associated state transshipment node in the NGMM network.

[34] While AD natural gas production in NGMM does not respond to natural gas prices in the short term, the evaluation of oil-directed projects, which are also expected to produce gas, takes into account the expected price of natural gas over the long term. This assessment occurs in OGSM where AD gas production is set.

Table 4.1 Mapping offshore regions to Lower 48 states

Offshore Region(s)	Designated State
North Atlantic Federal and State	Massachusetts
Mid Atlantic Federal and State	Virginia
South Atlantic Federal and State	Georgia
Northern Pacific Federal and State	Oregon
California Federal and State	California
Alabama State	Alabama
Louisiana State	Louisiana
Texas State	Texas

For the offshore regions, areas are first distinguished by state and federal waters in the Atlantic, Pacific, and Gulf of Mexico, and are further disaggregated into 2 or 3 regions (18 in total). State and federal waters for a given state are considered separate regions. Production from these regions must be tied to particular state nodes in the NGMM, as the QP is solved at the state level. With the exception of the federal waters of the Gulf of Mexico, the other offshore regions are mapped to a single state as shown in Table 4.1.

The federal waters of the Gulf of Mexico are represented as 3 regions (East, Central, and West) and are mapped to states in the NGMM (i.e., reflecting where the produced gas will flow) using average historical shares (*NumberOfYearsForAverage_GOMprod*); they are held constant throughout the projection. They are calculated using the following historical data: production in the 3 regions (as provided by the OGSM), flows from the Gulf of Mexico to the four adjoining states (Texas, Louisiana, Mississippi, and Alabama), and lease fuel consumed in the Gulf of Mexico. Lease fuel is allocated to the states throughout the projection using the proportions implied by the historical flows. Initial historical production estimates for the Gulf of Mexico by receiving state—not by the 3 regions represented in OGSM—are set to the historical flow levels plus the assigned state level lease fuel. These estimated production volumes are then allocated to the 3 Gulf of Mexico regions using the following assumptions:

- *All of the east Gulf of Mexico production is assumed to flow to Alabama*
- *Any remaining flow to Alabama is assumed to come from central Gulf of Mexico*
- *All of the flows to Mississippi are assumed to come from central Gulf of Mexico*
- *If the initial production estimates for flow to Texas exceed the west Gulf of Mexico production, all of the west Gulf of Mexico production is assumed to flow to Texas, the rest of the flow into Texas is assumed to come from central Gulf of Mexico, and all of the flow into Louisiana is assumed to come from central Gulf of Mexico*
- *If the initial production estimates for flow to Texas are less than the west Gulf of Mexico production, all of the flow to Texas is assumed to come from the west Gulf of Mexico and the remaining production from the west is assumed to flow to Louisiana, with any additional flows into Louisiana coming from the central Gulf of Mexico.*

From the historical data and the assumptions above, shares for the central and western Gulf of Mexico by state (α_{st} and β_{st}) are calculated and remain constant. In the projection period, production in the three Gulf of Mexico regions is allocated to supply nodes for each state in the NGMM using the following equations:

$$PROD\_GOM_{AL} = PROD_{EGOM} + \alpha_{AL} * PROD_{CGOM} \tag{30}$$

$$PROD\_GOM_{MS} = \alpha_{MS} * PROD_{CGOM} \tag{31}$$

$$PROD\_GOM_{LA} = \alpha_{LA} * PROD_{CGOM} + \beta_{LA} * PROD_{WGOM} \tag{32}$$

$$PROD\_GOM_{TX} = \alpha_{TX} * PROD_{CGOM} + \beta_{TX} * PROD_{WGOM} \tag{33}$$

where

$PROD\_GOM_{st}$ = implied historical natural gas production from the federal waters of the Gulf of Mexico for a given state (billion cubic feet, Bcf)

$PROD_G$ = natural gas production assigned to Gulf of Mexico region $G \in \{ECOM, CGOM, WGOM\}$ (Bcf)

α_{st} = historical share of central Gulf of Mexico production assigned to state $st \in \{AL, MS, LA, TX\}$ (Bcf)

β_{st} = historical share of western Gulf of Mexico production assigned to state $st \in \{LA, TX\}$ (Bcf)

Supplemental Supplies

Existing sources for synthetically produced pipeline-quality natural gas (SNG) and other supplemental supplies are assumed to continue to produce at historical levels. These include the following supply types:

- *Synthetic natural gas from coal (SNG_coal): SNG from the Great Plains Coal Gasification Plant in North Dakota, which is assumed to operate indefinitely throughout the projection; the NGMM does not allow for building of new coal-to-gas facilities.*
- *Synthetic natural gas from liquid hydrocarbons (SNG_liq): SNG is no longer produced from liquid hydrocarbons in the continental United States, although small amounts were produced in Illinois in some historical years; the small amount produced in Hawaii is included in California. supply/demand balancing as implicit consumption (i.e., they are not subtracted out of the Pacific Census division totals as is done with Alaska).*
- *Other supplemental supplies (SNG_oth): EIA defines other supplemental fuels as propane-air, coke oven gas, refinery gas, or biomass gas that is British thermal unit (Btu)-stabilized with steam or oxygen to manufacture pipeline-quality gas that enters the distribution network.*

Projected values for all three types of supplemental supplies are set at historical averages and held constant over the projection period. The number of years used to calculate this average is an input parameter (*NumberOfYearsForAverage_SNG_*). These volumes are assumed constant throughout the year when setting monthly levels.

LNG imports

In the prior natural gas model, the Natural Gas Transmission and Distribution Module (NGTDM), LNG imports were endogenously calculated. The algorithm for determining LNG import volumes was designed prior to the growth of domestic natural gas production from shale gas and tight oil formations. While LNG imports peaked at 770 Bcf in 2007, since 2013 they have remained relative flat and only averaged 84 Bcf per year through 2017. Cargoes primarily have gone to Everett, Massachusetts, where the LNG terminal operates in conjunction with the Mystic Generating Station power plant; sporadic deliveries have also gone to Cove Point, Maryland and Elba Island, Georgia during periods of peak natural gas demand. Therefore, the NGMM currently sets LNG imports for the projection period exogenously at the prior year's levels.

Demand

End-use consumption

Within the NGMM, natural gas demand in the United States is represented for the five primary consuming sectors—residential, commercial, industrial, electric generators, and transportation—based on projected consumption values set in the NEMS demand modules. For each NEMS iteration and projection year, the demand modules in the NEMS determine the level of natural gas consumption for each region and customer class given the delivered natural gas prices for the sector, as calculated by the NGMM in the previous NEMS iteration, and relevant outputs from other NEMS modules. In turn, the projected prices from the NGMM to supply these consumption levels are passed back to the appropriate demand module during the next NEMS iteration to reevaluate the consumption levels. The NEMS run is declared converged when the delivered prices and quantities for all fuels are within a user-specified tolerance from one iteration to the next.

In theory, the NGMM could represent demand using demand curves (i.e., approximate the demand response to a change in the price). Currently, domestic consumption is held constant in the NGMM in each NEMS iteration. If demand curves were to be used within the NGMM in NEMS, they would be built off of the price/quantity pairs from the previous NEMS iteration and included in the objective function of the QP.

For all but the electric sector, the NGMM disaggregates annual Census division consumption levels into the state and monthly representation that the NGMM requires. The regional representation for the electric generation sector differs from the other NEMS sectors because the Electricity Market Module (EMM) solves internally by North American Electric Reliability Corporation (NERC)-based regions for three seasons in each year, enabling a more disaggregated representation of consumption in the NGMM. Within the EMM, assumptions are made to translate their NERC-based regions, which do not always align with state borders and generally do not share common borders with the Census divisions, to 17 regions that do. This is done based on the relative geographic location (and natural gas-fired generation capacity) of the current and proposed electricity generation plants within each region. Within the NGMM, electric consumption by these 17 regions (the last of which is Alaska) and 3 seasons (peak, offpeak, and shoulder) is disaggregated to the state and monthly representation that NGMM requires.

NGMM disaggregates regional demands annually/seasonally using historical state and monthly shares. These shares remain constant throughout the projection period. The number of years used to calculate this average is an input parameter (*NumberOfYearsForAverage_Demand_*). For the Pacific Division in all sectors except electric power generation, natural gas consumption estimates for Alaska are first subtracted to establish a consumption level for just the contiguous Pacific Division before the historical share is applied. The consumption of gas in Hawaii is not handled separately because it is considered negligible.

Lease and plant fuel

The consumption of lease and plant fuel is calculated in the NGMM. Lease and plant fuel is defined as natural gas used in well, field, and lease operations (such as gas used in drilling operations, heaters, dehydrators, and field compressors) and fuel used in processing plants. For lease fuel, the NGMM calculates the average percentage of dry gas production that is consumed in lease operations for all OGSM regions over a defined number of years (*NumberOfYearsForAverage_LeaseFuel_*). These region-specific factors (*LeaseFuelFactor*) are then applied to the realized dry gas production in the QP's objective function to account for lease fuel consumption. They are also used to set lease fuel consumption for report writing purposes during postprocessing. All of the subregions in a given state or within the Gulf of Mexico are assigned the same factor. Offshore regions outside of the Gulf of Mexico (e.g., California) are included with onshore production for these calculations and are assigned the same factor as the rest of the state. These factors remain constant throughout the projection.

Plant fuel in the NGMM is assumed to be related to the volume of natural gas plant liquids (NGPL) processed in a given state. Historical and projected NGPL production is provided by the NEMS (via OGSM) by the OGSM regions and assigned to a state for processing in NGMM based on exogenously specified average historical shares (*PercentOfProductionMovedForPlantFuel*). Using historical NGPL production and these shares, the NGMM calculates the average plant fuel consumed per unit of total NGPL processed over an assumed number of years (*NumberOfYearsForAverage_PlantFuel_*) in a given state. In the projection years, state-level average factors (*PlantFuelFactor*) are multiplied by the total NGPL processed in-state to project plant fuel consumption. To control for some anomalies in the data, these factors are limited to a user-specified range.[35] States that historically do not process NGPL are assigned a national average factor. Within the QP, plant fuel is a fixed consumption level.

Pipeline fuel

Natural gas consumed in the operation of pipelines is accounted for in the NGMM. The module assumes this volume has 4 components:

- *natural gas used in the distribution pipeline network (i.e., local distribution companies, or LDCs)*
- *natural gas used in injecting and withdrawing volumes from storage*
- *natural gas used in intrastate transmission*
- *natural gas used in interstate transmission*

Because the NGMM represents the pipeline network at the state level, it only solves for flows on interstate pipelines. To account for the other three components of pipeline fuel consumption, the NGMM uses assumed factors (see Appendix E) and projects their consumption as follows:

[35] In AEO 2018, the allowed range for plant fuel factors was between 0.005 and 1.7.

$$PIP\_DIST_{mon,st} = \left(Q\_RES_{mon,st} + Q\_COM_{mon,st}\right) * PCT\_DIST_{st} \tag{34}$$

$$PIP\_STORE_{mon,st} = \left(STORE\_INJ_{mon,st} + STOR\_WTH_{mon,st}\right) * PCT\_STORE_{st} \tag{35}$$

$$PIP\_INTRA_{mon,st} = PROD\_DRY_{mon,st} * PCT\_INTRA_{st} \tag{36}$$

where

$PIP\_DIST_{mon,st}$ = natural gas consumed for distribution pipeline transportation in state st and month mon (Bcf)

$PIP\_STORE_{mon,st}$ = natural gas consumed injecting and withdrawing volumes from storage in transportation in state st and month mon (Bcf)

$PIP\_INTRA_{mon,st}$= natural gas consumed for intrastate pipeline transportation in state st and month mon (Bcf)

$Q\_RES_{mon,st}$ = consumption of natural gas by residential customers in state st and month mon (Bcf)

$Q\_COM_{mon,st}$ = consumption of natural gas by commercial customers in state st and month mon (Bcf)

$STORE\_INJ_{mon,st}$ = volume of natural gas injected into storage for state st and month mon (Bcf)

$STOR\_WTH_{mon,st}$ = volume of natural gas withdrawn from storage for state st and month mon (Bcf)

$PROD\_DRY_{mon,st}$ = total dry production of natural gas in state st and month mon (Bcf)

PCT_DIST_{st} = fraction of natural gas consumption by residential and commercial customers consumed or lost during transportation in distribution pipeline network for state st, based on respondent-level historical data from EIA's Form EIA-176 survey

PCT_STORE_{st} = fraction of total natural gas injections and withdrawals consumed or lost during transportation to or from storage facility for state st, based on respondent-level historical data from EIA's Form EIA-176 survey

PCT_INTRA_{st} = fraction of total dry production consumed during transportation in intrastate pipeline network in state st (this assumes intrastate pipelines primarily serve as gathering lines and to transport natural gas from processing plants to interstate pipeline system)[36]

mon = projection month

[36] Currently not assigned a value, so effectively set to zero pending the development of a basis for assigning or assuming a value.

st = Lower 48 states

To account for pipeline fuel used in the transportation of natural gas during transmission on the interstate pipeline system (as well as the associated cost), a loss factor needs to be assigned to each arc in the network. However, since data are available for pipeline fuel use on a state level, and flows are reported as one arc between states and measured at the border between them, the model breaks up the arcs from state x to state y into two arcs (state x to the border between x and y and from the border to state y). This representation assumes that the percentage lost as a function of the starting flow (i.e. the volume prior to fuel loss) on each arc segment within a state has the same value (P_LOSS_{st}). So for the arcs flowing from a state's border into the state transshipment node, the associated interstate pipeline fuel use is P_LOSS_{st} multiplied by the flow entering at the border ($FLOW_{st\_from,st}$). Whereas, for the arcs flowing from a state's transshipment node to the state's border, the associated interstate pipeline fuel use is the flow exiting at the border ($FLOW_{st,st\_to}$) multiplied by a factor that corrects for the fuel lost prior to being measured at the border [P_LOSS_{st} / (1 - P_LOSS_{st})]. Thus, the total interstate pipeline fuel used in each state equals:

$$PIPE\_TRANS_{st} = P\_LOSS_{st} * FLOW\_IN_{st} + \frac{P\_LOSS_{st}}{1 - P\_LOSS_{st}} * FLOW\_OUT_{st} \tag{37}$$

where

$$FLOW\_IN_{st} = \sum_{st\_from}^{lower\ 48} FLOW_{st\_from,st} \tag{38}$$

$$FLOW\_OUT_{st} = \sum_{st\_to}^{lower\ 48} FLOW_{st,st\_to} \tag{39}$$

$lower\ 48$ = Lower 48 states

st_from = Lower 48 state from which natural gas is flowing into a given state

st_to = Lower 48 state into which natural gas is flowing from a given state

And it therefore follows mathematically that the interstate pipeline loss factor for each state, based on historical data, can be calculated as

$$TOTAL_{st} = FLOW\_IN_{st} + FLOW\_OUT_{st} + PIP\_TRANS_{st} \tag{40}$$

After substituting $FLOW\_OUT_{st}$ with its equivalent expression according to equation (37), the following quadratic eqation with respect to pipeline fuel loss results:

$$FLOW\_IN_{st} * \left(P_{LOSS_{st}}\right)^2 - TOTAL_{st} * P_{LOSS_{st}} + PIP\_TRANS_{st} = 0$$

(41)

The quadratic formula is then used to solve for pipeline fuel loss:

$$P\_LOSS_{st} = \frac{TOTAL_{st} + \sqrt{(TOTAL_{st})^2 - 4 * FLOW\_IN_{st} * PIP\_TRANS_{st}}}{2 * FLOW\_IN_{st}}$$

(42)

where

P_LOSS_{st} = pipeline fuel loss factor for state st

$FLOW_{st\_from,st}$ = flow into state st from state st_from along an arc (Bcf)

$FLOW_{st,st\_to}$ = flow out of state st into state st_to along an arc (Bcf)

$FLOW\_IN_{st}$ = total of all flows into state st (Bcf)

$FLOW\_OUT_{st}$ = total of all flows out of state st (Bcf)

PIP_TRANS_{st} = total pipeline fuel used in state st (Bcf)

$TOTAL_{st}$ = total volumes flowing into and out of state st plus total pipeline fuel used (Bcf)

st = Lower 48 states

st_from = Lower 48 state from which natural gas is flowing into state st

st_to = Lower 48 state into which natural gas is flowing from state st

Pipeline fuel loss factors for historical data (after subtracting distribution, storage, and intrastate transportation losses) are averaged over a specified number of years to arrive at those used in the NGMM (*NumberOfYearsForAverage_Trans_)*. These loss factors are applied in the QP to account for the quantity lost on the interstate arcs and to effectively account for the cost of the fuel used. In addition, the loss factors are applied to the flows during post-processing to arrive at the total pipeline fuel used in interstate transmission. The four components are then added together, giving the total pipeline fuel used. Pipeline fuel is not independently accounted for in Canada or Mexico due to lack of historical data.

Balancing item

The NGMM also includes a balancing item, or discrepancy, to reflect the average historical difference between supply and demand. The historical balancing item is often consistently positive or negative, indicating a segment of the natural gas market is not being captured in the data. By excluding it, the projected supply values would not align properly with historical values. In the NGMM, the balancing item is held constant throughout the projection and is set for all demand hubs in the model according to the average difference in supply and demand over an assumed number of historical years. This number is a user-defined value specified in an input file (*NumberOfYearsForAverage_Discrepancy_* , *NumberOfYearsForAverage_Discrepancy_CN_):*

$$DISC_{hyr,h} = Q\_TOT_{hyr,h} + Q\_PIP_{hyr,h} + Q\_LAP_{hyr,h} + (1 + PCT\_LIQ) * LNG\_EXP_{hyr,h} + STORE\_INJ_{hyr,h} - STOR\_WTH_{hyr,h} - SUP\_TOT_{hyr,h} + FLOW\_OUT_{hyr,h} - FLOW\_IN_{hyr,h} \quad (43)$$

where

$DISC_{hyr,h}$ = discrepancy, or balancing item, for natural gas hub h and historical year hyr (Bcf)

$Q\_TOT_{hyr,h}$ = total end use consumption for natural gas hub h and historical year hyr (Bcf)

$Q\_PIP_{hyr,h}$ = total pipeline fuel use for natural gas hub h and historical year hyr (Bcf)

$Q\_LAP_{hyr,h}$ = total lease and plant fuel use for natural gas hub h and historical year hyr (Bcf)

$LNG\_EXP_{hyr,h}$ = total LNG exports out of natural gas hub h for historical year hyr (Bcf)

PCT_LIQ = percent of fuel used for liquefaction in export facilities

$STORE\_INJ_{hyr,h}$ = total storage injections for natural gas hub h and historical year hyr (Bcf)

$STOR\_WTH_{hyr,h}$ = total storage withdrawals for natural gas hub h and historical year hyr (Bcf)

$SUP\_TOT_{hyr,h}$ = total supply, including production, supplemental supplies, and LNG imports, for natural gas hub h and historical year hyr (Bcf)

$FLOW\_IN_{hyr,h}$ = total flow into natural gas hub h from pipeline network during historical year hyr (Bcf)

$FLOW\_OUT_{hyr,h}$ = total flow out of natural gas hub h from pipeline network during historical year hyr (Bcf)

hyr = historical year

h = natural gas hub in the NGMM that defines Lower 48 states or Canadian region

As data are not available for flows within Mexico, the balancing item is calculated for the whole country using imports and exports as the flows. That value is then divided by the number of Mexican regions, applying the same balancing item to all of them. Monthly values are then assigned according to the number of days in the month. For Alaska, the balancing item is not calculated since it is set outside of the QP. It is set to the historical average over a given number of years (*NumberOfYearsForAverage_Discrepancy*) as reported by the *Natural Gas Annual*.

Storage

Storage is represented in the NGMM for all Lower 48 states and Canada. While storage is an integral part of balancing natural gas markets in the short term to mitigate price increases during periods of peak demand, over the long term it is not expected to play a role in setting prices. The NGMM assumes that net storage withdrawals over a projection year equal zero, i.e., storage injections equal storage withdrawals at each hub.

To establish historically-based storage injections and withdrawals by month and storage region (state and Canada region) for use in the projection period, the NGMM starts by calculating the average injections and withdrawals (*NumberOfYearsForAverage_Storage_*) over a user-specified number of years for each month and region, with the intent of arriving at normalized levels. However, since the average injections do not exactly equal the average withdrawals, an adjustment is made to these monthly/regional averages to insure that net storage withdrawals over the year equal zero for each storage region:

$$\alpha_{storage} = \frac{AVE\_YR\_INJ_{storage} - AVE\_YR\_WTH_{storage}}{AVE\_YR\_INJ_{storage} + AVE\_YR\_WTH_{storage}} \tag{44}$$

$$\overline{AVE\_INJ}_{mon,storage} = \left(1 - \alpha_{storage}\right) * AVE\_INJ_{mon,storage} \tag{45}$$

$$\overline{AVE\_WTH}_{mon,storage} = \left(1 + \alpha_{storage}\right) * AVE\_WTH_{mon,storage} \tag{46}$$

where

$AVE\_INJ_{mon,storage}$ = average monthly (Jan-Dec) storage injections for storage region $storage$ over assumed range of historical years for month mon (Bcf)

$\overline{AVE\_INJ}_{mon,storage}$ = corrected average monthly (Jan-Dec) storage injections for storage region $storage$ over assumed range of historical years for month mon (Bcf)

$AVE\_WTH_{mon,storage}$ = average monthly (Jan-Dec) storage withdrawals for storage region $storage$ over assumed range of historical years for month mon (Bcf)

$\overline{AVE\_WTH}_{mon,storage}$ = corrected average monthly (Jan-Dec) storage withdrawals for storage region $storage$ over assumed range of historical years for month mon (Bcf)

$AVE\_YR\_INJ_{storage}$ = average total annual storage injections for storage region $storage$ over assumed range of historical years for month mon (Bcf)

$AVE\_YR\_WTH_{storage}$ = average total annual storage withdrawals for storage region $storage$ over assumed range of historical years for month mon (Bcf)

$\alpha_{storage}$ = storage adjustment factor for storage region $storage$

$storage$ = storage region (Lower 48 states plus Canadian regions)

mon = month of the year (Jan-Dec)

The resulting storage withdrawals and injections by month and storage region are assumed constant throughout the projection period with the exception of the STEO years. The assumption that historical storage injection and withdrawal patterns will continue throughout the projection could be modified in the pre-processing step if warranted. However, the QP as currently formulated requires preset storage activity levels and cannot dynamically solve for storage injections and withdrawals.

Canada

The NGMM respresents two hubs in Canada: eastern Canada (Ontario, Quebec, Manitoba, and the four Atlantic provinces) and western Canada (Saskatchewan, Alberta, British Columbia, and all three territories). These regions have the same representation as those for the Lower 48 states within the model code; however, lease, plant, and pipeline fuel are not explicitly calculated because of lack of available historical data. Each region has average monthly storage injections and withdrawals and LNG imports based on historical data. LNG exports from western Canada are also included as an exogenous assumption according to projections from the latest *International Energy Outlook*; the additional supply required to produce these exports is assumed to be exclusively reserved for export and not able to flow directly into the larger North American pipeline network. There are also hubs that represent the U.S. border crossing for each state and the associated pipeline capacity. The flows through these hubs reflect the projected import and export levels.

Canadian production is being modeled in the OGSM (both AD and NA). See the *Oil and Gas Supply Module – NEMS Documentation* for a discussion of how expected natural gas production is determined. After solving the QP, the NGMM sends supply prices for eastern and western Canada back to the OGSM, which it uses as a basis for setting expected natural gas production for the two regions in Canada.

Canadian demand is largely an exogenous assumption; however, there is the option to calculate the natural gas consumption during oil sands production endogenously using assumed values for oil sands production, which vary by world oil price case, and the ratio of syncrude to dilbit/synbit[37] that is produced in Canada in response to global demand (including that of U.S. refineries). Both the projected Canadian consumption by sector and the oil sands production by world oil price case reflect the most recently published *International Energy Outlook*[38] while global demand for upgraded and diluted bitumen are obtained from the International Energy Module (IEM) and the Liquid Fuels Market Module (LFMM) in the NEMS. The marketed natural gas consumed in oil sands production relative to the oil produced is set at an assumed ratio based on historical data for oil sands production by type[39] (mined, in situ) and the percentage of bitumen per barrel of oil type[40] (i.e., whether bitumen is upgraded or diluted for transport). This volume of natural gas consumed is then added to the exogenous projection for western Canadian industrial demand. The monthly shares for Canadian consumption by sector are calculated using historical data[41] in the same manner as U.S. consumption.

[37] Syncrude refers to synthetic crude from oil sands; dilbit/synbit refers to bitumen diluted with lighter petroleum products or synthetic crude.

[38] See *Assumptions to the Annual Energy Outlook (Natural Gas Market Module)* for any details or updated methodology.

[39] Alberta Energy Regulator, *ST98—Alberta's Energy Reserves and Supply/Demand Outlook*

[40] Canadian Associate of Petroleum Producers, *Canada's Oil Sands Overview and Bitumen Blending Primer*

[41] Canadian historical data is obtained from *Statistics Canada*.

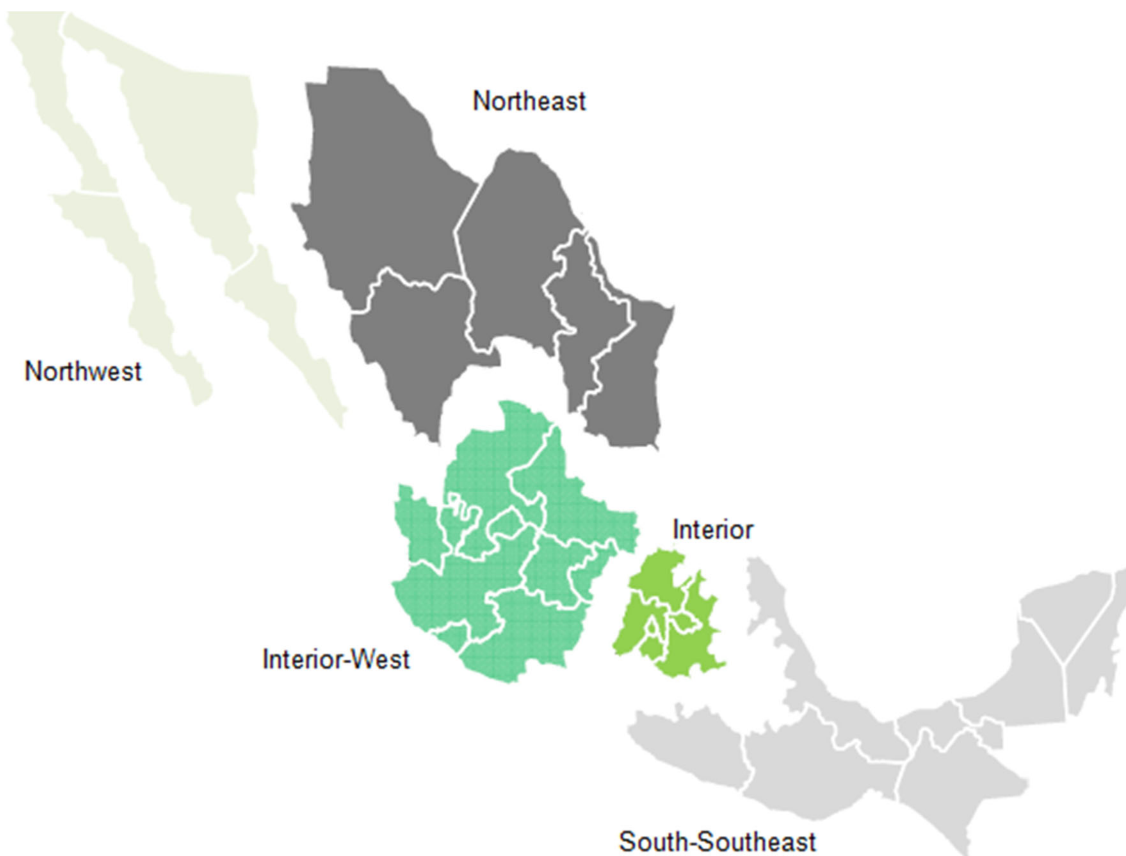
Mexico

The NGMM represents five hubs in Mexico. These regions, shown in Figure 4.2, correspond to the reporting regions for Mexico's Secretaría de Energía (SENER): Northeast, Northwest, Interior-West, Central, and South Southeast. There are also 4 hubs representing U.S. border crossings, the flows at which represent import and export volumes. Similarly to Canada, these hubs share the same representation in the mathematical program as the United States, resulting in a North American natural gas pipeline network that is used to model natural gas transmission. Mexican lease and plant fuel, pipeline fuel, and storage in Mexico are not separately accounted for in the NGMM. However, unlike Canada, natural gas production in Mexico is not modelled in OGSM and is represented within the NGMM.

Mexican supply

Three supply types are represented for Mexico: NA gas production, AD gas production, and LNG imports. While it is expected that the production of AD gas, which is co-produced with crude oil, will depend on world oil price, NA gas production is expected to respond to natural gas prices. LNG imports, on the other hand, are needed to meet demand in regions that do not have sufficient access to pipeline gas. Therefore, each of the three supply types are represented independently, using different methodologies, throughout the projection.

Figure 4.2 Mexico market regions in NGMM



All AD gas is assumed to be produced in the South-Southeast region, where Gulf of Mexico reserves have historically been developed and drilled. Using historical data,[42] an estimation is generated relating Mexico's crude oil production and related AD gas production to world oil price (see Appendix G):

$$PROD\_MX\_Oil_{yr} = \alpha\_Oil * PROD\_MX\_Oil_{yr-1} + \beta\_Oil * WOP_{yr-2} \tag{47}$$

$$PROD\_MX\_AD_{yr} = \alpha\_AD * PROD\_MX\_AD_{yr-1} + \beta\_AD * WOP_{yr-1} + \gamma\_\text{AD} * PROD\_MX\_Oil_{yr} \tag{48}$$

where

$PROD\_MX\_AD_{yr}$ = Mexican AD dry natural gas production in the South-Southeast region for projection year yr (minus lease fuel, plant fuel, and reinjected volumes) (Bcf)

$PROD\_MX\_Oil_{yr}$ = Mexican crude oil production for projection year yr (million barrels)

WOP_{yr-2} = average annual Brent crude oil price for year $yr - 2$ (1987$/MMBtu)

α_Oil = estimated coefficient for crude oil production in year $yr - 1$ when projecting crude oil production in year yr

β_Oil = estimated coefficient the Brent crude oil price in year $yr - 2$ when projecting crude oil production in year yr

α_AD = estimated coefficient for AD natural gas production in year $yr - 1$ when projecting AD natural gas production in year yr

β_AD = estimated coefficient the Brent crude oil price in year $yr - 2$ when projecting AD natural gas production in year yr

γ_AD= estimated constant term for crude oil production in year yr when projecting AD natural gas production in year yr

yr = projection year

As for the U.S. and Canadian supply regions, these supply volumes are fixed and are not allowed to change in response to price. NA gas production, however, is variable, with expected production as a function of the Henry Hub natural gas spot price. The assumption is that gas from these plays will be in direct competition with exports from the United States (i.e., not drilled for natural gas liquids); therefore, lower-priced U.S. natural gas will suppress NA gas production, and higher-priced U.S. natural gas will spur additional development. All NA gas production is assigned to the Northeast in the projection periods as this region accounts for a majority of historical production; furthermore, it is where most of Mexico's shale gas resources, such as the Burgos basin, are located.

Using an exogenously specified projection for natural gas production as well as assumptions for the share of NA gas,[43] parameters were estimated relating production to the Henry Hub price for two time

[42] Petróleos Mexicanos (PEMEX), *Production of Natural Gas by Region and Type*
[43] *Assumptions to the Annual Energy Outlook 2020: Natural Gas Market Module*

periods: the start of the projection until the onset of shale gas production and the period over which shale gas production occurs. The first year of shale gas production is therefore an exogenous assumption set in the model code (*MX_FirstShaleYear*). The estimation, and equation used in the NGMM, is as follows (see Appendix G):

$$PROD\_MX\_NA_{yr} = \alpha 1_{NA,t} * PROD\_MX\_NA_{yr-1} + \alpha 2_{NA,t} * PROD\_MX\_NA_{yr-2} + \beta_{NA,t} * HH\_PRICE_{yr-1} \quad (49)$$

where

$PROD\_MX\_NA_{yr}$ = Mexican NA dry natural gas production in the Northeast region for projection year yr (minus lease and plant fuel) (Bcf)

HH_PRICE_{yr-1} = average annual Henry Hub natural gas spot price for year $yr - 1$ (1987$/MMBtu)

$\alpha 1_{NA,t}$ = estimated coefficient for the last year's NA natural gas production for time range t in the projection period

$\alpha 2_{NA,t}$ = estimated coefficient for the two year's prior NA natural gas production for time range t in the projection period

$\beta_{NA,t}$ = estimated coefficient for last year's average Henry Hub spot price for time range t in the projection period

yr = projection year

t = range of time in the projection period before or after the onset of shale gas production

LNG imports into Mexico are set exogenously based on recent historical data. They are only allowed at existing LNG import facilities: Altamira in the Northeast, Costa Azul in the Northwest, and Mazanillo in the Interior-West. These volumes are expected to decline in the short term as the pipelines under construction are completed; this will bring natural gas via pipeline to demand markets that are currently pipeline constrained.

Mexican demand

Mexican demand is based on an exogenous consumption projection in EIA's most recent *International Energy Outlook*. Projections for consumption in the electric power sector are augmented in order to align with natural gas combined cycle power plants under construction in Mexico and announced plans to convert existing fuel oil generators to natural gas.[44] Industrial-sector natural gas consumption in Mexico is assumed to have two components: natural gas consumed in oil and natural gas exploration and production activities by Petróleos Mexicanos (PEMEX) and other industrial natural gas consumption. It is estimated as a function of historical data, crude oil production, and Henry Hub price:

$$Cons\_MX\_Ind_{yr} = Cons\_MX\_PEMEX_{yr} + Cons\_MX\_Ind\_other_{yr}$$

[44] *Assumptions to the Annual Energy Outlook 2020: Natural Gas Market Module*

(50)

where

$$Cons\_MX\_PEMEX_{yr} = \alpha\_PEMEX * Cons\_MX\_PEMEX_{yr-1} + \beta\_PEMEX * PROD\_MX\_Oil_{yr} +$$

(51)

$$Cons\_MX\_Ind\_other_{yr} = \alpha\_Ind * Cons\_MX\_Ind\_other_{yr-1} + \beta\_Ind * HH\_PRICE_{yr-1} + C\_Ind$$

(52)

where

$Cons\_MX\_Ind_{yr}$ = Total Mexican industrial sector consumption of natural gas in projection year yr (Bcf)

$Cons\_MX\_PEMEX_{yr}$ = Mexican consumption of natural gas by PEMEX for projection year yr (Bcf)

$Cons\_MX\_Ind\_other_{yr}$ = Mexican consumption of natural gas for all other industrial use for year yr (Bcf)

$Cons\_MX\_PEMEX_{yr-1}$ = Mexican consumption of natural gas by PEMEX for $yr - 1$ (Bcf)

$PROD\_MX\_Oil_{yr}$ = Mexican crude oil production as calculated in Eq. (47) for projection year yr (million barrels)

$Cons\_MX\_Ind\_other_{yr-1}$ = Mexican consumption of natural gas for all other industrial use for $yr - 1$ (Bcf)

HH_PRICE_{yr-1} average annual Henry Hub natural gas spot price for year $yr - 1$ (1987$/MMBtu)

α_PEMEX = estimated coefficient for PEMEX consumption in year $yr - 1$

β_PEMEX = estimated coefficient for crude oil production in year yr

α_Ind = estimated coefficient for all other industrial sector consumption in year $yr - 1$

β_Ind = estimated coefficient for the Henry Hub price in year $yr - 1$

C_Ind= estimated constant term for other industrial consumption

yr = projection year

The monthly shares for Mexican demand by sector are calculated in the same manner as U.S. demand using historical data published by Mexico.[45]

[45] Secretaría de Energía de México (SENER), *Sistema de Información Energética*

Lease and plant fuel are not explicitly calculated in the NGMM; rather, Mexican dry production is modeled excluding these volumes. Pipeline fuel is also not included in solving for the flows to and from Mexican hubs or explicitly solved for in the model code as data are not available.

LNG exports

LNG exports of domestically-sourced natural gas are projected endogenously in the NGMM. The pre-processing step involves projecting liquefaction capacity additions beyond that of the existing facilities and those that are under construction while the QP ultimately determines the utilization of this capacity. The basic approach in pre-processing is to evaluate the long-term economic viability of adding (or expanding) a generic LNG liquefaction facility consisting of up to three large trains of a specified capacity (*LNG_Increment*) in each projection year. This is done independently for each of the allowed coastal regions of the United States before selecting the most economically profitable region for construction, if any, and accounting for any assumed restrictions, such as earliest start year or maximum allowable volume. An underlying assumption is that facilities will be built if consumers are interested in signing long-term contracts at a price that allows cost recovery, so the economic viability is evaluated from the perspective of potential consumers. Once built,[46] the liquefaction facility is assumed to be able to operate at full capacity (accounting for some operational down-time) throughout the rest of the projection period.

To effectively assess the economic viability to consumers in representative world destinations of signing a contract with a new U.S. liquefaction facility versus an assumed alternative, the NGMM calculates a net present value over the assumed lifetime of a contract with the LNG liquefaction facility (*NumberOfYearsForLookAhead_LNG_*). This net present value corresponds to the cost of purchasing from the United States versus another global supplier. The price of the alternative supplier, or the world price of LNG at a given destination, over that time period is compared to the price of U.S. LNG at these destinations, which includes a sunk cost to recover the initial investment required to build the facility, the operational costs (including regasification), and the shipping costs to a cargo's destination.

World LNG prices

The model projects a representative price of LNG for each world destination represented.[47] These prices are calculated based on the projections from EIA's most recent *International Energy Outlook*, with updates to account for recent market events as well as additional nonpublished information and analyses based on EIA's International Natural Gas Model (INGM) results. The world natural gas prices are assumed to start at their recent historical ratio to the world oil price. Over time, the price of LNG becomes less tied to the world oil price as the ratio of flexibly-priced LNG to the representative regional net natural gas demand increases relative to its base year level. The concept is that the ratio reflects the tightness or looseness of the world LNG market pushing or pulling, respectively, world natural gas prices toward or away from the world oil price. The specific form of the price equation follows:

---

[46] The number of years between the decision to build new LNG export capacity and the beginning of its operations is a model assumption (*LNG_YrsUntilBuild*). New liquefaction facility builds are assumed to have a completion date three years from the current projection year in AEO 2018.

[47] In AEO 2018, the NGMM assesses global demand and prices of LNG in Europe and Asia.

$$PRICE\_LNG_{(yr+lookyr),d} = \left(WOP_{(yr+lookyr)}\right)^{\alpha_d} * \left(\frac{\left(\frac{FLEX_{(yr+lookyr)} + LNG\_USA_{(yr+lookyr)} + LNG\_ADD}{Q\_LNG_{(yr+lookyr),d}}\right)}{\left(\frac{FLEX_{lhyr} + LNG\_USA_{lhyr} + LNG\_ADD}{Q\_LNG_{lhyr,d}}\right)}\right)^{\beta_d} \quad (53)$$

where

$PRICE\_LNG_{(yr+lookyr),d}$ = price of LNG in global demand region d and year $yr + lookyr$, where yr is the current projection year (1987$/MMBtu)

$WOP_{(yr+lookyr)}$= world oil price in year $yr + lookyr$ (1987$/MMBtu)

$FLEX_{(yr+lookyr)}$ = exogenously set projected level of flexibly priced LNG on the world market, excluding any volumes from the United States, in year $yr + lookyr$ (Bcf)

$FLEX_{lhyr}$ = exogenously set historical level of flexibly priced LNG on the world market, excluding any volumes from the United States, in year $lhyr$ (Bcf)

$LNG\_USA_{(yr+lookyr)}$ = projected LNG exports from the United States from liquefaction facilities constructed for year $yr + lookyr$, where yr is the current projection year (Bcf)

LNG_USA_{lhyr} = historical LNG exports from the United States in year $lhyr$ (Bcf)

LNG_ADD= LNG exports from the United States from liquefaction facility under consideration for construction (*LNG_Increment*) (Bcf)

$Q\_LNG_{(yr+lookyr),d}$ = exogenously set projected LNG imports/consumption for global demand region d in year year $yr + lookyr$, where yr is the current projection year (Bcf) [48]

$Q\_LNG_{lhyr,d}$ = exogenously set historical LNG imports/consumption in year $lhyr$ for global demand region d (Bcf)[49]

α_d = an assumed coefficient representing the value necessary to align the oil price to the natural gas price in global demand region d (i.e. when the β_d term in the LNG price equation equals 1) (Appendix E)

β_d = an assumed coefficient that drives the movement of the natural gas price away from (or to) the oil price as the market loosens (or tightens) as defined by the ratio of flexible and U.S. LNG supply available globally to global demand region d for both historical and projection years (Appendix E)

yr = current projection year

$lhyr$ = last year of historical data

[48] Details for these assumptions and can be found in the *Assumptions to the Annual Outlook 2018: Natural Gas Market Module*.
[49] Details for these assumptions and can be found in the *Assumptions to the Annual Outlook 2018: Natural Gas Market Module*.

$lookyr$ = number of years after the current projection year to consider in assessing the net present value of signing a contract for U.S. LNG versus an alternative

d = demand regions considered as destinations for LNG exports (currently Europe and Asia)

Price of U.S. LNG

Lower 48 states

In order to assess whether the next incremental amount of U.S. LNG exports from an additional LNG train will be competitive in global markets over the look-ahead period, a U.S. LNG price is first estimated for each relevant domestic region. After this, costs for liquefaction (including fixed charges), regasification, and transport overseas are added to arrive at a potential price for this U.S. LNG in global markets.

The projected natural gas supply price for each potential U.S. export region is used for the current projection year; for all future years being assessed, the supply price for that year from the last NEMS cycle is used and adjusted to account for the difference between the natural gas production in the last NEMS cycle and the estimated production in that year in the current cycle. This difference equals the difference in existing and potential LNG export capacity in a given year between the two cycles as well as any domestic changes to demand. The adjustment is made using the following scaling factor:

$$FACTOR\_LNG_{(yr+lookyr),r} = \left(\frac{LAST\_PROD_{(yr+lookyr),r} + LNG\_CAP_{(yr+lookyr),r} - LAST\_LNG\_CAP_{(yr+lookyr),r}}{LAST\_PROD_{(yr+lookyr),r}}\right)^{\gamma} \tag{54}$$

where γ is calculated using the ratio between actual and expected production in the current projection year:

$$\gamma = \gamma_1 + \gamma_2 * \frac{PROD\_ACT_{yr}}{PROD\_EXP_{yr}} \tag{55}$$

where

γ = exponent used to approximate how a difference in production translates into a difference in price as derived by a series of offline test runs of the model

γ_1, γ_2 = assumed coefficients used in the calculation of γ in the current projection year yr being evaluated

The difference in the comparable estimate and realized price in the previous projection year (the portion of the equation indexed below to "$yr - 1$") is used as a basis for further adjusting the LNG supply price. Thus, in all projection years that are used to determine whether additional LNG export capacity would be economical, the U.S. supply price equals the following:

$$PRICE\_SUP_{(yr+lookyr),r} = LAST\_PRICE\_SUP_{(yr+lookyr),r} * FACTOR\_LNG_{(yr+lookyr),r} * \left(\frac{PRICE\_SUP_{yr-1,r}}{LAST\_PRICE\_SUP_{yr-1,r} * FACTOR\_LNG_{yr-1,r}}\right) \quad (56)$$

where

$PRICE\_SUP_{(yr+lookyr),r}$ = supply price of natural gas in region r and projection year $yr + lookyr$, where yr is the current projection year (1987$/MMBtu)

$LAST\_PRICE\_SUP_{(yr+lookyr),r}$= supply price of natural gas in region r for projection year $yr + lookyr$, where yr is the current projection year, for the last NEMS cycle (1987$/MMBtu)

$FACTOR\_LNG_{(yr+lookyr),r}$ = factor to scale the supply price from the last NEMS cycle to account for the difference in non-associated production from the last NEMS cycle and the expected non-associated production associated with the volume of LNG being evaluated for projection year $yr + lookyr$, where yr is the current projection year, and region r

$LAST\_PROD_{(yr+lookyr),r}$ = total NA natural gas production in region r and projection year $yr + lookyr$, where yr is the current projection year, for the last NEMS cycle (Bcf)

$LNG\_CAP_{(yr+lookyr),r}$= total LNG export capacity in region r for projection year $yr + lookyr$, where yr is the current projection year (Bcf)

$LAST\_LNG\_CAP_{(yr+lookyr),r}$ = total LNG export capacity in region r and projection year $yr + lookyr$, where yr is the current projection year, for the last NEMS cycle (Bcf)

$PROD\_ACT_{yr}$ = total U.S. realized NA natural gas production as solved for by the NGMM in the current projection year yr (Bcf)

$PROD\_EXP_{yr}$ = total U.S. expected NA natural gas production in the current projection year yr as provided by NEMS (OGSM) (Bcf)

$lookyr$ = number of years after the current projection year to consider in assessing the net present value of signing a contract for U.S. LNG versus an alternative

yr = current projection year

r = U.S. region being assessed for the economic feasibility of added LNG export capacity

For years that extend beyond the last model year, the price is set by applying a growth rate to regional supply price in the last model year that is consistent to the growth trend in the last years of the model.

The total fully-loaded[50] price of U.S. LNG that is supplied to the global market over the range of the lifespan of the LNG export facility is calculated:

$$PRICE\_USLNG_{(yr+lookyr),r,d} = PRICE\_SUP_{(yr+lookyr),r} * (1 + PCT\_LIQ_r) + COST\_LIQ_r + COST\_REGAS_r + COST\_SHIP_{r,d} \quad (57)$$

where

$PRICE\_USLNG_{(yr+lookyr),r,d}$ = price of U.S. LNG from region r to global demand market d for projection year $yr + lookyr$, where yr is the current projection year (1987$/MMBtu)

$PRICE\_SUP_{(yr+lookyr),r}$ = supply price of natural gas in U.S. region r in projection year $yr + lookyr$, where yr is the current projection year (1987$/MMBtu)

PCT_LIQ_r = percent of fuel used in the transport of natural gas to the export facility from the supply hub for U.S. region r and fuel used to liquefy natural gas

$COST\_LIQ_r$ = assumed cost for U.S. region r to liquefy natural gas, including any capacity charges or capital investment charges applied to the per unit cost (1987$/MMBtu)

$COST\_REGAS_r$ = assumed fixed charges for U.S. region r to regasify the LNG after it reaches its destination (1987$/MMBtu)

$COST\_SHIP_{r,d}$ = assumed shipping costs to transport LNG from U.S. region r to a specified world demand region (1987$/MMBtu)

yr = current projection year

$lookyr$ = number of years after the current projection year to consider in assessing the net present value of signing a contract for U.S. LNG versus an alternative

r = U.S. region assessed for the economic feasibility of added LNG export capacity

d = demand regions considered as destinations for LNG exports (currently Europe and Asia)

Alaska

The potential of building an LNG export facility in Alaska, bringing trapped North Slope natural gas to market, is assessed for its economic viability in comparison to other U.S. LNG export projects. Instead of considering each additional train individually, Alaska is assumed to have a fixed number of trains, all of which will be built if the project goes forward (*LNG_AKTrainTotal*). Additionally, the Alaskan natural gas supply price is an exogenous assumption based on the estimated cost of extracting gas that was previously produced and reinjected into the formation with oil produced in north Alaska; the total LNG price includes this per unit cost of extracting, the combined cost of transporting it via pipeline to the south coast and liquefying it, and the international shipping costs.

[50] The fully-loaded price includes the price of natural gas feedstock, losses during liquefaction and transportation, regasification, shipping costs, and any charges applied by a liquefaction terminal to cover the capital expenditure required to build the facility.

Net present value of LNG export capacity

Once prices are established for Europe and Asia over the assumed lifespan of liquefaction plant, a comparison is made to the expected future prices for LNG exports from the United States to these destinations. The differences in these two prices represents the added value to the consumer (or to whoever is able to capture the economic return) of purchasing LNG from the United States over other potential supply options. These price differences are accumulated over the lifetime of the plant and set in terms of the present projection year using an assumed discount rate to reflect the time value of money (*LNG_DCFDiscountRate*) :

$$NPV\_USLNG_{yr,r,d} = \sum_{yr+1}^{yr+lookyr} \frac{PRICE\_LNG_{(yr+lookyr),d} - PRICE\_USLNG_{(yr+lookyr),r,d}}{(1 + DCF\_RATE)^{lookyr}} \tag{58}$$

where

$NPV\_USLNG_{yr,r,d}$ = net present value of LNG sourced from region r in the United States relative to the market cost of LNG in global demand region d assuming the decision to build LNG export capacity is made in projection year yr (1987$/MMBtu)

$PRICE\_LNG_{(yr+lookyr),d}$ = price of LNG in global demand region d for projection year $yr + lookyr$, where yr is the current projection year (1987$/MMBtu)

$PRICE\_USLNG_{(yr+lookyr),r,d}$ = price of U.S. LNG from region r to global demand region d for projection year $yr + lookyr$, where yr is the current projection year (1987$/MMBtu)

DCF_RATE = the discount rate, i.e. the return that could be earned per unit of time on an investment with similar risk

$lookyr$ = number of years after the current projection year yr to consider in assessing the net present value of signing a contract for U.S. LNG versus an alternative

yr = current projection year

r = U.S. region being assessed for the economic feasibility of added LNG export capacity

d = demand regions considered as destinations for LNG exports

The region with the resulting highest net present economic value is assumed to be the location of the next liquefaction train, presuming other assumptions are not limiting factors. These include:[51]

- *earliest potential start year in that region (LNG_FirstYear)*
- *maximum allowed export volume in that region (LNG_MaxExports)*
- *maximum number of trains built in a year in the United States, reflecting practical limits on the necessary resources/manpower for such specialized construction (LNG_MaxTrainsYr)*
- *another LNG export facility has already been built that is high-risk, which is defined as having a net present value lower than the risk threshold (LNG_RiskThreshold)*

---

[51] *Assumptions to the Annual Energy Outlook 2018: Natural Gas Market Module*

Construction is assumed to take a specified number of years (*LNG_YrsUntilBuild*) before the train(s) are operational, and these additional volumes are phased in over time (*LNG_PhaseInYrs*).

Alaska

As Alaska is not part of the North American natural gas transmission system, it is modeled outside of the quadratic program (QP). The NGMM is responsible for projecting both Alaskan production and consumption in NEMS. Using historical data, the model code projects demand by sector. It then calculates Alaskan natural gas production by assuming it fulfills the projected demand.

The NEMS demand modules provide a projection of natural gas consumption for the total Pacific Census division, which includes Alaska. Therefore, the NGMM derives annual estimates of contiguous Pacific Division consumption levels by first estimating Alaska natural gas consumption for all sectors and then subtracting these from the core market consumption levels in the Pacific division provided. The series of equations for specifying the consumption of gas by Alaska residential and commercial customers follows (see Appendix G):

$$AK\_Q\_RES_{yr} = \alpha_{RES} * AK\_POP_{yr} + \beta_{RES} * UNEMP_{yr} + \gamma_{RES} * AK\_PRICE\_CG_{yr} \tag{59}$$

$$\begin{aligned} AK\_Q\_COM_{yr} = {} & \alpha_{COM} * AK\_POP_{yr} + \beta_{COM} * UNEMP_{yr} + \gamma_{COM} \\ & * \left(AK\_Q\_COM_{yr-1} - \alpha_{COM} * AK\_POP_{yr-1} + \beta_{COM} * UNEMP_{yr-1}\right) \end{aligned} \tag{60}$$

where

$AK\_Q\_RES_{yr}$ = consumption of natural gas by residential customers in Alaska in projection year yr(Bcf)

$AK\_Q\_COM_{yr}$ = consumption of natural gas by commercial customers in Alaska in projection year r (Bcf)

AK_POP_{yr} = exogenously specified projection of the population in Alaska[52] in projection year yr

$UNEMP_{yr}$ = U.S. unemployment rate (percent) from NEMS Macroeconomic Activity Module in year yr

$AK\_PRICE\_CG_{yr}$ = natural gas citygate price in Alaska in projection year yr ($1987/Mcf)

α_{SEC} = estimated coefficient for Alaska population for sector $sec \in \{res, com\}$

β_{SEC} = estimated coefficient for unemployment for sector $sec \in \{res, com\}$

γ_{RES} = estimated coefficient for Alaska citygate price

γ_{COM} = estimated year-to-year autocorrelation coefficient for variable $AK\_Q\_COM_{yr}$

[52] State of Alaska, Department of Labor and Workforce Development, *Alaska Population Projections*

yr = projection year

Alaska natural gas consumption for the industrial sector is an exogenous assumption signifying small volumes of natural gas, and it remains constant across the projection. The use of natural gas in compressed or liquefied natural gas vehicles in Alaska is assumed to be negligible or nonexistent. The Electricity Market Module (EMM) provides a value for natural gas consumption in Alaska by electric generators. While this approach projects the total gas consumption in the state given the current pipeline infrastructure in Alaska, if a pipeline is built to bring North Slope gas to the South, it is possible that the projected volumes could be higher. This could be particularly true for the industrial sector because consumption growth is currently hindered by declining supplies in South Alaska. This potential growth is not currently modeled.

The production of gas in Alaska is set equal to the sum of the volumes consumed in and transported out of Alaska plus what is consumed for lease, plant, and pipeline operations, and the balancing item. Furthermore, if a new LNG export facility is built (we assume Kenai will no longer export LNG), production also includes the volume of exported gas plus any related liquefaction fuel that is consumed. Lease and plant fuel is primarily consumed in north Alaska during crude oil extraction and is estimated as follows (see Appendix G):

$$AK\_LAP_{yr} = \alpha_{LAP} * AK\_PROD\_OIL_{yr} + \beta_{LAP} * AK\_LAP_{yr-1} + C_{LAP} \tag{61}$$

where

AK_LAP_{yr} = quantity of gas consumed for lease and plant operations in year yr, excluding that related to pipeline fuel (Bcf)

$AK\_PROD\_OIL_{yr}$ = crude oil production in Alaska (thousand barrels per day—Mbpd) in year yr, from OGSM

α_{LAP} = estimated coefficient for Alaska crude oil production

β_{LAP} = estimated coefficient for prior year's lease and plant fuel consumption in Alaska

C_{LAP} = estimated constant term

yr = projection year

Pipeline fuel, along with total consumption, is expected to be consumed in southern Alaska and is set as a percent of total consumption. Total production is assigned for both North and South Alaska according to the equations below:

$$AK\_PROD\_N_{yr} = AK\_LAP_{yr} + (1 + PCT\_LIQ) * AK\_LNG\_EXP_{yr} \tag{62}$$

$$AK\_PROD\_S_{yr} = (1 + PCT\_PIP) * AK_{Q_{TOTAL}}{}_{yr} + AK\_DISC_{yr}$$

(63)

where, for year yr,

$AK\_PROD\_N_{yr}$ = dry gas production in North Alaska (Bcf)

$AK\_PROD\_S_{yr}$ = dry gas production in South Alaska (Bcf)

AK_LAP_{yr} = quantity of gas consumed for lease and plant operations in projection year yr, excluding that related to either pipeline (Bcf)

$AK\_LNG\_EXP_{yr}$= total LNG exports out of Alaska in projection year yr (see LNG Exports)

$AK\_Q\_TOTAL_{yr}$ = total Alaska end use consumption in projection year yr

AK_DISC_{yr} = balancing item (discrepancy) calculated for Alaska (see Balancing Item)

PCT_LIQ = percent of fuel used for liquefaction in export facilities

PCT_PIP = pipeline fuel as a percent of gas consumption

yr = projection year

5. Post-processing routines

After the quadratic program (QP) solves, the solution values can be pulled directly to set output variables for the NGMM to pass to other modules or to the report writer in the NEMS. There are several NGMM output variables that require further calculations. Most volumes are set by simply aggregating to derive annual values at the regional level required, or they require a relatively simple calculation. Volumes that are set include: nonassociated (NA) and total dry gas production, natural gas import and export volumes, region-to-region flows, lease and plant fuel, pipeline fuel, and fuel used for liquefaction. Some volumes that are reported to the NEMS are already set in the pre-processing routines, such as: supplemental supplies, LNG imports, and associated-dissolved (AD) production. A more extensive process with further assumptions is required for setting module output prices, which include: spot prices, wellhead prices, citygate prices, import and export prices, and delivered prices to residential, commercial, industrial, electric generator, and natural gas vehicle (including rail and marine) customers.

For all variables defined in this chapter, Appendix C provides a reference to the full identifier name used in the NGMM, Appendix D identifies where to find specific equations within the code, and Appendix E indicates which input files contain input assumptions or parameters.

Production and supply prices

Production by month and supply region is an assigned using the QP decision variable *QProduction*. This volume is assigned to a demand region using mapping parameters (Appendix E) and aggregated to an annual total for reporting to NEMS.

For all supply regions where NA supply volumes exist in a given model year, the monthly supply price is set to the shadow price[53] of the constraint *SupplyAccounting*. Unlike the supply mass balance constraint, which ensures that the total supply—independent of its supply step and associated price—equals the flow from a supply region to its hub, the supply accounting constraint defines total production as the sum underneath all supply curve steps (e.g. the area under the curve). Therefore, its shadow price represents the marginal price corresponding to this constructed supply curve. The annual supply price is the average of all monthly prices weighted by production volume.

For supply regions without supply volumes, last year's price is assigned. The Henry Hub spot price is assigned the supply price corresponding to its location (*HenryHubRegion*, or onshore south Louisiana) plus an assumed gathering charge (*GatheringCharge*).

LNG exports

LNG export volumes are solved for endogenously within the QP (*QLNGexp*); however, this decision variable also includes the natural gas consumed during liquefaction. These two volumes are calculated as follows:

$$LNGEXP_{yr,lngexp\_qp} + Q\_LIQ_{yr,lngexp\_qp} = \sum_{step}^{LSMAX} \sum_{mon} LNG_{mon,step,lngexp\_qp} \qquad \forall \left(mon, step, lngexp_{qp}\right) | (mon \subseteq yr) \tag{64}$$

[53] The difference between the optimized value of the objective function and the value of the objective function, evaluated at the optional basis, when the right hand side of a constraint is increased by one unit.

$$Q\_LIQ_{yr,lngexp\_qp} = LNGEXP_{yr,lngexp\_qp} * PCT\_LIQ \qquad \forall(yr, lngexp_{qp})$$

(65)

where

$LNGEXP_{yr,lngexp\_qp}$ = total annual LNG export volumes in projection year yr from LNG export region $lngexp\_qp$, solved for in the QP (Bcf)

$Q\_LIQ_{yr,lngexp\_qp}$ =total volume of natural gas consumed during the liquefaction process in projection year yr in LNG export region $lngexp\_qp$ (Bcf)

$LNG_{mon,step,lngexp\_qp}$ = decision variable containing the total volume of LNG associated with a given LNG export volume associated with a given step of the LNG export demand curve in projection month mon for LNG export region $lngexp\_qp$ (Bcf)

$LSMAX$ = maximum step defining the LNG export demand curve (*LNGExpCrv_MaxStep*)

PCT_LIQ = percentage of LNG volume consumed during liquefaction

mon = projection month within projection year yr

yr = projection year

$step$ = price-quantity pair that defines the LNG export demand curve

$lngexp\_qp$ = region that contains a LNG export capacity whose utilization is determined within the QP

For the two regions where LNG exports are not considered part of the QP, western Canada and Alaska, the NGMM assumes that the LNG export capacity is fully utilized and assigns this value to the total exports. In both of these cases, it is not the supply price of natural gas that determines market competitiveness; rather, it is the comparatively high capital cost of the liquefaction projects, including the new pipeline infrastructure required, that would make building new LNG export capacity uneconomic. Their locations on the Pacific coast also mean that shipping costs to Asia are much less than LNG exports from the Gulf Coast. Therefore, once LNG export facilities are built, LNG from Alaska or western Canada would be expected to out-compete all other global LNG supplies on a variable cost basis. The corresponding fuel used for liquefaction in Alaska is included in the U.S. total; fuel used for liquefaction is not explicitly calculated for western Canada.

Imports and exports

The NGMM reports to the NEMS total pipeline imports and exports to/from Canada and Mexico, LNG imports, and LNG exports at the annual level. LNG imports are set in pre-processing as a historical average that is held constant, and LNG exports are calculated as described above using the LNG demand curve representation in the QP. Described below is the procedure for determining pipeline imports and exports as well as import and export prices (where applicable).

Pipeline import and export volumes

Pipeline import and export volumes are assigned using the decision variable *FlowHubToHub*. By treating the border crossings between individual states and Canadian or Mexican regions as hubs, and only allowing flows to and from the state and Canadian or Mexican region into or out of the border crossing hub, the flows directly correspond to how imports and exports are defined (i.e., not as volumes sent from or two a given state at its market hub, but as volumes as measured at a physical point on the border). The equations corresponding to total annual imports and exports to Canada and Mexico are given below.

$$IMP\_CN_{yr} = \sum_{bx_{cn}} \sum_{cn} \sum_{mon} FLOWH2H_{mon,cn,bx\_cn} \qquad \forall(mon, cn, bx\_cn)|(mon \subseteq yr) \tag{66}$$

$$EXP\_CN_{yr} = \sum_{bx\_cn} \sum_{st} \sum_{mon} FLOWH2H_{mon,st,bx\_cn} \qquad \forall(mon, st, bx\_cn)|(mon \subseteq yr) \tag{67}$$

$$IMP\_MX_{yr} = \sum_{bx\_mx} \sum_{mx} \sum_{mon} FLOWH2H_{mon,mx,bx\_mx} \qquad \forall(mon, mx, bx\_mx)|(mon \subseteq yr) \tag{68}$$

$$EXP\_MX_{yr} = \sum_{bx\_mx} \sum_{st} \sum_{mon} FLOWH2H_{mon,st,bx\_mx} \qquad \forall(mon, st, bx\_mx)|(mon \subseteq yr) \tag{69}$$

where

IMP_CN_{yr} = total imports from Canada to the United States for projection year yr (Bcf)

EXP_CN_{yr} =total annual exports from the United States to Canada for projection year yr (Bcf)

IMP_MX_{yr} = total annual imports from Mexico to the United States for projection year yr (Bcf)

EXP_MX_{yr} = total annual exports from the United States to Mexico for projection year yr (Bcf)

$FLOWH2H_{mon,h,h1}$ = flow from hub h to hub $h1$ in projection month mon (Bcf); in equations above, h and $h1$ one refer to flows to or from border-crossing hubs

yr = projection year

mon = projection month

cn = Canadian hub (western or eastern Canada)

mx = Mexican hub (one of 5 regions)

st = lower 48 state (+ DC)

bx_cn = Canadian border-crossing hub

bx_mx = Mexican border-crossing hub

Import and export prices

The NGMM assigns import and export prices for pipeline volumes to the shadow prices of the constraints requiring mass balance at the border crossing hub: *HubBalanace_BXtoUS* and *HubBalance_UStoBX*. Thus, imports and exports are priced at the marginal cost of natural gas at that hub. Because flows to or from the border crossing are uniquely defined in the NGMM, the hub balancing constraints must be formulated differently from those representing a supply/demand region. Unlike all other flows between hubs, flows to or from the border crossing are not defined as flow measured at the boundary delineating the two regions. They are defined as the flow from that boundary to a specific region. Therefore, for imports to the United States, the flow out of the international region into the border crossing minus pipeline fuel losses must equal the flow out of the border crossing – the import volume. For exports from the United States, the flow into the border crossing, minus pipeline fuel losses incurred as defined by the pipeline fuel loss factor for the state exporting natural gas, must equal the flow out of the border crossing – the export volume. Thus, the shadow prices of the mass balance constraints directly correspond to the volumes at the border.

The LNG export prices are calculated in the pre-processing routine. Because the LNG imports are fixed to a historical average and not determined endogenously, no LNG import price is currently being reported.

Delivered end-use prices

Spot prices

At each hub or node in the simplified pipeline network represented in the NGMM, the natural gas flows into and out of the node must balance, as forced by the constraint labeled as *HubBalance* in the QP. The shadow prices[54] associated with this constraint represents the marginal price at hub h, which is the variable cost of supplying one more unit to the node. The assumption in the NGMM is that this price is indicative of the spot price at this representative node. This is also supported by the construction of the variable tariff curves and the pipeline fuel loss factors, which together are intended to reflect historically observed basis differentials between reported spot prices as a function of the pipeline utilization rate. Balancing constraints, and therefore spot prices, are set for each state and month, as well as at each supply point and border crossing. At production nodes, these prices are assumed to reflect the wellhead or supply price. The NGMM does not report prices at state hubs, with the exception of the Henry Hub price, but uses these prices to generate citygate and delivered prices. The Henry Hub price (*NGTDMREP_OGHHPRNG*) is set at the wellhead price in South Louisiana plus an assumed gathering charge (*GatheringCharge*).

Citygate prices

Citygate prices are the prices local distribution companies (LDCs) or utilities pay for natural gas from the pipeline transmission system. They include the cost of the commodity (spot or contract price) as well as any additional costs of transporting natural gas in the pipeline system, applicable taxes, storage fees, and net losses from hedging. The NGMM calculates citygate prices for each projection year by

[54] See footnote (51).

state/month using econometrically-estimated equations. With several exceptions (described below), the average monthly spot price is a reasonable approximation for a commodity cost at the citygate. Other components of the citygate price are fixed at a constant monthly fee (loosely estimated as β in the equation below), and unitized by dividing by the sum of residential and commercial consumption (the bulk of LDC deliveries). Any other variable fees (e.g., storage injection/withdrawals costs) should be captured in the constant term, as follows:

$$PRICE\_CG\_ST_{mon,st} = \alpha_{per(mon),st} * PRICE\_SPOT_{mon,st} + \frac{\beta_{per(mon),st}}{\left(Q\_RES_{mon,st} + Q\_COM_{mon,st}\right)} + C_{per(mon),st} \quad (70)$$

where

$PRICE\_CG\_ST_{mon,st}$ = citygate price in state st and projection month mon (1987$/Mcf)

$PRICE\_SPOT_{mon,st}$ = spot price in state st and projection month mon (1987$/Mcf)

$Q\_RES_{mon,st}$= residential sector consumption of natural gas in state st and projection month mon (Bcf)

$Q\_COM_{mon,st}$= commercial sector consumption of natural gas in state st and projection month mon (Bcf)

$\alpha_{per(mon),st}$ = estimated coefficient for spot price for state st and the period of the year that includes month mon, expected to be close to 1.0 (unitless)

$\beta_{per(mon),st}$ = estimated coefficient, reflecting fixed monthly charges for state st and the period of the year that includes month mon (MM 1987$)

$C_{per(mon),st}$ = estimated constant term for state st and the period of the year that includes month mon (1987$/Mcf)

mon = projection month

st = state (including DC, excluding AK and HI)

$per(mon)$ = maps parameter values for each month to correspond to one of three periods of the year: either all months of the year, the winter months, or the non-winter months

Historical monthly citygate prices, spot prices, and consumption in the residential and commercial sectors are used to estimate the parameters in the above equation, as described in Appendix G. For most states the estimated parameters do not vary by month or season. However, this simplification did not always produce reasonable results.

Due to regulations requiring utilities to be able to meet demand for natural gas for their customers during peak periods of consumption, natural gas volumes are typically contracted; therefore, while citygate prices will rise during periods of high demand, they will often not see the same volatility as spot prices during extreme conditions. This is particularly evident in places such as New England, where pipeline constraints limit flows into the area. For four states (Arizona, Oregon, Nevada, and New York) the equation was estimated separately for the winter months (November through February) and the non-winter months to improve the estimation. For the states in New England, Utah, and Delaware, an estimation by season still did not provide a reasonable predictor, so prices for the winter months were estimated by setting the November and December price to October's value and setting the January and February price to the average of the year's March price and the previous year's October price.

End-use natural gas prices

Delivered natural gas prices are set by adding a markup to the average citygate or the average spot price at the appropriate regional level. An annual quantity-weighted average city gas price is calculated for each Census division, averaging across all months and relevant states using the residential plus commercial sector consumption levels as weights. The residential and commercial prices, as well as some of the vehicle fuel prices, are based on these average citygate prices. Prices to the industrial and electric generator sectors are based on average spot prices using the industrial and electric generator consumption levels, respectively, as weights.[55] The residential and commercial prices are benchmarked to the annual Census division price forecasts from STEO for the first (*NumberOfSTEOYears*) years of the projection, while the prices to electric generators are benchmarked to align with the national annual prices in the STEO. This is done by multiplying the initially calculated price by a factor that will align the result to the STEO value. The STEO factors calculated in the last STEO year are phased out over an assumed number of years (*NumberOfSTEOPhaseOutYears*) to a value of 1.0 after the last STEO year.

Residential sector

Prices charged to residential customers are set annually for each Census division to the average regional citygate price plus an estimated residential distribution markup, multiplied by a calculated STEO benchmark factor. The markup is a function of consumption per household, which is intended to capture fixed distribution charges, and has a constant term to capture variable charges, as follows:

$$MARKUP\_RES_{yr,r} = C_r^{RES} + \alpha^{RES} * \frac{QCD\_RES_{yr,r}}{HOUSES_{yr,r}} + \beta^{RES} * \frac{QCD\_RES_{yr,r}}{HDD_{yr,r}} \tag{71}$$

$$PRICE\_RES_{yr,r} = (PRICE\_CG\_CD_{yr,r} + MARKUP\_RES_{yr,r}) * STEO\_RES_{yr,r} \tag{72}$$

where

$MARKUP\_RES_{yr,r}$ = Markup from citygate price to delivered price to residential customers for Census division r in projection year yr (1987$/Mcf)

$PRICE\_RES_{yr,r}$ = Delivered price to residential customers for Census division r in projection year yr (1987$/Mcf)

$PRICE\_CG\_CD_{yr,r}$ = Quantity-weighted average citygate prices for Census division r in projection year yr, set using state/month level citygate prices and residential plus commercial consumption as weights (1987$/Mcf)

[55] While the model is structured to allow the user to calculate delivered prices using different markups and different base prices, as relevant, the particular options used for AEO2018 are generally the only ones described in the documentation.

$QCD\_RES_{yr,r}$ = natural gas consumed by residential customers in Census division r in projection year yr (Bcf)

$HOUSES_{yr,r} =$ Number of residential households in Census division r that consume natural gas in projection year yr

$HDD_{yr,r} =$ Number of heating degree days in Census division r in projection year yr

$STEO\_RES_{yr,r}$ = factor to align initially calculated residential prices to prices forecasted in the STEO for Census division r in projection year yr, phased to 1.0 after a user-specified number of years after the last STEO year (*NumberOfSTEOPhaseOutYears*)

C_r^{RES} = estimated constant term for Census division r (1987$/Mcf) (*DIV_RES*(option1))

α^{RES} = estimated parameter (*PAR1_RES*(option1))

β^{RES} = estimated parameter (*PAR2_RES*(option1))

r = Census division

yr = projection year

Historical annual average residential and citygate prices, residential consumption, and the number of residential households using natural gas were used to estimate the parameters in the above equation by Census division, as described in Appendix G.

Commercial sector

Average annual prices charged to commercial customers are set similarly to the residential sector prices, using the same average citygate prices, but with the following equation for commercial distribution markups:

$$MARKUP\_COM_{yr,r} = C_r^{COM} + \alpha^{COM} * \frac{QCD\_COM_{yr,r}}{FLOORSPACE_{yr,r}} + \beta^{COM} * QCD\_COM_{yr,r} \tag{73}$$

$$PRICE\_COM_{yr,r} = (AVG\_PRICE\_CG_{yr,r} + MARKUP\_COM_{yr,r}) * STEO\_COM_{yr,r} \tag{74}$$

where

$MARKUP\_COM_{yr,r}$ = Markup from citygate price to delivered price to commercial customers in Census division r in projection year y (1987$/Mcf)

$PRICE\_COM_{yr,r}$ = Delivered price to commercial customers in Census division r in projection year y (1987$/Mcf)

$AVG\_PRICE\_CG_{yr,r}$ = Quantity-weighted average citygate prices in Census division r in projection year y, set using state/month level citygate prices and residential plus commercial consumption as weights (1987$/Mcf)

$QCD\_COM_{yr,r}$ = natural gas consumed by commercial customers in Census division r in projection year y (Bcf)

$FLOORSPACE_{yr,r}$ = total commercial floorspace in Census division r in projection year y (million square feet)

$STEO\_COM_{yr,r}$ = factor to align initially calculated commercial prices to prices forecasted in the STEO in Census division r in projection year y, phased to 1.0 after a user-specified number of years after the last STEO year (NumberOfSTEOPhaseOutYears)

C_r^{COM} = estimated constant term for Census division r (1987$/Mcf) (*DIV_COM*(Option 1))

α^{COM} = estimated parameter (*PAR1_COM*(Option 1))

β^{COM} = estimated parameter (*PAR2_COM*(Option 1))

r = Census division

yr = projection year

Historical annual average commercial and citygate prices, commercial consumption, and commercial floorspace were used to estimate the parameters in the above equation by Census division, as described in Appendix G.

Industrial sector

The average annual prices charged to the industrial sector are set based on the quantity-weighted average spot price in each Census division, averaged from state/monthly spot prices using industrial consumption as weights. Average markups by Census division are set based on the historical difference between delivered prices to the industrial sector and this average spot price, and held constant through the projection period. Historical prices for the industrial sector are estimated rather than extracted directly from annual/state level published EIA prices. These prices only reflect revenues received from industrial customers who purchase gas from local distribution companies, or about 15% of the sector's consumption. However, price data from EIA's Manufacturing Energy Consumption Survey (MECS) are assumed to better approximate prices seen by the whole sector, even though they do not include nonmanufacturing industries. Since the survey only provides prices every four years and by the four Census regions, an estimation (see Appendix G) was necessary to fill in the missing years and regional detail. Furthermore, since prices from the STEO are based on EIA's annual/state level prices, the NGMM did not benchmark the industrial prices to align with STEO. Industrial prices are set as follows:

$$PRICE\_IND_{yr,r} = (AVGind\_PRICE\_SPOT_{yr,r} + MARKUP\_IND_{yr,r}) * STEO\_IND_{yr} \quad (75)$$

where

$PRICE\_IND_{yr,r}$ = Delivered price to industrial customers in Census division r in projection year yr (1987$/Mcf)

$MARKUP\_IND_{yr,r}$ = Historically based markup from quantity-weighted average spot price to delivered price to industrial customers in Census division r in projection year yr (1987$/Mcf), set as average over user-specified historical years (*Year_IND*).

$AVGind\_PRICE\_SPOT_{yr,r}$ = Quantity-weighted average spot prices in Census division r in projection year yr, using state/month level spot prices and industrial consumption as weights (1987$/Mcf)

$STEO\_IND_{yr}$ = factor to align industrial prices to STEO results in projection year yr (set to 1.0 since not used)

$r =$ Census division

$yr =$ projection year

While the price to the industrial sector in NEMS is separately categorized for core and noncore customers, this distinction is no longer being used, and the same price is assigned to both NEMS variables.

Electric generation sector

The NGMM provides delivered prices to electric generators to, and receives consumption levels by electric generators from, the Electricity Market Module (EMM) in the NEMS by 17 regions (one of which is Alaska) and three seasons. For the regions in the Lower 48 states, these prices are based on the average regional/seasonal spot price, calculated by averaging over state/month spot prices, with state/month electric generator consumption levels as weights. The base markup or the lagged markup in the first projection year is set to an historical average difference between the delivered price and spot price in each region/season. The projected markup in each year is allowed to increase/decrease depending on how much the electric generator consumption increases/decreases compared to consumption in the other sectors.[56] This is intended to reflect that electric generators will likely need to reserve more space on the pipeline system as their market share increases. Because these markups can theoretically be negative, the spot price is added to the markup to ensure it is positive and then subtracted after the scaling is applied. Since the STEO only forecasts a single national price for electric generators, the model code only sets/uses one STEO benchmark factor for each STEO year to ensure that the quantity-weighted average annual/national price to electric generators aligns with the annual/national STEO value. These factors are phased to 1.0 after the last STEO year as is done for the residential and commercial sectors. The relevant equations follow:

$$MARKUP\_ELEC_{yr,p,e} = (MARKUP\_ELEC_{yr-1,p,e} + AVGelec\_PRICE\_SPOT_{yr,p,e}) * \left[\frac{1 + \frac{QEMM\_ELEC_{yr,p,e} - QEMM\_ELEC_{yr-1,p,e}}{QEMM\_ELEC_{yr,p,e}}}{1 + \frac{QEMM\_TOT_{yr,p,e} - QEMM\_TOT_{yr-1,p,e}}{QEMM\_TOT_{yr,p,e}}}\right]^{Factor\_EL} - AVGelec\_PRICE\_SPOT_{yr,p,e} \tag{76}$$

$$PRICE\_ELEC_{yr,p,e} = (AVGelec\_PRICE\_SPOT_{yr,p,e} + MARKUP\_ELEC_{yr,p,e}) * STEO\_ELEC_{yr} \tag{77}$$

where

[56] This ratio is represented in brackets in the equation below and is limited to fall between 0.5 and 2.0.

$PRICE\_ELEC_{yr,p,e}$ = delivered price to electric generators in projection year yr, season p, and NGEMM region e (1987$/Mcf)

$MARKUP\_ELEC_{yr,p,e}$ = historically based markup from quantity-weighted average spot price to delivered price to electric generators in projection year yr, season p, and NGEMM region e (1987$/Mcf), set as average over user-specified historical years (*Year_EL*).

$AVGelec\_PRICE\_SPOT_{yr,p,e}$ = quantity-weighted average spot prices in projection year yr, season p, and NGEMM region e, using state/month level spot prices and electric generator consumption as weights (1987$/Mcf)

$QEMM\_ELEC_{yr,p,e}$ = electric generator consumption in projection year yr, season p, and NGEMM region e (Bcf)

$QEMM\_TOT_{yr,p,e}$ = total delivered consumption across all sectors in projection year yr, season p, and NGEMM region e (Bcf)

$STEO\_ELEC_{yr}$ = factor to align national average electric generator prices to STEO results in projection year yr

$Factor\_EL$ = assumed parameter, set exogenously

p = seasonal period (peak – December to March, offpeak – June to September, shoulder – remaining months)

e = sixteen NGEMM regions in the Lower 48 states

yr = projection year

The price to electric generators in Alaska does not vary by season and is set by adding a historically based markup to an estimated citygate price for Alaska (see Appendix G), as follows:

$$AK\_PRICE\_CG_{yr} = e^{\alpha} * \left(WOP_{yr}\right)^{\beta} \tag{78}$$

$$PRICE\_ELEC_{yr,p,AK} = (AK\_PRICE\_CG\_AK_{yr} + AK\_MARKUP\_ELEC_{yr}) * STEO\_ELEC_{yr} \tag{79}$$

where

$PRICE\_ELEC_{yr,p,AK}$ = delivered price to electric generators in Alaska in projection year yr and season p (1987$/Mcf)

$AK\_PRICE\_CG_{yr}$ = citygate price in Alaska in projection year yr (1987$/Mcf)

$AK\_MARKUP\_ELEC_{yr}$ = historically based markup from city gas price in Alaska to delivered price to electric generators for projection year yr (1987$/Mcf), exogenously specified (*PriceMarkup*)

$STEO\_ELEC_{yr}$ = factor to align national average electric generator prices to STEO results for projection year yr

WOP_{yr} = U.S. crude oil imported refinery acquisition cost for projection year yr (1987$/barrel)

α = estimated parameter, constant term in log-log regression (*x_AK_Citygate1*)

β = estimated parameter (*x_AK_Citygate2*)

p = seasonal period (peak – December to March, offpeak – June to September, shoulder – remaining months)

AK = NGEMM region 17, representing Alaska

yr = projection year

Transportation sector

End-use, or delivered, natural gas prices to the transportation sector (i.e., to natural gas fueled vehicles) are calculated for two fuel types (compressed natural gas-CNG, liquefied natural gas-LNG) and 4 different modes of transportation: personal vehicles (cars and trucks), fleet vehicles (cars, trucks, and buses), rail, and marine. These prices, 8 in total, have 4 different components:

- *Price of natural gas delivered to the dispensing station or a LNG facility (either citygate price plus historical markup, industrial gas price, or electric gas price)*
- *For LNG, the cost of liquefying and transporting fuel to the dispensing station*[57]
- *Retail markup, or the cost of delivered CNG or LNG at the dispensing station above the base price (includes per-unit cost of dispensing fuel)*
- *Federal and state motor fuels taxes*

The base price of natural gas is a model assumption; all three options can be used. Using the citygate price as the basis for fuel prices to vehicles implies that dispensing stations buy from a local distribution company (LDC) and have the additional cost of reserving firm capacity on pipelines as part of the end use price. Personal and fleet vehicles use the citygate price as their base price. The historical markup is calculated based on the historical difference between the price of CNG from either public stations (i.e. personal vehicles) or private stations (i.e. fleet vehicles) reported in the Office of Energy Efficiency and Renewable Energy's quarterly *Clean Cities Alternative Fuels Price Report*[58] and the historical citygate price. On the other hand, using the industrial or electric prices[59] as a base price indicates that stations or LNG facilities buy natural gas and reserve pipeline space similarly to these sectors (i.e. on an interruptible basis and in large volumes).

The fuel cost, represented as a loss factor, associated with liquefying and transporting LNG to the dispensing station is assumed to be the same as that assumed for LNG export facilities plus an additional loss factor similar to that for CNG.

[57] For AEO2018, the cost associated with the fuel used to liquefy and transport LNG to a dispensing station was captured in the NGMM via a loss factor (LOSS); however, additional charges associated with providing this service (e.g., capital cost recovery) were inadvertently excluded. This will be corrected in the future.
[58] U.S. Office of Energy Efficiency & Renewable Energy, *Clean Cities Alternative Fuel Price Report*.
[59] For AEO2018, all but CNG vehicles are assumed to see prices based on the industrial price.

Retail markups at dispensing stations for the eight categories of natural gas vehicle fuel were calculated based on assumed sizes and costs of generic dispensing facilities, short of motor fuel taxes.[60] The series of equations to derive these retail markups follow:

$$CAPEX\_YR_{f,v} = WACC_{f,v} * \frac{\left(1 + WACC_{f,v}\right)^{YRS\_INVEST_{f,v}}}{\left[\left(1 + WACC_{f,v}\right)^{YRS\_INVEST_{f,v}}\right] - 1} \tag{80}$$

$$CAPEX\_MCF_{f,v} = \frac{CAPEX\_TOT_{f,v}}{CAP\_DAY_{f,v} * 365 * CAP\_UTIL_{f,v}} * CAPEX\_YR_{f,v} \tag{81}$$

$$COST\_RETAIL_{f,v} = CAPEX\_MCF_{f,v} + OPEX\_MCF_{f,v} \tag{82}$$

where, for fuel f and vehicle type v,

$CAPEX\_YR_{f,v}$ = cost that must be recovered each year in order to recover the capital expenditures necessary to build a dispensing station (1987$)

$CAPEX\_MCF_{f,v}$ = cost added to the price per unit of fuel dispensed in order to recover the capital costs (1987$/Mcf)

$COST\_RETAIL_{f,v}$ = total markup, or cost added, to the retail price in order to recover all capital and operational costs of a dispensing station (1987$/Mcf)

$WACC_{f,v}$ = weighted average cost of capital for a the construction of a dispensing station, representing the discount rate for calculating net present value (%); represents the minimum rate of return required to satisfy investors

$YRS\_INVEST_{f,v}$= number of years over which the total capital expenditures are expected to be fully recovered for the dispensing station

$CAPEX\_TOT_{f,v}$ = total capital expenditure required to construct a dispensing station (1987$)

$CAP\_DAY_{f,v}$ = daily capacity, or total volume of fuel able to be dispensed, of a dispensing station (Mcf/d)

$CAP\_UTIL_{f,v}$= expected utilization of a dispensing station (%)

$OPEX\_MCF_{f,v}$ = cost required to operate a dispensing station per unit of fuel dispensed (1987$/Mcf)

f = fuel type (CNG, LNG)

[60] *Assumptions to the Annual Energy Outlook 2018: Natural Gas Market Module*

v = vehicle type (personal, fleet, rail, marine)

Finally, appropriate federal and state motor fuels taxes, net of credits, are added to the price. Federal taxes are held constant in nominal dollars throughout the projection period, consistent with the federal tax code. While the laws for adjusting state taxes vary, a simplifying assumption was applied in the NGMM that state taxes are constant in real dollars of the first model year; therefore, it is assumed only federal taxes rise with inflation while state taxes do not.

The following equations are used to set transportation prices:

For markups from industrial sector price, the following equation is used:

$$PRICE\_TRANS_{yr,f,v,r} = COST\_RETAIL_{f,v} + \frac{TAX\_FED_{yr,f}}{GDP\_87_{yr}} + \frac{TAX\_STATE_{f,r}}{GDP\_87_{2016}} + \left[\left(1 + LOSS_{f,v}\right) * PRICE\_IND_{yr,r}\right] \tag{83}$$

For markups from the electric power sector price:

$$PRICE\_TRANS_{yr,f,v,r} = COST\_RETAIL_{f,v} + \frac{TAX\_FED_{yr,f}}{GDP\_87_{yr}} + \frac{TAX\_STATE_{f,r}}{GDP\_87_{2016}} + \left[\left(1 + LOSS_{f,v}\right) * PRICE\_ELEC_{yr,r}\right] \tag{84}$$

For markups from the citygate price, used for CNG fleet and personal vehicles:

$$PRICE\_TRANS_{yr,f,v,r} = COST\_RETAIL_{f,v} + HIST\_TARIFF_{f,v,r} + \frac{TAX_{FED_{yr,f}}}{GDP\_87_{yr}} + \frac{TAX\_STATE_{f,r}}{GDP\_87_{2016}} + \left[\left(1 + LOSS_{f,v}\right) * PRICE\_CG_{yr,r}\right] \tag{85}$$

where

$PRICE\_TRANS_{yr,f,v,r}$ = delivered price of transportation fuel to consumers at dispensing station in projection year yr for fuel type f, vehicle type v, and Census division r (1987$/Mcf)

$COST\_RETAIL_{f,v}$ =assumed additional charge related to dispensing fuel f to customers for vehicle type v (1987$/Mcf)

$TAX\_FED_{yr,f}$ = Federal motor vehicle fuel tax in year yr for fuel f, excluded when setting prices for marine vehicles (nominal$/Mcf)

$TAX\_STATE_{f,r}$ = average state motor vehicle fuel tax in year yr for Census division r excluded when setting prices for marine and rail vehicles (2016$/Mcf)

GDP_87_{yr} = GDP conversion from year yr dollars to 1987 dollars (from the NEMS macroeconomic module)

$LOSS_{f,v}$ = fuel loss associated with converting natural gas to fuel f and transporting it to dispensing station for vehicle type v (*Trans_PctFuelLoss*)[61]

$PRICE\_IND_{yr,r}$ = delivered price natural gas to industrial sector in Census division r and projection year yr (1987$/Mcf)

$PRICE\_ELEC_{yr,r}$ = average annual delivered price of natural gas to the electric power sector in Census division r and projection year yr (1987$/Mcf)

$HIST\_TARIFF_{f,v,r}$ = average historical tariff for the transportation sector to deliver natural gas from the citygate to the station for fuel f and vehicle type v in Census division r over the last user- specified number of years (*NumberofYearsforAverage_Trans*) (1987$/ Mcf)

$PRICE\_CG_{yr,r}$ = citygate price in Census division r and projection year yr, (1987$/Mcf)

yr = projection year

f = fuel type (CNG, LNG)

v = vehicle type (personal, fleet, rail, marine)

r = Census division

Reporting to the NEMS

During post-processing, several additional values must be calculated for reporting to the NEMS. These include the following:

- *Lease fuel consumption*
- *Plant fuel consumption*
- *Pipeline fuel consumption*
- *Annual and regional flows of natural gas*

For details as to how they are assigned, please refer to the pre-processing section of the NGMM documentation.

The final step in the NGMM (for all iterations except for the NEMS reporting loop) is filling the NEMS global arrays (within the NGMM, all NEMS variables are renamed and mapped to NGMM indexes to adhere to the model code's naming conventions and units). In the procedure *Write_to_NEMS*, NGMM parameters are assigned to the corresponding NEMS global variable; additionally, any aggregations unique to NEMS variables (e.g. assigning a total U.S. value to the final position of an array) are calculated here.

[61] Currently this is only associated with LNG, so the value for CNG is zero.

Appendix A. Model abstract

Model Name

Natural Gas Market Module

Acronym

NGMM

Description

The NGMM models the North American natural gas transmission and distribution network that links the suppliers and consumers of natural gas (including global LNG markets), and in so doing determines the regional market clearing natural gas end-use and supply prices. Model outputs include the following:

- *Average annual natural gas end-use price levels by sector and Census division*
- *Average annual natural gas production volumes and prices by OGSM region*
- *Average annual natural gas import and export volumes (pipeline and LNG) and prices (pipeline) by Census division*
- *Annual pipeline fuel consumption by Census division*
- *Annual lease and plant fuel consumption by Census division*
- *Annual flow of gas between regions*
- *Annual pipeline capacity additions and utilization levels by arc*

Purpose

The NGMM is the component of the National Energy Modeling System (NEMS) that represents the mid-term natural gas market. The purpose of the NGMM is to derive natural gas supply and end-use prices and flow patterns for movements of natural gas through the regional interstate network. The prices and flow patterns are derived by obtaining a market equilibrium across the three main components of the natural gas market: the supply component, the demand component, and the transmission and distribution network that links them.

Date of last update

This is the first updated documentation published for the NGMM. The first version of this documentation was published in October 2018. Documention for its predecessor, the Natural Gas Transmission and Distribution Module, was last published in 2014.

Part of another model

NEMS

Model Interfaces

Model receives input from the Macroeconomic Activity Module, the International Energy Module, the Liquid Fuels Market Module, the Oil and Gas Supply Module, the Residential Demand Module, the Commercial Demand Module, the Industrial Demand Module, the Transportation Demand Module, the Integrating Module, and the Electricity Market Module.

The model provides outputs to the Macroeconomic Activity Module, Liquid Fuels Market Module, the Oil and Gas Supply Module, the Residential Demand Module, the Commercial Demand Module, the Industrial Demand Module, the Transportation Demand Module, the Integrating Module, and the Electricity Market Module.

Official Model Representative

Office of Energy Analysis
Office of Petroleum, Natural Gas, and Biofuels Analysis, EI-33
Model Contact: Katie Dyl
Telephone: (202) 287-5862
Email: Kathryn.dyl@eia.gov

Documentation

Natural Gas Market Module of the National Energy Modeling System: Model Documentation 2020

Archive Media and Manuals

- *The NGMM is archived as a component of NEMS, which is available from the Annual Energy Outlook 2020 website as a zip file:* https://www.eia.gov/outlooks/aeo/info_nems_archive.php.
- *For more information and detailed instructions, go to the* Annual Energy Outlook Archive *site.*

Energy system described

North American natural gas market transmission and distribution

Coverage

- **Geographic***: The NGMM represents the natural gas transmission system within the United States, Canada, and Mexico. The Lower 48 states are represented at the state level. Canada is represented by an eastern and western region. Mexico is represented by 5 regions: northwest, northeast, interior-west, central, and south-southeast. Destinations for LNG exports are represented as Atlantic basin (Europe) or Pacific basin (Asia). Supply and demand regions are defined by their respective NEMS modules.*
- **Time Unit/Frequency:** *Model is solved on a monthly level from user-defined first model year (i.e. 2016) to user-defined last model year (2050). Time is represented from 1990 through 2080.*
- **Product(s):** *Natural gas, liquefied natural gas (LNG)*
- **Economic Sector(s):** *Residential, commercial, industrial, electric generators and transportation*

Modeling Features

- **Model structure:** *quadratic mathematical program that maximizes consumer plus producer surplus, minus transportation costs, subject to linear mass balance and capacity constraints*
- **Model technique:** *Natural gas supply and demand markets are balanced with the marginal price setting market prices. Demands are held constant and disaggregated according to historical data. Expected variable supply is modeled as a short-term supply curve. Storage injections and withdrawls are set to historical levels, scaled to equal zero net injections in all projection years. Liquefied natural gas export capacity is build if it is determined its net present value over a planning horizon is economically favorable given global LNG prices.*

- **Special features:** *Can run stand-alone in AIMMS or within NEMS, but should always be run with the Oil and Gas Supply Module for best results. Report pages have been developed in AIMMS wthin a graphical user interface to visualize model results at a disaggregated level. Complete run results are saved in case files for each NEMS cycle and can be loaded into AIMMS independent of NEMS. The model can be run either keeping AIMMS open throughout the entire run (if sufficient AIMMS licenses are available) or by opening and closing the AIMMS each time the model is called (lengthening runtime considerably).*

Non-DOE Input Sources

- *Natural Gas Intelligence*
  - *Historical spot prices*
- *Secretaría de Energía de México/Sistema de Información Energética*
  - *Historical annual Mexico gross natural gas production by supply type and field*
  - *Historical Mexico consumption by month, sector, and region*
  - *Historical dry natural gas balance and distribution*
  - *Historical annual LNG imports to Mexico by terminal*
  - *Historical annual natural gas pipeline capacities*
- *Comisión Reguladora de Energía (Mexico)*
  - *Historical natural gas spot prices by month for Mexican market regions*
- *Statistics Canada*
  - *Historical Canadian storage injections and withdrawals by month and province*
- *National Energy Board of Canada*
  - *Historical annual natural gas pipeline capacities and flows within Canada by region*
  - *Historical Canadian consumption by month, sector, and province*
  - *Historical LNG imports into Canada by month*
- *Alberta Energy Regulator*
  - *Historical annual bitumen production from oil sands in Alberta by mining type*
  - *Historical annual natural gas produced, consumed, and purchased for oil sands production by mining type*
- *Internal Revenue Service*
  - *Federal natural gas vehicle taxes by fuel type*
- *State of Alaska, Department of Labor and Workforce Development*
  - *Alaska population projections by year*

DOE Input Sources

- *Energy Information Administration, Natural Gas Annual/Natural Gas Monthly*
  - *Natural gas consumption and delivered prices by month, state, and sector*
  - *Natural gas pipeline import and export volumes and prices by month, state, and border crossing*
  - *Natural gas storage injections and withdrawals by month and state*
  - *Balancing item by state and year*
  - *Interstate flows of natural gas by state and year*
  - *Citygate prices by state and month*
  - *Number of residential customers for natural gas by state and eyar*

  - *Supplemental supply volumes by state and year*
  - *Pipeline fuel consumption, lease fuel consumption, and plant fuel consumption by state and year*
  - *Natural gas plant liquid volumes processed, extraction losses, and total condensate by state and year*
- *Energy Information Administration, natural gas pipeline data*
  - *Historical U.S. state-to-state natural gas pipeline capacity by year and state*
  - *Planned natural gas pipeline projects by year and state*
- *Energy Information Administration, Electric Power Monthly*
  - *Natural gas consumption and prices to electric generators by state and month*
- *Energy Information Administration, EIA-846, Manufacturing Energy Consumption Survey*
  - *Base year core and non-core industrial end use prices by Census region*
- *Energy Information Administration, Short-Term Energy Outlook*
  - *Natural gas delivered end use price forecasts by Census division for the the first 2 years beyond history*
  - *National natural gas market forecast for the first two years beyond history*
- *Energy Information Administration, International Energy Outlook*
  - *Natural gas consumption projections for Canada and Mexico by sector and year*
  - *Natural gas production projection for Mexico by year*
  - *Projected flexible liquefied natural gas supplies (i.e. liquefied natural gas volumes not sold under contracts) available to the global market by year*
  - *Liquefied natural gas imports into Europe and Asia by year*
- *Office of Fossil Energy*
  - *Liquefied natural gas export capacity planned and under construction by facility*
  - *Import and export volumes and prices by border crossing*
- *Office of Energy Efficiecy and Renewable Energy, Clean Cities Alternative Fuel Price Report*
  - *Delivered compressed natural gas prices to the transportation sector at public and private dispensing stations*
- *Office of Energy Efficiecy and Renewable Energy, Alternative Fuels Data Center*
  - *State natural gas vehicle taxes by fuel type*

Computing Environment

- Hardware Used: PC
- Operating System: UNIX simulation (in NEMS), Windows (Stand-alone)
- Language/Software Used: AIMMS
- Memory Requirement: Unknown
- Storage Requirement: Unknown
- Estimated Run Time: 45 minutes (within NEMS running with the Oil and Gas Supply module and opening and closing AIMMS each time the model is called)

Independent Expert Review Conducted

Lauren K. Busch, Leidos*Review of Natural Gas Models in support of U.S. Energy Information Administration Natural Gas Transmission and Distribution Module (NGTDM) Redesign Effort*. Washington, DC, September 4, 2014.

EIA Network Modeling Workshop. Washington, DC, September 4, 2014. Participants and commentary from the following organizations: U.S. Energy Information Administration, OnLocation, Leidos, RBAC, ICF, NERA, DOE, GA Tech, UMD, Chevron.

Reginald Sanders, OnLocation. *Review of Natural Gas Transmission and Distribution Module — Component Design Report*. Washington, DC, June 17, 2015.

Andy Kydes, *Review of Natural Gas Transmission and Distribution Module — Component Design Report*. Washington, DC, June 17, 2015.

Joseph Benneche, U.S. Energy Information Administration. U.S. Energy Information Administration. *Natural Gas Transmission and Distribution Module Component Design Report: Discussion of Model Design (Review Meeting 1)*. Washington, DC, May 27, 2015. Participants from the following organizations: U.S. Energy Information Administration, OnLocation, Leidos.

Joseph Benneche, U.S. Energy Information Administration. *Natural Gas Transmission and Distribution Module Component Design Report: Discussion of Model Design (Review Meeting 2)*. Washington, DC, July 21, 2015. Participants from the following organizations: U.S. Energy Information Administration, OnLocation, Leidos, RBAC.

Status of Evaluation Efforts by Sponsor

EIA continues to evaluate and improve historical calibration. Future goals include running the NGMM in historical years to set model input parameters.

Appendix B. References

AIMMS, *AIMMS—The Language Reference*, 2012.

AIMMS, *AIMMS—The User's Guide*, July 8, 2015.

Energy Information Administration, Office of Energy Analysis, Model Documentation Report: Natural Gas Transmission and Distribution Module, July 2014.

Energy Information Administration, Office of Energy Analysis, *Natural Gas Transmission and Distribution Module Component Design Report*, August 2015.

Energy Information Administration, Office of Energy Analysis, *Requirements for a Redesigned Natural Gas Transmission and Distribution Module in the National Energy Modeling System*, August 2014.

ICF International, "Changes in Mexico's Gas Markets and Implications for Investment and Trade," report submitted to Energy Information Administration, November 29, 2016.

Leidos, Review of Natural Gas Models: In Support of U.S. Energy Information Administration Natural Gas Transmission and Distribution Module (NGTDM) Redesign Effort, September 2014.

National Energy Board, *Canada's Energy Future*, January 2016.

Appendix C. Documentation variables mapped to model identifiers

Section	Name in Documentation	Name in Model
Capacity Expansion		
	Q_CAPEX	ConsumptionForCapExp
	Q_MONTH	Cons_State_Mon
	f^1	WeatherFactor1
	f^2	WeatherFactor2
Supply		
	PROD_GOM	ImpliedStateGOMproduction
	PROD	AnnualSupply
	α(st)	GOM_OGDIST2StateShare
	β(st)	GOM_OGDIST2StateShare
Pipeline Fuel		
	PIP_DIST	DistributionLosses
	Q_RES	Cons_State_Mon
	Q_COM	Cons_State_Mon
	PCT_DIST	DistributionLossFactor
	PIP_STORE	StorageLosses
	STORE_INJ	StorageInjections
	STORE_WTH	StorageWithdrawals
	PCT_STORE	StorageLossFactor
	PIP_INTRA	IntrastatePipeFuelLosses
	PROD_DRY	ActualProductionMonthly
	PCT_INTRA	IntrastatePipeFuelFactor
	PIP_TRANS	TranFuelLosses
	P_LOSS	Ploss
	FLOW	HistoricalFlowAnnual
	FLOW_IN	HistoricalFlowIn
	FLOW_OUT	HistoricalFlowOut
	TOTAL	Intersum
Balancing Item		
	DISC	BalanceItem
	Q_TOT	TotalConsumption
	Q_PIP	HistoricalAnnualPipeFuel
	Q_LAP	LeaseAndPlantFuelAnnual
	PCT_LIQ	(historically in pipe fuel)
	LNG_EXP	LNGExports
	STORE_INJ	StorageInjections
	STORE_WTH	StorageWithdrawals
	SUP_TOT	Supply

Section	Name in Documentation	Name in Model
	FLOW_IN	HistoricalFlowAnnual
	FLOW_OUT	HistoricalFlowAnnual
Storage		
	AVE_YR_INJ	AnnualAverageStorageInjections
	AVE_YR_WTH	AnnualAverageStorageWithdrawals
	α(storage)	StorageScalingParameter
	AVE_INJ	AverageStorageInjections
	AVE_WTH	AverageStorageWithdrawals
Mexico		
	PROD_MX_AD	MX_Oil_Production
	WOP	WorldOilPrice
	α_Oil	x_MX_Oil_Prod1_2
	β_Oil	x_MX_Oil_WOP2_2
	PROD_MX_AD	AnnualSupply
	α_AD	x_MX_AD_Prod1
	β_AD	x_MX_AD_WOP1
	γ_AD	x_MX_AD_OilProd
	PROD_MX_NA	AnnualSupply
	HH_PRICE	HenryHubPrice
	α(NA,t)	x_MX_NA_Prod1_1, x_MX_NA_Prod1_2
	β(NA,t)	x_MX_NA_HH1_1, x_MX_NA_HH1_2
	Cons_MX_PEMEX	Cons_MX_PEMEX
	α_PEMEX	x_MX_Cons_PEMEX_1
	β_PEMEX	x_MX_Cons_PEMEX_Oil
	Cons_MX_Ind_other	Cons_MX_Industrial
	α_Ind	x_MX_Cons_Industrial_1
	β_Ind	x_MX_Cons_Industrial_HH
	C_Ind	c_MX_Cons_Industrial
	Cons_MX_Ind	IEO_MX_Consumption
LNG Exports		
	PRICE_LNG	WorldLNGPrice
	FLEX	LNG_WorldFlex
	LNG_USA	LNG_USFlex
	LNG_ADD	LNG_Increment
	Q_LNG	LNG_Demand
	α_d	LNG_ExpOil
	β_d	LNG_ExpFlexLNGYr
	PRICE_SUP	USLNGSupplyPrice
	LAST_PRICE_SUP	LastCycle_USLNGSupplyPrice
	FACTOR_LNG	LNG_ConvergenceFactor

Section	Name in Documentation	Name in Model
	LAST_PROD	LastCycle_Production
	LNG_CAP	LNG_ExportCapacity
	LAST_LNG_CAP	LastCycle_LNGExports
	γ	LNG_Gamma_Adj
	γ_1	x_LNG_Gamma1
	γ_2	x_LNG_Gamma2
	PROD_ACT	ActualProductionAnnual
	PROD_EXP	AnnualSupply
	PRICE_USLNG	LNG_USLookAheadPrice
	PCT_LIQ	LNG_PctFuelCharge
	COST_LIQ	LNG_Liquefaction
	COST_REGAS	LNG_Regasification
	COST_SHIP	LNG_ShippingCost
	NVP_USLNG	LNG_USDiscount
	DCF_RATE	LNG_DCFDiscountRate
Alaska		
	AK_PROD_N	AK_Production
	AK_PROD_S	AK_Production
	AK_LAP	AK_LeasePlant
	AK_LNG_EXP	AK_LNG_Exports
	AK_Q_TOTAL	AK_Cons_EndUse
	PCT_LIQ	Pct_Liquifaction_Fuel
	PCT_PIP	AK_PIP_Percent
	α^{LAP}	x_AK_N_LAP2
	β^{LAP}	x_AK_N_LAP3
	C^{LAP}	x_AK_N_LAP1
	AK_PROD_OIL	AK_Prod_Crude_Total
	AK_Q_RES	AK_Cons_EndUse
	AK_Q_COM	AK_Cons_EndUse
	AK_POP	Population
	UNEMP	Unemployment
	α^{RES}	x_AK_Cons_Residential1
	β^{RES}	x_AK_Cons_Residential2
	γ^{RES}	x_AK_Cons_Residential3
	α^{COM}	x_AK_Cons_Commercial1
	β^{COM}	x_AK_Cons_Commercial2
	γ^{COM}	x_AK_Cons_Commercial3
	AK_DISC	AK_Discrepancy
Imports and exports		

Section	Name in Documentation	Name in Model
	LNGEXP	AnnualLNGExports
	Q_LIQ	LNGFuelForLiquefaction
	LNG	QTotalLNGExports
	PCT_LIQ	Pct_Liquifaction_Fuel
	LSMAX	LNGExpCrv_MaxStep
	IMP_CN	Imports_Canada
	EXP_CN	Exports_Canada
	IMP_MX	Imports_Mexico
	EXP_MX	Exports_Mexico
	FLOWH2H	FlowHubToHub
Delivered prices		
	PRICE_CG_ST	Price_Citygate
	PRICE_CG_CD	PriceCitygateAnnualforMarkups
	PRICE_SPOT	Price_Spot
	Q_RES	Cons_State_Mon
	Q_COM	Cons_State_Mon
	$\alpha_{per(mon),st}$	pspot_peak, pspot_offpeak
	$\beta_{per(mon),st}$	PQ_peak, PQ_offpeak
	$C_{per(mon),st}$	Inter_peak, inter_offpeak
Residential	PRICE_RES	Price_Enduse
	MARKUP_RES	Markups_Enduse
	QCD_RES	NEMS_Consumption
	HOUSES	Households
	C^{RES}_{r}	DIV_RES
	α^{RES}	PAR1_RES
	ρ^{RES}	PAR2_RES
	STEO_RES	STEOEndUsePriceFactor
Commercial	PRICE_COM	Price_Enduse
	MARKUP_COM	Markups_Enduse
	QCD_COM	NEMS_Consumption
	FLOORSPACE	Floorspace
	C^{COM}_{r}	DIV_COM
	α^{COM}	PAR1_COM
	ρ^{COM}	PAR3_COM
	β^{COM}	PAR2_COM
	STEO_COM	STEOEndUsePriceFactor
Industrial	PRICE_IND	Price_Enduse

Section	Name in Documentation	Name in Model
	MARKUP_IND	AverageMarkupIND
	AVGind_PRICE_SPOT	PriceCitygateAnnualforMarkups
	STEO_IND	STEOEndUsePriceFactor
Electric	PRICE_ELEC	Price_EnduseElectric
	MARKUP_ELEC	Markups_Enduse_Electric
	AVGelec_PRICE_SPOT	PriceSpotAnnualElectric
	QEMM_ELEC	NEMS_Consumption_EMM
	QEMM_TOT	QOTHER
	FACTOR_EL	Factor_EL
	STEO_ELEC	STEOElectricPriceFactor
	AK_PRICE_CG	AK_Citygate
	WOP	WorldOilPrice
	α	x_AK_Citygate1
	β	x_AK_Citygate2
	AK_MARKUP_ELEC	PriceMarkup
Transportation	PRICE_TRANS	Price_EndUseTransportation
	CAPEX_YR	Trans_PctCapexPerYr
	WACC	Trans_WACC
	YRS_INVEST	Trans_YrsRecover
	CAPEX_MCF	Trans_CapexPerMcf
	CAPEX_TOT	Trans_Capex
	CAP_DAY	Trans_DailyCapacity
	CAP_UTIL	Trans_Utilization
	COST_RETAIL	Trans_CostMarkup
	OPEX_MCF	Trans_Opex
	COST_RETAIL	Trans_CostMarkup
	TAX_FED	Tax_Federal
	TAX_STATE	Tax_State
	GDP_87	GDPPriceDeflator87
	HIST_TARIFF	Trans_Tariff

Appendix D. Documentation equations mapped to procedures in the NGMM code

Equation in Documentation	Procedure in the NGMM
Model Structure and Design	
(1)	CurveDefinitions, SupplyCurveParameters
(2)-(5)	SupplyCurveParameters
(6)	CurveDefinitions
(7)-(28)	individual identifiers
(29)	DefineConsumptionForCapExp
Pre-processing	
(30)-(33)	Calculate_GOMProductionShares
(34)-(36)	Losses_InitializeData
(37)	PipefuelFactors_Initialize
(38),(39)	Flow_InitializeData
(40), (41)	PipefuelFactors_Initialize
(42)	BalancingItem_Initialize
(43)-(45)	Storage_InitializeData
(46)-(52)	Mexico_Run
(53)	Calculate_WorldLNGPrices
(54)-(58)	Calculate_USLNGExportPrices
(59)-(63)	Alaska_Subroutine_Run
Post-processing	
(64)-(65)	LNGAnnualExports_PostProcess
(66)-(69)	Import_Export_Run
(70)	CityGatePrice
(71)	EndUseMarkups_Residential
(72)	EndUsePrice
(73)	EndUseMarkups_Commercial
(74),(75)	EndUsePrice
(76)	EndUsePrice_Electric
(77)	EndUsePrice
(78),(79)	Alaska_Subroutine_Run
(80)-(82)	EndUsePrice_TransCost_Initialize
(83)-(85)	EndUsePrice_Transportation

Appendix E. Model variables mapped to input files

Input file	Identifier	Identifier type
ngassumptions	AK_ANGTS_Min_WHPrice	Parameter
ngassumptions	AK_Cons_Industrial	Parameter
ngassumptions	AK_LSE_Percent	Parameter
ngassumptions	AK_PIP_Percent	Parameter
ngassumptions	AK_PLT_Percent	Parameter
ngassumptions	Calibration_Option	Parameter (binary)
ngassumptions	c_MX_Cons_Industrial	Parameter
ngassumptions	DistributionLossFactor	Parameter
ngassumptions	Error_PriceCheck	Parameter (binary)
ngassumptions	FlowsZeroOut	Parameter
ngassumptions	GatherCharge_Add	Parameter
ngassumptions	GatheringCharge	Parameter
ngassumptions	IEO_CN_Bitumen	Parameter
ngassumptions	IEO_CN_Consumption	Parameter
ngassumptions	IEO_LastHistoricalYear	Parameter
ngassumptions	IEO_LNGExport	Parameter
ngassumptions	IEO_MX_Consumption	Parameter
ngassumptions	IEO_MX_Production	Parameter
ngassumptions	LastDataYear	Element Parameter
ngassumptions	LastNGAYear	Element Parameter
ngassumptions	map_SupCrvOption	Parameter
ngassumptions	MaxSteoIter	Parameter
ngassumptions	MX_OilRecoveryYear	Element Parameter
ngassumptions	MX_Sector_PEMEX	Element Parameter
ngassumptions	MX_Sector_Industrial	Element Parameter
ngassumptions	NoBuildYear	Element Parameter
ngassumptions	NumberOfSTEOPhaseOutYears	Parameter
ngassumptions	NumberOfSTEOYears	Parameter
ngassumptions	NumberOfYearsForAverage_Demand_	Parameter
ngassumptions	NumberOfYearsForAverage_Discrepancy_	Parameter
ngassumptions	NumberOfYearsForAverage_Discrepancy_CN_	Parameter
ngassumptions	NumberOfYearsForAverage_GOMprod_	Parameter
ngassumptions	NumberOfYearsForAverage_LeaseFuel_	Parameter
ngassumptions	NumberOfYearsForAverage_PlantFuel_	Parameter
ngassumptions	NumberOfYearsForAverage_SNG_	Parameter
ngassumptions	NumberOfYearsForAverage_LNG_	Parameter
ngassumptions	NumberOfYearsForAverage_Storage_	Parameter
ngassumptions	NumberOfYearsForAverage_Trans_	Parameter

Input file	Identifier	Identifier type
ngassumptions	NumberOfYearsForLookAhead_LNG_	Parameter
ngassumptions	OilSandsSwitch	Parameter (binary)
ngassumptions	Param_SupCrv	Parameter
ngassumptions	Param_SupElas	Parameter
ngassumptions	Parameter_LNGExpPrc	Parameter
ngassumptions	Parameter_LNGExpQty	Parameter
ngassumptions	Parameter_PrcElasticity	Parameter
ngassumptions	Parameter_SupCrv	Parameter
ngassumptions	Parameter_SupElasticity	Parameter
ngassumptions	Pct_Liquifaction_Fuel	Parameter
ngassumptions	PercentOfProductionMovedForPlantFuel	Parameter
ngassumptions	Population	Parameter
ngassumptions	PriceMarkup	Parameter
ngassumptions	PriceSpot_Add	Parameter
ngassumptions	SaveEachCycle_Switch	Parameter (binary)
ngassumptions	SaveEachIteration_Switch	Parameter (binary)
ngassumptions	STEOCNExportPercent	Parameter
ngassumptions	STEOCNImportPercent	Parameter
ngassumptions	STEONGIND	Parameter
ngassumptions	STEOScaleNG	Parameter
ngassumptions	StorageLossFactor	Parameter
ngassumptions	WeatherFactor1	Parameter
ngassumptions	WeatherFactor2	Parameter
ngassumptions	x_AK_Citygate1	Parameter
ngassumptions	x_AK_Citygate2	Parameter
ngassumptions	x_AK_Cons_Commercial1	Parameter
ngassumptions	x_AK_Cons_Commercial2	Parameter
ngassumptions	x_AK_Cons_Commercial3	Parameter
ngassumptions	x_AK_Cons_Residential1	Parameter
ngassumptions	x_AK_Cons_Residential2	Parameter
ngassumptions	x_AK_Cons_Residential3	Parameter
ngassumptions	x_AK_N_LAP1	Parameter
ngassumptions	x_AK_N_LAP2	Parameter
ngassumptions	x_AK_N_LAP3	Parameter
ngassumptions	x_MX_AD_OilProd	Parameter
ngassumptions	x_MX_AD_Prod1	Parameter
ngassumptions	x_MX_AD_WOP1	Parameter
ngassumptions	x_MX_Cons_Industrial_1	Parameter
ngassumptions	x_MX_Cons_Industrial_HH	Parameter
ngassumptions	x_MX_Oil_Prod1_2	Parameter

Input file	Identifier	Identifier type
ngassumptions	x_MX_Oil_WOP2_2	Parameter
ngassumptions	x_MX_Cons_PEMEX_1	Parameter
ngassumptions	x_MX_Cons_PEMEX_Oil	Parameter
ngassumptions	x_MX_NA_HH1_1	Parameter
ngassumptions	x_MX_NA_HH1_2	Parameter
ngassumptions	x_MX_NA_Prod2_1	Parameter
ngassumptions	x_MX_NA_Prod1_2	Parameter
ngcanada	Historical Demand	Parameter
ngcanada	HistoricalCapacity	Parameter
ngcanada	HistoricalCNFlowThru	Parameter
ngcanada	HistoricalFlowAnnual	Parameter
ngcanada	HistoricalStorageInjections	Parameter
ngcanada	HistoricalStorageWithdrawals	Parameter
ngcanada	HistoricalSupply	Parameter
ngcapacity	HistoricalCapacity	Parameter
ngcapacity	PlannedCapacity	Parameter
ngeia	HistoricalAnnualCitygatePrice	Parameter
ngeia	HistoricalAnnualDemand	Parameter
ngeia	HistoricalAnnualEndUsePrice	Parameter
ngeia	HistoricalAnnualLeaseFuel	Parameter
ngeia	HistoricalAnnualPipeFuel	Parameter
ngeia	HistoricalAnnualPlantFuel	Parameter
ngeia	HistoricalAnnualSupply	Parameter
ngeia	HistoricalBalanceItem	Parameter
ngeia	HistoricalCitygatePrice	Parameter
ngeia	HistoricalDemand	Parameter
ngeia	HistoricalEndUsePrice	Parameter
ngeia	HistoricalExports	Parameter
ngeia	HistoricalExportsPrice	Parameter
ngeia	HistoricalFlowAnnual	Parameter
ngeia	HistoricalImports	Parameter
ngeia	HistoricalImportsPrice	Parameter
ngeia	HistoricalLNGExports	Parameter
ngeia	HistoricalLNGExportsPrice	Parameter
ngeia	HistoricalLNGImportsPrice	Parameter
ngeia	HistoricalStorageInjections	Parameter
ngeia	HistoricalStorageWithdrawals	Parameter
nglngexp	Cons_EuropeOECD	Parameter
nglngexp	Cons_Japan	Parameter
nglngexp	HistoricalLNGPrice	Parameter

Input file	Identifier	Identifier type
nglngexp	LNG_AKTrainTotal	Parameter
nglngexp	LNG_CostsYrDollars	Element Parameter
nglngexp	LNG_DCFDiscountRate	Parameter
nglngexp	LNG_Demand	Parameter
nglngexp	LNG_ExpFlexLNG_adj	Parameter
nglngexp	LNG_ExpFlexLNGYr	Parameter
nglngexp	LNG_ExpOil	Parameter
nglngexp	LNG_ExportCapacity	Parameter
nglngexp	LNG_FirstYear	Element Parameter
nglngexp	LNG_Gamma_Adj	Parameter
nglngexp	LNG_GrowthRateYr	Element Parameter
nglngexp	LNG_HighPriceRatio	Parameter
nglngexp	LNG_Increment	Parameter
nglngexp	LNG_Liquefaction	Parameter
nglngexp	LNG_LowPriceRatio	Parameter
nglngexp	LNG_MaxExports	Parameter
nglngexp	LNG_MaxTransYr	Parameter
nglngexp	LNG_PctCapacityYr1	Parameter
nglngexp	LNG_PctCapacityYr2	Parameter
nglngexp	LNG_PctFuelCharge	Parameter
nglngexp	LNG_PctLossShipping	Parameter
nglngexp	LNG_PeakExports	Parameter
nglngexp	LNG_PhaseInYrs	Parameter
nglngexp	LNG_Regasification	Parameter
nglngexp	LNG_RiskThreshold	Parameter
nglngexp	LNG_ShippingCost	Parameter
nglngexp	LNG_ShippingCost	Parameter
nglngexp	LNG_step_OilPrice	Parameter
nglngexp	LNG_SunkCost	Parameter
nglngexp	LNG_WorldFlex	Parameter
nglngexp	LNG_YrsUntilBuild	Parameter
nglngexp	LNGLastHistoricalYear	Element Parameter
nglngexp	x_LNG_Gamma1	Parameter
nglngexp	x_LNG_Gamma2	Parameter
ngmarkups	ADJ_flag_	Parameter
ngmarkups	CommercialOption	Element Parameter
ngmarkups	Conv_dge_Mcf	Parameter
ngmarkups	DIV_COM	Parameter
ngmarkups	DIV_EL	Parameter
ngmarkups	DIV_IND	Parameter

Input file	Identifier	Identifier type
ngmarkups	DIV_RES	Parameter
ngmarkups	DIV_TRANS	Parameter
ngmarkups	ElectricOption	Element Parameter
ngmarkups	Factor_EL	Parameter
ngmarkups	HistoricalAnnualEndUseNNGEMMPrice	Parameter
ngmarkups	HistoricalAnnualPriceEERE	Parameter
ngmarkups	HistoricalAnnualRoadPriceEERE	Parameter
ngmarkups	HistoricalIndustrialPrice_MESC	Parameter
ngmarkups	IndustrialOption	Element Parameter
ngmarkups	Inter_offpeak	Parameter
ngmarkups	Inter_peak	Parameter
ngmarkups	Inter_peak, inter_offpeak	Parameter
ngmarkups	Inter_year	Parameter
ngmarkups	LAG_TRANS	Parameter
ngmarkups	PAR1_COM	Parameter
ngmarkups	PAR1_EL	Parameter
ngmarkups	PAR1_IND	Parameter
ngmarkups	PAR1_MESC_IND	Parameter
ngmarkups	PAR1_RES	Parameter
ngmarkups	PAR2_COM	Parameter
ngmarkups	PAR2_IND	Parameter
ngmarkups	PAR2_MESC_IND	Parameter
ngmarkups	PAR2_RES	Parameter
ngmarkups	PAR3_COM	Parameter
ngmarkups	PAR3_MESC_IND	Parameter
ngmarkups	PQ_offpeak	Parameter
ngmarkups	PQ_peak	Parameter
ngmarkups	PQ_peak, PQ_offpeak	Parameter
ngmarkups	PQ_year	Parameter
ngmarkups	pspot_offpeak	Parameter
ngmarkups	pspot_peak	Parameter
ngmarkups	pspot_peak, pspot_offpeak	Parameter
ngmarkups	pspot_year	Parameter
ngmarkups	ResidentialOption	Element Parameter
ngmarkups	Tax_Federal	Parameter
ngmarkups	Tax_State	Parameter
ngmarkups	Trans_Capex	Parameter
ngmarkups	Trans_CostsYrDollars	Element Parameter
ngmarkups	Trans_DailyCapacity	Parameter
ngmarkups	Trans_Opex	Parameter

Input file	Identifier	Identifier type
ngmarkups	Trans_PctFuelLoss	Parameter
ngmarkups	Trans_TaxYrDollars	Element Parameter
ngmarkups	Trans_Utilization	Parameter
ngmarkups	Trans_WACC	Parameter
ngmarkups	Trans_YrsRecover	Parameter
ngmarkups	TransportationOption	Element Parameter
ngmexico	HistoricalAnnualSupply	Parameter
ngmexico	HistoricalCapacity	Parameter
ngmexico	HistoricalConsumptionSENER	Parameter
ngmexico	HistoricalSpotPrice	Parameter
ngmexico	PlannedCapacity	Parameter
ngsetmap	AK_	Set
ngsetmap	AK_HI_	Set
ngsetmap	AKRegion_	Set
ngsetmap	AKRegion_South	Element Parameter
ngsetmap	AKState	Element Parameter
ngsetmap	AKSupply_	Set
ngsetmap	Alabama_	Set
ngsetmap	AlabamaGOM	Element Parameter
ngsetmap	BitumenFraction	Parameter
ngsetmap	BorderCrossings_	Set
ngsetmap	BorderCrossingsMX_	Set
ngsetmap	CaliforniaState	Element Parameter
ngsetmap	Canada_	Set
ngsetmap	CanadaEast	Set
ngsetmap	CanadaWest	Set
ngsetmap	CD_ENCentral	Element Parameter
ngsetmap	CD_ESCentral	Element Parameter
ngsetmap	CD_MidAtlantic	Element Parameter
ngsetmap	CD_Mountain	Element Parameter
ngsetmap	CD_NewEngland	Element Parameter
ngsetmap	CD_Pacific	Element Parameter
ngsetmap	CD_SAtlantic	Element Parameter
ngsetmap	CD_WNCentral	Element Parameter
ngsetmap	CD_WSCentral	Element Parameter
ngsetmap	CentralGOM	Element Parameter
ngsetmap	Correct_NGFLOWS	Parameter (binary)
ngsetmap	CrudeType_	Set
ngsetmap	DomesticSupply_	Set
ngsetmap	EasternGOM	Element Parameter

Input file	Identifier	Identifier type
ngsetmap	FederalGOM_	Set
ngsetmap	FederalOffshore_	Set
ngsetmap	FedGOM_OGDIST_	Set
ngsetmap	Fuel_CNG	Element Parameter
ngsetmap	Fuel_LNG	Element Parameter
ngsetmap	GOMpriceRegion	Element Parameter
ngsetmap	GOMRegion	Element Parameter
ngsetmap	HawaiiState	Element Parameter
ngsetmap	HenryHubRegion	Element Parameter
ngsetmap	HighWOPCase	Element Parameter
ngsetmap	Latitude_center	Parameter
ngsetmap	LNG_OilPriceSteps_	Set
ngsetmap	LNGDestination_	Set
ngsetmap	LNGTerminals_QP_	Set
ngsetmap	Longitude_center	Parameter
ngsetmap	Louisiana_	Set
ngsetmap	LouisianaGOM	Element Parameter
ngsetmap	LowWOPCase	Element Parameter
ngsetmap	map_capexp_season	Parameter (binary)
ngsetmap	map_citygate_season	Parameter (binary)
ngsetmap	map_DemandArcs	Parameter (binary)
ngsetmap	map_GOMregions	Parameter (binary)
ngsetmap	map_hub_Region_OilGas	Parameter (binary)
ngsetmap	map_hubs_borderXings	Parameter (binary)
ngsetmap	map_MXsector_sector	Parameter (binary)
ngsetmap	map_season_mn	Parameter (binary)
ngsetmap	map_Sector_Subsector	Parameter (binary)
ngsetmap	map_State_CensusRegion	Parameter (binary)
ngsetmap	map_State_NNGEMM	Parameter (binary)
ngsetmap	map_substate_state	Parameter (binary)
ngsetmap	map_Supply	Parameter (binary)
ngsetmap	map_supply_lngexp	Parameter (binary)
ngsetmap	map_SupplyArcs	Parameter (binary)
ngsetmap	Mexico_	Set
ngsetmap	Mexico_NE	Element Parameter
ngsetmap	Mexico_NW	Element Parameter
ngsetmap	Mexico_SS	Element Parameter
ngsetmap	MexicoNorthEast	Set
ngsetmap	MexicoSouth_	Set
ngsetmap	MichiganState	Element Parameter

Input file	Identifier	Identifier type
ngsetmap	MinnesotaState	Element Parameter
ngsetmap	Mississippi_	Set
ngsetmap	MississippiGOM	Element Parameter
ngsetmap	MX_Sector_	Set
ngsetmap	NA_AD_	Set
ngsetmap	NEMScase_	Set
ngsetmap	NEMSmap_LNGTER_lngexp	Parameter (binary)
ngsetmap	NEMSmap_M2_d_lng	Parameter (binary)
ngsetmap	NEMSmap_M2_Units	Parameter (binary)
ngsetmap	NEMSmap_M3_season	Parameter (binary)
ngsetmap	NEMSmap_MNCRUD_CrudeType	Parameter (binary)
ngsetmap	NEMSmap_MNUMCR_CensusReg	Parameter (binary)
ngsetmap	NEMSmap_MNUMOR_AKreg	Parameter (binary)
ngsetmap	NEMSmap_NGFLOWS_BX	Parameter (binary)
ngsetmap	NEMSmap_NGFLOWS_OilGasRegions_M12_M6	Parameter (binary)
ngsetmap	NEMSmap_NNGEMM_CD	Parameter (binary)
ngsetmap	NEMSmap_OGDIST_r_ak	Parameter (binary)
ngsetmap	NEMSmap_OGDIST_SupplyNode	Parameter (binary)
ngsetmap	NEMSmap_reg_MNUMCR	Parameter (binary)
ngsetmap	NEMSmap_SupplyNode_MNUMOR	Parameter (binary)
ngsetmap	NewMexico_	Set
ngsetmap	OhioState	Element Parameter
ngsetmap	PennState	Element Parameter
ngsetmap	ReferenceCase	Element Parameter
ngsetmap	Region_Census_	Set
ngsetmap	Region_OilGas_	Set
ngsetmap	Season_	Set
ngsetmap	Sector_	Set
ngsetmap	Sector_Commercial	Element Parameter
ngsetmap	Sector_Electric	Element Parameter
ngsetmap	Sector_Industrial	Element Parameter
ngsetmap	Sector_ResCom_	Set
ngsetmap	Sector_Residential	Element Parameter
ngsetmap	Sector_Transportation	Element Parameter
ngsetmap	ShaleGasRegion	Element Parameter
ngsetmap	SNG_	Set
ngsetmap	StateOffshore_	Set
ngsetmap	States_	Set
ngsetmap	Supply_AD	Element Parameter
ngsetmap	Supply_Curve_Step_	Set

Input file	Identifier	Identifier type
ngsetmap	Supply_LNG	Element Parameter
ngsetmap	Supply_NA	Element Parameter
ngsetmap	Supply_SNGcoal	Element Parameter
ngsetmap	SupplyNode_	Set
ngsetmap	SupplyType_	Set
ngsetmap	Tariff_Curve_Step_	Set
ngsetmap	Texas_	Set
ngsetmap	TexasGOM	Element Parameter
ngsetmap	TexasState	Element Parameter
ngsetmap	TransFuel_	Set
ngsetmap	Units_	Set
ngsetmap	VariableSupply_	Set
ngsetmap	Vehicle_Fleet	Element Parameter
ngsetmap	Vehicle_Marine	Element Parameter
ngsetmap	Vehicle_Personal	Element Parameter
ngsetmap	Vehicle_Rail	Element Parameter
ngsetmap	VehicleRoad_	Set
ngsetmap	VehicleType_	Set
ngsetmap	WesternGOM	Element Parameter
ngspotprc	HistOffshorePriceAdjustment	Parameter
ngspotprc	HistoricalAnnualSpotPrice	Parameter
ngspotprc	HistoricalAnnualWellhdPrice	Parameter
ngspotprc	HistoricalSpotPrice	Parameter
ngspotprc	HistoricalSubstateSpotPricecDifferential	Parameter
ngvartar	GenericTariffCurve	Parameter
ngvartar	VariableTariffCurve	Parameter

Appendix F. Global data transferred between the NEMS and the NGMM

The table below lists all global data transferred between NEMS and the NGMM. Variables In bold identify those that are both sent from NEMS and sent back to NEMS by the model (i.e. output from the NGMM that is then sent to other NEMS modules). The table lists the following:

- *the name of the global datum identifier in the NGMM, which consists of the common block name, an underscore, and then the variable name itself*
- *a brief description of the variable*
- *the main NEMS module communicating with the NGMM for that variable (if applicable); full names can be found on the Abbrevations page*

Unless otherwise stated, all variables are annual totals or consumption-weighted annual average prices.

Identifer in the NGMM	Description	NEMS module
COMMREP_CMTotalFlspc	Commercial floor space	CDM
COMMREP_DegreeDays	Heating degree days	CDM
CONVFACT_CFNGC	Average conversion factor (Bcf to TBtus)	
CONVFACT_CFNGCL	Conversion factor (Bcf to TBtus) for synthetic natural gas from coal	
CONVFACT_CFNGN	Conversion factor (Bcf to TBtus) for non-utility consumption	
CONVFACT_CFNGU	Conversion factor (Bcf to TBtus) for utility consumption	
INTOUT_IT_WOP	Brent crude oil price ($/bbl, $/MMBtu)	IEM
INTOUT_Q_NON_US_DEMAND	Total non-United States demand for crude oil by type	IEM
LFMMOUT_Q_CRUDE_IMPORTS	Crude oil delivered into PADDs by type	LMFF
MACOUT_MC_JPGDP	Chain-linked price index- gross domestic product (base year=2009)	MAM
MACOUT_MC_RUC	Unemployment rate	MAM
MPBLK_PGFIN	**Delivered price to firm industrial customers by Census division**	**IDM**
MPBLK_PGIIN	**Delivered price to interruptible industrial customers by Census division**	**IDM**
MPBLK_PNGCM	**Delivered price to the commercial sector by Census division**	**CDM**
MPBLK_PNGEL	**Delivered price to the electric power sector by Census division**	**EMM**
MPBLK_PNGIN	**Delivered price to the industrial sector by Census division**	**IDM**

Identifer in the NGMM	Description	NEMS module
MPBLK_PNGRS	**Delivered price to the residential sector by Census division**	**RDM**
MPBLK_PNGTR	**Delivered price to the transportation sector by Census division**	**TDM**
MXPBLK_XSTART_PRICE	Brent crude oil prices through 2080	IEM
NCNTRL_CURCALYR	Current NEMS calendar year	
NCNTRL_CURITR	Current NEMS iteration	
NCNTRL_CURIYR	Current NEMS year by index (1-61)	
NCNTRL_FCRL	NEMS flag indicating convergence check	
NCNTRL_NCRL	NEMS flag indicating report loop	
NGRPT_NGCAPS	**Total capacity between regions**	
NGRPT_NGFLOWS	**Regional flows of natural gas**	
NGTDMOUT_PGFTRFV	**Delivered price of CNG to fleet vehicles by Census division**	**TDM**
NGTDMOUT_PGFTRPV	**Delivered price of CNG to personal vehicles by Census division**	**TDM**
NGTDMOUT_PGFTRRAIL	**Delivered price of CNG to rail vehicles by Census division**	**TDM**
NGTDMOUT_PGFTRSHIP	**Delivered price of CNG to marine vessels by Census division**	**TDM**
NGTDMOUT_PGLTRFV	**Delivered price of LNG to fleet vehicles by Census division**	**TDM**
NGTDMOUT_PGLTRPV	**Delivered price of LNG to personal vehicles by Census division**	**TDM**
NGTDMOUT_PGLTRRAIL	**Delivered price of LNG to rail vehicles by Census division**	**TDM**
NGTDMOUT_PGLTRSHIP	**Delivered price of LNG to marine vessels by Census division**	**TDM**
NGTDMOUT_PNGELGR	**Annual average delivered price to electric power sector by NGEMM region**	**EMM**
NGTDMOUT_QGFTRFV	Consumption of CNG by fleet vehicles by Census division	TDM
NGTDMOUT_QGFTRPV	Consumption of CNG by personal vehicles by Census division	TDM
NGTDMOUT_QGFTRRAIL	Consumption of CNG by rail vehicles by Census division	TDM
NGTDMOUT_QGFTRSHIP	Consumption of CNG by marine vessels by Census division	TDM

Identifer in the NGMM	Description	NEMS module
NGTDMOUT_QGLTRFV	Consumption of LNG by fleet vehicles by Census division	TDM
NGTDMOUT_QGLTRPV	Consumption of LNG by personal vehicles by Census division	TDM
NGTDMOUT_QGLTRRAIL	Consumption of LNG by rail vehicles by Census division	TDM
NGTDMOUT_QGLTRSHIP	Consumption of LNG by marine vessels by Census division	TDM
NGTDMOUT_SPNGELGR	**Seasonal delivered price to electric power sector by NGEMM region**	**EMM**
NGTDMREP_ACGPR_RESCOM	**Citygate price by Census division**	
NGTDMREP_EL_MRKUP_BETA	**Seasonal consumption coefficient for electric power**	**EMM**
NGTDMREP_NALNGEXP	**Total LNG exports by LNG export region**	
NGTDMREP_NGBAL	**Balancing item by Census division**	
NGTDMREP_NGEXPPRC	**Export prices (Canada, Mexico, LNG, Total)**	
NGTDMREP_NGEXPVOL	**Export volumes (Canada, Mexico, LNG, Total)**	
NGTDMREP_NGIMPPRC	**Import prices (Canada, Mexico, LNG, Total)**	
NGTDMREP_NGIMPVOL	**Import volumes (Canada, Mexico, LNG, Total)**	
NGTDMREP_NGSCRV_ELAS	**Elasticity for approximated national supply curve**	**EMM**
NGTDMREP_NGSCRV_MAX	**National supply curve maximum quantity**	**EMM**
NGTDMREP_NGSCRV_MIN	**National supply curve minimum quantity**	**EMM**
NGTDMREP_NGSCRV_P	**National supply curve solution price**	**EMM**
NGTDMREP_NGSCRV_P0	**National supply curve base price**	**EMM**
NGTDMREP_NGSCRV_PER	**National supply curve segment percent deviation**	**EMM**
NGTDMREP_NGSCRV_Q	**National supply curve solution quantity**	**EMM**
NGTDMREP_NGSCRV_Q0	**National supply curve base quantity**	**EMM**
NGTDMREP_NGSPOT_EMM	**Electric power sector consumption-weighted spot price by NGEMM region**	**EMM**
NGTDMREP_OGHHPRNG	**Henry Hub spot price**	
NGTDMREP_OGPRCNG	**Supply price by oil and gas region**	**OGSM**
NGTDMREP_OGPRDNG	**Total production by oil and gas region**	**OGSM**
NGTDMREP_OGPRSUP	**Total supplemental supplies**	
NGTDMREP_OGSUPGAS	**Supplemental supplies by type and Census division**	
NGTDMREP_OGWPRNG	**Supply price by oil and gas district**	**OGSM**
NGTDMREP_PINTLNG	**World LNG price by destination**	

Identifer in the NGMM	Description	NEMS module
NGTDMREP_PTRANSNG	**Total markup from supply price to LNG export price by destination and LNG export region**	
NGTDMREP_UDTAR	**Distributor tariff applied to electric power delivered price by NGEMM region**	
OGSMOUT_CNADGPRD	Canadian AD production (East, West Canada)	OGSM
OGSMOUT_CNENAGPRD	Expected Canadian NA production (East, West Canada)	OGSM
OGSMOUT_CNRNAGPRD	**Realized Canadian NA production (East, West Canada)**	**OGSM**
OGSMOUT_OGADGPRD	AD production by oil and gas district	OGSM
OGSMOUT_OGCNPPRD	Canadian supply price (East, West Canada)	OGSM
OGSMOUT_OGENAGPRD	Expected NA production by oil and gas district	OGSM
OGSMOUT_OGNGPLPRD	NGPL production by oil and gas district	OGSM
OGSMOUT_OGPRCOAK	Alaskan crude oil production by Alaska region	OGSM
OGSMOUT_OGRNAGPRD	**Realized NA production by oil and gas district**	**OGSM**
OGSMOUT_OGSHALENG	Production from oil shale plants (WY)	OGSM
QBLK_QGFIN	Firm industrial consumption by Census division	IDM
QBLK_QGIIN	Interruptible industrial consumption by Census division	IDM
QBLK_QGPTR	**Pipeline fuel consumption by Census division**	**TDM**
QBLK_QLPIN	**Lease and plant fuel consumption by Census division**	**IDM**
QBLK_QNGCM	Commercial sector consumption by Census division	CDM
QBLK_QNGEL	Electric power sector consumption by Census division	EMM
QBLK_QNGHM	Fuel consumed for hydrogen production by Census division	IDM
QBLK_QNGRS	Residential sector consumption by Census division	RDM
QBLK_QNGTR	Transportation sector consumption by Census division	TDM
QMORE_QGTLRF	Gas-to-liquids consumption (production) by Census division	LFMM
QMORE_QGTLSN	Gas-to-liquids consumption (heat and power) by Census division	LFMM
QMORE_QNGLQ	**Fuel used for liquefaction by Census division**	**IDM**

Identifer in the NGMM	Description	NEMS module
RESDREP_RSGASCUST	Number of residential customers by Census division	RDM
UEFDOUT_SQNGELGR	Seasonal electric power sector consumption by NGEMM region	EMM

Appendix G. Documentation of estimations

The following is a description, including regression statistics, for all econometric estimations used in the NGMM. They are presented in the order in which they appear in the documentation and include the following:

G-1	Mexico crude oil production
G-2	Mexico associated-dissolved natural gas production
G-3	Mexico non-associated natural gas production
G-4	Mexico industrial natural gas consumption
G-5	Alaska residential sector consumption, commercial sector consumption, and citygate prices
G-6	Alaska lease and plant fuel consumption
G-7	Citygate prices by month for the Lower 48 states
G-8	Distributor tariff markup for delivered prices to the residential sector by Census division
G-9	Distributor tariff markup for delivered prices to the commercial sector by Census division
G-10	Historical delivered prices to the industrial sector by Census division

G-1. Mexico crude oil production

Data: Parameter estimates for Mexican crude oil production from 2025-2050 for use in associated dissolved natural gas production and industrial consumption

Author: Kathryn Dyl, EIA, 2019

Source: Crude oil production in Mexico – International Energy Outlook 2019 (unpublished input data); Oil price – AEO 2020

Variables:

P_Oil_{yr} = Mexican crude oil production for the historical year yr (million barrels)

WOP_{yr-2} = average annual Brent crude oil price for the year that is two prior to the historical year yr $(yr - 2)$ (1987$/MMBtu)

yr = historical year

Derivation:

Initial upstream inputs into EIA's International Energy Outlook 2019 for Mexico were obtained and smoothed for the period from 2025-2050 using a catmull-rom spline interpolation. This production path was then estimated as a logarithmic function of Brent world oil price and the prior year's crude oil price in order for Mexico's crude oil production to respond to varying oil price paths in AEO 2020 side cases. Mexican crude oil production through 2025 was assumed continue to decline from historical levels.

Projected Brent crude oil prices from 2025-2050 and Mexican crude oil production from 2025-2050 were used to estimate crude oil production as a function of crude oil price from two years prior and last year's crude oil production. Years where there was an inflection in the data (2033, 2044) were excluded. The equation follows:

$$\log P\_Oil_{yr} = \alpha * \log P\_Oil_{yr-1} + \beta * \log WOP_{yr-2}$$

Regression diagnostics and parameter estimates:

Dependent Variable: LOG_OIL_IEO_SM2
Method: Least Squares
Date: 09/05/19 Time: 13:45
Sample: 2025 2032 2034 2043 2046 2050
Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG_OIL_IEO_SM2(-1)	1.003	0.020	50.359	0.000
LWOP(-2)	2.187	2.388	0.916	0.370

R-squared	0.997902	Mean dependent var	844.2377
Adjusted R-squared	0.997802	S.D. dependent var	179.4289
S.E. of regression	8.4114	Akaike info criterion	7.179994
Sum squared resid	1485.785	Schwarz criterion	7.278733
Log likelihood	-80.5699	Hannan-Quinn criter.	7.204827
Durbin-Watson stat	0.524254		

Data:

year	P_Oil (initial)	P_Oil (smoothed)	WOP
2016	708.530	708.530	3.444
2017	698.730	698.730	4.279
2018	664.700	664.700	5.477
2019	617.200	617.200	5.517
2020	619.610	619.610	5.451
2021	647.820	623.140	5.536
2022	667.940	626.370	5.724
2023	686.300	629.570	5.866
2024	664.110	633.000	6.062
2025	644.510	636.920	6.204
2026	625.620	641.600	6.090
2027	647.280	647.280	6.289
2028	669.830	669.830	6.301
2029	693.660	693.660	6.392
2030	698.170	698.170	6.449
2031	702.120	702.120	6.536
2032	703.210	703.210	6.585
2033	704.260	704.260	6.808
2034	694.160	726.980	6.804
2035	691.010	750.050	6.887
2036	707.300	773.520	6.998
2037	737.630	797.400	7.144
2038	763.460	821.730	7.234
2039	781.000	846.540	7.349
2040	787.190	871.880	7.494
2041	749.610	897.760	7.598
2042	826.130	924.220	7.830
2043	951.290	951.290	7.890
2044	1005.550	1005.550	7.916
2045	1070.390	1070.390	8.015
2046	1063.460	1095.100	8.107
2047	1000.690	1116.160	8.198
2048	934.780	1134.490	8.272
2049	955.040	1150.990	8.377
2050	1166.580	1166.580	8.468

G-2. Mexico associated-dissolved natural gas production

Data: Parameter estimates for Mexican production of associated-dissolved natural gas production, assigned to the South-Southeast region of Mexico

Author: Kathryn Dyl, EIA, 2019

Source: Associated-dissolved production of natural gas in Mexico and crude oil production in Mexico – PEMEX; Oil price – Thompson Reuters

Variables:

P_SS_{yr} = Mexican AD dry natural gas production in the South-Southeast region for historical year yr (minus lease fuel, plant fuel, and reinjected volumes) (Bcf)

P_Oil_{yr} = Mexican crude oil production for historical year yr (million barrels)

WOP_{yr-1} = average annual Brent crude oil price for the year prior to the historical year yr $(yr-1)$ (1987$/MMBtu)

yr = historical year

Derivation:

Brent crude oil prices from 1993 to 2017, Mexico's crude oil production from 1993 to 2017, and Mexico's total annual associated-dissolved natural gas production, minus natural gas reinjected into wells, lease fuel, and plant fuel, from 1993 to 2017 were used to estimate total dry associated-dissolved natural gas production—or total production for the South-Southeast region of Mexico over that time frame. The year 2008 was excluded as an outlier due to the global economic downturn and commodity price spike. The equation follows:

$$P\_SS_{yr} = \alpha * P\_SS_{yr-1} + \beta * WOP_{yr-1} + \gamma * P\_Oil_{yr}$$

Regression diagnostics and parameter estimates:

Dependent Variable: P_SS
Method: Least Squares
Date: 09/04/19 Time: 13:27
Sample: 1993 2007 2009 2017
Included observations: 24

Variable	Coefficient	Std. Error	t-Statistic	Prob.
P_SS(-1)	0.723	0.067	10.836	0.000
WOP(-1)	21.642	6.462	3.349	0.003
P_Oil	0.119	0.037	3.205	0.004

R-squared	0.90796	Mean dependent var	820.1104
Adjusted R-squared	0.899194	S.D. dependent var	195.7051
S.E. of regression	62.13625	Akaike info criterion	11.213
Sum squared resid	81079.19	Schwarz criterion	11.36026
Log likelihood	-131.5561	Hannan-Quinn criter.	11.25207
Durbin-Watson stat	1.155303		

Data:

year	WOP	P_SS	P_Oil
1993	2.243	618.490	975.791
1994	2.116	613.250	980.062
1995	2.291	600.412	955.278
1996	2.703	698.438	1043.291
1997	2.379	770.087	1103.110
1998	1.536	786.990	1120.716
1999	2.172	730.967	1060.700
2000	3.399	741.033	1099.367
2001	2.630	764.765	1141.368
2002	2.795	684.651	1159.641
2003	3.209	697.713	1230.377
2004	4.038	661.060	1234.758
2005	5.338	589.017	1216.672
2006	6.253	595.421	1188.286
2007	6.923	768.892	1122.634
2008	9.352	1085.408	1018.925
2009	5.920	1143.785	949.541
2010	7.514	1142.093	940.612
2011	9.927	1047.380	931.705
2012	9.353	1039.158	929.988
2013	9.150	1048.447	920.576
2014	8.171	1102.842	886.500
2015	4.161	1109.171	827.394
2016	3.444	966.245	786.037
2017	4.279	762.342	711.116

G-3. Mexico non-associated natural gas production

Data: Parameter estimates for Mexican production of nonassociated natural gas, both before and after the onset of shale gas development, assigned to the Northeast region of Mexico

Author: Kathryn Dyl, EIA, 2019

Source: Historical non-associated natural gas production in Mexico – Petróleos Mexicanos. Projected Henry Hub spot price – EIA's Annual Energy Outlook 2020 Reference case.

Variables:

P_NE_{yr} = Mexican NA dry natural gas production in the Northeast region for historical year yr (minus lease and plant fuel) (Bcf)

HH_{yr-1} = average annual Henry Hub natural gas spot price for the year prior to historical year yr (1987$/MMBtu)

yr = historical year

Derivation:

Total non-associated natural gas production in the projection—or total production for the Northeast region of Mexico—was estimated as a function of the last two years' non-assoicated natural gas production and last year's Henry Hub spot price. The estimation for non-associated natural gas production from shale gas development remained unchanged from its prior estimation in AEO 2018 (Appendix G-2). The equation follows:

$$P\_NE_{yr} = \alpha_{yr-1} * P\_NE_{yr-1} + \alpha_{yr-2} * P\_NE_{yr-2} + \beta * HH_{yr-1}$$

Regression diagnostics and parameter estimates:

Non-associated natural gas production without shale gas development

Dependent Variable: P_NE
Method: Least Squares
Date: 09/05/19 Time: 09:00
Sample: 1993 2017
Included observations: 25

Variable	Coefficient	Std. Error	t-Statistic	Prob.
P_NE(-1)	1.453	0.148	9.792	0.000
P_NE(-2)	-0.585	0.127	-4.604	0.000
HH(-1)	27.588	7.802	3.536	0.002

R-squared	0.980895	Mean dependent var	555.8746
Adjusted R-squared	0.979158	S.D. dependent var	240.7349
S.E. of regression	34.75448	Akaike info criterion	10.04666
Sum squared resid	26573.23	Schwarz criterion	10.19293
Log likelihood	-122.583	Hannan-Quinn criter.	10.08723
Durbin-Watson stat	1.739416		

Non-associated natural gas production with shale gas development

Dependent Variable: P_NE
Method: Least Squares
Date: 01/26/18 Time: 12:22
Sample: 2026 2050
Included observations: 25

Variable	Coefficient	Std. Error	t-Statistic	Prob.
HH(-1)	6.99758	2.591596	2.700104	0.0128
P_NE(-1)	1.036203	0.008235	125.8324	0

R-squared	0.998853	Mean dependent var	797.1868
Adjusted R-squared	0.998803	S.D. dependent var	326.7448
S.E. of regression	11.30372	Akaike info criterion	7.764759
Sum squared resid	2938.804	Schwarz criterion	7.862269
Log likelihood	-95.05949	Hannan-Quinn criter.	7.791805
Durbin-Watson stat	2.474468		

Data:

year	HH	P_NE
2015	1.477	401.135
2016	1.406	317.322
2017	1.680	285.430
2018	1.769	271.195
2019	1.841	262.435
2020	1.953	245.586
2021	1.951	258.420
2022	1.973	262.435
2023	2.041	265.720
2024	2.154	279.624
2025	2.266	294.555
2026	2.385	321.930

year	HH	P_NE
2033	2.486	569.765
2034	2.484	605.170
2035	2.495	646.780
2036	2.520	688.446
2037	2.536	724.890
2038	2.569	762.120
2039	2.583	797.890
2040	2.580	834.480
2041	2.560	877.825
2042	2.581	921.260
2043	2.623	964.695
2044	2.650	1054.080

2027	2.477	354.050	2045	2.696	1086.605
2028	2.539	391.986	2046	2.744	1137.705
2029	2.571	426.685	2047	2.781	1218.005
2030	2.575	455.520	2048	2.833	1283.928
2031	2.547	494.940	2049	2.871	1351.230
2032	2.515	534.726	2050	2.909	1424.960

G-4. Mexico industrial natural gas consumption

Data: Parameter estimates for Mexican consumption of natural gas in the industrial sector, split into consumption by PEMEX and by other industries

Author: Kathryn Dyl, EIA, 2019

Source: Historical consumption and production data for Mexico – SENER, Sistema de Información Energética. Historical Henry Hub spot price – Thompson Reuters.

Variables:

Q_PEMEX_{yr} = Mexican consumption of natural gas by PEMEX (e.g., oil and natural gas production, exporation, refining) for historical year yr (Bcf)

P_Oil_{yr} = Mexican crude oil production for historical year yr (million barrels)

Q_IND_{yr} = Mexican industrial consumption for all other industries outside of PEMEX for historical year yr (Bcf)

HH_{yr-1} = average annual Henry Hub natural gas spot price for the year prior to historical year yr (1987$/MMBtu)

yr = historical year

Derivation:

Mexican natural gas consumption in its industrial sector is assumed to have two components: natural gas consumed by Petroleos Mexicanos (PEMEX) in its expoloration, production, refining, and petrochemical activities (excluding natural gas reinjected into oil wells); and natural gas consumed in other industries, such as manufacturing.

These components were estimated separately using historical data from SENER. Coefficients were estimated using historical data from 1996 through 2017 relating PEMEX consumption and oil production (see Appendix G-1):

$$Q\_PEMEX_{yr} = \alpha * Q\_PEMEX_{yr-1} + \beta * P\_Oil_{yr}$$

Meanwhile, other industrial natural gas use in Mexico was estimated as a function of the prior year's natural gas use, the Henry Hub spot price, and a constant term. Because of the recent (2015-2018) rapid growth in the industrial sector, the estimation was done using an extended data series for other industrial natural gas consumption through 2023 that allows for slower growth moving forward (see data below).

A unit root test indicated that the econometric estimation be determined as a first-difference of the equation below:

$$Q\_IND_{yr} = \alpha * Q\_IND_{yr-1} + \beta * HH_{yr-1} + C$$

Regression diagnostics and parameter estimates:

Industrial consumption associated with oil and natural gas exploration and production (PEMEX)

Dependent Variable: Q_PEMEX
Method: Least Squares
Date: 09/04/19 Time: 17:17
Sample: 1996 2017
Included observations: 22

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Q_PEMEX(-1)	0.918	0.032	28.333	0.000
P_OIL	0.066	0.023	2.892	0.009

R-squared	0.925154	Mean dependent var	728.4173
Adjusted R-squared	0.921412	S.D. dependent var	94.68294
S.E. of regression	26.54307	Akaike info criterion	9.481922
Sum squared resid	14090.69	Schwarz criterion	9.581108
Log likelihood	-102.3011	Hannan-Quinn criter.	9.505288
Durbin-Watson stat	2.411247		

Other industrial consumption

Dependent Variable: DQ_IND
Method: Least Squares
Date: 09/05/19 Time: 16:05
Sample (adjusted): 2001 2023
Included observations: 23 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DQ_IND(-1)	0.973	0.051	19.149	0.000
DHH(-1)	-19.148	9.138	-2.095	0.049
C	95.552	51.175	1.867	0.077

R-squared	0.973825	Mean dependent var	644.3747
Adjusted R-squared	0.971207	S.D. dependent var	251.0545
S.E. of regression	42.6002	Akaike info criterion	10.4627
Sum squared resid	36295.55	Schwarz criterion	10.61081
Log likelihood	-117.321	Hannan-Quinn criter.	10.49995
F-statistic	372.0371	Durbin-Watson stat	1.875929

Data:

year	Q_PEMEX	P_Oil	Q_IND	HH
1996	571.101	1043.291	—	—
1997	555.094	1103.110	—	—
1998	564.730	1120.716	—	—
1999	555.910	1060.700	—	—
2000	648.956	1099.367	414.073	3.106
2001	668.565	1141.368	335.118	2.809
2002	683.271	1159.641	396.753	2.335
2003	729.167	1230.377	407.368	3.748
2004	749.129	1234.758	430.457	3.920
2005	740.932	1216.672	420.546	5.743
2006	788.276	1188.286	454.952	4.223
2007	775.873	1122.634	469.091	4.263
2008	793.753	1018.925	463.889	5.327
2009	784.633	949.541	423.312	2.344
2010	816.396	940.612	481.849	2.569
2011	798.020	931.705	509.200	2.292
2012	829.685	929.988	528.055	1.550
2013	829.209	920.576	562.840	2.067
2014	830.509	886.500	610.813	2.394
2015	803.000	827.393	659.869	1.425
2016	774.391	786.037	758.723	1.340
2017	734.580	711.116	867.151	1.541
2018	—	—	979.544	1.620
2019	—	—	935.860	1.296
2020	—	—	986.960	1.360
2021	—	—	1015.795	1.518
2022	—	—	1045.725	1.612
2023	—	—	1076.750	1.668

G-5. Alaska residential and commercial sector consumption and Alaska citygate price

Data: Parameter estimates for the Alaskan natural gas consumption equations for the residential and commercial customers and the Alaskan natural gas citygate price

Author: Arthur Dai, EIA, 2015

Source: Consumption and citygate price − Natural Gas Annual (DOE/EIA-0131); Alaska population − U.S. Census Bureau, Population Division; Heating degree days − National Oceanic and Atmospheric Administration (Anchorage International Airport). National unemployment rate − NEMS Macroeconomic Activity Module; Oil price − Petroleum Marketing Annual, DOE/EIA-0487.

Variables:

For historical year yr,

$CONS\_R_{yr}$ = residential natural gas consumption in Alaska (MMcf)

$CONS\_\mathrm{C}_{yr}$ = commercial natural gas consumption in Alaska (Bcf)

POP_AK_{yr} = thousands of people in Alaska (Bcf)

HDD_DVN_{yr} = deviation from normal heating degree days (0 in projection)

$UNEMP_{yr}$ = U.S. civilian unemployment rate as a percent (MC_RUC in NEMS Macroeconomic Activity Module)

$PRICE\_\mathrm{CITYGATE}_{yr}$ = natural gas citygate price in Alaska (1987$/Mcf)

$IRAC1987_{yr}$ = U.S. crude oil imported acquisition cost by refiners (1987$/bbl)

L1982_4 = dummy variable with value of one for years 1982-1984, zero elsewhere

L1985_94 = dummy variable with value of one for years 1985-1994, zero elsewhere

L1995_00 = dummy variable with value of one for years 1995-2000, zero elsewhere

L2008 = dummy variable with value of one for year 2008, zero elsewhere

Derivation:

Residential sector consumption

Annual data for population, heating degree days, national unemployment rate, and citygate prices in the range 1985 to 2013 were used to estimate Alaska residential consumption, as follows:

$$CONS\_R_{yr} = C + \left(\alpha_1 * POP\_AK_{yr}\right) + \left(\alpha_2 * HDD\_DVN_{yr}\right) + \left(\alpha_3 * UNEMP_{yr}\right) + (\alpha_4 * PRICE\_CITYGATE_{yr})$$

Regression diagnostics and parameter estimates

Dependent Variable: CONS_R
Method: Least
Squares
Date: 07/13/15
Time: 09:00
Sample (adjusted): 1985 2013
Included observations: 29 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-321.577	1766.384	-0.182054	0.8571
POP_AK	27.02888	3.097776	8.725253	0
HDD_DVN	1.257146	0.157816	7.96588	0
UNEMP	-254.648	56.77627	-4.485115	0.0002
PRICE_CITYGATE	577.5681	139.5432	4.138991	0.0004

R-squared	0.981609	Mean dependent var	16478.72
Adjusted R-squared	0.978543	S.D. dependent var	2845.665
S.E. of regression	416.8361	Akaike info criterion	15.05885
Sum squared resid	4170056	Schwarz criterion	15.29459
Log likelihood	-213.353	Hannan-Quinn criter.	15.13268
F-statistic	320.2384	Durbin-Watson stat	1.860195
Prob(F-statistic)	0		

Commercial sector consumption

Annual data for population, heating degree days, and the national unemployment rate in the range 1984 to 2013 were used to estimate Alaska commercial consumption. A visual display of the consumption data showed clear discontinuities in the series. The particular reasons were not identified, but dummy variables were used in the estimation to account for the shifts, as follows:

$$CONS\_C = C + (\alpha_1 * POP\_AK) + (\alpha_2 * HDD\_DVN) + (\alpha_3 * UNEMP) +$$
$$(\alpha_4 * L1982\_4) + (\alpha_5 * L1985\_94) + (\alpha_6 * L1995\_00) + (\alpha_7 * L2008) + AR(1)$$

Regression diagnostics and parameter estimates

Dependent Variable: CONS_C
Method: ARMA Conditional Least Squares (Marquardt - EViews legacy)
Date: 07/13/15
Time: 10:09
Sample: 1984 2013
Included
observations: 30
Convergence achieved after 60 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3211.041	5415.784	0.592904	0.5596
POP_AK	24.06646	8.656774	2.780073	0.0112
HDD_DVN	0.914788	0.277139	3.300831	0.0034
UNEMP	-369.309	180.5575	-2.04538	0.0535
L1982_4	11846.88	1737.303	6.81912	0
L1985_94	6444.171	1134.411	5.680631	0
L1995_00	10353.29	740.55	13.98055	0
L2008	-1759.78	809.041	-2.175141	0.0412
AR(1)	0.374087	0.212914	1.756986	0.0935

R-squared	0.965492	Mean dependent var	20752.33
Adjusted R-squared	0.952346	S.D. dependent var	3667.556
S.E. of regression	800.6205	Akaike info criterion	16.45198
Sum squared resid	13460857	Schwarz criterion	16.87234
Log likelihood	-237.78	Hannan-Quinn criter.	16.58645
F-statistic	73.44406	Durbin-Watson stat	1.747117
Prob(F-statistic)	0		

Inverted AR Roots	0.37

Citygate price

Annual data for the international refinery acquisition price from 1994 to 2013 were used to estimate the citygate price in Alaska, as follows.

$$LOG\left(PRICE\_CITYGATE_{yr}\right) = C + (\alpha_1 * LOG\left(IRAC1987_{yr}\right))$$

Regression diagnostics and parameter estimates

Dependent Variable: LOG(PRICE_CITYGATE)
Method: Least Squares
Date: 01/15/17 Time:
19:05
Sample: 1994 2013
Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.82408	0.276787	-6.590174	0
LOG(IRAC1987)	0.807019	0.08474	9.52347	0

R-squared	0.834401	Mean dependent var	0.769816
Adjusted R-squared	0.825201	S.D. dependent var	0.526927

S.E. of regression	0.220303	Akaike info criterion	-0.09299
Sum squared resid	0.873597	Schwarz criterion	0.006582
Log likelihood	2.929908	Hannan-Quinn criter.	-0.07355
F-statistic	90.69648	Durbin-Watson stat	1.772905
Prob(F-statistic)	0		

Data:

	CONS_R	CONS_C	POP_AK	HDD_DVN	UNEMP	PRICE_ CITYGATE	IRAQ1987
1984	11833	24654	513.7	-625	7.50		31.18
1985	13256	20344	532.5	351	7.20	0.35	28.24
1986	12091	20874	544.3	-431	7.00	0.34	14.36
1987	12256	20224	539.3	-475	6.20	0.33	18.13
1988	12529	20842	542.0	-236	5.50	0.32	14.07
1989	13589	21738	547.2	343	5.30	0.31	16.81
1990	14165	21622	550.0	657	5.62	0.30	19.52
1991	13562	20897	569.3	24	6.85	0.28	16.23
1992	14350	21299	587.1	477	7.49	0.29	15.44
1993	13858	20003	597.0	-867	6.91	0.27	13.38
1994	14895	20698	600.6	92	6.10	1.31	12.59
1995	15231	24979	601.3	-191	5.59	1.33	13.63
1996	16179	27315	604.9	784	5.41	1.23	16.12
1997	15146	26908	608.8	-475	4.94	1.39	14.22
1998	15617	27079	615.2	334	4.50	1.31	9.14
1999	17634	27667	619.5	1398	4.22	0.99	12.91
2000	15987	26485	628.0	170	3.97	1.17	20.26
2001	16818	15849	633.7	405	4.74	1.66	15.73
2002	16191	15691	642.3	-234	5.78	1.66	16.70
2003	16853	17270	648.4	-30	5.99	1.61	19.13
2004	18200	18373	659.3	103	5.54	2.05	24.12
2005	18029	16903	666.9	-95	5.08	2.43	31.81
2006	20616	18544	675.3	1074	4.61	3.32	37.28
2007	19843	18756	680.3	620	4.62	4.15	41.25
2008	21439	17025	687.5	1435	5.80	4.07	55.98
2009	19978	16620	698.9	830	9.28	4.92	35.43
2010	18714	15920	710.2	431	9.63	3.95	44.88
2011	20262	19399	723.4	716	8.93	3.79	59.49
2012	21380	19898	730.3	1509	8.07	3.50	55.87
2013	19215	18694	735.1	204	7.35	3.38	50.19

G-6. Alaska lease and plant fuel consumption

Data: Alaska lease and plant fuel consumption

Author: Arthur Dai, EIA, 2015

Source: Lease and plant fuel – Natural Gas Annual, DOE/EIA-0131; oil production – Petroleum Supply Annual, DOE/EIA-0340

Variables:

For historical year yr,

$CONS\_LP_{yr}$ = annual historical lease and plant fuel consumption in Alaska (Bcf)

$OILPROD_{yr}$ = annual historical crude oil production in in Alaska (Mbl/d)

Derivation: Annual data for lease and plant fuel consumption and crude oil production in Alasa were used to estimate Alaska lease and plant fuel consumption, as follows:

$$CONS\_LP_{yr} = C + \left(\alpha_1 * OILPROD_{yr}\right) + \left(\alpha_2 * CONS\_LP_{yr-1}\right)$$

Regression diagnostics and parameter estimates:

Dependent Variable: CONS_LP
Method: Least Squares
Date: 07/13/15 Time: 14:45
Sample: 2003 2013
Included observations: 11

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	259.8324	48.58907	5.347548	0.0007
OILPROD	0.432398	0.089434	4.834852	0.0013
CONS_LP(-1)	-0.39235	0.242663	-1.616836	0.1446

R-squared	0.827496	Mean dependent var	265.6656
Adjusted R-squared	0.78437	S.D. dependent var	20.55004
S.E. of regression	9.542621	Akaike info criterion	7.576414
Sum squared resid	728.493	Schwarz criterion	7.684931
Log likelihood	-38.6703	Hannan-Quinn criter.	7.50801
F-statistic	19.18782	Durbin-Watson stat	1.525552
Prob(F-statistic)	0.000886		

Data:

	CONS_LP	OILPROD
2002	285.477	359.382
2003	300.463	355.603
2004	281.546	332.441
2005	303.215	315.387
2006	257.091	270.481
2007	268.571	263.595
2008	252.164	249.874
2009	258.608	235.491
2010	249.234	218.904
2011	243.87	204.829
2012	251.732	192.368
2013	255.828	187.954

G-7. Lower 48 states natural gas citygate prices by month

Data: Equations for citygate prices

Author: Katie Dyl, EIA, 2017

Source: The source for monthly citygate prices and residential and commercial consumption was EIA's Natural Gas Monthly. Monthly spot prices were aggregated from daily reported spot prices from selected hubs used to represent each state. Daily spot prices at the hubs listed in Table G.5.5 were taken from Natural Gas Intelligence. Data for January and February 2014 are excluded as outliers resulting from extreme cold weather.

Variables:

$PRICE\_CG_{mon,st}$ = citygate price in state st and historical month mon (1987$/Mcf)

$PRICE\_SPOT_{mon,st}$ = spot price in state st and historical month mon (1987$/Mcf)

$Q\_RES_{mon,st}$ = residential sector consumption of natural gas in state st and historical month mon (Bcf)

$Q\_COM_{mon,st}$ = commercial sector consumption of natural gas in state st and historical month mon (Bcf)

α_{st} = estimated coefficient for spot price for state st (unitless)

β_{st} = estimated coefficient for reciprocal of consumption for state st (Bcf*1987$/Mcf)

C_{st} = estimated constant term for state st (1987$/Mcf)

mon = historical month

st = state (including DC, excluding AK and HI)

Derivation:

Citygate prices by state and month were estimated for 48 states plus DC over the 2013-2018 time period. The equation below was first estimated in Eviews as a balanced pool estimation with cross-section weights using the estimated generalized least squares (EGLS) method:

$$PRICE\_CG_{mon,st} = \alpha_{st} * \; PRICE\_SPOT_{mon,st} + \frac{\beta_{st}}{\left(Q\_RES_{mon,st} \; + \; Q\_COM_{mon,st}\right)} + C_{st}$$

After the estimation for these states and DC (results in Table G.7.1) was analyzed, states that had either an α-coefficient less than 0.4, indicating a weak correlation, or a β-coefficient that was either less than –2.5 or greater than 10, indicating an anomalous correlation to consumption, were re-estimated using the same equation; however, different coefficients were determined for winter months (November-March) and non-winter months (April-October) (Group 2, results in Table G.7.2). States in New England, where there are known to be specific issues in correlating spot and city gate prices, and states where the overall estimation is of poor quality (DC, DE, UT) were reestimated only using non-winter months (Group 3, results in Table G.7.3). For November and December in all projection years, citygate prices are set to equal the October value. For January through March, prices are extrapolated between the December citygate price of the last year and the estimated April price of the current projection year.

There are five states that do not follow this prescription. In the case of MN, MS, and OR, these states used a seasonal (summer and winter) set of estimations project citygate price—despite having estimated coefficients within the thresholds—due to prior identified issues in their data. And in the case of FL and TX, which do have estimated coefficients that do not meet the threshold, there are known market issues that explain the estimation. It was also deemed that seasonal variations would not explain markets in these states or improve the projections.

Regression diagnostics and parameter estimates:

Table G.7.1 Regression results for Lower 48 states and DC (estimated using all months) from Eviews
Dependent Variable: P_CG?
Method: Pooled EGLS (Cross-section weights)
Date: 08/06/19 Time: 14:01
Sample: 2013M01 2013M12 2014M03 2018M12
Included observations: 70
Cross-sections included: 49
Total pool (balanced) observations: 3430
Iterate weights to convergence
White cross-section standard errors & covariance (d.f. corrected)
Convergence achieved after 1 weight iterations
WARNING: estimated coefficient covariance matrix is of reduced rank
Total panel (balanced) observations: 4606

state	C	std. error	t-stat	P-value	α	std. error	t-stat	P-value	β	std. error	t-stat	P-value
AL	**0.718**	0.089	8.054	0.000	**0.803**	0.049	16.543	0.000	**0.282**	0.132	2.147	0.032
AZ	**1.591**	0.192	8.281	0.000	**0.491**	0.114	4.308	0.000	**-0.577**	0.472	-1.222	0.222
AR	**0.257**	0.153	1.685	0.092	**0.863**	0.075	11.573	0.000	**6.133**	0.322	19.020	0.000
CA	**1.010**	0.332	3.038	0.002	**0.450**	0.154	2.914	0.004	**0.341**	6.793	0.050	0.960
CO	**0.371**	0.133	2.788	0.005	**0.965**	0.073	13.167	0.000	**4.656**	0.490	9.494	0.000
CT	**1.440**	0.185	7.796	0.000	**0.151**	0.037	4.108	0.000	**5.195**	0.725	7.163	0.000
DE	**2.107**	0.194	10.869	0.000	**0.024**	0.031	0.781	0.435	**1.778**	0.198	8.985	0.000
DC	**1.142**	0.281	4.058	0.000	**0.373**	0.142	2.623	0.009	**2.119**	0.151	14.008	0.000
FL	**3.048**	0.468	6.518	0.000	**0.052**	0.119	0.437	0.662	**-2.177**	2.105	-1.034	0.301
GA	**1.450**	0.274	5.294	0.000	**0.357**	0.163	2.187	0.029	**1.676**	0.591	2.837	0.005
ID	**0.912**	0.176	5.174	0.000	**0.553**	0.095	5.811	0.000	**0.153**	0.183	0.836	0.403
IL	**0.584**	0.089	6.530	0.000	**0.791**	0.047	16.944	0.000	**7.356**	1.264	5.820	0.000
IN	**0.974**	0.093	10.461	0.000	**0.579**	0.051	11.422	0.000	**3.047**	0.224	13.586	0.000
IA	**0.778**	0.211	3.682	0.000	**0.835**	0.110	7.608	0.000	**0.978**	0.355	2.754	0.006

state	C	std. error	t-stat	P-value	α	std. error	t-stat	P-value	β	std. error	t-stat	P-value
KS	**0.702**	0.142	4.944	0.000	**0.873**	0.096	9.061	0.000	**3.457**	0.303	11.414	0.000
KY	**0.884**	0.092	9.609	0.000	**0.788**	0.050	15.734	0.000	**-0.256**	0.114	-2.247	0.025
LA	**1.061**	0.096	11.047	0.000	**0.656**	0.043	15.406	0.000	**-0.351**	0.197	-1.785	0.074
ME	**3.199**	0.387	8.268	0.000	**0.236**	0.061	3.902	0.000	**-0.037**	0.143	-0.262	0.793
MD	**1.333**	0.257	5.193	0.000	**0.341**	0.131	2.605	0.009	**9.496**	0.656	14.467	0.000
MA	**1.171**	0.186	6.299	0.000	**0.256**	0.033	7.821	0.000	**19.180**	1.646	11.654	0.000
MI	**1.180**	0.157	7.535	0.000	**0.586**	0.080	7.355	0.000	**-2.613**	1.013	-2.581	0.010
MN	**0.800**	0.101	7.920	0.000	**0.772**	0.039	19.948	0.000	**2.081**	0.443	4.697	0.000
MS	**1.145**	0.131	8.758	0.000	**0.681**	0.065	10.444	0.000	**-0.106**	0.126	-0.841	0.400
MO	**0.650**	0.129	5.055	0.000	**0.752**	0.076	9.962	0.000	**7.717**	0.299	25.800	0.000
MT	**0.515**	0.206	2.505	0.012	**0.818**	0.129	6.362	0.000	**0.296**	0.149	1.993	0.046
NE	**0.989**	0.136	7.288	0.000	**0.766**	0.070	10.902	0.000	**0.346**	0.139	2.485	0.013
NV	**0.604**	0.117	5.157	0.000	**0.957**	0.062	15.430	0.000	**0.798**	0.390	2.047	0.041
NH	**1.784**	0.284	6.272	0.000	**0.380**	0.045	8.502	0.000	**0.603**	0.124	4.876	0.000
NJ	**2.201**	0.237	9.271	0.000	**0.200**	0.103	1.928	0.054	**2.153**	2.228	0.967	0.334
NM	**0.540**	0.073	7.409	0.000	**0.857**	0.037	23.022	0.000	**0.224**	0.105	2.136	0.033
NY	**0.430**	0.172	2.498	0.013	**0.311**	0.056	5.504	0.000	**78.426**	3.779	20.755	0.000
NC	**0.993**	0.309	3.212	0.001	**0.380**	0.173	2.194	0.028	**4.938**	0.385	12.828	0.000
ND	**0.783**	0.230	3.401	0.001	**0.902**	0.128	7.029	0.000	**0.096**	0.066	1.452	0.147
OH	**1.149**	0.105	10.946	0.000	**0.652**	0.055	11.801	0.000	**-3.630**	0.893	-4.065	0.000
OK	**1.009**	0.108	9.326	0.000	**0.702**	0.058	12.136	0.000	**2.106**	0.256	8.213	0.000
OR	**1.085**	0.116	9.325	0.000	**0.533**	0.073	7.305	0.000	**2.329**	0.219	10.646	0.000
PA	**1.160**	0.195	5.954	0.000	**0.539**	0.134	4.014	0.000	**16.291**	0.986	16.523	0.000
RI	**1.156**	0.137	8.468	0.000	**0.116**	0.031	3.770	0.000	**0.127**	0.103	1.225	0.221
SC	**1.428**	0.310	4.610	0.000	**0.368**	0.172	2.144	0.032	**0.745**	0.232	3.212	0.001
SD	**0.683**	0.146	4.677	0.000	**0.863**	0.067	12.841	0.000	**0.237**	0.052	4.517	0.000
TN	**0.677**	0.077	8.829	0.000	**0.887**	0.041	21.557	0.000	**0.601**	0.196	3.061	0.002
TX	**1.505**	0.128	11.784	0.000	**0.718**	0.059	12.236	0.000	**-4.377**	1.491	-2.937	0.003
UT	**2.768**	0.191	14.532	0.000	**0.294**	0.106	2.788	0.005	**-2.473**	0.420	-5.885	0.000
VT	**2.029**	0.228	8.908	0.000	**0.351**	0.069	5.113	0.000	**0.126**	0.083	1.512	0.131
VA	**1.077**	0.254	4.236	0.000	**0.420**	0.145	2.892	0.004	**9.683**	0.592	16.361	0.000
WA	**0.696**	0.335	2.075	0.038	**0.699**	0.223	3.136	0.002	**4.529**	0.780	5.804	0.000
WV	**0.784**	0.096	8.160	0.000	**0.676**	0.052	12.964	0.000	**1.575**	0.114	13.840	0.000
WI	**0.567**	0.149	3.794	0.000	**0.909**	0.072	12.606	0.000	**4.582**	0.674	6.802	0.000
WY	**1.035**	0.119	8.663	0.000	**0.738**	0.066	11.220	0.000	**-0.034**	0.067	-0.516	0.606
Threshold					<0.45				<-2.5 or >10			

Table G.7.2 Regression results Group 2 Lower 48 states from Eviews using 2 estimations (non-winter months, winter months)

Non-winter months

Dependent Variable: P_CG?

Method: Pooled EGLS (Cross-section weights)

Date: 08/06/19 Time: 14:03

Sample: 2013M01 2013M12 2014M03 2018M12 IF @MONTH>=4 AND @MONTH<=10

Included observations: 42

Cross-sections included: 49

Total pool (balanced) observations: 2058

Iterate weights to convergence

Convergence achieved after 1 weight iterations

Winter months

Dependent Variable: P_CG?

Method: Pooled EGLS (Cross-section weights)

Date: 08/06/19 Time: 14:05

Sample: 2013M01 2013M12 2014M03 2018M12 IF @MONTH<=3 OR @MONTH>=11

Included observations: 28

Cross-sections included: 49

Total pool (balanced) observations: 1372

Iterate weights to convergence

Convergence achieved after 1 weight iterations

Non winter months (Apr-Oct)

state	C	std. error	t-stat	P-value	α	std. error	t-stat	P-value	β	std. error	t-stat	P-value
GA	0.772	0.150	5.136	0.000	0.830	0.053	15.569	0.000	0.522	0.779	0.669	0.503
MI	0.411	0.125	3.275	0.001	0.939	0.056	16.779	0.000	-0.960	0.919	-1.045	0.296
MN	0.113	0.117	0.963	0.336	0.990	0.057	17.487	0.000	3.859	0.429	9.000	0.000
MS	0.956	0.203	4.705	0.000	0.810	0.071	11.467	0.000	-0.174	0.242	-0.721	0.471
NJ	1.040	0.284	3.662	0.000	1.019	0.115	8.824	0.000	2.661	2.718	0.979	0.328
NY	-1.148	0.228	-5.030	0.000	1.161	0.090	12.938	0.000	85.589	4.620	18.525	0.000
NC	-0.064	0.155	-0.412	0.681	1.075	0.061	17.655	0.000	4.255	0.451	9.437	0.000
OH	0.887	0.204	4.350	0.000	0.793	0.094	8.414	0.000	1.858	0.525	3.539	0.000
OR	0.970	0.164	5.928	0.000	0.661	0.078	8.474	0.000	2.108	0.236	8.926	0.000
PA	0.802	0.186	4.315	0.000	0.833	0.090	9.261	0.000	16.367	1.491	10.974	0.000
SC	0.354	0.189	1.868	0.062	1.193	0.071	16.707	0.000	-0.059	0.297	-0.199	0.842
VA	0.246	0.252	0.975	0.330	0.950	0.109	8.745	0.000	9.789	0.976	10.026	0.000

Winter months (Jan-Mar, Nov-Dec)

state	C	std. error	t-stat	P-value	α	std. error	t-stat	P-value	β	std. error	t-stat	P-value
GA	1.978	0.484	4.091	0.000	0.128	0.118	1.088	0.277	-0.076	6.680	-0.011	0.991
MI	1.358	0.354	3.832	0.000	0.465	0.075	6.216	0.000	2.748	18.515	0.148	0.882
MN	0.655	0.252	2.602	0.009	0.697	0.059	11.908	0.000	14.064	5.936	2.369	0.018
MS	1.289	0.287	4.483	0.000	0.550	0.106	5.182	0.000	0.323	0.751	0.430	0.667
NJ	2.657	0.483	5.498	0.000	0.089	0.068	1.299	0.194	-11.419	18.666	-0.612	0.541
NY	0.913	0.325	2.812	0.005	0.204	0.035	5.769	0.000	60.811	24.340	2.498	0.013
NC	1.885	0.446	4.226	0.000	0.124	0.092	1.356	0.175	-1.405	4.974	-0.283	0.778
OH	1.568	0.367	4.269	0.000	0.482	0.121	3.981	0.000	-11.231	13.666	-0.822	0.411
OR	0.658	0.303	2.168	0.030	0.449	0.098	4.560	0.000	7.500	1.977	3.794	0.000
PA	1.189	0.443	2.683	0.007	0.366	0.120	3.060	0.002	29.412	15.582	1.888	0.059

SC	2.236	0.458	4.884	0.000	0.111	0.090	1.234	0.218	-1.198	2.118	-0.566	0.572
VA	1.849	0.406	4.555	0.000	0.212	0.078	2.733	0.006	2.305	6.086	0.379	0.705

Table G.7.3 Regression results Group 3 Lower 48 states from Eviews estimating only non-winter months

Non-winter months

Dependent Variable: P_CG?

Method: Pooled EGLS (Cross-section weights)

Date: 08/06/19 Time: 14:03

Sample: 2013M01 2013M12 2014M03 2018M12 IF @MONTH>=4 AND @MONTH<=10

Included observations: 42

Cross-sections included: 49

Total pool (balanced) observations: 2058

Iterate weights to convergence

Convergence achieved after 1 weight iterations

Non winter months (Apr-Oct)

state	C	std. error	t-stat	P-value	α	std. error	t-stat	P-value	β	std. error	t-stat	P-value
CT	0.409	0.557	0.734	0.463	0.764	0.200	3.815	0.000	4.979	1.451	3.432	0.001
DE	-0.026	0.348	-0.076	0.940	-0.076	0.325	-0.234	0.815	1.124	0.325	3.458	0.001
DC	3.853	0.926	4.160	0.000	0.827	0.135	6.124	0.000	2.527	0.288	8.766	0.000
ME	1.481	0.573	2.584	0.010	0.415	0.153	2.716	0.007	0.516	0.184	2.799	0.005
MA	0.738	0.577	1.280	0.201	0.498	0.202	2.463	0.014	19.294	2.694	7.161	0.000
NH	0.048	0.835	0.057	0.955	0.869	0.291	2.990	0.003	1.003	0.214	4.692	0.000
RI	0.229	0.303	0.754	0.451	0.703	0.120	5.874	0.000	0.086	0.120	0.714	0.475
UT	2.188	0.329	6.659	0.000	0.531	0.160	3.324	0.001	-1.885	0.726	-2.597	0.010
VT	0.944	0.438	2.155	0.031	0.909	0.203	4.470	0.000	0.130	0.152	0.856	0.392

Data:

For the data used in these estimations, please contact kathryn.dyl@eia.gov.

Spot price data from Natural Gas Intelligence were used for the following hubs to generate a monthly average spot price by state. Up to 3 hubs per state were chosen using analysts' judgements and are given in the table below.

Table G.7.4 *Natural Gas Intelligence* trading hub prices used for Lower 48 states.

state	hub 1	hub 2	hub 3
AL	Florida Gas Zone 3	Transco Zone 4	
AZ	SoCal Border Avg.	El Paso S. Mainline/N. Baja	El Paso non-Bondad
AR	Texas Gas Zone 1	Texas Eastern, M1, 24	Enable South
CA	PG&E Citygate	SoCal Citygate	
CO	Northwest S. of Green River	White River Hub	

state	hub 1	hub 2	hub 3
CT	Algonquin Citygate	Iroquois Zone 2	
DE	Transco Zone 6 non-NY		
DC	Columbia Gas	Transco Zone 5	
FL	FGT Citygate		
GA	Transco Zone 4	Transco Zone 5	Florida Gas Zone 3
ID	Stanfield	Northwest Wyoming Pool	Kingsgate
IL	Chicago Citygate		
IN	Chicago Citygate	Lebanon	Michigan Consolidated
IA	Northern Natural Ventura	Chicago Citygate	
KS	Panhandle Eastern	NGPL Midcontinent	
KY	Texas Gas Zone 1	Lebanon	Columbia Gas
LA	Henry Hub	Texas Gas Zone 1	Transco Zone 3
ME	Dracut		
MD	Columbia Gas	Transco Zone 5	
MA	Algonquin Citygate	Tenn Zone 6 200L	Dracut
MI	Consumers Energy	Michigan Consolidated	ANR ML7
MN	Northern Natural Ventura	Emerson	
MS	Tennessee Line 500	Florida Gas Zone 3	Texas Eastern M-1, 30
MO	Texas Eastern, M1, 24	Chicago Citygate	Panhandle Eastern
MT	Empress	Stanfield	Opal
NE	NGPL Amarillo Mainline	Cheyenne Hub	
NV	Opal	Malin	Kern Delivery
NH	Tenn Zone 6 200L	Dracut	
NJ	Texas Eastern M-3, Delivery	Transco Zone 6 non-NY	
NM	El Paso non-Bondad	El Paso Permian	Transwestern
NY	Transco Zone 6 NY	Tenn Zone 5 200L	
NC	Transco Zone 5	Columbia Gas	
ND	Emerson	Cheyenne Hub	Northern Natural Ventura
OH	Lebanon	Columbia Gas	Clarington (non-Tenn)
OK	NGPL Midcontinent	OGT	Panhandle Eastern
OR	Stanfield	Malin	
PA	Tennessee Zn 4 Marcellus	Dominion South	Texas Eastern M-3, Delivery
RI	Algonquin Citygate		
SC	Transco Zone 4	Transco Zone 5	
SD	Emerson	Cheyenne Hub	Northern Natural Ventura
TN	Texas Eastern M-1, 30	Transco Zone 4	Texas Gas Zone 1
TX	Carthage	Houston Ship Channel	Waha
UT	Questar	Northwest Wyoming Pool	
VT	Iroquois Waddington		
VA	Transco Zone 5	Dominion South	
WA	Northwest Sumas	Stanfield	Kingsgate
WV	Dominion South	Columbia Gas	

state	hub 1	hub 2	hub 3
WI	ANR ML7	Emerson	
WY	Opal	Cheyenne Hub	CIG

G-8. Distributor tariff markup for delivered prices to the residential sector by Census division

Data: Residential distributor tariffs

Author: Katie Dyl, EIA, 2019

Source: Residential delivered prices, citygate prices, and residential and commercial consumption from EIA's Natural Gas Monthly (NGM), DOE/EIA-0130. Number of households and heading degree days (HDD) were provided via the NEMS Residential Module by Census division (RSGASCUST), with the original data based on EIA's Residential Energy Consumption Survey. The NEMS Residential Module data was supplemented by the number of residential gas consumers from the Natural Gas Annual (NGA).

Variables:

For historical year yr and Census division r,

$HOUSEHOLDS_{yr,r}$ = Number of households

$NUM\_GASCONSUMERS_{yr,r}$ = Number of residential gas consumers from the NGA

$R\_MARKUP_{yr,r}$ = Residential distributor markups or Census division level average residential delivered prices (averaged over states using residential consumption as weights) minus the average citygate prices (averaged over states using residential plus commercial consumption as weights) (1987$/Mcf) [MARKUP_RES in documentation]

$QPERRESNUM_{yr,r}$ = Consumption per household or residential consumption [QCD_RES in documentation] divided by the number of residential natural gas customers [HOUSES in documentation] (Bcf/household)

$QPERHDD_{yr,r}$ = Consumption per heating degree day (Bcf/HDD)

C_r= estimated constant term for Census division r (1987$/Mcf)

Derivation:

The household data in NEMS for the nine Census divisions for 2009 and subsequent years were used. For years prior to 2009, the NEMS data was supplemented by the number of residential gas consumers from the NGA. Ratio of households to residential gas consumers was assumed to be constant each year:

$$HOUSEHOLDS_{yr<2009,r} = NUM\_GASCONSUMERS_{yr<2009,r} * \frac{HOUSEHOLDS_{yr=2009,r}}{NUM\_GASCONSUMERS_{yr=2009,r}}$$

The estimated equation is used to generate plausible markups for the residential sector based on available price data, namely the residential end use prices from the NGM and citygate prices in the division. The estimated equation follows:

$$R\_MARKUP_{yr,r} = \alpha * QPERRESNUM_{yr,r} + \beta * QPERHHD_{yr,r} + C$$

The equation was estimated for each Census division individually.

Regression diagnostics and parameter estimates:

For every Census division X:
Dependent Variable: R_MARKUP
Method: Panel Least Squares
Date: 07/31/19 Time: 09:03
Sample: 1990 2017 IF @CROSSID=X
Periods included: 28
Cross-sections included: 1
Total panel (balanced) observations: 28

Parameter estimates:

Cross ID	Census division	Variable	Coefficient	Std. Error	t-Statistic	Prob.
1	East North Central	QPERRESNUM	-10716.350	2150.174	-4.984	0.000
1	East North Central	QPERHDD	-14.924	2.774	-5.379	0.000
1	East North Central	C	6.667	0.504	13.236	0.000
2	East SouthCentral	QPERRESNUM	-33121.400	2819.843	-11.746	0.000
2	East SouthCentral	QPERHDD	-38.008	10.518	-3.613	0.001
2	East SouthCentral	C	8.236	0.479	17.176	0.000
3	Middle Atlantic	QPERRESNUM	-11562.360	7440.487	-1.554	0.133
3	Middle Atlantic	QPERHDD	-10.485	7.444	-1.409	0.171
3	Middle Atlantic	C	6.237	1.543	4.042	0.000
4	Mountain	QPERRESNUM	-28839.590	8747.375	-3.297	0.003
4	Mountain	QPERHDD	-4.984	10.087	-0.494	0.626
4	Mountain	C	5.073	1.379	3.677	0.001
5	New England	QPERRESNUM	-9364.297	8520.240	-1.099	0.282
5	New England	QPERHDD	69.787	25.551	2.731	0.011
5	New England	C	2.821	1.037	2.720	0.012
6	Pacific	QPERRESNUM	-44234.410	14292.300	-3.095	0.005
6	Pacific	QPERHDD	-0.056	12.549	-0.004	0.997
6	Pacific	C	5.779	2.838	2.036	0.053
7	South Atlantic	QPERRESNUM	-39960.300	7926.799	-5.041	0.000
7	South Atlantic	QPERHDD	10.446	6.695	1.560	0.131
7	South Atlantic	C	5.582	1.659	3.365	0.003
8	West North Central	QPERRESNUM	-17818.190	1731.943	-10.288	0.000
8	West North Central	QPERHDD	-41.888	10.435	-4.014	0.001
8	West North Central	C	6.866	0.637	10.783	0.000
9	West South Central	QPERRESNUM	-50154.140	5786.311	-8.668	0.000
9	West South Central	QPERHDD	9.639	5.922	1.628	0.116
9	West South Central	C	4.391	0.877	5.009	0.000

Regression Diagnostics:

Cross-ID Census division	1 ENC	2 ESC	3 MIDATL	4 MTN	5 NE	6 PAC	7 SATL	8 WNC	9 WSC
R-squared	0.863	0.931	0.118	0.711	0.245	0.339	0.772	0.911	0.776
Adjusted R-squared	0.852	0.925	0.047	0.688	0.185	0.286	0.754	0.904	0.759
S.E. of regression	0.134	0.168	0.266	0.182	0.297	0.409	0.336	0.115	0.272
Sum squared resid	0.447	0.707	1.765	0.825	2.207	4.187	2.822	0.333	1.849
Log likelihood	18.200	11.780	-1.030	9.622	-4.163	-13.126	-7.601	22.296	-1.687
F-statistic	78.823	168.556	1.667	30.817	4.065	6.415	42.296	127.539	43.410
Prob(F-statistic)	0.000	0.000	0.209	0.000	0.030	0.006	0.000	0.000	0.000
Mean dependent var	2.053	2.979	3.487	2.341	4.094	3.139	3.845	2.210	2.844
S.D. dependent var	0.348	0.616	0.272	0.325	0.329	0.484	0.677	0.372	0.553
Akaike info criterion	-1.086	-0.627	0.288	-0.473	0.512	1.152	0.757	-1.378	0.335
Schwarz criterion	-0.943	-0.484	0.431	-0.330	0.654	1.295	0.900	-1.236	0.477
Hannan-Quinn criter.	-1.042	-0.584	0.331	-0.429	0.555	1.195	0.801	-1.335	0.378
Durbin-Watson stat	1.3476	1.76962	1.396	2.360	1.467	0.730	0.7016	1.99392	0.738

Data:

year	Census division	R_MARKUP	QPERRESNUM	QPERHDD
1990	East North Central	1.7707613	0.000128232	0.2374135
1991	East North Central	1.7820364	0.000132095	0.2385928
1992	East North Central	1.7138418	0.000135193	0.2319445
1993	East North Central	1.8589532	0.000139175	0.2299349
1994	East North Central	2.0211115	0.000132739	0.2325066
1995	East North Central	1.6656922	0.000136805	0.2341077
1996	East North Central	1.525236	0.000143796	0.2404529
1997	East North Central	1.7904426	0.000133203	0.2370003
1998	East North Central	1.9169999	0.000110065	0.2486064
1999	East North Central	1.7651257	0.00011732	0.2392757
2000	East North Central	1.6001902	0.000122126	0.2404053
2001	East North Central	1.8039611	0.000110812	0.2375676
2002	East North Central	1.9311637	0.000115538	0.2399462
2003	East North Central	1.7133112	0.000120918	0.2368166
2004	East North Central	1.8896326	0.000112047	0.2333398
2005	East North Central	1.9625456	0.000107707	0.2296336
2006	East North Central	2.1529381	9.63715E-05	0.2209236
2007	East North Central	2.0828961	0.000103734	0.2246008
2008	East North Central	2.2650821	0.000109263	0.2153676
2009	East North Central	2.5278436	0.000103775	0.2090175
2010	East North Central	2.3255826	9.82213E-05	0.2081002
2011	East North Central	2.5081204	9.96237E-05	0.2113826

year	Census division	R_MARKUP	QPERRESNUM	QPERHDD
2012	East North Central	2.7213388	8.68169E-05	0.2123102
2013	East North Central	2.3131021	0.000106963	0.2118478
2014	East North Central	2.1441504	0.000114721	0.2094543
2015	East North Central	2.4748894	9.86802E-05	0.2124294
2016	East North Central	2.570595	9.33738E-05	0.2177848
2017	East North Central	2.6815236	9.2644E-05	0.2171866
1990	East SouthCentral	2.1711388	0.000107108	0.0607747
1991	East SouthCentral	2.352662	0.000109191	0.0580403
1992	East SouthCentral	2.25485	0.00011272	0.0564068
1993	East SouthCentral	2.1719513	0.000117599	0.0562381
1994	East SouthCentral	2.4963455	0.000109962	0.0598538
1995	East SouthCentral	2.383704	0.000111318	0.0583342
1996	East SouthCentral	1.9479228	0.000120308	0.0621862
1997	East SouthCentral	2.5162729	0.000106995	0.0585539
1998	East SouthCentral	2.6575737	9.45879E-05	0.0639781
1999	East SouthCentral	2.6006621	9.3311E-05	0.0613211
2000	East SouthCentral	2.3421351	0.00010155	0.0608024
2001	East SouthCentral	2.7941651	9.71672E-05	0.0628035
2002	East SouthCentral	2.8832089	9.61383E-05	0.0604386
2003	East SouthCentral	2.6944278	9.69323E-05	0.059458
2004	East SouthCentral	3.074118	8.89594E-05	0.0594502
2005	East SouthCentral	3.2990372	8.80346E-05	0.0576279
2006	East SouthCentral	3.7059001	7.81682E-05	0.0539539
2007	East SouthCentral	3.2310557	7.89497E-05	0.0549812
2008	East SouthCentral	3.1842848	8.67481E-05	0.0532852
2009	East SouthCentral	3.8024821	8.28996E-05	0.0516283
2010	East SouthCentral	3.2618163	9.18779E-05	0.0513061
2011	East SouthCentral	3.4151104	8.24798E-05	0.0544628
2012	East SouthCentral	3.8352892	6.63391E-05	0.0510914
2013	East SouthCentral	3.3720367	8.52176E-05	0.0518687
2014	East SouthCentral	3.2047344	9.37162E-05	0.0530407
2015	East SouthCentral	3.7804525	8.11634E-05	0.0551468
2016	East SouthCentral	3.8113759	7.15687E-05	0.0510365
2017	East SouthCentral	4.1653597	6.74508E-05	0.0529432
1990	Middle Atlantic	3.3319547	9.92317E-05	0.1471641
1991	Middle Atlantic	3.4184438	9.33438E-05	0.1438507
1992	Middle Atlantic	3.4189488	0.000102747	0.1391629
1993	Middle Atlantic	3.4110606	0.000109619	0.1411921
1994	Middle Atlantic	3.8075648	0.000111744	0.1450217
1995	Middle Atlantic	3.8831815	0.000105681	0.1409911
1996	Middle Atlantic	3.410511	0.000113979	0.1489228
1997	Middle Atlantic	3.5764641	0.000106556	0.1448153
1998	Middle Atlantic	4.0125062	9.26654E-05	0.1542371

year	Census division	R_MARKUP	QPERRESNUM	QPERHDD
1999	Middle Atlantic	3.735435	0.000100737	0.1518544
2000	Middle Atlantic	2.857753	0.000106351	0.152182
2001	Middle Atlantic	3.3376977	9.83996E-05	0.155807
2002	Middle Atlantic	3.1207607	9.62491E-05	0.1521956
2003	Middle Atlantic	3.3287477	0.000107016	0.1517591
2004	Middle Atlantic	3.4807385	0.000101471	0.1531857
2005	Middle Atlantic	3.0552968	0.000102566	0.1533227
2006	Middle Atlantic	3.4787034	8.67996E-05	0.1500882
2007	Middle Atlantic	3.4831354	9.7171E-05	0.1539274
2008	Middle Atlantic	3.4994556	9.60064E-05	0.1501954
2009	Middle Atlantic	4.0773159	9.71312E-05	0.1490496
2010	Middle Atlantic	3.6108428	9.44524E-05	0.1541609
2011	Middle Atlantic	3.7090016	9.41042E-05	0.1552653
2012	Middle Atlantic	3.6324895	8.53741E-05	0.1552266
2013	Middle Atlantic	3.5104315	0.000100764	0.1559812
2014	Middle Atlantic	3.2023286	0.00011049	0.1611742
2015	Middle Atlantic	3.2645655	0.000100757	0.1666019
2016	Middle Atlantic	3.386443	9.1381E-05	0.1632023
2017	Middle Atlantic	3.5854266	9.45929E-05	0.1697113
1990	Mountain	1.9767465	9.72209E-05	0.0476465
1991	Mountain	1.988692	0.000101561	0.0513659
1992	Mountain	1.9227291	9.48678E-05	0.0503797
1993	Mountain	1.9253767	0.000102882	0.0517224
1994	Mountain	1.9588682	9.38988E-05	0.0542249
1995	Mountain	2.1966264	9.13354E-05	0.0554891
1996	Mountain	1.8088947	9.49702E-05	0.0600448
1997	Mountain	1.9210336	9.62492E-05	0.0609531
1998	Mountain	2.4243292	9.24343E-05	0.0621792
1999	Mountain	2.4851398	8.7399E-05	0.0652585
2000	Mountain	1.9478578	8.61856E-05	0.0660197
2001	Mountain	2.5622402	8.55762E-05	0.0674221
2002	Mountain	2.5702037	8.63105E-05	0.0675822
2003	Mountain	2.1836586	8.09357E-05	0.0704733
2004	Mountain	2.3671548	8.1013E-05	0.0701816
2005	Mountain	2.4738993	7.82066E-05	0.0721017
2006	Mountain	2.674742	7.50494E-05	0.0709784
2007	Mountain	2.3894948	7.74442E-05	0.0745951
2008	Mountain	2.1536206	7.93481E-05	0.0733569
2009	Mountain	2.6762654	7.5584E-05	0.0717648
2010	Mountain	2.3985108	7.6852E-05	0.0741273
2011	Mountain	2.3408848	7.82957E-05	0.0724361
2012	Mountain	2.7616186	6.95575E-05	0.0755409
2013	Mountain	2.3149175	7.92174E-05	0.0755028

year	Census division	R_MARKUP	QPERRESNUM	QPERHDD
2014	Mountain	2.6336701	7.21695E-05	0.0774175
2015	Mountain	3.0277643	6.41008E-05	0.077264
2016	Mountain	2.7762378	6.54344E-05	0.0791474
2017	Mountain	2.6736673	6.47688E-05	0.0798376
1990	New England	3.9259777	9.25899E-05	0.0294272
1991	New England	4.0584964	8.94586E-05	0.0282129
1992	New England	3.848764	0.000102258	0.0284033
1993	New England	3.7982949	0.000102195	0.0291749
1994	New England	4.2030947	9.73717E-05	0.0284663
1995	New England	4.2056144	8.92441E-05	0.0266432
1996	New England	3.6768178	9.64169E-05	0.0286035
1997	New England	4.0821058	9.09786E-05	0.0274324
1998	New England	4.0233291	8.12047E-05	0.0287997
1999	New England	3.8637275	8.5841E-05	0.029469
2000	New England	3.372084	8.96858E-05	0.0291518
2001	New England	3.8937886	8.4582E-05	0.0295035
2002	New England	3.535314	8.38873E-05	0.0294419
2003	New England	3.7196126	9.41472E-05	0.0300373
2004	New England	4.2690615	8.71585E-05	0.0287558
2005	New England	3.6746266	8.91105E-05	0.0297532
2006	New England	4.5307905	7.59046E-05	0.0293782
2007	New England	4.559466	8.34632E-05	0.0292965
2008	New England	4.1896915	9.71334E-05	0.0323679
2009	New England	4.2774574	9.15618E-05	0.0319509
2010	New England	4.1814974	8.72034E-05	0.0340795
2011	New England	4.1060076	8.96697E-05	0.0340262
2012	New England	4.4569099	8.13556E-05	0.0339919
2013	New England	4.2967337	8.57159E-05	0.0311198
2014	New England	4.3570205	9.26089E-05	0.0325233
2015	New England	4.3155764	0.000103134	0.0334464
2016	New England	4.635812	9.14173E-05	0.0326989
2017	New England	4.5719891	9.82681E-05	0.0346334
1990	Pacific	2.6728995	6.8735E-05	0.1696581
1991	Pacific	3.0000965	6.72072E-05	0.1721809
1992	Pacific	2.8043307	6.29172E-05	0.1795167
1993	Pacific	2.8295035	6.71386E-05	0.1763825
1994	Pacific	3.007491	6.75053E-05	0.1737043
1995	Pacific	3.3947626	6.14515E-05	0.1777317
1996	Pacific	2.9271581	6.28498E-05	0.1815129
1997	Pacific	2.8357896	6.15468E-05	0.1857315
1998	Pacific	3.3127746	7.0184E-05	0.1834078
1999	Pacific	2.9830482	7.00881E-05	0.196054
2000	Pacific	2.8624243	6.06897E-05	0.1795033

year	Census division	R_MARKUP	QPERRESNUM	QPERHDD
2001	Pacific	2.7939411	6.32694E-05	0.187457
2002	Pacific	2.9070519	6.12227E-05	0.1857809
2003	Pacific	2.7077186	5.8715E-05	0.189538
2004	Pacific	2.6226587	5.94914E-05	0.195
2005	Pacific	2.7461312	5.62955E-05	0.1865562
2006	Pacific	3.099355	5.63085E-05	0.1811439
2007	Pacific	3.0818839	5.6393E-05	0.1864186
2008	Pacific	2.9073192	5.6921E-05	0.1847239
2009	Pacific	3.2271983	5.52747E-05	0.182805
2010	Pacific	2.9756135	5.51073E-05	0.1777141
2011	Pacific	3.2049268	5.82017E-05	0.1777382
2012	Pacific	3.2632756	5.45107E-05	0.1864319
2013	Pacific	3.2182506	5.53312E-05	0.192319
2014	Pacific	3.3903052	4.71E-05	0.1988004
2015	Pacific	4.2244608	4.77368E-05	0.1899948
2016	Pacific	4.4911928	4.93048E-05	0.1879455
2017	Pacific	4.3926817	5.29906E-05	0.1923832
1990	South Atlantic	3.150255	9.10036E-05	0.1484726
1991	South Atlantic	3.0840673	9.33722E-05	0.1394001
1992	South Atlantic	3.0377147	9.94494E-05	0.13641
1993	South Atlantic	3.083436	0.000103662	0.137799
1994	South Atlantic	3.2435117	9.7204E-05	0.1473089
1995	South Atlantic	3.1612883	9.8354E-05	0.1414396
1996	South Atlantic	2.8633428	0.000107327	0.1528645
1997	South Atlantic	3.280077	9.59296E-05	0.1533322
1998	South Atlantic	3.2331936	8.39149E-05	0.1639933
1999	South Atlantic	3.4154057	8.25188E-05	0.1544389
2000	South Atlantic	3.2023353	9.85625E-05	0.1671956
2001	South Atlantic	3.5239823	8.52233E-05	0.1662404
2002	South Atlantic	3.6727044	8.81631E-05	0.1706358
2003	South Atlantic	3.5173452	9.4233E-05	0.1711401
2004	South Atlantic	3.9132797	8.95354E-05	0.1751662
2005	South Atlantic	3.8294457	8.78016E-05	0.1724233
2006	South Atlantic	4.6919389	7.54998E-05	0.1670162
2007	South Atlantic	4.6096357	7.86893E-05	0.1747797
2008	South Atlantic	4.1972745	8.04825E-05	0.1673521
2009	South Atlantic	4.6533156	8.17136E-05	0.1632694
2010	South Atlantic	4.336531	8.89955E-05	0.1580234
2011	South Atlantic	4.4922676	7.71151E-05	0.1693113
2012	South Atlantic	4.9797536	6.83672E-05	0.1674837
2013	South Atlantic	4.3102357	8.26894E-05	0.171314
2014	South Atlantic	4.0004991	9.05103E-05	0.1741847
2015	South Atlantic	4.5249151	6.68741E-05	0.1862039

year	Census division	R_MARKUP	QPERRESNUM	QPERHDD
2016	South Atlantic	4.5942779	6.30363E-05	0.1786823
2017	South Atlantic	5.0510074	6.03352E-05	0.1893502
1990	West North Central	1.6989114	0.000112759	0.0697931
1991	West North Central	1.7240521	0.00012006	0.071776
1992	West North Central	1.7986918	0.00011245	0.068789
1993	West North Central	1.7597729	0.000123988	0.0683741
1994	West North Central	1.8335409	0.000115158	0.0705651
1995	West North Central	1.8982782	0.000117996	0.0699942
1996	West North Central	1.8595225	0.000128188	0.0717791
1997	West North Central	1.789809	0.000114428	0.0707462
1998	West North Central	2.0968946	9.74046E-05	0.0732453
1999	West North Central	2.049099	0.000100758	0.073256
2000	West North Central	2.0153952	0.000103954	0.0709028
2001	West North Central	2.1429921	0.000101889	0.0732535
2002	West North Central	2.0142349	0.000102131	0.0715125
2003	West North Central	1.8890511	0.000102056	0.071022
2004	West North Central	2.2117481	9.55583E-05	0.0699589
2005	West North Central	2.1538251	9.27416E-05	0.0699115
2006	West North Central	2.5823773	8.34034E-05	0.0678722
2007	West North Central	2.5579359	9.08267E-05	0.0674362
2008	West North Central	2.2663035	9.95327E-05	0.066698
2009	West North Central	2.5448288	9.47839E-05	0.0663138
2010	West North Central	2.4528512	9.04117E-05	0.0660448
2011	West North Central	2.6497689	8.92216E-05	0.065373
2012	West North Central	2.8267491	7.37731E-05	0.0645774
2013	West North Central	2.5008266	9.57207E-05	0.0651564
2014	West North Central	2.1871639	0.000100931	0.0675263
2015	West North Central	2.8231261	7.3289E-05	0.0665722
2016	West North Central	2.7467765	7.01247E-05	0.067356
2017	West North Central	2.7980337	7.10561E-05	0.0661937
1990	West South Central	2.4099973	7.63176E-05	0.1939126
1991	West South Central	2.4060221	7.90727E-05	0.1821185
1992	West South Central	2.3304219	7.67019E-05	0.1806623
1993	West South Central	2.2093882	8.20185E-05	0.1691247
1994	West South Central	2.4722259	7.55434E-05	0.1851468
1995	West South Central	2.4435842	7.24914E-05	0.178507
1996	West South Central	2.2684711	7.89894E-05	0.1854893
1997	West South Central	2.2980458	7.70334E-05	0.1723471
1998	West South Central	2.7426692	6.69692E-05	0.1836355
1999	West South Central	2.6042638	6.007E-05	0.1827747
2000	West South Central	2.3728277	6.53096E-05	0.1688388
2001	West South Central	2.4336558	6.57898E-05	0.1700231
2002	West South Central	2.5465802	6.6548E-05	0.1643567

year	Census division	R_MARKUP	QPERRESNUM	QPERHDD
2003	West South Central	2.5383399	6.44922E-05	0.1670567
2004	West South Central	2.900507	5.8584E-05	0.165561
2005	West South Central	2.9304065	5.6772E-05	0.1653273
2006	West South Central	3.4200479	5.01769E-05	0.1621809
2007	West South Central	2.7312684	5.67725E-05	0.1604794
2008	West South Central	2.7974867	5.67204E-05	0.1602822
2009	West South Central	3.276333	5.49299E-05	0.1544046
2010	West South Central	2.9655673	6.35449E-05	0.1569286
2011	West South Central	2.8623598	5.67976E-05	0.1623132
2012	West South Central	3.6445818	4.70164E-05	0.1721472
2013	West South Central	3.1437661	5.86759E-05	0.1529888
2014	West South Central	2.8774424	6.544E-05	0.1644938
2015	West South Central	3.5429932	5.68801E-05	0.1687542
2016	West South Central	4.0574405	4.70176E-05	0.167335
2017	West South Central	4.3946700	0.0000446	0.1769668

G-9. Distributor tariff markup for delivered prices to the commercial sector by Census division

Data: Commercial distributor tariffs

Author: Samantha Calkins, EIA 2017

Source: Commercial delivered prices, citygate prices, and residential and commercial consumption from EIA's Natural Gas Monthly (NGM), DOE/EIA-0130. Volume of commercial floorspace was provided via the NEMS Commercial Module by Census division (CMTOTALFLSPC). The NEMS Commercial Module data was supplemented by the number of commercial gas consumers from the Natural Gas Annual (NGA).

Variables:

For historical year yr and Census division r,

$FLOORSPACE_{yr,r}$ = Commercial floorspace (million cubic feet)

$NUM\_GASCONSUMERS_{yr,r}$ = Number of commercial gas consumers from the NGA

$C\_MARKUP_{yr,r}$ = Commercial distributor markups or Census-division level average commercial delivered prices (averaged over states using commercial consumption as weights) minus the average citygate prices (averaged over states using residential plus commercial consumption as weights) (1987$/Mcf)

$QPERFLOR_{yr,r}$ = Consumption per commercial floorspace or commercial consumption [Q_COM_BCF] divided by the commercial floorspace [FLOORSPACE] (Bcf/million cubic feet)

$Q\_COM\_BCF_{yr,r}$ = Commercial consumption (Bcf)

C_r = estimated constant term for Census division r (1987$/Mcf)

Derivation:

The floorspace data in NEMS for the nine Census divisions were used starting in 2003. Prior to 2003, the NEMS data was supplemented by the number of commercial gas consumers from the NGA. Ratio of floorspace to commercial gas consumers was assumed to be constant year to year:

$$FLOORSPACE_{yr<2003,r} = NUM\_GASCONSUMERS_{yr<2003,r} * \frac{FLOORSPACE_{yr=2003,r}}{NUM\_GASCONSUMERS_{yr=2003,r}}$$

The estimated equation is used to generate plausible markups for the commercial sector based on available price data, namely the commercial end use prices from the NGM and citygate prices in the division. The estimated equation follows:

$C\_MARKUP_{yr,r} = C_r + (\alpha * QPERFLOOR_{yr,r}) + (\beta * Q\_COM\_BCF_{yr,r})$ equation was estimated using a basic ordinary least squares approach applied to the data provided below.

Regression diagnostics and parameter estimates:

For every Census division X:
Dependent Variable: C_MARKUP
Method: Panel Least Squares
Date: 07/31/19 Time: 09:22
Sample: 1990 2017 IF @CROSSID=X
Periods included: 28
Cross-sections included: 1
Total panel (balanced) observations: 28

Parameter estimates:

Cross ID	Census division	Variable	Coefficient	Std. Error	t-Statistic	Prob.
1	East North Central	QPERFLOOR	-23.355	6.280	-3.719	0.001
1	East North Central	Q_COM_BCF	0.001	0.001	0.940	0.356
1	East North Central	C	2.339	0.492	4.758	0.000
2	East SouthCentral	QPERFLOOR	-56.832	4.367	-13.013	0.000
2	East SouthCentral	Q_COM_BCF	0.006	0.003	1.717	0.098
2	East SouthCentral	C	3.426	0.437	7.835	0.000
3	Middle Atlantic	QPERFLOOR	20.165	13.873	1.453	0.159
3	Middle Atlantic	Q_COM_BCF	-0.003	0.001	-3.007	0.006
3	Middle Atlantic	C	2.570	0.578	4.444	0.000
4	Mountain	QPERFLOOR	-32.567	10.544	-3.089	0.005
4	Mountain	Q_COM_BCF	0.001	0.003	0.313	0.757
4	Mountain	C	3.017	1.160	2.602	0.015
5	New England	QPERFLOOR	-30.103	10.440	-2.883	0.008
5	New England	Q_COM_BCF	0.006	0.002	2.438	0.022
5	New England	C	2.883	0.390	7.401	0.000
6	Pacific	QPERFLOOR	-11.546	8.388	-1.376	0.181
6	Pacific	Q_COM_BCF	0.007	0.004	1.658	0.110
6	Pacific	C	0.290	1.415	0.205	0.839
7	South Atlantic	QPERFLOOR	-69.223	13.973	-4.954	0.000
7	South Atlantic	Q_COM_BCF	0.004	0.001	6.807	0.000
7	South Atlantic	C	2.337	0.468	4.995	0.000
8	West North Central	QPERFLOOR	-43.301	3.209	-13.495	0.000
8	West North Central	Q_COM_BCF	0.006	0.001	6.098	0.000
8	West North Central	C	1.847	0.218	8.467	0.000
9	West South Central	QPERFLOOR	-49.039	8.661	-5.662	0.000

9	West South Central	Q_COM_BCF	0.002	0.002	1.343	0.191
9	West South Central	C	2.303	0.414	5.558	0.000

Regression Diagnostics:

Cross-ID	1	2	3	4	5	6	7	8	9
Census division	**ENC**	**ESC**	**MIDATL**	**MTN**	**NE**	**PAC**	**SATL**	**WNC**	**WSC**
R-squared	0.380	0.880	0.284	0.493	0.273	0.125	0.807	0.889	0.574
Adjusted R-squared	0.330	0.870	0.227	0.453	0.215	0.055	0.792	0.881	0.540
S.E. of regression	0.163	0.147	0.263	0.188	0.294	0.327	0.136	0.077	0.160
Sum squared resid	0.661	0.542	1.732	0.888	2.162	2.676	0.463	0.148	0.640
Log likelihood	12.722	15.497	-0.769	8.588	-3.875	-6.860	17.689	33.673	13.163
F-statistic	7.648	91.291	4.959	12.167	4.691	1.790	52.394	100.540	16.843
Prob(F-statistic)	0.003	0.000	0.015	0.000	0.019	0.188	0.000	0.000	0.000
Mean dependent var	1.487	2.139	1.930	1.520	2.448	2.294	2.239	1.353	1.447
S.D. dependent var	0.199	0.408	0.299	0.255	0.332	0.337	0.298	0.223	0.236
Akaike info criterion	-0.694	-0.893	0.269	-0.399	0.491	0.704	-1.049	-2.191	-0.726
Schwarz criterion	-0.552	-0.750	0.412	-0.256	0.634	0.847	-0.907	-2.048	-0.583
Hannan-Quinn criter.	-0.651	-0.849	0.313	-0.356	0.535	0.748	-1.006	-2.147	-0.682
Durbin-Watson stat	1.421	1.852	1.804	2.068	1.071	0.786	1.810	2.229	1.485

Data:

year	Census division	C_MARKUP	QPERFLOOR	Q_COM_BCF
1990	East North Central	1.2899587	0.0633088	655.60101
1991	East North Central	1.3514716	0.0632875	667.30798
1992	East North Central	1.2680307	0.0648389	692.29199
1993	East North Central	1.4412	0.0664958	721.22699
1994	East North Central	1.5997646	0.065876	720.54303
1995	East North Central	1.2986358	0.0681678	761.185
1996	East North Central	1.1735077	0.0713212	811.09302
1997	East North Central	1.3887244	0.0671472	771.62299
1998	East North Central	1.5040035	0.0575576	669.09698
1999	East North Central	1.4174357	0.0600164	710.89697
2000	East North Central	1.2142611	0.0630417	759.84998
2001	East North Central	1.2958694	0.0589497	709.02698
2002	East North Central	1.7753179	0.0585594	726.04797
2003	East North Central	1.2788058	0.0624405	775.74298
2004	East North Central	1.3822199	0.0581267	732.81097
2005	East North Central	1.3367659	0.0563307	720.85101
2006	East North Central	1.4407466	0.0513617	668.38702
2007	East North Central	1.4627915	0.0535165	707.72699

year	Census division	C_MARKUP	QPERFLOOR	Q_COM_BCF
2008	East North Central	1.6372353	0.056515	759.95502
2009	East North Central	1.6759545	0.0538482	732.80298
2010	East North Central	1.4602517	0.0492455	676.03198
2011	East North Central	1.788127	0.0518872	715.39099
2012	East North Central	1.8885774	0.0496863	633.54999
2013	East North Central	1.5315808	0.0601395	770.117
2014	East North Central	1.5035663	0.0653043	839.52399
2015	East North Central	1.6812657	0.0578841	747.67297
2016	East North Central	1.7709096	0.0552263	718.20898
2017	East North Central	1.7891484	0.0556875	729.42548
1990	East SouthCentral	1.5804121	0.0423349	121.27
1991	East SouthCentral	1.6816195	0.0432066	125.39
1992	East SouthCentral	1.6688387	0.0441148	130.27901
1993	East SouthCentral	1.6406871	0.0457672	138.30901
1994	East SouthCentral	1.9010663	0.0445813	137.536
1995	East SouthCentral	1.8029282	0.0450697	142.47301
1996	East SouthCentral	1.4553385	0.0484489	156.198
1997	East SouthCentral	1.8305475	0.0463382	153.96201
1998	East SouthCentral	1.9021557	0.0391915	136.745
1999	East SouthCentral	1.8828288	0.0411694	140.575
2000	East SouthCentral	1.6575039	0.0411747	144.888
2001	East SouthCentral	1.9828783	0.0396629	140.84399
2002	East SouthCentral	2.1639033	0.0393123	140.42799
2003	East SouthCentral	1.8959186	0.0397821	147.94299
2004	East SouthCentral	2.0822545	0.0380317	144.16701
2005	East SouthCentral	2.4640082	0.0365871	141.403
2006	East SouthCentral	2.4390863	0.0334566	132.015
2007	East SouthCentral	2.1742501	0.0332259	133.743
2008	East SouthCentral	2.3582539	0.0343388	141.116
2009	East SouthCentral	2.6811087	0.0320874	134.53101
2010	East SouthCentral	2.3328497	0.0340364	144.534
2011	East SouthCentral	2.4236148	0.0314847	134.621
2012	East SouthCentral	2.608196	0.0239678	117.352
2013	East SouthCentral	2.5000616	0.0281366	138.72
2014	East SouthCentral	2.437166	0.0304063	151.01601
2015	East SouthCentral	2.7438782	0.0273782	137.215
2016	East SouthCentral	2.7005001	0.0254851	129.02
2017	East SouthCentral	2.8862721	0.0256305	131.03496
1990	Middle Atlantic	2.1060396	0.0520908	449.73001
1991	Middle Atlantic	2.1411384	0.0508628	459.23599
1992	Middle Atlantic	2.2342084	0.0546482	496.92899
1993	Middle Atlantic	2.169062	0.0555397	497.84
1994	Middle Atlantic	2.5383127	0.0557834	510.62799

year	Census division	C_MARKUP	QPERFLOOR	Q_COM_BCF
1995	Middle Atlantic	2.5209658	0.0569835	531.12598
1996	Middle Atlantic	2.1119403	0.0606107	575.84698
1997	Middle Atlantic	1.7301653	0.06864	653.375
1998	Middle Atlantic	1.8095216	0.0629108	633.11102
1999	Middle Atlantic	1.3707854	0.0677229	689.138
2000	Middle Atlantic	1.7815456	0.066037	692.35498
2001	Middle Atlantic	2.3262254	0.063296	639.28003
2002	Middle Atlantic	1.4295574	0.0652735	664.48798
2003	Middle Atlantic	2.1882314	0.0635569	670.07098
2004	Middle Atlantic	2.2727463	0.0648901	692.47699
2005	Middle Atlantic	1.9622751	0.0566753	610.52002
2006	Middle Atlantic	1.6733639	0.0513387	559.07898
2007	Middle Atlantic	1.6309992	0.0561151	618.05798
2008	Middle Atlantic	1.7954875	0.0557571	620.88501
2009	Middle Atlantic	1.6314168	0.0553237	622.29602
2010	Middle Atlantic	1.8649335	0.0554667	627.24701
2011	Middle Atlantic	1.8159202	0.0565709	642.42798
2012	Middle Atlantic	1.5247842	0.0525677	590.97601
2013	Middle Atlantic	1.8990692	0.0575016	647.76898
2014	Middle Atlantic	1.8993217	0.0628561	709.89001
2015	Middle Atlantic	1.8631231	0.0574608	651.85699
2016	Middle Atlantic	1.8341448	0.0543225	618.80499
2017	Middle Atlantic	1.9224833	0.0548279	628.46906
1990	Mountain	1.1796958	0.0574937	184.573
1991	Mountain	1.2232965	0.058723	196.32001
1992	Mountain	1.1164639	0.0566279	188.735
1993	Mountain	1.2420462	0.0615783	208.241
1994	Mountain	1.2583673	0.0587362	201.879
1995	Mountain	1.5051339	0.0583997	204.16499
1996	Mountain	1.1125651	0.0606239	216.093
1997	Mountain	1.1631508	0.0608757	221.16499
1998	Mountain	1.6786655	0.0575607	215.825
1999	Mountain	1.7174305	0.0538637	208.64101
2000	Mountain	1.1920122	0.0543444	216.40199
2001	Mountain	1.7888476	0.0536108	216.53799
2002	Mountain	1.780769	0.0538575	223.30099
2003	Mountain	1.4160222	0.0514769	216.54201
2004	Mountain	1.489863	0.0510515	220.48801
2005	Mountain	1.5542356	0.0501923	222.39301
2006	Mountain	1.827135	0.0484877	221.567
2007	Mountain	1.6807472	0.0481963	227.5
2008	Mountain	1.4261298	0.0488043	237.71201
2009	Mountain	1.8046458	0.0480187	241.33299

year	Census division	C_MARKUP	QPERFLOOR	Q_COM_BCF
2010	Mountain	1.5854456	0.0460957	235.201
2011	Mountain	1.5349621	0.0469408	241.69099
2012	Mountain	1.7263302	0.0459216	225.006
2013	Mountain	1.4503244	0.0507888	251.05
2014	Mountain	1.6632139	0.0482452	240.62601
2015	Mountain	2.0419882	0.0464796	234.539
2016	Mountain	1.7466736	0.0476196	243.755
2017	Mountain	1.6407867	0.0477442	247.56181
1990	New England	2.5216293	0.0389321	99.921997
1991	New England	2.505014	0.0379501	100.426
1992	New England	2.3247054	0.0445715	117.234
1993	New England	2.0786119	0.0414981	120.755
1994	New England	2.4772011	0.0544149	150.907
1995	New England	2.3908541	0.0494212	147.51401
1996	New England	2.0161421	0.054498	165.84801
1997	New England	2.2945824	0.057227	177.83299
1998	New England	2.1841015	0.0510549	159.08099
1999	New England	2.1379933	0.0439167	142.121
2000	New England	1.6302065	0.0429987	144.69
2001	New England	2.0669719	0.0402669	136.47501
2002	New England	2.0381589	0.0401501	137.42799
2003	New England	2.5327587	0.0387113	133.647
2004	New England	2.8255934	0.0353347	123.537
2005	New England	2.4975503	0.0347402	123.112
2006	New England	2.8478299	0.0311991	112.423
2007	New England	3.0170968	0.0353942	129.218
2008	New England	2.7087138	0.0382953	141.80701
2009	New England	2.5947859	0.0385128	143.922
2010	New England	2.1873288	0.0383071	143.989
2011	New England	2.2464137	0.0421866	159.15401
2012	New England	2.5256068	0.0343416	147.849
2013	New England	2.6982208	0.0428053	185.063
2014	New England	2.9659594	0.0457143	198.47
2015	New England	2.7401603	0.0460751	200.826
2016	New England	2.8263998	0.0443632	194.733
2017	New England	2.6756366	0.044741	197.77422
1990	Pacific	1.9376624	0.0500506	377.78
1991	Pacific	2.1690924	0.0493498	384.491
1992	Pacific	2.0256798	0.0492134	375.59201
1993	Pacific	2.4203045	0.04583	352.254
1994	Pacific	3.1848459	0.0464159	358.91101
1995	Pacific	2.9372289	0.0483282	377.035
1996	Pacific	2.2690053	0.0445834	349.10001

year	Census division	C_MARKUP	QPERFLOOR	Q_COM_BCF
1997	Pacific	2.2154991	0.0460425	362.68799
1998	Pacific	2.5818508	0.0382492	401.91901
1999	Pacific	2.3660772	0.0442665	361.492
2000	Pacific	2.053458	0.0424275	346.81601
2001	Pacific	1.8815636	0.0430176	355.26599
2002	Pacific	2.1521639	0.0403096	336.716
2003	Pacific	1.912608	0.0385585	332.12201
2004	Pacific	1.8093987	0.0380324	332.797
2005	Pacific	1.869813	0.0379257	337.18701
2006	Pacific	2.1956118	0.0389772	352.202
2007	Pacific	2.2525976	0.0396367	364.16199
2008	Pacific	2.264242	0.0392017	366.05801
2009	Pacific	2.3514874	0.0381046	360.867
2010	Pacific	2.1173134	0.0368538	351.65201
2011	Pacific	2.3704763	0.0377259	361.53601
2012	Pacific	2.2651345	0.0271198	364.78601
2013	Pacific	2.1722235	0.0274446	370.58401
2014	Pacific	2.2127067	0.0258408	350.267
2015	Pacific	2.6491866	0.0252881	344.90701
2016	Pacific	2.8727344	0.0252503	347.474
2017	Pacific	2.7207811	0.0254079	352.90063
1990	South Atlantic	1.8999492	0.0230792	246.345
1991	South Atlantic	1.8147944	0.0248233	274.702
1992	South Atlantic	1.8042799	0.0265597	299.57401
1993	South Atlantic	1.9100274	0.0266043	308.36401
1994	South Atlantic	1.9390849	0.025694	307.14499
1995	South Atlantic	1.9467523	0.0258049	316.46899
1996	South Atlantic	1.8097252	0.0262245	334.25101
1997	South Atlantic	2.0603988	0.0251899	326.20801
1998	South Atlantic	2.0284612	0.0249332	325.18701
1999	South Atlantic	1.9973893	0.0202069	320.685
2000	South Atlantic	2.0419032	0.0257924	357.508
2001	South Atlantic	2.2778693	0.0247666	343.22699
2002	South Atlantic	2.216186	0.0257987	355.108
2003	South Atlantic	2.1121199	0.0265753	372.01599
2004	South Atlantic	2.273521	0.026286	377.534
2005	South Atlantic	2.0556418	0.0260092	382.70901
2006	South Atlantic	2.47168	0.023396	353.17599
2007	South Atlantic	2.5119425	0.0236616	367.05099
2008	South Atlantic	2.2929865	0.0236184	375.914
2009	South Atlantic	2.599288	0.0233839	380.909
2010	South Atlantic	2.4825544	0.0239842	396.78101
2011	South Atlantic	2.6365432	0.0224071	374.55399

year	Census division	C_MARKUP	QPERFLOOR	Q_COM_BCF
2012	South Atlantic	2.7003795	0.0200125	359.36099
2013	South Atlantic	2.6202668	0.0221013	399.866
2014	South Atlantic	2.4573951	0.0232332	423.62701
2015	South Atlantic	2.6108377	0.0217632	400.32199
2016	South Atlantic	2.5218769	0.021439	398.54599
2017	South Atlantic	2.5919577	0.0215054	404.77023
1990	West North Central	0.932501	0.0612933	293.74799
1991	West North Central	1.0117042	0.0638051	317.345
1992	West North Central	1.071219	0.0582881	297.14099
1993	West North Central	1.1199411	0.062925	319.284
1994	West North Central	1.1600236	0.0594641	312.06799
1995	West North Central	1.1505925	0.061306	323.51401
1996	West North Central	1.2284949	0.0656062	351.51401
1997	West North Central	1.0545873	0.0590665	312.43701
1998	West North Central	1.2724218	0.0509309	280.36899
1999	West North Central	1.2359991	0.0529542	285.815
2000	West North Central	1.288448	0.0542881	297.478
2001	West North Central	1.2742501	0.0527474	292.74899
2002	West North Central	1.2487592	0.0539173	303.86401
2003	West North Central	1.206473	0.0530219	301.147
2004	West North Central	1.4020858	0.051065	294.505
2005	West North Central	1.3749692	0.0481573	281.94601
2006	West North Central	1.5638026	0.0447968	266.27301
2007	West North Central	1.6196171	0.0469421	283.15601
2008	West North Central	1.5081952	0.0517568	316.83801
2009	West North Central	1.517327	0.0493771	306.50299
2010	West North Central	1.5422755	0.0464527	290.83899
2011	West North Central	1.606853	0.0474024	298.633
2012	West North Central	1.5943613	0.0418249	257.85001
2013	West North Central	1.5492348	0.0525259	325.85699
2014	West North Central	1.361915	0.0554519	346.91299
2015	West North Central	1.7486575	0.0481984	304.09601
2016	West North Central	1.621516	0.0462427	294.57999
2017	West North Central	1.6176083	0.0464904	299.18054
1990	West South Central	1.222199	0.0333404	268.875
1991	West South Central	1.2319799	0.0348103	281.427
1992	West South Central	1.2044469	0.0335838	285.04599
1993	West South Central	1.2439698	0.0322863	278.03101
1994	West South Central	1.3467453	0.0317717	278.513
1995	West South Central	1.1790301	0.0342151	313.021
1996	West South Central	1.1649336	0.0314617	290.99899
1997	West South Central	1.2437555	0.0350777	327.112
1998	West South Central	1.5717055	0.0301957	276.79599

year	Census division	C_MARKUP	QPERFLOOR	Q_COM_BCF
1999	West South Central	1.4589985	0.0305075	272.57999
2000	West South Central	1.2009566	0.0335954	301.42899
2001	West South Central	0.8888666	0.0309096	275.20001
2002	West South Central	1.3294294	0.0374295	335.29001
2003	West South Central	1.4323264	0.0357125	322.19501
2004	West South Central	1.6311371	0.0317105	292.70599
2005	West South Central	1.717926	0.0279002	262.99399
2006	West South Central	1.6951013	0.0252439	243.24899
2007	West South Central	1.4212271	0.0268812	264.707
2008	West South Central	1.5399226	0.0271998	274.69601
2009	West South Central	1.7416343	0.026616	275.37399
2010	West South Central	1.5036384	0.0290677	306.32101
2011	West South Central	1.3716477	0.0279709	298.20901
2012	West South Central	1.6440114	0.0238117	271.67099
2013	West South Central	1.6443893	0.0261693	301.82101
2014	West South Central	1.5258819	0.0277196	323.35901
2015	West South Central	1.7713139	0.0257844	305.259
2016	West South Central	1.7650975	0.0236872	285.56601
2017	West South Central	1.817638	0.023640	290.025787

G-10. Historical delivered end use prices to the industrial sector

Data: Historical industrial sector natural gas prices by Census division, assigned exogenously to variable *HistoricalIndustrialPrice_MESC* in NGMM for the years 1990 through the last historical year in the model. Used as a basis for setting the markup to delivered industrial prices in the projection.

Author: Samantha Calkins, EIA, 2017

Source: Industrial prices by Census region for available years − EIA's Manufacturing Energy Consumption Survey (MECS); industrial prices by state as purchased from a local distribution company − EIA's Natural Gas Monthly (NGM), DOE/EIA-0130; wellhead prices by the 17 onshore supply regions −input data in the Natural Gas Transmission and Distribution Module.

Variables:

$MECSTOT87_{yr,cr}$ = historical industrial natural gas prices from MECS for available year yr and Census region cr (1987$/Mcf)

$SUPPLY87_{yr,cr}$ = historical natural gas wellhead prices averaged to Census division from 17 supply regions using industrial consumption as a weight for available year yr and Census region cr (1987$/Mcf). [Historical wellhead prices are a combination of published EIA wellhead prices (last provided in 2012) and average regional spot prices minus an assumed gathering charge.]

$NGAP87_{yr,cr}$ = historical industrial prices for natural gas purchased from local distribution companies for available year yr and Census region cr (1987$/Mcf).

cr = Census region

yr = available historical years (2002, 2006, 2010, 2014)

Derivation:

While the industrial prices from the NGM only reflect natural gas purchases through local distribution companies (about 15% of the market), prices from MECS represent the majority of the market, although they still do not include the nonmanufacturing portion. However, the MECS data are only available every four years by the four Census regions. The estimated equation is used to fill in the missing data with plausible prices for the industrial sector based on available price data, namely the industrial prices from the NGM and wellhead prices in the region. The estimated equation follows:

$$MECSTOT87_{yr,r} = C + \left(\alpha * \ SUPPLY87_{yr,r}\right) + (\beta * \ NGAP87_{yr,r})$$

The equations were estimated using a basic ordinary least squares approach applied to the data provided below.

Regression diagnostics and parameter estimates:

Dependent Variable: MECSTOT87
Method: Least Squares
Date: 09/05/17 Time: 16:57
Sample: 1 16
Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.112168	0.238582	-0.470142	0.646
$SUPPLY87$	0.647965	0.152617	4.245685	0.001
$NGAP87$	0.450337	0.090602	4.970492	0.0003

R-squared	0.955979	Mean dependent var	3.745
Adjusted R-squared	0.949207	S.D. dependent var	1.066171
S.E. of regression	0.240287	Akaike info criterion	0.153395
Sum squared resid	0.750591	Schwarz criterion	0.298255
Log likelihood	1.772843	Hannan-Quinn criter.	0.160813
F-statistic	141.1572	Durbin-Watson stat	2.219504
Prob(F-statistic)	0		

Data:

year	region	MECSTOT87	SUPPLY87	NGAP87
2002	Midwest	3	1.95	3.59
2006	Midwest	5.15	4.03	6.06
2010	Midwest	3.38	2.4	3.99
2014	Midwest	3.17	2.67	3.98
2002	Northeast	3.51	2.64	4.33
2006	Northeast	6.42	4.19	7.55
2010	Northeast	4.25	2.78	5.35
2014	Northeast	3.6	2.79	5.41
2002	South	2.64	2.28	2.83
2006	South	4.94	4.2	5.08
2010	South	3.05	2.39	3.23
2014	South	2.74	2.24	3.02
2002	West	2.84	2.08	3.56
2006	West	4.77	3.73	6.07
2010	West	3.35	2.31	4.12
2014	West	3.11	2.35	4.08