

Shared Freight Transportation and Energy Commodities Phase One: Coal, Crude Petroleum, & Natural Gas Flows



Shih-Miao Chin, Ph.D.
Hyeonsup Lim
Ho-Ling Hwang, Ph.D.
Diane Davidson

**Approved for public release
Distribution is unlimited**

July 2016

DOCUMENT AVAILABILITY

Reports produced after January 1, 1996, are generally available free via US Department of Energy (DOE) SciTech Connect.

Website <http://www.osti.gov/scitech/>

Reports produced before January 1, 1996, may be purchased by members of the public from the following source:

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone 703-605-6000 (1-800-553-6847)
TDD 703-487-4639
Fax 703-605-6900
E-mail info@ntis.gov
Website <http://www.ntis.gov/help/ordermethods.aspx>

Reports are available to DOE employees, DOE contractors, Energy Technology Data Exchange representatives, and International Nuclear Information System representatives from the following source:

Office of Scientific and Technical Information
PO Box 62
Oak Ridge, TN 37831
Telephone 865-576-8401
Fax 865-576-5728
E-mail reports@osti.gov
Website <http://www.osti.gov/contact.html>

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Energy and Transportation Science Division
Center for Transportation Analysis

**SHARED FREIGHT TRANSPORTATION AND ENERGY COMMODITIES
PHASE ONE: COAL, CRUDE PETROLEUM, & NATURAL GAS FLOWS**

Shih-Miao Chin, Ph.D.
Hyeonsup Lim
Ho-Ling Hwang, Ph.D.
Diane Davidson

Date: July 6, 2016

Prepared for
US Department of Energy

Prepared by
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, TN 37831-6283
managed by
UT-BATTELLE, LLC
for the
US DEPARTMENT OF ENERGY
under contract DE-AC05-00OR22725

CONTENTS

1.	INTRODUCTION.....	3
1.1	BACKGROUND.....	3
1.2	OBJECTIVE.....	3
1.3	ORGANIZATION OF THE REPORT	4
2.	COAL MOVEMENTS.....	5
2.1	DATA SOURCES.....	5
2.1.1	Energy Information Administration Data.....	5
2.1.2	American Coke and Coal Chemicals Institute	7
2.1.3	Carload Waybill Sample from Surface Transportation Board	8
2.2	ESTIMATION PROCEDURES FOR COAL.....	8
2.3	Estimating Domestic Movements of Coal.....	8
2.3.1	Imports/Exports of Coal	9
3.	NATURAL GAS MOVEMENTS.....	11
3.1	DATA SOURCES	11
3.1.1	EIA Data on Natural Gas.....	11
3.1.2	Census Data and Others	12
3.2	ESTIMATION PROCESS FOR NATURAL GAS	13
3.2.1	Estimating Domestic Movements of Natural Gas	13
3.2.2	Imports/Exports of Natural Gas	13
4.	CRUDE PETROLEUM MOVEMENTS	15
4.1	DATA SOURCES.....	15
4.1.1	EIA Petroleum Data	15
4.1.2	Census County Business Patterns.....	17
4.1.3	Carload Waybill Samples	17
4.1.4	DrillingEdge	17
4.2	ESTIMATION PROCEDURES FOR CRUDE PETROLEUM	18
4.2.1	Estimating Domestic Movements of Crude Petroleum.....	18
4.2.2	Imports of Crude Petroleum.....	18
5.	FINAL REMARKS.....	19
5.1	DATA VALDATION	19
5.2	ITERATIVE PROPORTIONAL FITTING	21

1. INTRODUCTION

1.1 BACKGROUND

The Freight Analysis Framework (FAF) integrates data from a variety of sources to create a comprehensive picture of nationwide freight movements among states and major metropolitan areas for all modes of transportation. It provides a national picture of current freight flows to, from, and within the United States, assigns selected flows to the transportation network, and projects freight flow patterns into the future. The latest release of FAF is known as FAF4 with a base year of 2012. The FAF4 origin-destination-commodity-mode (ODCM) matrix is provided at national, state, major metropolitan areas, and major gateways with significant freight activities (e.g., El Paso, Texas).

The U.S. Department of Energy (DOE) is interested in using FAF4 database for its strategic planning and policy analysis, particularly in association with the transportation of energy commodities. However, the geographic specification that DOE requires is a county-level ODCM matrix. Unfortunately, the geographic regions in the FAF4 database were not available at the DOE desired detail. Due to this limitation, DOE tasked Oak Ridge National Laboratory (ORNL) to assist in generating estimates of county-level flows for selected energy commodities by mode of transportation. The energy commodities of specific interest, as defined by the Standard Classification of Transported Goods (SCTG) in the FAF4, include:

- Coal (SCTG 15)
- Crude petroleum (SCTG 16)
- Fuel - gasoline, aviation, and ethanol (SCTG 17)
- Fuel oil - diesel, bunker C and biodiesel (SCTG 18)
- Natural gas (SCTG 19)

Due to resource considerations, DOE and ORNL agreed to a two-phase approach to conduct this research effort. County-level flows under the first phase (Phase-I) focused on shipments of three commodities: coal, crude petroleum (crude oil), and natural gas. The estimations of county-level flows of petroleum products (fuel and fuel oil) will be conducted in the Phase-II study. Note that the base year for this study is 2012, so that it is consistent with FAF4.

1.2 OBJECTIVE

In this phase, ORNL identified data sources and developed estimation methods to produce comprehensive county-to-county flows of the targeted energy commodities. All shipments of the selected energy commodities moved in the U.S. are covered in this phase. That is, flows of domestic shipments and foreign trades (imports/exports) moved within the U.S. were included. Furthermore, modes of transportation, such as truck, rail, water, multiple mode, and pipeline are specified in these county-to-county flows.

1.3 ORGANIZATION OF THE REPORT

This report documents the data sources and the estimation methods developed for generating county-level origin-destination (OD) flows for Phase-I selected commodities, i.e., coal, natural gas, and crude petroleum. Discussions associated with each of the commodities examined are presented in Sections 2, 3, and 4, respectively. Each Section details the data sources identified as well as the estimation procedures developed and applied in estimating county-level OD flows of the given energy commodity. The resulting OD matrix is then briefly described in Section 5 of this report with a brief explanation of the Iterative Proportional Fitting Method (IPF) included as well.

2. COAL MOVEMENTS

2.1 DATA SOURCES

2.1.1 Energy Information Administration Data

The U.S. Energy Information Administration (EIA) collects and publishes a broad range of statistics and information on coal production, consumption, and other related matters. The EIA provides state-to-state coal distribution based on data collected from surveys¹ (e.g., EIA-923, EIA-5, EIA-3, and EIA-7A). The EIA data contains coal distribution by origin state, consumer type, and transportation mode. This state-level coal distribution data is used as a control total in building the county-level flows. Specific EIA datasets identified for this study are briefly discussed below.

Annual Coal Report

The *Annual Coal Report*² provides annual data on U.S. coal production, number of mines, capacity, consumption, etc., by coal-producing state and mine type (underground or surface). In some cases, county-level statistics are also available. Based on this EIA data, county-level productions (in short tons) were ranked, and the top ten counties are listed in Table 2-1. These top ten counties accounted for over 50% of the total U.S. coal production in 2012.

Table 2-1. Top Ten Coal Production Counties in 2012

Coal Production County	Short Tons	Percentage
Campbell, Wyoming	354,059,807	34.8%
Converse, Wyoming	34,316,314	3.4%
Greene, Pennsylvania (Bituminous)	25,673,405	2.5%
Bighorn, Montana	22,671,797	2.2%
Boone, West Virginia (Southern)	16,467,029	1.6%
Mercer, North Dakota	15,289,024	1.5%
Logan, West Virginia (Southern)	15,114,224	1.5%
Marshall, West Virginia (Northern)	14,719,674	1.4%
Union, Kentucky (West)	13,474,902	1.3%
Saline, Illinois	13,453,213	1.3%
Rest of Counties	491,219,029	48.3%
Total U.S.	1,016,458,418	100.0%

Based on this data, there were 1,509 coalmines and 227 counties that produced coal in 2012.

¹ State level coal flow by survey: Form EIA-923, "Power Plant Operations Report"; Form EIA-5, "Quarterly Coal Consumption and Quality Report - Coke Plants"; Form EIA-3, "Quarterly Coal Consumption and Quality Report, Manufacturing and Transformation/Processing Coal Plants and Commercial and Institutional Coal Users" and Form EIA-7A, "Coal Production and Preparation Report"

² *Annual Coal Report*, U.S. EIA, available at: <http://www.eia.gov/coal/annual/>; data is also available from <http://www.eia.gov/coal/data.cfm#production>.

Annual Coal Distribution Report

The EIA also provides county-level consumption data by consumer sectors, which include the electric power sector, coke plants, industrial plants, commercial and institutional sectors. The EIA’s *Annual Coal Distribution Report*, specifically, provides domestic coal distribution by origin state, destination state, consumer type, and mode of transportation (rail, water, truck). It also provides a summary of foreign coal distribution by coal-producing state. Based on the EIA published data, about 93% of the total coal distributed in 2012 was consumed by the electric power sector³.

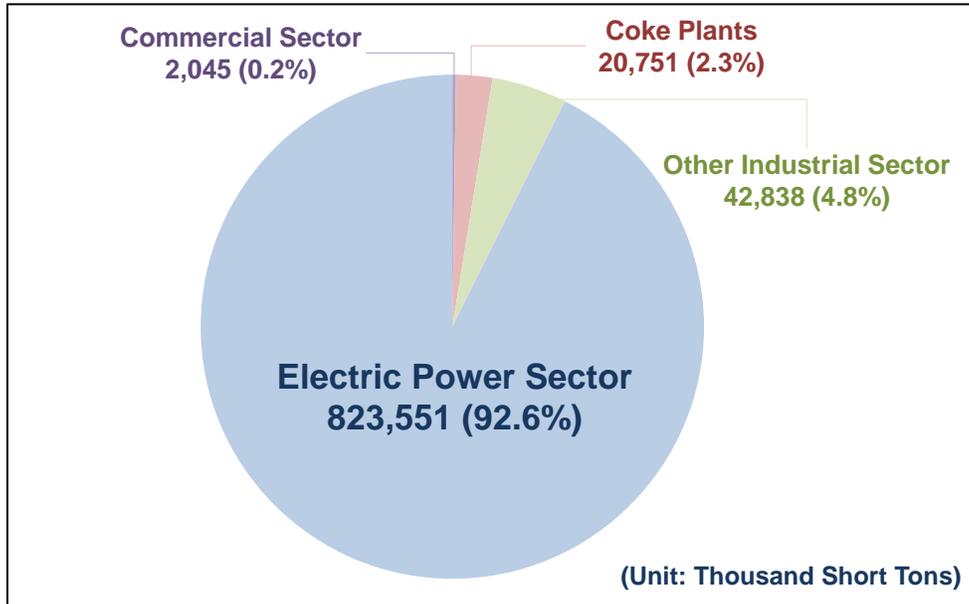


Figure 2-1. 2012 Coal consumption by consumer type based on EIA data.

U.S. Power Plant Operations Report (EIA-923)

The survey Form EIA-923⁴ collects detailed information on electric power plants in the United States on a monthly and annual basis. It includes data on electricity generation, fuel consumption, fossil fuel stocks, and receipts at the power plant and prime mover level. Over 1,900 electric power plants are included in the monthly survey and another 4,100 electric power plants are covered in the annual survey. Using the fuel receipts and costs data from EIA-923 annual data, the quantity of coal moved from a coalmine to a power plant by its primary transportation mode can be obtained.

U.S. Coal Imports by Country of Origin

The data associated with U.S. coal imports are compiled by the EIA Office of Oil, Gas, and Coal Supply Statistics and are based on the information from the “Monthly Report IM 145” of the U.S. Census. This data includes both stream coal and metallurgical coal. The available

³ EIA Beta, *Coal Data Browser, Total Consumption: Electric Power 2012*. Website [accessed in May 2016](http://www.eia.gov/beta/coal/data/browser/). <http://www.eia.gov/beta/coal/data/browser/>.

⁴ EIA 923 information can be obtained from: <https://www.eia.gov/electricity/data/eia923/>

information⁵ includes total tonnage and price per ton by year, country of coal origin, and customs district for 2002-2015. According to the EIA, in 2012, coal was imported from 15 countries through 28 customs districts.

U.S. Coal Exports by Country of Destination

Data associated with U.S. coal exports are also compiled by the EIA Office of Oil, Gas, and Coal Supply Statistics in the EIA, using the information obtained from Census “Monthly Report 545.” The available information includes total tonnage and price per ton by year, country of coal destination, and customs district for 2002-2015. In 2012, coal was exported to 88 countries through 33 customs districts.

2.1.2 American Coke and Coal Chemicals Institute

The American Coke and Coal Chemicals Institute (ACCCI) is an organization with members representing over 95% of the metallurgical coke produced in the United States and Canada. A list of operating U.S. coke plants (as of January 2012) was obtained from the ACCCI website⁶, which contains information on company names and city/state plant locations. However, no other attributes regarding these plants are provided. Figure 2-2 displays the locations of the 19 coke plants based on information from the ACCCI website.



Figure 2-2. Coke plants located in the U.S. based on ACCCI January 2012 data.

⁵ EIA data available at: http://www.eia.gov/coal/archive/coal_historical_imports.xls

⁶ U.S. & Canadian Coke Plants, as of January 2012, see: http://accii.org/documents/CokePlantListing_010112.pdf.

2.1.3 Carload Waybill Sample from Surface Transportation Board

The Surface Transportation Board (STB) publishes and maintains the Carload Waybill Sample⁷ database, which is a stratified sample of carload waybills for all U.S. rail traffic submitted by rail carriers terminating 4,500 or more revenue carloads annually. In order to protect sensitive shipping and revenue information of rail companies reporting the Waybill, STB provides a public-use version of the Waybill data that contains aggregated data, in addition to the more detailed limited-access confidential version. The 2012 *Carload Waybill Sample* database captured over 623,000 waybills that were submitted electronically in 2012.

2.2 ESTIMATION PROCEDURES FOR COAL

2.3 ESTIMATING DOMESTIC MOVEMENTS OF COAL

Step 1. Building State-to-State Coal Flow by Mode and Consumer Type

The data from the Annual Coal Distribution Report is used to build state-to-state coal movements. These state-level flows for each mode and consumer type are utilized as constraints in the subsequent steps to generate county level flows. The resulting county-level coal flows must sum up to the corresponding state-to-state flows.

Step 2. County Level Production

Total production from each coal producing location (e.g., coalmine) within each county is summed to obtain the total coal production for the given county.

Step 3. County Level Consumption

Similarly, total coal consumption for each receiving location (e.g., coke plant and electric power plant) within each county is summed to obtain the total coal consumption for a given county. For the locations where the total consumption information is not directly available, the associated attributes, such as capacity of coal in plants, are used to estimate the coal consumption for the county.

Step 4. Assigning EIA-923 Data

The county-level flows obtained from the EIA-923 data are assigned directly to associated county pairs in the OD matrix. That is, data obtained from the EIA-923 is assumed as actual coal flows for these county-level OD pairs. These “known flows” account for 91% of total domestic coal flows.

⁷ For access of Carload Waybill Sample data visit Surface Transportation Board website at: http://www.stb.dot.gov/stb/industry/econ_waybill.html, or access the 2012 documentation at: <http://www.stb.dot.gov/stb/docs/Waybill/2012%20STB%20Waybill%20Reference%20Guide%20-%20FINAL.pdf>.

If the total volume between a given state-OD-pair (from Step 1) is the same as the aggregated volumes from their “known” county OD pairs, no further processing is necessary for this state OD-pair. Otherwise, the IPF estimation procedure (next Step) is required in order to estimate the “missing cells” in the county-level OD matrix.

Step 5. IPF Estimation for the Coal Flow among Missing County-Pairs

The IPF approach is used to estimate county-level flows for the rest of the state combinations. Note that coal flows are estimated by consumer type and by transportation mode. More details of the IPF procedure are explained in Section 5.2.

2.3.1 Imports/Exports of Coal

Step 1. Imports/Exports by Ports

For the imports, the information of origin countries and ports of entry are given. Thus, the subsequent processes are implemented to estimate the true destination from the port. Likewise, for the exports, subsequent steps are required to estimate the true origin counties before the coal is transported to the port of exit.

Step 2. Distances from the Port to Possible Origin/Destination Counties

The second step is to calculate the distances from the ports to possible origin/destination counties. In the imports/exports estimation, the county-to-county distance matrix is used only when there is no domestic flow from/to the ports.

Step 3. Estimation of True Origin/Destination Counties

If there is domestic flow from the county of port, then true origin/destination counties are estimated based on the amount of the flow from the port to each possible destination county (for exports, from the each possible origin county to the port).

If there is no domestic flow from the county of port, then the true origin/destination counties are selected based on the distances from the port to possible origin/destination counties.

Step 4. Estimation of Imports/Exports to the Estimated True Origin/Destination Counties

The final step is to calculate the proportions of domestic flows of the estimated true origin/destination counties. Based on this ratio, the coal imports/exports of each port will be re-distributed to each true origin/destination county.

3. NATURAL GAS MOVEMENTS

3.1 DATA SOURCES

There are numerous data sources from two organizations that were used to create the flows of natural gas products. Data was collected from the EIA and from the U.S. Census Bureau (Census).

3.1.1 EIA Data on Natural Gas

Just like the estimation of coal movements, the estimation of natural gas flows relies mainly on EIA-published data. Other data that provides supplemental information is also utilized in the flow estimation process. Below is a brief discussion of the data sources identified for natural gas flow estimation uses.

Natural Gas Annual

This publication, which is released by the EIA annually, provides information on the supply and disposition of natural gas. The data presented in the *Natural Gas Annual* are primarily taken from surveys conducted by the EIA, particularly from the “Annual Report of Natural and Supplemental Gas Supply and Disposition” (Form EIA-176) and the “Monthly Natural Gas Production Report” (Form EIA-914).

For this study, specific information from Table 12⁸ of the *Natural Gas Annual* report, entitled “Interstate movements and movements across U.S. borders of natural gas by state, 2012,” are used. In addition, the information on dry production and withdrawals from underground storage by individual states is also available from this report. This information, along with the disposition by state data, is used to derive control totals for intrastate natural gas movements.

Natural Gas Consumption by End Use

This EIA data series provides annual statistics⁹ on total “volumes delivered to consumers” by state for end-use sectors of residential, commercial, industrial, vehicle fuel, and electric power plants..

U.S. Natural Gas Imports by Point of Entry

The major data source used in estimating imported natural gas flows is the *U.S. Natural Gas Imports by Point of Entry*, which can be obtained from the EIA website¹⁰ or from the Table 9 “Summary of U.S. natural gas imports by point of entry, 2008 – 2012” of the *Natural Gas Annual* report published by the EIA. This data series provides volume (in million cubic feet) and price for natural gas and liquefied natural gas (LNG) transported by pipeline. In 2012, 94.4% of

⁸ *Natural Gas Annual*, see: http://www.eia.gov/naturalgas/annual/archive/2012/pdf/table_012.pdf.

⁹ *Natural Gas Consumption by End Use*, at: http://www.eia.gov/dnav/ng/ng_cons_sum_a_EPG0_VC0_mmcf_a.htm.

¹⁰ *U.S. Natural Gas Imports by Point of Entry*,
https://www.eia.gov/dnav/ng/ng_move_poe1_a_EPG0_IRP_Mmcf_a.htm

the U.S. natural gas imports were from Canada and Mexico by pipeline; the imports from other countries were transported by LNG tankers.

U.S. Natural Gas Exports by Point of Exit

Similarly, a data source for estimating exported natural gas flows are obtained from the U.S. Natural Gas Exports and Re-Exports by Point of Exit website¹¹, as well as from Table 11 of the *Natural Gas Annual* report, entitled “Summary of U.S. Natural Gas Exports by Point of Exit, 2008-2012.”

U.S. Natural Gas Imports & Exports by State

In addition to imports by point of entry, the EIA also reports data on natural gas imports/exports by state¹², which provides total imported/exported volumes (in million cubic feet) and prices (dollars per thousand cubic feet) of natural gas at the state level. Total volumes by state from both tables (one by point of entry or exit and another by state) are examined to determine destinations of imported or origin of exports on natural gas flows. According to the EIA, natural gas is transported by pipeline domestically, and a “processing plant” is the place where pipeline-quality natural gas is produced.

Database on Natural Gas Receipt/Delivery Points

A database that contains natural gas processing plant locations is also used to determine origins and destinations of natural gas shipments. This database contains geographic information representing locations on interstate natural gas pipelines where natural gas gathering systems connect (receipt) to the pipeline, or where natural gas local distribution systems and other end-users (delivery) connect to the pipeline within North America. These receipt points are used to identify origins of domestic regions for exported flows.

3.1.2 Census Data and Others

Several Census data, such as population data and County Business Patterns (CBP) data are used with natural gas data obtained from the EIA. These supplemental data are mainly utilized in disaggregation models to distribute estimates from an aggregated geography level (e.g., state) to the desired county-level. Other data sources considered in the disaggregation process include vehicle population and data on electric generating units.

¹¹ *U.S. Natural Gas Exports and Re-exports by Point of Exit*, http://www.eia.gov/dnav/ng/ng_move_poe2_a_EPG0_ENP_Mmcf_a.htm.

¹² Statistics by state are available at http://www.eia.gov/dnav/ng/ng_move_state_a_EPG0_IM0_Mmcf_a.htm.

3.2 ESTIMATION PROCESS FOR NATURAL GAS

3.2.1 Estimating Domestic Movements of Natural Gas

Step 1. Building State-to-State Flows

Since the data sources for interstate and intrastate flows are different, they must be summed together to obtain the complete state-to-state flow for natural gas. Because all domestic flows of natural gas is transported by pipeline, only one matrix of state-level flow is needed.

Step 2. County Level Production

The state-level production information provided by the EIA is disaggregated into county level proportionally to the daily scheduled capacity of natural gas production from the database on natural gas receipt/delivery points.

Step 3. County Level Consumption

The natural gas consumption provided by the EIA is state level with consumer type information. For each consumer type, the CBP data and populations of each county are used to disaggregate the natural gas consumption by state into county level. For example, the total amount of natural gas consumed for industrial businesses within each state are distributed proportionally to each county's payroll of industrial businesses to estimate the county level natural gas consumption for industrial type.

Step 4. IPF Estimation

Unlike the domestic coal flow estimation, the "known flows," which can be directly assigned to county-level OD matrix, are not available in the natural gas flow estimation. Thus, the IPF approach is used to estimate county-level natural gas flows for all of the state OD pairs.

3.2.2 Imports/Exports of Natural Gas

The estimation process for imports/exports of natural gas is the same as that of coal flow except for the data sources. The U.S. natural gas imports and exports from/to each port are assigned to the county level, and then the true origin/destination counties are estimated based on the domestic flow from/to the county of port if it is known. In cases where there is no domestic flow from/to the county of port, the county-to-county distance matrix and inbound/outbound domestic natural gas flows of each county are used to estimate the flows of true origin/destination counties.

4. CRUDE PETROLEUM MOVEMENTS

4.1 DATA SOURCES

Basic information on crude petroleum production, imports, exports, and related activities at refineries can also be obtained from data collected and published by the EIA. Note that the geographic region of the EIA petroleum data is typically available at the Petroleum Administration for Defense Districts (PADD) level. Figure 4-1 displays a map of the PADDs as defined by the EIA.

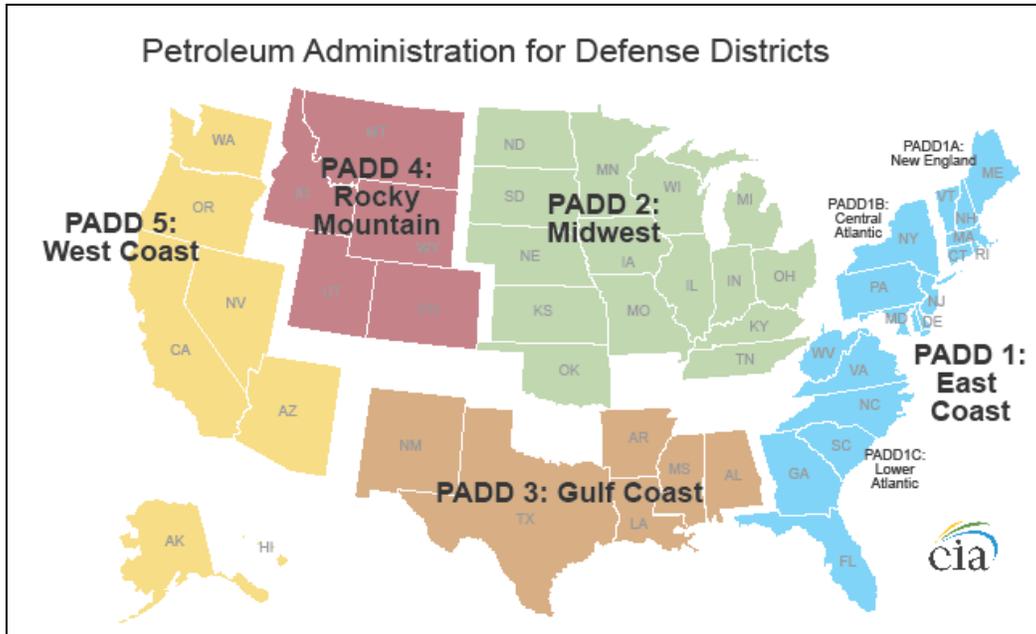


Figure 4-1. Definition of Petroleum Administration for Defense Districts (PADDs).

4.1.1 EIA Petroleum Data

Several sets of EIA-published statistics on crude petroleum movements (including domestic, imports, and exports), as well as other data sources, are briefly discussed below.

Movement of Crude Petroleum by Rail

This table contains monthly and annual crude petroleum movements specifically by rail between regions¹³. This crude petroleum by rail table provides detailed movements among PADDs, and trade between PADDs and Canada. It provides control totals on crude petroleum by rail for domestic (intra-U.S.), U.S. exports to Canada, and U.S. imports from Canada. The EIA also provides the crude petroleum movements by other modes, water (tanker and barge) and pipeline, but they do not have information on imports and exports from/to Canada.

¹³ Table "Movements of Crude Oil by Rail", Petroleum & Other Liquids, U.S. EIA, at: http://www.eia.gov/dnav/pet/PET_MOVE_RAILNA_A_EPC0_RAIL_MBBL_M.htm.

Company-Level imports

This database contains monthly statistics regarding imports of crude petroleum and petroleum products at the company level. Specific information provided in this database includes importing company name and country, product name, port of entry location (city, state, PADD), import quantity (in thousand barrels), and other information. This database can be obtained in Microsoft Excel spreadsheet format¹⁴ from the EIA website.

Crude Petroleum Production

This table contains the annual production of crude petroleum, in thousands of barrels, by state and PADD¹⁵. The statistics published in this table are based on estimates reported from states, and summed to PADD and the U.S. level. It is noted that state production estimates reported by the EIA were normally different from those reported by state agencies. One explanation for this difference can be found in *Today in Energy*, released on July 10, 2015¹⁶. That article pointed out that the reason for this difference was that EIA's estimates accounted for expected revisions to data collected by the states.

Exports by Destination

[In 2012, the only exported crude petroleum from the U.S. was a total of approximately 24.7 million barrels to Canada and about 5 thousand barrels to Mexico¹⁷. U.S. crude petroleum exports have been restricted to](#) (1) crude petroleum derived from fields under the State waters of Alaska's Cook Inlet; (2) Alaskan North Slope crude petroleum; (3) certain domestically produced crude petroleum destined for Canada; (4) shipments to U.S. territories; and (5) California crude petroleum to Pacific Rim countries. [Recently, the U.S. Congress has approved a measure to repeal the 40-year ban on oil exports, allowing energy companies to export U.S. crude petroleum.](#)

Exports by PADD

[In addition to the country-level total, annual volumes of crude petroleum \(in thousand barrels\) exported from each PADD¹⁸ is also available from the EIA. This provides control totals for exported crude petroleum by PADD.](#)

¹⁴ Data on "Company Level Imports" can be downloaded from the EIA website at:

http://www.eia.gov/oil_gas/petroleum/data_publications/company_level_imports/cli.html.

¹⁵ Annual crude oil production statistics is available from the EIA website at:

http://www.eia.gov/dnav/pet/pet_crd_crdpn_adc_mbbbl_a.htm.

¹⁶ "EIA's estimates for state crude oil production account for incomplete, lagged data," *Today in Energy*, EIA,

<http://www.eia.gov/todayinenergy/detail.cfm?id=22012>

¹⁷ Table "Exports by Destination", Petroleum & Other Liquids, U.S. EIA, at:

http://www.eia.gov/dnav/pet/pet_move_expc_a_EP00_EEX_mbbbl_m.htm.

¹⁸ Table on crude oil "Exports", Petroleum & Other Liquids, see the EIA website:

http://www.eia.gov/dnav/pet/pet_move_exp_dc_NUS-Z00_mbbbl_m.htm.

Movements by Mode between PADDs

Annual volumes of crude petroleum movements (measured in thousand barrels) by transportation mode (including pipeline, tanker, barge, and rail) between PADDs are published by the EIA¹⁹. Statistics published in this EIA table are based on information collected from state reported EIA forms, specifically the EIA-813 (*Monthly Crude Oil (Petroleum) Report*).

Refinery Net Input

The EIA also publishes annual data on refinery net input for crude petroleum by PADDs and refining regions (sub-PADD level)²⁰. These statistics represent total crude petroleum (domestic plus foreign) input to crude petroleum distillation units and other refinery processing units.

4.1.2 Census County Business Patterns

The Census published CBP²¹ is an annual series that provides county-level economic data by industry. It contains information on the number of establishments, employment size, and annual payroll. Employment payroll statistics from the CBP can be utilized to distributing state-level flow estimates to the county level.

4.1.3 Carload Waybill Samples

As in the estimation of coal flows, *Carload Waybill Sample* data is also used mainly in determining mode shares for transporting crude petroleum and for calibration of crude petroleum flows.

4.1.4 DrillingEdge

DrillingEdge provides some of the oil and gas data from respective state and local governments that were beneficial to this study. The county-level production information can be obtained directly from the *DrillingEdge* website for 21 States²².

¹⁹ Table on *Movements by Tanker, Pipeline, Barge, and Rail between PAD Districts*, can be found at EIA website: http://www.eia.gov/dnav/pet/pet_move_ptb_a_EPC0_TNR_mbb1_a.htm.

²⁰ Refinery Net Input, Petroleum & Other Liquids, see EIA website: http://www.eia.gov/dnav/pet/pet_pnp_inpt2_a_epc0_YIY_mbb1_a.htm.

²¹ County Business Patterns, U.S. Census, see website at: <http://www.census.gov/econ/cbp/>.

²² The 21 states which county level crude oil production is available: Alabama, Alaska, Arkansas, California, Colorado, Kansas, Louisiana, Michigan, Mississippi, Montana, Nebraska, New Mexico, New York, North Dakota, Ohio, Pennsylvania, South Dakota, and Texas
<http://www.drillingedge.com/>

4.2 ESTIMATION PROCEDURES FOR CRUDE PETROLEUM

4.2.1 Estimating Domestic Movements of Crude Petroleum

Step 1. Building PADD-to-PADD Crude Petroleum Flow by Mode

Unlike the coal and natural gas flow estimation, the domestic flow information of crude petroleum provided by the EIA is at PADD level, not state. Since the crude petroleum flows have different OD patterns by transportation mode, the PADD-to-PADD crude petroleum flows are constructed in different matrices by each transportation mode.

The data of movements by mode between PADDs, provided by the EIA, does not include the movements within the PADDs (intra-PADD movements). The intra-PADD movements are estimated by subtracting the inter-PADD movements from the crude petroleum amounts of refinery receipts for each transportation mode.

Step 2. County Level Production

Based on the EIA's crude petroleum production data, there are 31 states that produced crude petroleum in 2012. Among the 31 states, 21 states have county-level production information available from the *DrillingEdge* website. For the rest of the counties, the CBP data is used to estimate county-level crude petroleum production from the EIA's state-level production. Only the counties that have oil wells are included as producing counties.

Step 3. County Level Consumption (Refinery Input)

The refinery input data provided by the EIA is at the state level. To get the county-level consumption, the state-level consumption is distributed proportionally to the operation capacity of each crude petroleum refinery and summed to the county level.

Step 4. IPF Estimation

Like the natural gas flow estimation, the IPF approach is used to estimate county-level natural gas flows for all of the state OD pairs.

4.2.2 Imports of Crude Petroleum

The company-level imports data has 12,122 records of crude petroleum imports with the location information for processing companies. With this information, the destination counties from the ports can be obtained by referencing the locations of companies to county level.

5. FINAL REMARKS

5.1 DATA VALDATION

Note that data quality checks and benchmarking with published data were performed throughout the estimation processes in this research effort. The resulting OD matrix developed in this Phase-I study contains flow information at the county level for coal, natural gas, and crude petroleum. Specifically, the resulting database includes the following data elements:

- Origin county FIPS code
- Destination county FIPS code
- Trade type
- Foreign zone of origin/destination country
- Primary transportation mode
- Weight/value of commodity

Table 5-1 lists the variables contained in the output database, including field names and brief descriptions of their contents. The foreign countries involved in imports/exports of these commodities are aggregated into eight zones, according to those defined for FAF4, which are listed in Table 5-2. Furthermore, the transportation modes are categorized into eight types and listed in Table 5-3.

Table 5-1. Field Descriptions of Output File

Field Name	Description
Trade_Type	1 = Domestic, 2 = Imports, 3 = Exports
Dms_Orig	Origin county FIPS code
Dms_Dst	Destination county FIPS code
Fr_Area	Foreign zone of origin/destination country for imports and exports
Dms_Mode	Primary transportation mode
Fr_Mode	Transportation mode from origin country to port of entry (from port of exit to destination country, for exports)
Short Tons	Weight of commodity in short tons
Thousand Dollars	Value of commodity in thousand dollars

Table 5-2. Foreign Zone of Output File

Foreign Zone	Description
801	Canada
802	Mexico
803	Rest of Americas
804	Europe
805	Africa
806	SW & Central Asia
807	Eastern Asia
808	SE Asia & Oceania

Table 5-3. Transportation Mode

Transportation mode	Description
1	Truck
2	Rail
3	Water
4	Air
5	Multiple Mode
6	Pipeline
7	Other and Unknown
8	No domestic mode (for imports only)

As an example, using the estimates generated from this research effort, Figure 5-1 presents the coal flows originating from Campbell County, WY, by mode (truck, rail, and multiple modes). The coal flows of each transportation mode are displayed in different colors and the thickness of line reflects the volume of the flow.

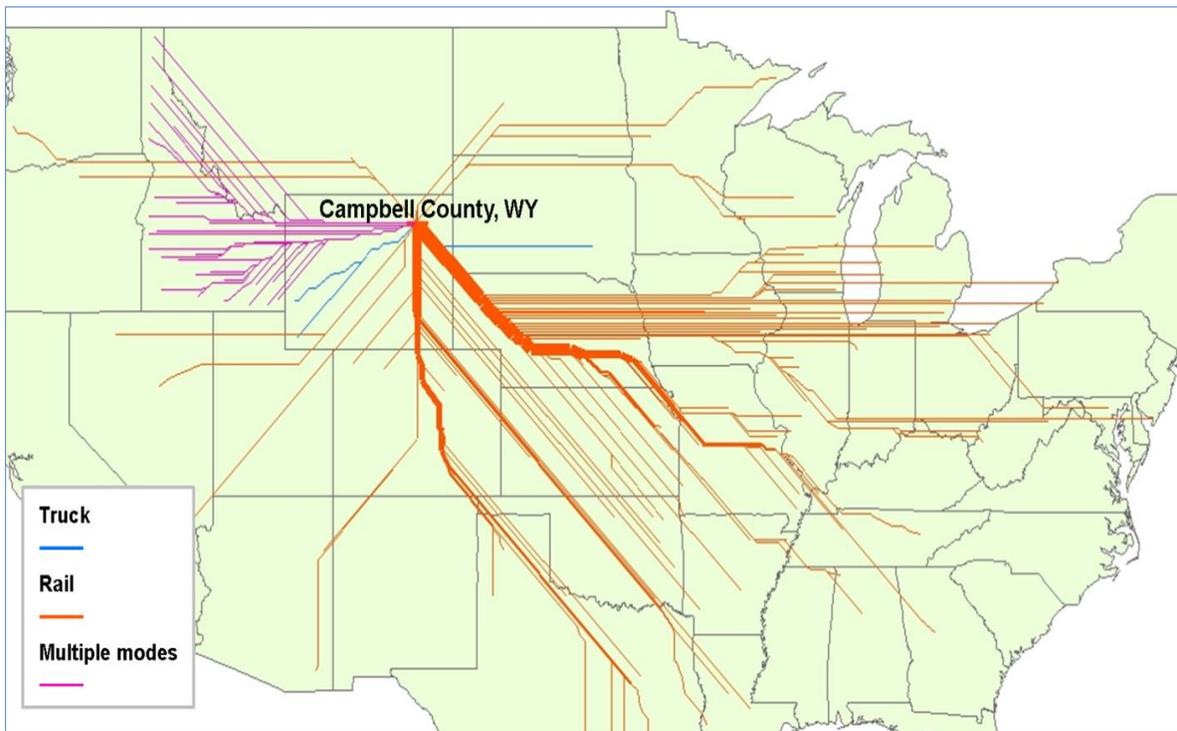


Figure 5-1. Example of county-level coal flow by mode.

The coal flow from Campbell County accounted for about 40% of total domestic coal flow in 2012. Based on the estimated results, there are 400 destination counties that received coal from Campbell County, with 98% transported by rail. Among the 400 destination counties, Monroe County, GA—which has one of the nation’s largest power plants, Plant Scherer—is the county that receives the largest amount of coal from Campbell County.

5.2 ITERATIVE PROPORTIONAL FITTING

The IPF is a procedure for adjusting a matrix of data cells, such that they sum up to selected totals for both the columns and rows of the matrix. The unadjusted data cells may be referred to as the ‘seed’ cells and the selected totals may be referred to as the ‘marginal’ totals. In this project, the ‘seed’ cells are county-to-county distances, and the ‘marginal’ totals are production and consumption totals. In these notations, this process may be referred to as the gravity model, which is one of the most popular methods used for trip distributions in transportation planning. Mathematically, the model can be expressed as:

$$T_{ij} = \frac{K_i K_j T_i T_j}{f(C_{ij})}$$

where,

T_{ij} is flows from origin i to j

T_i is flows originating at i and T_j is flows destined for j

C_{ij} is distance from origin i to j

K_i, K_j is balancing factors solved iteratively in IPF

In this study, IPF is used as the last procedure to fill in the “unknown” cells of the matrix during the estimation of energy commodity flows. Using coal flow as the example, Figure 5-2 displays how IPF was used in the estimation processes. As seen in Figure 5-2, many of the cells in the matrix can be filled using information provided from the EIA-923 data (box 1 on top left). These state-level data are then converted into county-level coal flow information for available counties (boxes 2 and 3). The IPF process then uses the information of county-level productions and consumptions, with the state-to-state flows as control totals, to estimate the coal flows for the remaining cells (box 4 on bottom right).



Figure 5-2. Iterative Proportional Fitting (IPF) process flow diagrams.