The Freight Analysis Framework, Version 3: Overview of the FAF^3 National Freight Flow Tables

PREPARED FOR:
FEDERAL HIGHWAY ADMINISTRATION
Office of Freight Management and Operations
Federal Highway Administration
U.S. Department of Transportation
1200 New Jersey Ave, SE,
Washington, DC 20590

October 28, 2010

PREPARED BY
Frank Southworth
Diane Davidson
Ho-Ling Hwang
Bruce E. Peterson
and
Shih-Miao Chin
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37831-6285
managed by
UT-BATTELLE, LLC
for the
U.S. DEPARTMENT OF ENERGY
Under Contract No. DE-AC05-00OR22725

TABLE OF CONTENTS

1. Introduction 1
1.1 Purpose of This Document 1
1.2 FAF^3 Data Products 1
1.3 Links to More In-Depth Technical Documentation 2
1.4 Improvements in Reporting Introduced with FAF^3 3
2 FAF3 Geography, Commodity and Modal Classes 4
2.1 Geography 4
2.2 Commodity Classes 5
2.3 Transportation Modes 6
3. The Flow Matrix Construction Process 7
3.1 Overview 7
3.2 Modeling to Enhance CFS In-Scope Flows 8
3.2.1 CFS Data Gaps and Data Tables 8
3.2.2 Log-Linear Modeling of Missing Cell Values 10
3.3 Data and Modeling of Non-CFS Flows 13
3.3.1 Domestic Flows 13
3.3.2 Import and Export Flows 22
References 24
Appendix A: Differences in the FAF^3 and FAF^2 Freight Flow Matrices 26

1. Introduction

1.1 Purpose of This Document

This document provides an overview of how the origin-destination-commodity-mode (ODCM) annual freight flows matrix developed under the Freight Analysis Framework, Version 3 (FAF^3) program. FAF^3 is a Federal Highway Administration (FHWA) funded and managed data and analysis program that provides estimates of the total volumes of freight moved into, out of and within the United States, between individual states, major metropolitan areas, sub-state regions, and major international gateways. The FAF^3 database is constructed by Oak Ridge National Laboratory (ORNL). Staff at MacroSys contributed to the development of a number of industrial sector-specific commodity flow estimates. Staff at Battelle Memorial Institute, and at IHS Global Insight have also developed FAF^3 data products that derive from the 2007 freight flow matrices described in this report.

This present document is devoted to describing how the base year, 2007 annual tonnage and dollar valued flows are estimated in the FAF^3 ODCM matrix. The document is labeled an overview because a detailed description of the flow matrix building procedure is very lengthy. This present document should suffice the majority of readers interested in knowing the basics of where the flow estimates come from. More detailed descriptions of specific flow estimation components are provided for those wishing to go further into the process. Separate FAF^3 documents also describe how these flows are projected into future years, and how these base and forecast year flows are then converted into vehicle/vessel traffic volumes and assigned to (i.e. routed over) individual links and routes within the US national highway, rail and waterway networks.

1.2 FAF^3 Data Products

FAF^3 data products are the result of merging datasets from a large number of different sources. The principal data products developed under the FAF^3 umbrella are the following:

- A set of annual freight flow matrices, reported in annual tonnages and annual dollar value of goods transported, for calendar year 2007 for the United States,
- Based on these base year flow estimates, a set of forecast year freight flow matrices, projected out to calendar year 2040,
- A set of annual freight tonnage and vehicle/vessel movement volumes assigned to specific links and routes over the United States multimodal truck-rail-waterways transportation network, based on these base year 2007 and forecast year 2040 flow estimates.

Based on these estimated freight flows and their network assignments, a set of annual freight tonnage, dollar value, and ton-mileage statistics, broken down by mode of transport and commodity class are also developed.

Figure 1.1 show the functional linkage between these various FAF^3 data products, starting with the creation of the calendar year 2007 FAF^3 national freight flows matrix. Also shown in Figure 1.1 is a new data product coming out of the FAF^3 effort. This is not a data set per se, but an on-line, web-based tool for extracting data elements from the FAF^3 database and constructing useful data tables on a regional, modal and/or commodity specific basis.

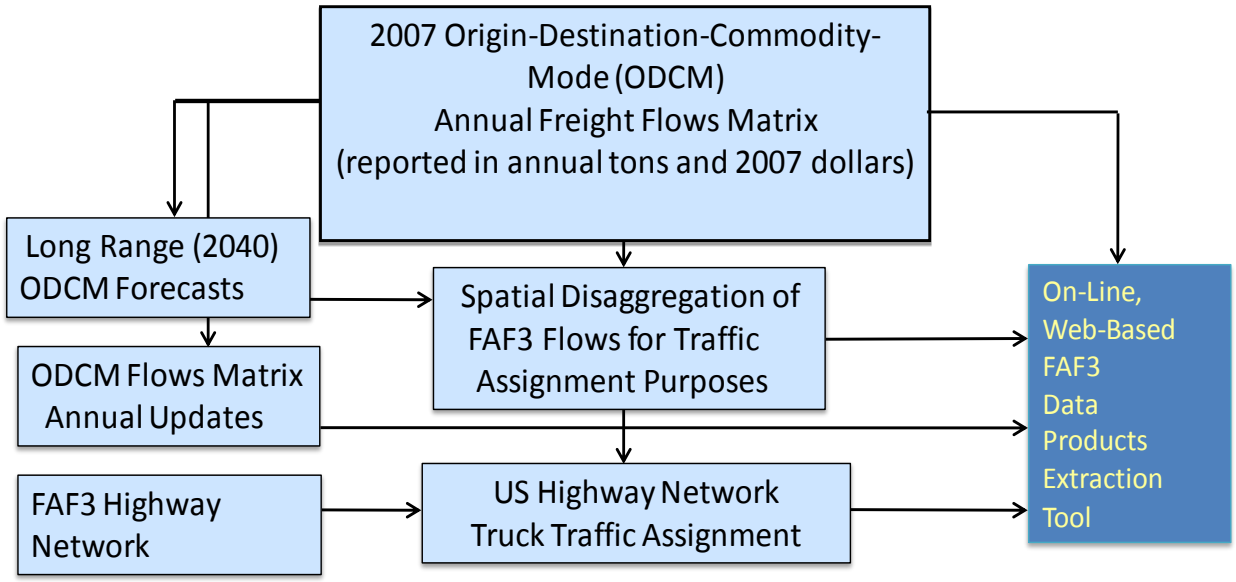

Figure 1.1 Principal FAF^3 Data Products

Freight origin-to-destination (O-D) movements are estimated in FAF^3 on both an annual tonnage and annual dollar value basis, for calendar year 2007. These estimates are then used as the basis for developing both annual *provisional* updates and as the starting point for a set of longer-range freight movement forecasts, reported at five year intervals from 2015 out to year 2040. The principal dimensions of these *FAF^3 Freight Flow Matrices* are:

- Shipment origination region (O),
- Shipment destination region (D),
- The class of commodity being transported (C), and
- The mode of transportation used (M).

The FAF^3 freight flows matrix is made up of 131 Origin (O) x 131 Destination (D) x 43 Commodity Class (C) x 8 Modal Category (M) data cells, for each of 2 reporting metrics, annual tons and annual dollar values.

1.3 Links to Technical Documentation

FAF^3 is the third database of its kind, with the FAF^1 database providing similar freight data products based on calendar year 1997 data, and FAF^2 providing freight data products based on calendar year 2002 data. Since the very first FAF effort, a number of changes in both data

products and in the sources of the data used to produce them have taken place. A description these earlier data products, along with the FAF^3 data products, can be found at the following FHWA website:

http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/index.htm

This site also guides the user to the *FAF^3 on-line Data Extraction Tool*, which can also be accessed directly at:

http://cta-gis.ornl.gov/faf/

At this site a user can customize and download a variety of fright flow tables directly from the FAF^3 database. Interactive links are also provided to FAF^3 Data Documentation, Data Summary, and maps.. Users can also download the entire FAF^3 2007 regional database in either Microsoft Access 2003 (125MB) or in CSV (100MB) format.

1.4 Improvements in Reporting Introduced with FAF^3

With this latest version of the FAF a number of improvements to the commodity flow matrix have been possible over previous versions. These include:

- A roughly doubling of the number of U.S. shipping establishments sampled as part of the 2007 U.S. Commodity Flow Survey (from some 50,000 establishments in 2002, to approximately 100,000 establishments surveyed in 2007);[1]
- The use of PIERS data to support improved estimates of the internal to the U.S. allocations of imports and exports to FAF domestic zones of freight origination (for U.S. exports) and destinations (for U.S. imports);
- Incorporation of additional federal datasets within an improved FAF^3 log-linear modeling/iterative proportional fitting algorithm, as well as the development of the Out-of-Scope estimates;
- Greater use of U.S. inter-industry input-output ('use' and 'make') coefficients in the development of the FAF^3 out-of-scope (to the 2007 CFS) commodity flow estimates;
- FAF^3 provides an O-D specific treatment of natural gas products, which were evaluated only at the level of national or broad regional activity totals in FAF^2; and
- The ability to access FAF^3 data products via a user friendly web-based data set construction and download tool (cf. Section 1.3 above).

[1] For changes in the CFS between 2002 and 2007 see the following Bureau of Transportation Statistics website: http://www.bts.gov/help/commodity_flow_survey.html#diff_2007_2002

2. FAF^3 Geography, Commodity and Modal Classes

2.1 Geography

The 2007 CFS commodity flow tables are based on a revised geography that contains 11 additional traffic analysis regions, for a total of 123 domestic regions in all. FAF^3 uses the same geography. Figure 2.1 shows the boundaries of the 123 domestic FAF^3 flow analysis regions, also referred to as FAF^3 analysis zones.

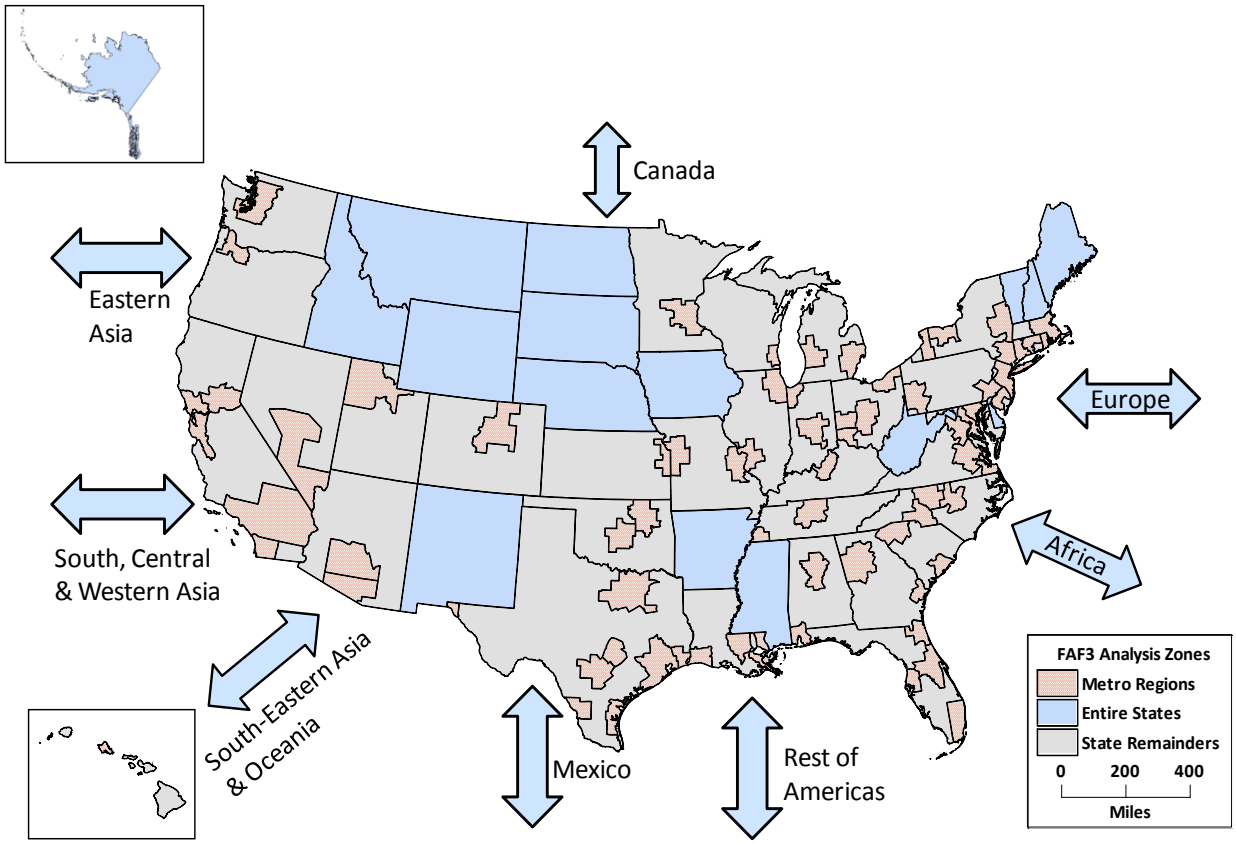

Figure 2.1 FAF^3 Geography

Three subsets of regions are highlighted: 74 metropolitan area determined regions, 33 regions made up of state remainders, representing a state's territory outside these metropolitan regions, and 16 regions identified as entire states, within which no FAF^3 metropolitan regions exist.

Note that metropolitan regions do not cross State boundaries: so that the Chicago, Kansas City, Philadelphia, and St. Louis metros are split into two state-specific FAF^3 regions, while the New York and Washington metropolitan areas are split into three distinct zones. To avoid crossing State boundaries the metropolitan areas of Atlanta (GA), Boston (MA), Charlotte (NC), Louisville (KY), Memphis (TN), Minneapolis-St. Paul (MN), Portland (OR), Providence (RI), Sacramento (CA), and Virginia Beach (VA) are each defined by the state in which most of the

metro areas' population resides and economic activity takes place. Also shown in Figure 2.1 are the 8 world regions that act as the origination and destination points for U.S. exported and imported freight. In addition to flows between the U.S. and Canada and the U.S. and Mexico, flows between the U.S. and the remaining six foreign FAF^3 regions are based on an allocation of countries to their respective United Nations geographic region.[2]

2.2 Commodity Classes

FAF3 reports annual tonnage and dollar valued freight flows using the same 43 2-digit Standard Classification of Transported Goods (SCTG) classes used by the 2007 U.S. Commodity Flow Survey (CFS).

Table 2.1 FAF^3 Commodity Classes

SCTG	Commodity	SCTG	Commodity	SCTG	Commodity
01	Live animals/fish	**15**	Coal	**29**	Printed products
02	Cereal grains	**16**	Crude petroleum	**30**	Textiles/leather
03	Other agricultural products.	**17**	Gasoline	**31**	Nonmetal mineral products
04	Animal feed	**18**	Fuel oils	**32**	Base metals
05	Meat/seafood	**19**	Coal-n.e.c.	**33**	Articles-base metal
06	Milled grain prods.	**20**	Basic chemicals	**34**	Machinery
07	Other foodstuffs	**21**	Pharmaceuticals	**35**	Electronics
08	Alcoholic beverages	**22**	Fertilizers	**36**	Motorized vehicles
09	Tobacco prods.	**23**	Chemical prods.	**37**	Transport equipment
10	Building stone	**24**	Plastics/rubber	**38**	Precision instruments
11	Natural sands	**25**	Logs	**39**	Furniture
12	Gravel	**26**	Wood products	**40**	Misc. mfg. products.
13	Nonmetallic minerals	**27**	Newsprint/paper	**41**	Waste/scrap
14	Metallic ores	**28**	Paper articles	**43**	Mixed freight
				99	Commodity unknown

[2] See http://unstats.un.org/unsd/methods/m49/m49regin.htm for these country-to-region allocations.

2.3 Transportation Modes

FAF^3 flows are also broken down by 8` modes of transportation. Table 2.2 lists these mode and commodity classes.

The "multiple modes and mail" category includes truck-rail, truck-water, and rail-water intermodal shipments involving one or more end-to-end transfers of cargo between two different modes. Detailed SCTG code definitions can be downloaded at either of the following Census and Bureau of Transportation Statistics websites:

http://www.census.gov/svsd/www/cfsdat/2002data/cfs021200.pdf

http://www.bts.gov/publications/commodity_flow_survey/survey_materials/pdf/sctg_booklet.pdf

Appendix A describes how these CFS-based regional, modal, and commodity class definitions differ from those used by FAF^2.

Table 2.2 FAF^3 Mode Classes

Mode Identification	Mode Name	Mode Description
1	Truck	Includes private and for-hire truck. Private trucks are owned or operated by shippers, and exclude personal use vehicles hauling over-the-counter purchases from retail establishments.
2	Rail	Any common carrier or private railroad.
3	Water	Includes shallow draft, deep draft and Great Lakes shipments.
4	Air (includes truck-air)	Includes shipments typically weighing more than 100 pounds that move by air or a combination of truck and air in commercial or private aircraft. Includes air freight and air express. Shipments typically weighing 100 pounds or less are classified with *Multiple Modes and Mail*
5	Multiple Modes and Mail	Includes shipments by multiple modes and by parcel delivery services, U.S. Postal Service, or couriers. This category is not limited to containerized or trailer-on-flatcar shipments.
6	Pipeline	Includes flows from offshore wells to land, which are counted as water moves by the U.S. Army Corps of Engineers.
7	Other and Unknown	Includes flyaway aircraft, vessels, and vehicles moving under their own power from the manufacturer to a customer and not carrying any freight, unknown, and miscellaneous other modes of transport.
8	No Domestic Mode	A 'No Domestic Mode' category is used to capture petroleum imports that go directly from foreign, inbound ships to an on-shore US refinery. This is done to ensure a proper accounting when foreign and domestic flows are summed, while avoiding assigning flows to the domestic transportation network that do not use it.

3. The Flow Matrix Construction Process

3.1 Overview

The FAF[3] modeling process draws from many data sources but the most important is the U.S. Commodity Flow Survey.(CFS). Figure 3.1 shows the principal types of data used to construct the FAF[3] ODCM freight flows matrix. This matrix construction process begins with the data reported by the 2007 CFS[3], adopting both the CFS definitions for the 123 internal to the U.S. freight analysis zones and the same 43 SCTG 2-digit commodity classes, but using a modification of CFS modal definitions. Each of these three data dimensions is elaborated on below.

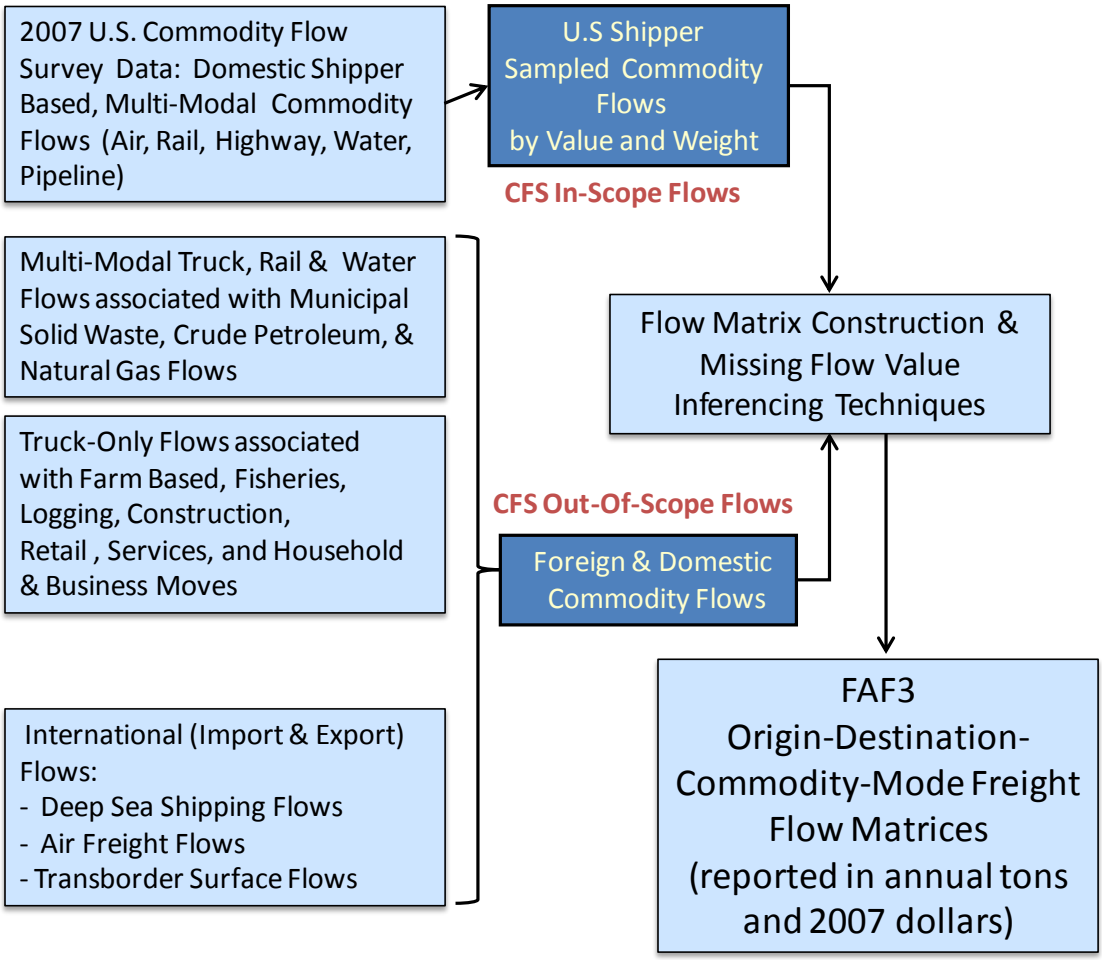

Figure 3.1 Overview of the FAF[3] Freight Flow Matrix Construction Process

[3]For the details of how the 2007 CFS survey methodology, and for on-line access to the public domain CFS data products , go to: http://www.bts.gov/publications/commodity_flow_survey/

The CFS itself is conducted every 5 years as part of the U.S. Economic Census, with major funding for the survey provided by the Bureau of Transportation Statistics (BTS). Data are collected on all shipments from the surveyed establishment for an entire week in each of the four quarters of the census year. In 2007, about twice as many establishment samples were recorded as in 2002.

The CFS represents the best basis for FAF construction because it provides shipper sampled, and subsequently expanded estimates of both tons shipped and dollar value trades within and between all US regions for all modes of freight transportation. However, the CFS has a number of well researched weaknesses that require considerable additional effort in order to construct a complete accounting of freight movements within the United States (see TRB, 2006). First, the CFS does not report imports, while CFS reporting of export flows is also subject to data quality issues resulting from limited sample size. Second, the CFS also either does not collect data from the following freight generating and receiving industries, or collects insufficient data to cover the industries in a comprehensive manner:

- Truck, rail and pipeline flows of crude petroleum, and natural gas,
- Truck freight shipments associated with farm-based, fishery, logging, construction, retail, services, municipal solid waste, and household and business moves, and.
- Imported and exported goods transported by ship, air, and trans-border land (truck, rail) modes.

In FAF^3 these industries produce what are referred to in Figure 3.1 as Non-CFS or Out-Of-Scope (OOS) to the CFS freight flows. Their estimation requires a good deal of data collection and integration into the larger flow matrix generation process. The data sources for these OOS flows are for the most part derived from freight carrier reported data sources, in some cases requiring the use of secondary or indirect data sources, such as location specific measures of industrial activity, employment or population, to allocate flows to specific geographic regions. These OOS flows represent some 32% of all U.S. freight movements measured on an annual tonnage basis. Developing OOS flow estimates represents a considerable effort, with different commodity classes requiring very different, typically multi-step treatments: including the use of both spatial and commodity class "crosswalks" that convert mode and industry class specific estimates from their native coding categories into FAF^3 regional and commodity class breakdowns.

3.2 Modeling to Enhance CFS In-Scope Flows

3.2.1 CFS Data Gaps and Data Tables

The 2007 CFS is a large and very sparse matrix of annual tonnage and dollar valued freight shipment volumes, with many individual cells assigned a value of value of zero tons and zero dollars of freight shipped during the calendar year. The complete set of 2007 CFS data products

includes a large number of different data matrices.[4] This includes the most detailed of the published matrices, Table CF0700A25, which reports annual tons, dollar values, and also ton-miles shipped by state of origin, state of destination, mode and 2-digit commodity class.[5] Although these are the four flow dimensions needed for the FAF this matrix contains many data gaps, and reports only state-to-state shipment totals that need to be assigned in some manner to FAF region-to-region flows. Fortunately, other CFS tables provide 1, 2 and 3 dimensional looks at this same data, including marginal totals at the FAF regional level that do not suffer to the same extent from data suppression. Without going through the contents of each CFS data table in turn, these gaps in the 2007 CFS coverage can be summarized as follows:

- Annual O-D commodity flow estimates exist but some are missing either a modal or commodity breakdown, or both,

- Modal share estimates exist but lack the geographic and/or commodity detail required of the FAF^3 flows matrix, and

- Data on shipment lengths exists, by mode and/or commodity, but with little or no linkage to either State or FAF^3 regional O-D geography.

In many instances data is missing or suppressed at the 2- or 3-, as well as 4-dimensional level of flow resolution. That is, we have a flow matrix that contains a variety of levels of coverage, with many data gaps needing to be filled.

While many of these zero valued cells are accurate, CFS sample size limitations may also be responsible for missing some of these flows at the origin-destination-commodity-mode level of resolution sought by the FAF; or for creating flow estimates that have such high variability (sampling error) that the US Census Bureau chose to suppress their values. Where such suppression occurs in the CFS a cell value has been replaced by the letter 'S'. In some cases 'S' reported cells may represent quite large freight flows in the real world, because a large coefficient of variation does not necessarily mean that we have only small O-D flows to deal with. For FAF reporting purposes an estimate is desired for these suppressed cell values, and also for any zero valued cells where limited CFS sampling has failed to produce a positive flow estimate, but where freight is likely being shipped.[6] The question the FAF has to answer is not

---

[4] http://www.bts.gov/publications/commodity_flow_survey/ Click on "Interactive tables."
[5] http://www.census.gov/svsd/www/02CFSdata.html
[6]Reporting of individual CFS cell values may also be suppressed to avoid disclosing information about an individual company's activity. For the CFS, the primary method of disclosure avoidance is Noise Infusion: Noise infusion is a method of disclosure avoidance in which values for each shipment are perturbed prior to tabulation by applying a random noise multiplier. Disclosure protection is accomplished in a manner that causes the vast majority of cell values to be perturbed by at most a few percentage points. In certain circumstances, some individual cells may be suppressed on a case by case basis for additional disclosure avoidance purposes. Such cell values have their flow values replaced by the letter 'D' in published CFS tables.
http://www.bts.gov/publications/commodity_flow_survey/def_terms/index.html#samplingerror

only what size each of these flows should be, but also, which of the many zero valued cells ought to contain a positive flow at all.

3.2.2 Log-Linear Modeling of Missing Cell Values

The procedure used for estimating these missing cell values is shown in Figure 3.2. This figure is a high level treatment of the problem. The following description provides an overview of the major data steps in this data modeling process.

In FAF^3, missing 2007 CFS cell values are first of all estimated using a six-dimensional log-linear model. The first four of these dimensions are the above-defined FAF origin region (O), FAF destination region (D), FAF commodity class (C) and FAF mode of transport (M). To this are added two additional dimensions:

- A 'freight metrics' dimension, U, defined by the two classes of metric reported by the CFS, i.e. tonnage (u =1) and dollar value of freight moved (u = 2); and

- A data source' dimension, S, that captures four different classes (= sources) of freight flow estimates, i.e. the 2007 CFS (s = 1), the 2002 CFS (s =2), the 2007 Railcar Waybill dataset (s = 3), and the 2007 Waterborne Commerce dataset (s = 4).

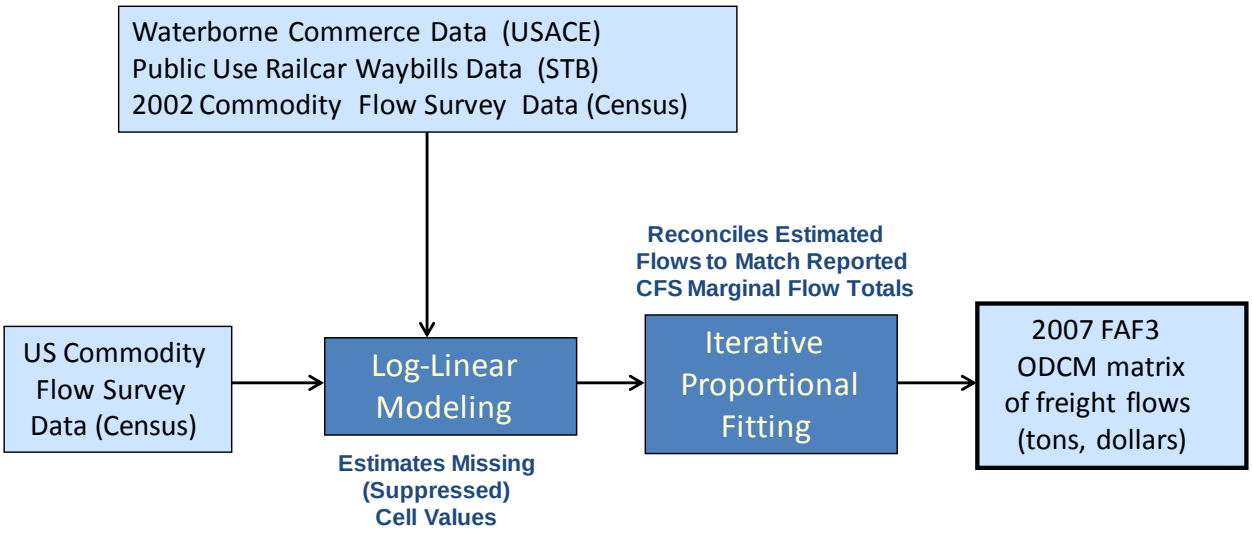

Figure 3.2 Estimation of Missing Cell Values in the 2007 US Commodity Flow Survey

Zero valued cells in the 2007 CFS can be categorized as either "structural" or sampling zeros. For example, truck commodity flows between Hawaii and mainland US regions is an obvious structural zero. Sampling zeros are divisible two types:

1. Cells where no sample data was obtained by the 2007 CFS, but flows may exist; and

2. Cells where the volume of freight sampled was so small that it fell below the CFS reporting threshold, i.e. below 500 tons, or below half a million dollars, and was therefore rounded down to '0' in the CFS published tables.

In particular, a large number of CFS cells have had their value suppressed, for either confidentiality or statistical robustness reasons. For example, cell values are suppressed reported in the 2007 CFS if the coefficient of variation associated with the cell estimate exceeds 50%. The method used for estimating these suppressed, and therefore, missing cells values in the CFS flow matrix is a combination of log-linear modeling (LLM) and iterative proportional fitting (IPF). This LLM/IPF procedure was selected because it has the following characteristics:

1. It makes extensive use of existing data within the matrix in the estimation of missing cell values,

2. It offers the ability to fill in missing cell values while maintaining reported marginal flow totals and observed cell values across all dimensions of the matrix,

3. It has the ability to handle missing values at multiple levels of data aggregation, and

4. It offers the ability to bring different, including non-CFS sources of flow estimates, into the solution, including completely new one, two, and three-dimensional data tables, as needed.

This last characteristic has been exploited extensively for the first time in developing the FAF^3 freight flows matrix, and represents a major enhancement to the modeling process used in the previous flow matrix generation process. Specifically, flows reported by two carrier-reported, mode specific datasets are used to help the FAF^3 flows matrix capture potentially missing or under-represented flow estimates. These are:

1. Calendar year 2007 annual rail flow volumes (tonnages) reported in the Surface Transportation Board's (STB) public use railcar waybills[7], and

2. Calendar year 2007 annual flow volumes (tonnages) reported in the US Army Corps of Engineers Waterborne commerce dataset.[8]

In addition, data from the 2002 CFS is also used to look for potentially positive, but zero valued (i.e. sampling zero) flow cells.

In practice, each of these data sources is treated as a component of a sixth dimension in an expanded FAF^3 freight flows matrix.[9] Where a positive cell value is reported in any of these data

---

[7] Accessible via http://www.stb.dot.gov/stb/industry/econ_waybill.html
[8] Accessible via http://www.iwr.usace.army.mil/ndc/data/data1.htm
[9] By housing these alternative modal data sources within a single dimension of the matrix in this manner we are also allowing, without loss of generality, for the application of more sophisticated across the board CFS + non-CFS weighting schemes in the future.

sources, these cells are subsequently assigned a positive value by the LLM/IPF routine, from which a maximum likelihood estimate of that flow's volume is estimated.

The complete FAF^3 commodity flow model, referred to as the "Log-Linear Model" in Figure 3.2, has the following form:

$$\begin{aligned}\mathrm{Ln}(F^{ODCMUS}) = {} & \lambda_0 + \lambda^{O} + \lambda^{D} + \lambda^{M} + \lambda^{C} + \lambda^{U} + \lambda^{S} + \lambda_j^{OD} + \lambda^{OC} + \lambda^{OM} + \lambda^{OU} + \lambda^{DC} + \\ & \lambda^{DM} + \lambda^{DU} + \lambda^{CM} + \lambda^{CU} + \lambda^{MU} + \lambda^{OS} + \lambda^{DS} + \lambda^{CS} + \lambda^{MS} + \lambda^{US} + \lambda^{ODC} + \lambda^{ODM} + \lambda^{ODU} + \lambda^{OCM} \\ & + \lambda^{OCU} + \lambda^{OMU} + \lambda^{DCM} + \lambda^{DCU} + \lambda^{DMU} + \lambda^{CMU} + \lambda^{ODS} + \lambda^{OCS} + \lambda^{OMS} + \lambda^{OUS} + \lambda^{DCS} + \lambda^{DMS} + \\ & \lambda^{DUS} + \lambda^{CMS} + \lambda^{CUS} + \lambda^{MUS} + \lambda^{ODCM} + \lambda^{ODCU} + \lambda^{ODCS} + \lambda^{ODMU} + \lambda^{ODMS} + \lambda^{ODUS} + \lambda^{OCMU} + \lambda^{OCMS} \\ & + \lambda^{OCUS} + \lambda^{OMUS} + \lambda^{DCMU} + \lambda^{DCMS} + \lambda^{DCUS} + \lambda^{DMUS} + \lambda^{CMUS} + \lambda^{ODCMU} + \lambda^{ODCMS} + \lambda^{ODMUS} + \\ & \lambda^{ODCUS} + \lambda^{OCMUS} + \lambda^{DCMUS} + \lambda^{ODCMUS}\end{aligned}$$

where $\mathrm{Ln}(F^{ODCMUS})$ is the model estimated natural log (log to the base e) annual volume of commodity 'C' moved by mode 'M' between FAF^3 origin zone 'O' and FAF^3 destination zone 'D' in 2007, measured in units 'U' (i.e U=1 for annual tons, U=2 for annual dollar value of the freight moved), and found in data source 'S' (e.g. S = 1 for CFS 2007, S=2 for CFS 2002, S= 3 for 2007 Railcar Waybills, and S = 4 for 2007 Waterborne Commerce).

The λ's represent the model parameters to be estimated, often termed the (natural log of the) *effects* of the different dimensions, or combinations of dimensions, on the resulting flow estimates. For example, λ^{OM} represents the effect of shipment origin O and mode M, λ^{ODCM} represents a four-way, O,D,C,M interaction effect, and λ_0 represents the grand mean of all these effects. Parameters representing all possible levels and combinations of the matrix dimensions O,D,C,M,U and S are used to fit the data to what is usually termed a *saturated* model that tries to get the most out of the statistical relationships represented by the data sources. This equation is translated into an additive, natural log form for solution (i.e. for computational) purposes. In practice, many of the λ's are set to a value of 0.0. For example, since both the 2007 railcar waybill and waterborne commerce flows are only reported in tons, all dollar valued λ's associated with these two data sources = 0.0 and play no further part in the estimation process.

3.2.3 Iterative Proportional Fitting (IPF) to CFS Marginal Totals

Once all of the log-linear model's λ *effects* have been computed, they are used to generate a positive value of each zero valued flow cell in the original 2007 CFS commodity flow matrix. In each case, where a zero valued cell is found it is replaced with an estimate based on the above multiplicative log-linear model. Three additional steps are then taken:

1) Cells considered to be structural zeros are returned to a value of 0.0.

2) To further assist with filling in of missing CFS cell values, an additional dataset was provided by the U.S. Bureau of the Census. This is a matrix containing the number of establishments

sampled within each ODCM cell in the matrix, i.e. a set of raw sample responses. If one or more positive responses are identified for a specific cell, then this is taken to imply the presence of some freight movement activity, and it is therefore treated as a sampling zero for the purpose of cell value estimation.

3) A third modification to process then involves the removal of unreasonable dollar per ton estimates caused by biased or limited sampling, in which either the tonnage or the dollar value allocated to a particular cell by the log-linear/IPF modeling process creates a dollar-per-ton ratio that exceeds expected values for the commodity class in question by a significant amount. To prevent this from occurring, a check is made every ten iterations of the IPF to look for such outliers. If one or more are found, an adjustment is made to either the tonnage or dollar value in such a cell and the iterative process re-commenced.

The resulting matrix (now with no missing values) is then adjusted through IPF to comply with known control totals from numerous CFS marginal tables. It is important to note here that after the full LLM/IPF procedure is completed, no 2007 CFS ODCM or higher (3 or 2 dimensional) marginal cell value has been changed if it contained a positive flow value to begin with. Only potentially missing valued cells (of which there are many) are altered by the process.

3.3 Data and Modeling of Non-CFS (Out-of-Scope) Flows

3.3.1 Domestic Flows

U.S. freight shipping establishments in the following industrial sectors were not surveyed as part of the 2007, or previous, US Commodity Flow Surveys. The following out-of-scope (OOS) industries therefore had to be assigned commodity and mode specific O-D flows using other methods:

1. Farm Based	6. Retail
2. Fishery	7. Household and Business Moves
3. Logging	8. Municipal Solid Waste
4. Construction	9. Crude Petroleum
5. Services	10. Natural Gas Products

OOS flows were estimated using commodity specific datasets and different computational methods for each industrial class. Where an industrial sector produces O-D flows in more than one commodity class, data from national inter-industry input-output “use” and “make” tables was used to determine how much freight each sector contributes to a specific set of SCTG 2-digit commodity flows. State and county level data on volume of production, industrial or commodity specific sector sales, or industrial sector employment is then used to allocate flows between origins and destinations. Spatial allocation formulas are then used to produce O-D flow volumes. Where truck movements were concerned this occurred in one of two ways. Either county level origin and destination activity totals were determined, and then a spatial interaction model was applied to these county productions and attractions, with subsequent aggregation of inter-county

flows back up to FAF^3 region-to-region flow totals. Or county Os and Ds are first of all estimated and aggregated to their FAF^3 regional supply and demand totals. These regional totals are then used to estimate O-to-D flows directly at the FAF^3 region-to-region level.

The specific form of spatial interaction model used also varied by commodity class. Either a distance decay coefficient is calibrated against an empirically derived average shipping distance, or a simple allocation is made based on market potentials (i.e. on the relative size of a county's or region's demand for a specific commodity). County-level spatial interaction modeling here allows for cross-county flows to be captured that are also cross-FAF^3 adjacent regional flows. Use of regional O and D shipment totals prior to spatial interaction modeling occurred where data sources proved more reliable at this less detailed level or geography.

Figure 3.3 shows the general idea. In practice, each industrial sector has its own data gaps and idiosyncrasies that needed to be dealt with.

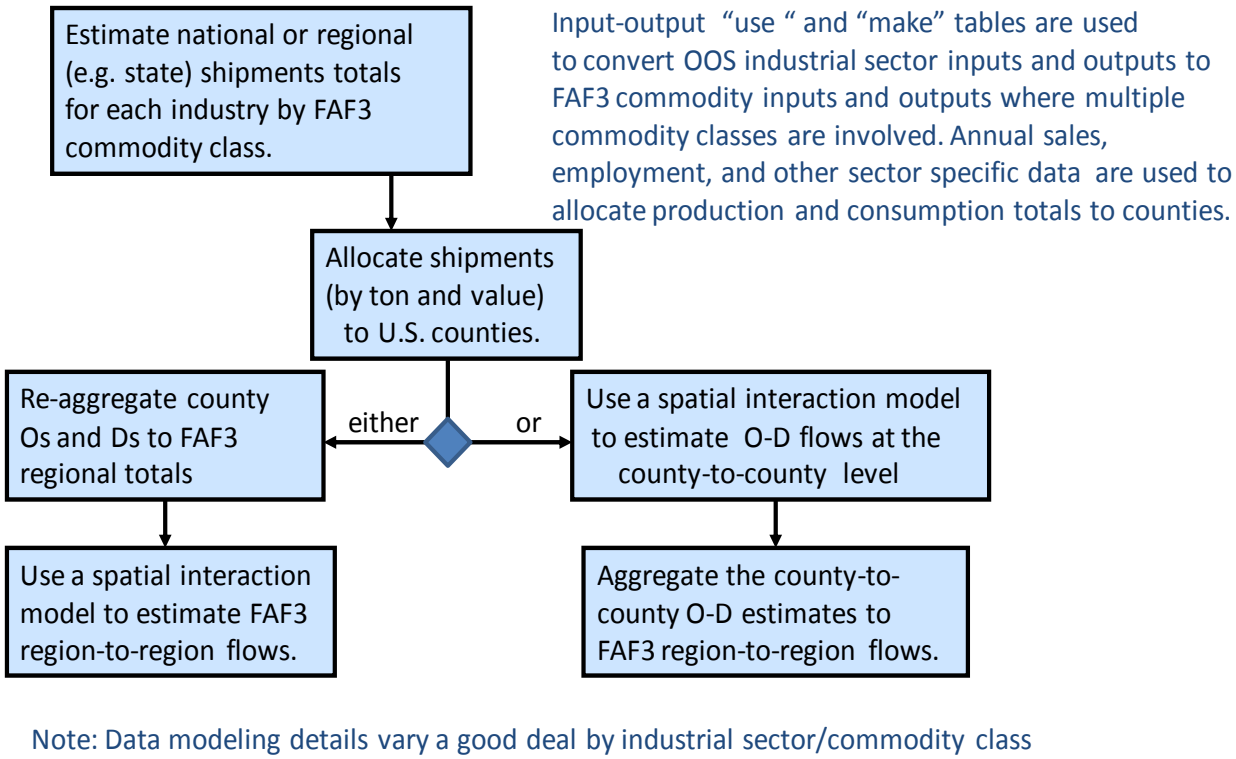

Figure 3.3 Four Step Process for Generating OOS Truck Freight Flows

The following sections focus on summarizing the datasets used to produce the FAF^3 flow estimates. For greater detail on estimation methods, the reader should consult FAF^3 industry sector-specific write-ups.

Farm Based Flows

Farm-based agricultural shipments represent one of the most significant out-of-scope areas for CFS. These shipments are almost entirely moved by truck. The vast majority of these shipments represent farm-to-storage elevator (e.g., grains) or farm-to-distribution/processing center (e.g., fruit, livestock) trips, at which point further transportation of these products is captured as part of the CFS sample frame. At the fully national level, the total tonnage of farm-based agricultural shipments constitutes nearly 7% of the 2007 total tonnage moved within the nation, and over 9% of all truck tons shipped. County and state level data published by in U.S. Department of Agriculture's (USDA) 2007 *Census of Agriculture* and the *2008 Agricultural Statistics* were used to generate FAF^3 tons and dollars shipped estimates, supplemented with data from several of USDA's *Statistical Bulletins*.

The dollar value of these farm originating agricultural products were estimated using information obtained from the *2007 Census of Agriculture* and related publications. Specifically, data provided under the category of "Market value of agricultural products sold"[10] was used as an estimate for total farm-based agricultural shipments. The estimation of tonnages for these out-of-scope shipments was less straightforward. Commodity statistics published in the USDA's *2007 Census of Agriculture* use a variety of commodity specific units of measurement (e.g., pounds, bushels, hundredweight, barrels, tons, etc). In some cases, different conversion factors, all based on information obtained from *Agriculture Statistics 2008,* were also needed for different commodities using the same basic unit of measurement. For example, the approximate net weight for a bushel of wheat is 60 pounds, while a bushel of husked corn on the ear weights 70 pounds, and shelled corn weighs in at 56 pounds per bushel on the average. Following these unit conversions, each farm-based agricultural commodity is then placed within its 2-digit SCTG commodity class.

Where a State is divided into more than one FAF^3 region, USDA county level data was used and subsequently re-aggregated to FAF^3 regional totals. This was done after filling gaps in this county-specific data, by using acreages devoted to a specific crop-growing activity as a surrogate for gaps in direct reporting of crop yields. O-D flows are then estimated, first by summing these county originations to their FAF^3 regional totals, then sharing these totals to FAF^3 destination regions on the basis of a) truck trip length distributions reported by the 2002 VIUS, and b) using the volumes of agricultural commodity originations reported by the 2007 CFS to allocate these flows. That is, these CFS originations (from the distribution centers, grain elevators, processing centers, etc. located within a CFS region) constitute the first non-farm stop in the agricultural product's supply chain. Hence they represent a good surrogate for the destinations of farm-based shipments. Separate allocations are made on the basis of tons shipped and dollar valued trades.

---

[10] The "market value of agricultural products sold" category represents the value of products sold which combines total sales not under production contract and total sales under production contract. It is equivalent to total sales. See Appendix B, General Explanation and Census of Agriculture Report Form, in the *2007 Census of Agriculture* report for further explanation (http://www.agcensus.usda.gov/Publications/2007/Full_Report/Volume_1,_Chapter_1_US/usappxb.pdf)

As a result of this process, the annual tons and dollar valued flows between any two FAF^3 regions are consistent with both VIUS truck trip length distributions for a specific FAF^3 freight originating region and commodity class, and also create a consistency between OOS farm-based flows and the non-farm based agricultural commodity flows reported in the 2007 CFS.

Construction Industry Flows

Shipments originating from activities in the construction sector, including companies or establishments engaged in construction of residential and non-residential buildings, utility systems, roadways and bridges, and from specific trade contractors, are not in-scope for the CFS. It is estimated that this industry transported just under 1.08 billion tons of freight over the course of 2007, valued at $905.7 million. However, putting a dollar value on such freight is not straight-forward. The primary commodity shipped was debris (included in SCTG 41 under Waste and Scrap), for which the value would be relatively small unless recyclable materials are separated and sold. An estimate of the amount of debris generated by the construction industry was developed based on publications by the U.S. Environmental Protection Agency (EPA) publications,[11] the National Demolition Association, Construction Materials Recycling Association, and Gershman, Brickner & Bratton, Inc. Similar dollar to ton conversions for other commodity classes are drawn from the CFS or other industry specific sources.

Data on shipment distances for the industry are limited at best for 2007, and in FAF^3 all of these shipments are assumed to be short distance truck movements, most occurring within a single county, and all within the same FAF3 zone. Shipment volumes were assigned to FAF3 regions using sales data from the 2007 Economic Census (EC) where available, and using a combination of 2007 county level employment data from the Census Bureau's County Business Patterns (CBP) dataset, multiplied by Census developed labor productivity rates by industry class at the state level.

Fishery Flows

The CFS omits fishery shipments that move from vessels at the dock/port to the first point of processing or distribution centers. Establishments involved in this data gap are within the NAICS category 114 (fishing, hunting and trapping). Industries in this NAICS sector harvest fish and other wild animals from their natural habitats and are dependent upon a continued supply of the natural resource. Based on statistics published in the *Fisheries of the United States 2008*[12], an annual report prepared by the National Marine Fisheries Service (NMFS) of the National

[11] http://www.epa.gov/osw/conserve/rrr/imr/cdm/pubs/cd-meas.pdf.

[12] Information obtained from the *Fisheries of the United States 2008* report, published by National Marine fisheries Service, Office of Science and Technology in July 2009, was used to supplement its 2007 report under this analysis. Although 2007 statistics are available in the *Fisheries of the United States 2007*, many are in preliminary forms. The 2008 report provides more updated information on statistics for 2007.

Oceanic and Atmospheric Administration (NOAA), commercial landings by U.S. fishermen at ports in the 50 states were totaled at approximately 4.7 million tons and valued at over $4 billion in 2007. In addition, catches of Alaska Pollock, Pacific whiting, and other Pacific ground fish that are processed at-sea aboard U.S. vessels in the northeastern Pacific (off Washington, Oregon, and Alaska) are credited as landing to the state nearest to the area of capture. According to NMFS, these at-sea processed fishery products accounted for a total about 1.4 million tons and valued approximately $19 million in 2007. It is assumed that this freight activity is mostly local, and that all shipments involve intra-regional FAF^3 truck-only movements.[13]

Retail Industry Flows

The 2007 CFS also does not cover shipping activities originating from the vast majority of the nation's retail stores. It is estimated that 378.6 million tons of freight were shipped by the U.S. retail industry in 2007, valued at $624 billion. Based on the U.S. Bureau of Economic Analysis's National Input-Output Make and Use Tables, the retail industry generates commodity flows in most of the FAF^3 commodity classes.

Although most of the shipments from retail stores are within the same county, there is a possibility that retailers may transport large items purchased by customers from their warehouses, which may be located in other counties. At the county level this would be an issue, but is less likely to be of concern when aggregating O-D flows from counties up to FAF^3 regional totals. An issue with retail industry flows is whether some of these shipments are originated from retailer-owned warehouses that serve retail stores not covered by the CFS. In this case some inter-regional flows might be missing from FAF^3 totals. These volumes are believed to be quite small in percentage terms.

Service Industry Flows

This sector covers a wide range of services, including finance and insurance, real estate, rental and leasing, professional, scientific and technical services, administrative support, waste management and remediation services, education services, and health care and social assistance. These industries are typically involved in providing services to the general public, local business establishments, and branches of government, and in toto originate freight shipments in a large number of FAF^3 commodity classes. Also not covered by the 2007 CFS are the mail shipments by these service industries. The sector as a whole is estimated to have generated 378.6 million tons of commodity freight in 2007, worth just under an estimated $504.7 billion. To this is added some 11.4 million tons of mail, valued at $525.6 billion.

---

[13] Based on NMFS published statistics, total imported edible and non-edible fishery products were over 2.4 million tons and worth about $28.8 billion in 2007. Because imports are categorized as a separate out-of-scope area of the CFS (see Section 3.3.2 in this report), to avoid double counting, imported fishery is not included under this fishery shipment data gap study.

The availability of county level sales data varies by type of service offered. For example, the county level sales data for educational services are released for only 10 states. For real estate and food services, the sales data at the county level are available for 20 states. A first step was therefore to fill in this data gap for those service industries, then sum the sales of individual types of services to obtain an overall sales statistic for each county. Shipment volumes between counties were then estimated as follows (MacroSys, 2010):

- For non-mail shipments, the county level demand for service sector products (i.e. the market potentials for these destination counties) was determined by two factors: (i) the amount of a commodity used by industries according to the Use table in the U.S. I-O model and (ii) industrial employment at counties. Next, a spatial interaction ("gravity") model was used to distribute flows from each freight generating county to surrounding counties within our across FAF^3 regional boundaries.

- For mail shipments, total employment in services at the county level served as a surrogate for market potentials. Since mail is known to be shipped over long as well as short distances across the county, and lacking any empirical data on this distribution, no distance decay effect was applied to this sharing process in FAF^3.

Household and Business Move Flows

It is estimated that some 254.3 million tons of freight were moved by the industrial sector, nearly all of it by truck. The value of the goods moved is estimated at just $30.9 billion. Several sources of data on the volumes of U.S. household and business moves were examined, including the U.S. Census Bureau's Annual Services Survey and related studies conducted by the American Trucking Association and the American Moving and Storage Association.

All of these shipments are assumed to be truck moves in FAF^3. These truck shipments were allocated to counties on the basis of CBP-reported sector employment totals. The shipments are then allocated spatially between county O-D pairs based on IRS reported county level in-migration and out-migration totals. (In the absence of available data on trip length distributions, a distance decay effect was not used in this allocation process).

Logging Flows

Some 372.3 million tons of logs, totaling almost $9.5 billion by value, are estimated to have been transported in the U.S. as a whole in 2007, of which the vast majority are transported by truck from domestic forests to nearby sawmills and other local sites. County level logging products were estimated by multiplying the year 2007 employment in logging industries,, by an average tons per employee multiplier. To allow for logging products being transported across FAF^3 regional boundaries, these products were assigned to counties located within a 75 mile radius of the producing county, based on the employment in wood product industries within each county, and upon data collected on the average haul to market distance of logging products (e.g. sawlogs, peeler logs, OSB, pulpwood and rustic fencing).

Municipal Solid Waste Flows

Municipal solid waste (MSW) is not covered in the CFS, and also does not have a specific code in NAICS. The main data sources used for estimating 2007 MSW shipments came from information compiled by Franklin Associates[14] in collaboration with the U.S. EPA,[15] supplemented by information in the *BioCycle* journal[16]. Additional, mode specific data was also obtained from the U.S Army Corps of Engineers Waterborne Commerce statistics, and from the Surface Transportation Board's Railcar Waybills sample. As defined by the U.S. EPA, MSW includes the following 'Subtitle D wastes':

- Containers and packaging, such as soft drink bottles and cardboard boxes,
- Durable goods, such as furniture and appliances,
- Nondurable goods, such as newspapers, trash bags, and clothing, and
- Other wastes, such as food scraps and yard trimmings.

It is estimated that 413 million tons of MSW, as defined above, were transported within the U.S. in calendar year 2007. All of this MSW is collected at the source and transported to one of four types of processing facility: local landfills, local incineration facilities, local material recovery facilities, and waste transfer stations where garbage trucks unload MSW for accumulation and transfer to larger transport vehicles (truck, rail, or barge), for more economical long-distance hauling to a final disposal site (Curlee, 2009).

Data on the flows between states was based on work done by McCarthy (2007) for the Congressional Research Service. Combining this work with data from other sources, it is estimated that more than 42% of total state-to-state transfers (i.e. state exports) come from three states—New York, New Jersey, and Illinois, whole several other states export more than 10% of the U.S. total across state lines. The District of Columbia exports all of its total MSW generation, while New Jersey exports over 45%, New York exports over 33%, and Maryland over 29%. Additional states that export more than 10% of their MSW include Connecticut, Illinois, Kansas, Massachusetts, Missouri, North Carolina, Vermont, Washington, and West Virginia. More than 46% of all these state exports go to three states—Pennsylvania, Virginia, and Michigan Only five additional states account for more than 4% of the national total shipments of inter-state MSW—Georgia, Illinois, Indiana, New Jersey, Ohio, and Oregon. Based on ORNL discussions with local officials for the previous, FAF^2 effort, it appears that the large majority of shipments to adjoining states are essentially local shipments. For example, the city of Memphis ships MSW to Mississippi. Chicago ships tons to Indiana. The District of Columbia ships to Virginia. Also, small to medium sized towns near a state line may ship to an adjoining county across the state line. While these are truck movements, some longer distance shipments are by rail or (much less so) by inland waterway (i.e. by barge). It is estimated that just under 40% of inter-state

[14] http://www.fal.com/solid-waste-management.html
[15] http://www.epa.gov/epawaste/nonhaz/municipal/msw99.htm
[16] http://www.jgpress.com/biocycle.htm

shipments of MSW are by rail (mostly) or barge. This represents less than 4% of all MSW shipments.

The FAF3 MSW estimates also include significant tonnages moving from Maine to New Brunswick, Canada, from Ontario, Canada to Michigan, and a from Ontario to New York state (Curlee, 2009). Allocation of (truck-only) MSW between FAF^3 regions below the state level then used county populations to distribute inter-state flows, with subsequent re-aggregation from counties to FAF^3 regions. County-to-county O-D flows were estimated using a spatial interaction model, using an average O-D distance of just under 32 miles, derived from the MSW literature. These inter-county flows were then aggregated to their FAF^3 region-to-region totals.

Crude Petroleum

It is estimated that the US transported some 744.4 million tons of crude petroleum (crude oil) in 2007, using a variety of modes. This crude was valued at some $336.4 trillion dollars. These crude oil shipments begin either at domestic oil fields, or from large marine terminals that act as the first domestic storage and transfer point for foreign oil imports. The crude is delivered either to refineries or to long-term storage facilities such as the Strategic Petroleum Reserve.. A great deal of this transport is accomplished by pipeline, and by marine vessels (inland barge and oceangoing tanker), with significant tonnages also moved by rail tanker car and locally by tank truck.

National level crude oil shipment information by transportation mode is based on *Shifts in Petroleum Transportation* published annually by the Association of Oil Pipelines. This report's modal information is in turn based on several other data sources, including:

- Oil Pipelines: Annual Report of oil pipeline companies provided to the Federal Energy Regulatory Commission (FERC Form 6);
- Water Carriers: Waterborne Commerce of the United States, U.S. Army Corps of Engineers, (Part 5, Table 2-2);
- Motor Carriers: Petroleum Tank Truck Carriers Annual Report, American Trucking Association, Inc. and Petroleum Supply Annual, Energy Information Administration (EIA) (Volume 1, Table 46); and
- Railroads: Carload Waybill Statistics, Report TD-1, USDOT, Federal Railroad Administration, and Freight Commodity Statistics, Association of American Railroads (Table A3).

O-D flows of crude petroleum were derived using US DOE/EIA supplied data at various levels of geographic detail, ranging from five broad multi-state PADDs (Petroleum Administration for

Defense Districts)[17] and individual States, to specific refinery locations. This includes data from EIA's Petroleum Supply Annual (EIA, 2010) on:

- Production of Crude oil by PAD District and State,
- Refinery Input of Crude Oil by Refining Districts, and
- Refinery Receipts of Crude Oil by Method of Transportation, by PADD.

Spatial interaction (e.g. "gravity") models were then used to disaggregate flows down to a State-to-State and FAF region-to-FAF region level. First, U.S. Census' County Business Pattern data for 2007 was used to share total crude production by state down to the county level. This allocation was based on a county's reported total annual payroll for industries classified under NAICS code 211111 – 'Crude Petroleum and Natural Gas Extraction'.[18] These county activity totals were then aggregated to their respective FAF^3 regions. This resulted in 80 different petroleum sourcing regions, serving 50 petroleum refining FAF^3 regions. O-to-D allocations between these pairs of regions were then estimated using a distance-decay based spatial interaction model, applied at this broader regional level of resolution.

Natural Gas Products

Delivering natural gas (principally methane, but also smaller volumes of ethane, propane, butane and pentane) is an enormous enterprise. This gas is transported to consumers through more than 300,000 miles of transmission pipelines with the help of vast storage reservoirs and thousands of compressors. This gas is sold to marketers, large commercial and industrial consumers, and distribution companies for delivery to consumers over a network of more than 1.1 million miles of local distribution pipelines.

National Natural Gas flow totals, and O-D region-to-region flows were derived from the EIAs' Natural Gas Annual (EIA, 2010)[19], making use of data at various levels of geographic detail, including:

- Gross Withdrawals and Marketed Production of Natural Gas by State and the Gulf of Mexico,
- Offshore Gross Withdrawals of Natural Gas by State and the Gulf of Mexico,
- Summary of U.S. Natural Gas Imports By Point of Entry, and
- Summary of U.S. Natural Gas Exports By Point of Exit, Natural Gas Annual.

Spatial interaction models were then used, where necessary, to disaggregate flows down to a

---

[17] The New England, Midwest, East Coast, Gulf Coast, and West Coast PADDs. For specific state allocations to APDDs see: http://www.eia.gov/glossary/index.cfm?id=P#PADD_def

[18] The data is obtained by county level from the County Business Pattern at the U.S. Census Bureau - http://www.census.gov/econ/cbp/intro.htm.

[19] See http://tonto.eia.doe.gov/dnav/ng/ng_pub_publist.asp

State-to-State and a FAF region-to-FAF region level.

3.3.2 Import and Export Flows

Imported as well as exported freight flows in FAF^3 are constructed from a variety of data sources, each of which must have its flows converted from agency specific commodity codes to FAF^3's 2-digit SCTG codes, as well as have its flows either spatially aggregated or disaggregated to match FAF^3 analysis zones. Figure 3.4 provides a top-down view of this process.[20] The following sections describe each source data-specific procedure in more detail.

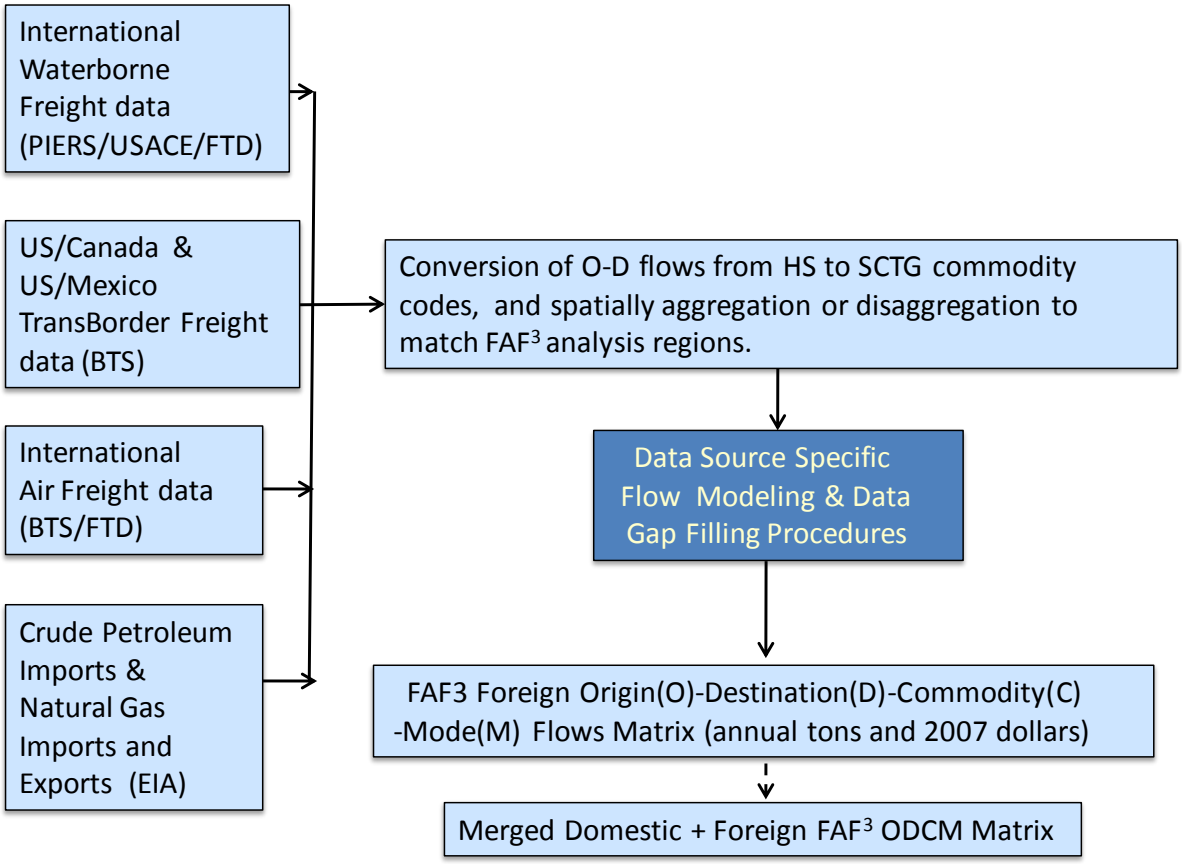

Figure 3.4 FAF3 International (Import/Export) Data Modeling

Waterborne Imports and Exports are derived in FAF^3 using four different data sets, each of which provides a different look at the nation's international freight movements by ocean vessels:

[20] Although the 2007 CFS does also collect data on export shipments by US establishments, both coverage and statistical accuracy is limited by sample size issues and this data was not used as a source for FAF3 export flow estimates.

- The U.S. Army Corps of Engineers International Waterborne Commerce Database[21]
- The U.S. Census Bureau's Foreign Trade Database[22]
- A FAF^3-specific extraction of data from the PIERS Import/Export Database[23]
- Imported & Exported Petroleum & Natural Gas data from the U.S. Department of Energy's Energy Information Agency (EIA)

The availability of these last two data sources represents a significant enhancement in FAF^3, and especially the PIERS dataset, which provided estimates of the internal to the U.S distribution of imported and exported goods. In 2002, the distribution of domestic CFS shipments was used to impute domestic trip ends and modes used in FAF^2 for every commodity that passed through a seaport. In 2007, information from PIERS was used to impute many of these domestic trip ends, with 2007 CFS data being used to impute the modes used between U.S. seaports and their internal U.S. destinations or origins.

International Air Freight Flows: Data published by the U.S. DOT's Office of Airline Information (OAI), Bureau of Transportation Statistics provided the FAF^3 estimates of total tons shipped annually between originating airports (where the cargo is first loaded onto an aircraft) and destination airports (where the cargo is unloaded for final land-based delivery, usually by truck).[24] This data is combined with data collected by U.S. Customs on the commodity class and value of international air shipments, as reported by the Foreign Trade Division (FTD) of the U.S. Department of Commerce's Bureau of the Census.[25] This FTD dataset includes information on the value,[26] quantity, method of transportation, and shipping weights for 9,000 export commodities, 17,000 imported commodities, 240 trading partners, and 45 U.S. Customs Districts.

The OAI and FTD data are combined into a single FAF^3 air freight dataset by reconciling differences in the level of spatial and commodity detail to match those required by the FAF. First each airport was assigned to its U.S. county, and each county to both its appropriate U.S. Customs District and FAF3 region, using geographic coordinates data files available from OAI and the Census Bureau. Commodities are reported in the FTD dataset using the 10-digit Harmonized Tariff Schedule (HS Schedule B for exports). This data is aggregated and translated to FAF3's 43 2-digit SCTG commodity classes using a crosswalk specifically developed for the purpose. Where differences exist between the OAI and FTD flow totals, the OAI database was taken to be definitive for total tons shipped, and the FTD database was used to control the

---

[21] http://www.iwr.usace.army.mil/ndc/data/dataimex.htm
[22] http://www.census.gov/foreign-trade/reference/products/index.html
[23] Special tabulations prepared for the FAF3 project by PIERS staff. (http://www.piers.com/)
[24] T-100 (foreign) market data. http://www.bts.gov/publications/freight_transportation/
[25] http://www.census.gov/foreign-trade/reference/products/index.html
[26] Export values are reported free-alongside-ship (F.A.S.) Import values are reported as customs-insurance-freight (C.I.F) values.

allocation of freight shipments to commodity classes, and to assign value-to-weight ratios to these flows.

U.S.-Canada and U.S.-Mexico Transborder Freight Flows: Truck and rail freight movements between the United States and its NAFTA neighbors Canada and Mexico are derived in FAF^3 from the Bureau of Transportation Statistics (BTS) Transborder Freight Database, itself constructed from data collected at border crossings by the U.S. Customs Service. After converting the Harmonized Tariff Schedule (HS) commodity classes in this dataset to FAF^3 SCTG classes, County Business Patterns are used to allocate flows reported at the State level to their most likely FAF3 regions within the United States.

Imports and Exports of Natural Gas and Imports of Crude Petroleum: Liquefied Natural Gas (LNG) is imported or exported to/from the U.S. by large tanker ships. The US Department of Energy's Energy Information Administration (EIA) reports annual LNG imports/exports in millions of cubic feet by U.S. seaport of entry/exit. The EIA also reports the annual trade in pipeline supplied natural gas (NG) between the U.S. and Canada and the U.S. and Mexico, also in millions of cubic feet. Reporting here is both by State and by specific U.S. seaport of entry/exit, requiring assignment of flows to seaport-inclusive FAF^3 regions.[27]

EIA databases were also used to estimate crude petroleum imports in FAF^3, taking advantage of the fact that crude petroleum imports are reported to the EIA monthly at the company, U.S. seaport of entry/exit, and foreign country level[28], allowing the complete movement of imported crude oil from the foreign country (source of commodity), passing through the port (domestic origin), to the refinery (domestic destination) to be estimated. The allocation of these flows to specific modes of transportation was then based on EIA data on crude oil refinery receipts, broken down by mode of transportation (ship, pipeline, rail, barge, truck), and further broken down by domestic versus foreign sources of production.[29]

References

Curlee, T. R. (2009) The Municipal Solid Waste (MSW) Management Supply Chain: FAF3 Out-of-Scope Commodity Flows. Oak Ridge National Laboratory. January, 2009 (Draft).

DOT/DOC (2006) Standard Classification of Transported Goods (SCTG). SCTG Commodity Codes. U.S, Department of Transportation/U.S. department of Commerce. 2007 Commodity Flow Survey. CFS-1200 (10-24-2006). Washington, D.C.

---

[27] Both the EIA's LNG and NG data sources for US Imports/exports can be found at: http://tonto.eia.doe.gov/oil_gas/natural_gas/info_glance/natural_gas.html

[28] *http://www.eia.doe.gov/oil_gas/petroleum/data_publications/company_level_imports/cli.html*

[29] *http://www.eia.doe.gov/oil_gas/petroleum/data_publications/refinery_capacity_data/refcapacity.html*

EIA (2010) Natural Gas Annual 2008. Office of Oil and Gas, Energy Information Administration, U.S. Department of Energy Washington, D.C.

MacroSys (2010) Outline of Methodologies for Estimating Freight Flows of Four Out-of-Scope Industries and International Air Cargo. Report submitted to Oak Ridge National Laboratory, April, 2010.

McCarthy, J. E. (2007) Interstate Shipments of Municipal Solid Waste: 2007 Update. Congressional Research Service, Order Code RL-34043, Washington, D.C.

TRB (2006) Commodity Flow Survey Conference. *Transportation Research Circular E-2088*: 29-46. January 2006.

Appendix A: Differences in the FAF^3 and FAF^2 Freight Flow Matrices

The FAF^3 ***Analysis Zones*** are different from the FAF^2 zones. Since the FAF freight flow matrix is developed around the data supplied by the U.S Commodity Flow Surveys (CFS) the geography has changed with CFS geography. In 2007 the use of more CFS analysis zones (made possible by the much larger size of the CFS sample) allows the FAF to adopt these CFS zones while maintaining its focus on U.S. coastal analysis zones that both receive and pass on most U.S. imports and exports. This compatibility with the CFS geography should make future development of FAF flow estimates not only less time consuming but also prone to one fewer sources of possible estimation bias.

The FAF^3 ***Mode Classes*** have also changed since 2002. Table A1 below shows the differences. Note that, due to the redefinition and changed reporting of intermodal/multimodal categories between the 2002 and 2007 CFS on which the FAF is based, there is no direct equivalence in the modal classes implied between these two sets of definitions,. Differences in the way the 2007 versus the 2002 CFS assigned water-only versus water-inclusive intermodal shipments (typically, truck-water combinations) also means that direct comparisons of water only traffic volumes and modal shares is problematic.

Table A1. Modal Class Changes 2002 – 2007

FAF2 Modes (2002)	FAF3 Modes (2007)
Truck	Truck
Rail	Rail
Water	Water
Air, air and truck	Air,air and truck
Truck and rail	Multiple modes and Mail
Other intermodal[1]	Pipeline
Pipeline and Unknown	Other and Unknown

FAF^2 "Other intermodal" includes U.S. Postal Service and courier shipments and all intermodal combinations except air and truck.

FAF^3 Modal definitions are given below:

Table A2. FAF^3 Modal Definitions

Mode Identification	Mode Name	Mode Description
1	Truck	Includes private and for-hire truck. Private trucks are owned or operated by shippers, and exclude personal use vehicles hauling over-the-counter purchases from retail establishments.
2	Rail	Any common carrier or private railroad.
3	Water	Includes shallow draft, deep draft and Great Lakes shipments.
4	Air (includes truck-air)	Includes shipments typically weighing more than 100 pounds that move by air or a combination of truck and air in commercial or private aircraft. Includes air freight and air express. Shipments typically weighing 100 pounds or less are classified with *Multiple Modes and Mail*
5	Multiple Modes and Mail	Includes shipments by multiple modes and by parcel delivery services, U.S. Postal Service, or couriers. This category is not limited to containerized or trailer-on-flatcar shipments.
6	Pipeline	Includes flows from offshore wells to land, which are counted as water moves by the U.S. Army Corps of Engineers.
7	Other and Unknown	Includes flyaway aircraft, vessels, and vehicles moving under their own power from the manufacturer to a customer and not carrying any freight, unknown, and miscellaneous other modes of transport.
8	No Domestic Mode	A 'No Domestic Mode' category is used to capture petroleum imports that go directly from foreign, inbound ships to an on-shore US refinery. This is done to ensure a proper accounting when foreign and domestic flows are summed, while avoiding assigning flows to the domestic transportation network that do not use it.

FAF^2 modal definitions are as follows:

1 – 4. Truck, Rail, Water and Air (including truck-air) definitions are the same as those used in FAF^3.

5. **Truck-Rail Intermodal**—Shipments that use a combination of truck and rail.

6. **Other Multiple Modes**—Includes Parcel (U.S. Postal Service or Courier), truck-water, and water-rail.

7. **Other and Unknown Modes**—Includes Pipeline and any mode not listed above.

The FAF3 Commodity Classes, like those in FAF^2, mirror the 43, 2-digit (i.e. most aggregate) SCTG classes reported by the 2007 CFS. Differences in the composition of these classes between 2002 and 2007 are relatively minor, with two exceptions:

- Printed product flows, which were absent from the 2002 CFS and hence modeled as OOS flows in FAF^2 were covered in the 2007 CFS.
- A second change for FAF^3 was the O-D specific treatment of natural gas products, which were evaluated only at the level of national or broad regional activity totals in FAF^2.