

BEST PRACTICES FOR INCORPORATING COMMODITY FLOW SURVEY AND RELATED DATA INTO THE MPO AND STATEWIDE PLANNING PROCESSES

Requested by:

American Association of State Highway
and Transportation Officials (AASHTO)

Standing Committee on Planning

Prepared by:

Rick Donnelly
Parsons Brinckerhoff, Inc.
6100 Uptown Boulevard, N.E., Suite 700
Albuquerque, New Mexico 87110

August, 2010

The information contained in this report was prepared as part of NCHRP Project 08-36, Task 65, National Cooperative Highway Research Program, Transportation Research Board.

SPECIAL NOTE: This report **IS NOT** an official publication of the National Cooperative Highway Research Program, Transportation Research Board, National Research Council, or The National Academies.

Acknowledgements

This study was requested by the American Association of State Highway and Transportation Officials (AASHTO), and conducted as part of the National Cooperative Highway Research Program (NCHRP) Project 08-36. The NCHRP is supported by annual voluntary contributions from the state Departments of Transportation. Project 08-36 is intended to fund quick response studies on behalf of the AASHTO Standing Committee on Planning. The report was prepared by Rick Donnelly of Parsons Brinckerhoff, Inc. The work was guided by a task group consisting of Paul Agnello, Daniel Beagan, Ronald Duych, Tom McQueen, Amelia Regan, Rolf Schmitt, and Michael Sprung. The project was managed by Nanda Srinivasan, NCHRP Senior Program Officer.

Disclaimer

The opinions and conclusions expressed or implied are those of the research agency that performed the research and are not necessarily those of the Transportation Research Board or its sponsors. The information contained in this document was taken directly from the submission of the author(s). This document is not a report of the Transportation Research Board or of the National Research Council.

The Transportation Research Board of the National Academies, the National Research Council, the Federal Highway Administration, the American Association of State Highway and Transportation Officials, and the individual states participating in the National Cooperative Highway Research Program do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of the report.

Contents

Introduction	1
Overview of the Commodity Flow Survey	4
Survey Structure and Content	4
State and Urban Applications Involving the CFS	6
Assessments of the CFS	8
Advanced CFS Applications	10
Michigan Statewide Truck Model	10
Wisconsin Statewide Model	13
Freight Flows of Indiana	14
The Freight Analysis Framework	16
Reported Use of the FAF	18
Users Perception and Use of the FAF	20
FAF2 Adaptation Success Stories	22
FAF2 in Urban Freight Modeling	22
Statewide Modeling Uses of the FAF2	29
Key Research Findings	32
References	35
Appendix A: Standard Classification of Transportable Goods (SCTG) Codes	37
Appendix B: Comparison of CFS, FAF, and Transearch Data Products	38

Figures

Figure 1: Example of a first generation FAF flow map	3
Figure 2: Schematic of the second generation Michigan statewide model	11
Figure 3: Truck trip generation in Wisconsin	14
Figure 4: FAF2 domestic zones	17
Figure 5: Assignment of synthetic county-to-county truck flows to the national network	23
Figure 6: FAF2 assignment flows in the Atlanta region	24
Figure 7: Comparison of observed versus estimated external station flows	25
Figure 8: 2002 external trip matrix for the Atlanta region	26
Figure 9: Structure of the SANDAG truck model	27
Figure 10: Original and modified overall growth rates	28

Figure 11: Structure of the Maryland Statewide Travel Model 30
 Figure 12: Assignment of modified FAF2 origin-destination truck data to Maryland 30

Tables

Table 1: Summary tables included in the 2002 CFS state reports 5
 Table 2: Modes of transportation coded in the CFS 6

Abbreviations

BEA	Bureau of Economic Analysis
BTS	Bureau of Transportation Statistics
CFS	Commodity Flow Survey
FAF	Freight Analysis Framework
FHWA	Federal Highway Administration
IO	Input-output (accounts or matrix)
MPO	Metropolitan planning organization
NAFTA	National American Free Trade Agreement
NAICS	North American Industry Classification System
NTAR	National Transportation Analysis Region
PUMS	Public Use Microdata Sample
RMSE	Root mean squared error
SCTG	Standard Classification of Transportable Goods
SIC	Standard Industrial Code
STB	Surface Transportation Board
TRB	Transportation Research Board
USDOT	U.S. Department of Transportation
VIUS	Vehicle Inventory and Use Survey
VMT	Vehicle miles of travel

Introduction

The Commodity Flow Survey (CFS) is an important data program of the U.S. Department of Transportation (USDOT), providing information about the extent and broad characteristics of freight movement throughout the country. The CFS is conducted as part of the U.S. Economic Census, allowing it to reach a selected subset of all American firms. Data are collected on all shipments from the surveyed firm for an entire week in each of the four quarters of the census year. In 2002 data on approximately 2.6 million individual shipments were recorded. These data are used in a variety of federal transportation and energy programs, and aggregate CFS tabulations for individual states have long been used by federal and state policy analysts. What is less well understood is how state departments of transportation and metropolitan planning organizations (MPOs) use the CFS. This study was designed to fill that gap through a review of the literature and practice, highlighting innovative applications of the CFS that others could learn from.

It was imagined at the outset of the research that the CFS was widely used by state and urban transportation planners across the country. Part of that perception was based upon the large amount of interest in the CFS evident among practitioners in the freight community. The author has also used the CFS to build several models that overcome some of its key limitations, and supposed that many others had done the same. With a resurgent interest in freight prompted by its prominent discussion in recent federal legislation, it was thought that use of the CFS had grown nationwide.

Alas, the research failed to find a large number of such users. It is likely that many use the tabular summaries published for each state to support broad policy analyses or document trends. Such casual uses of the CFS products are difficult to estimate. Moreover, such uses would extend much wider than to the relatively small number of users who work full-time in freight planning and analyses, making them harder to identify and reach. Instead, the focus of this research was upon more focused use of the CFS, where the products were regularly used in the conduct of project or policy planning. Such uses were thought to be both more significant and demanding, and likely to encompass the majority of regular users of the CFS at the state and urban level.

A total of 22 state and 15 MPO planners from across the country were interviewed as part of this research to ascertain their use of the CFS. They were not chosen at random, but rather represented those most likely to be involved in the use of these data. They were identified through a number of means:

- TRB staff and members of the oversight committee for the project advanced the names of individuals they knew were active users of the CFS.
- Staff at the Bureau of Transportation Statistics (BTS), the USDOT managers and custodians of the survey, were asked for names of known users.
- The co-chairs of the TRB Task Force on Innovations in Freight Modeling (AT016T) surveyed their membership for leads on agencies known to be using the survey.
- A literature review was conducted to identify research that employed the CFS.
- Participants at a TRB peer review exchange on statewide travel models were questioned about their known use of the CFS in statewide or regional models.

The interviews were conducted by telephone or in person, and sought to obtain insight into the nature and extent of the use of the CFS, and to collect available documentation.

The research revealed some unexpected findings. Many respondents reported a strong desire to use the CFS, but their inability to do so. Some of the limitations most often cited are recounted in the next chapter. Despite these limitations what the survey had going for it was its relatively transparent nature, low cost¹, and coverage of all modes of transportation. Its official status lent authority to the published findings, and some states expressed a desire or requirement to remain consistent with federal estimates of total freight activity in their state. But perhaps the biggest motivation for using the CFS was that it was a singular product. As the only publicly available source of multimodal freight flow data available when this research began it filled a niche it was never designed to, in large part because planners had no alternative source of data to fill their needs.²

Some of these adaptations of the CFS are quite creative, stretching the envelope on how aggregate data can be used for much more disaggregate analyses by fusing them with other data sources and intuition about the underlying industry and economic relationships. Such uses form the basis for the next two chapters. However, such clever uses of the CFS proved more the exception than the rule, While such techniques are probably widely transferable, the fact that there are so few instances of their use to question their broad applicability. If a key goal of the research was to demonstrate the value of the CFS to the states and metropolitan areas the evidence did not appear to support it.

These preliminary findings were reported to the oversight committee and NAS staff in the Spring of 2007. By that time it was apparent that the potential for direct use of the CFS was overtaken by the Freight Analysis Framework (FAF), a FHWA-sponsored initiative initially designed to provide objective information to support the debate about corridors of national significance and reauthorization of the federal transportation legislation. Although not currently used by transportation planners, the original FAF flow maps were widely used in many studies, with informal reports of some users attempting to measure the bandwidth of the flows to estimate truck volumes on specific highways. An example is shown in Figure 1. The measure was doomed to failure, as the bandwidths depicted ranges of volumes rather than scaled flows. However, it underscored the widespread need among users for accessible, credible, and relevant data on freight flows.

Although the funds for the research were nearly exhausted in the conduct of the original scope of work, it was felt that the story of how the FAF can and is being used by state and urban planners is too compelling and exciting to pass up. To the extent that the CFS is an important ingredient in the FAF a story about the success of the latter tells the same about the former. That was the motivation for the original research. Thus, refocusing this work on the practical and useful applications that have been made of the FAF finally tells the story that could not be told when focusing narrowly on the CFS. Chapters 4 and 5 tell that story. Chapter 4 provides a brief overview of the FAF version 2 (FAF2), the current tool at this writing. The technical details of the FAF are not described in detail, as excellent and abundant documentation about it are readily available online. The following chapter describes several innovative uses of the FAF2. The report closes with a discussion of issues relating

¹ As a federal data program the CFS products are free. They can be downloaded from the Internet or obtained on CDROM or printed documents. However, most analysts felt that to say that the CFS was free was a bit misleading. The cost of learning about the data and how to use them was seen as a significant cost. Thus, it might be more accurate to describe the data as “low cost” rather than free.

² The Reebie Transearch data, provided by a commercial vendor, has been widely used as a substitute. Drawing upon the CFS, other publicly-available data sources, and their own market research and surveying, Reebie Associates provides value-added extensions. Examples of how it has been used to augment the CFS are presented later in this report, but its description and evaluation is beyond the scope of this research.

to the transferability of these models, steps to their adoption elsewhere, and suggestions about needs for further research and development in this area.

It must be emphasized that the statewide and urban adaptations of the FAF2 described in this report are but a few of the earlier ones. They are profiled because they are completed works, from which the key lessons learnt can be discerned, and for which enough certainty about their structure and success can be gleaned. There are many more adaptations underway at this writing. Some are quite similar to the completed adaptations reported here, while others are more sophisticated. A brief annotated summary of known work in progress is presented in Appendix B. These stand in contrast to the few comparable uses of the CFS reported herein. No ongoing applications of the CFS by states and MPOs are known. In large measure this amply testifies to the greater utility of the FAF2 for meeting the needs of state and urban planners. While far from complete or offering all of the functionality desired by users, it represents a significant and highly cost-effective advance in capabilities and information for a wide spectrum of users.

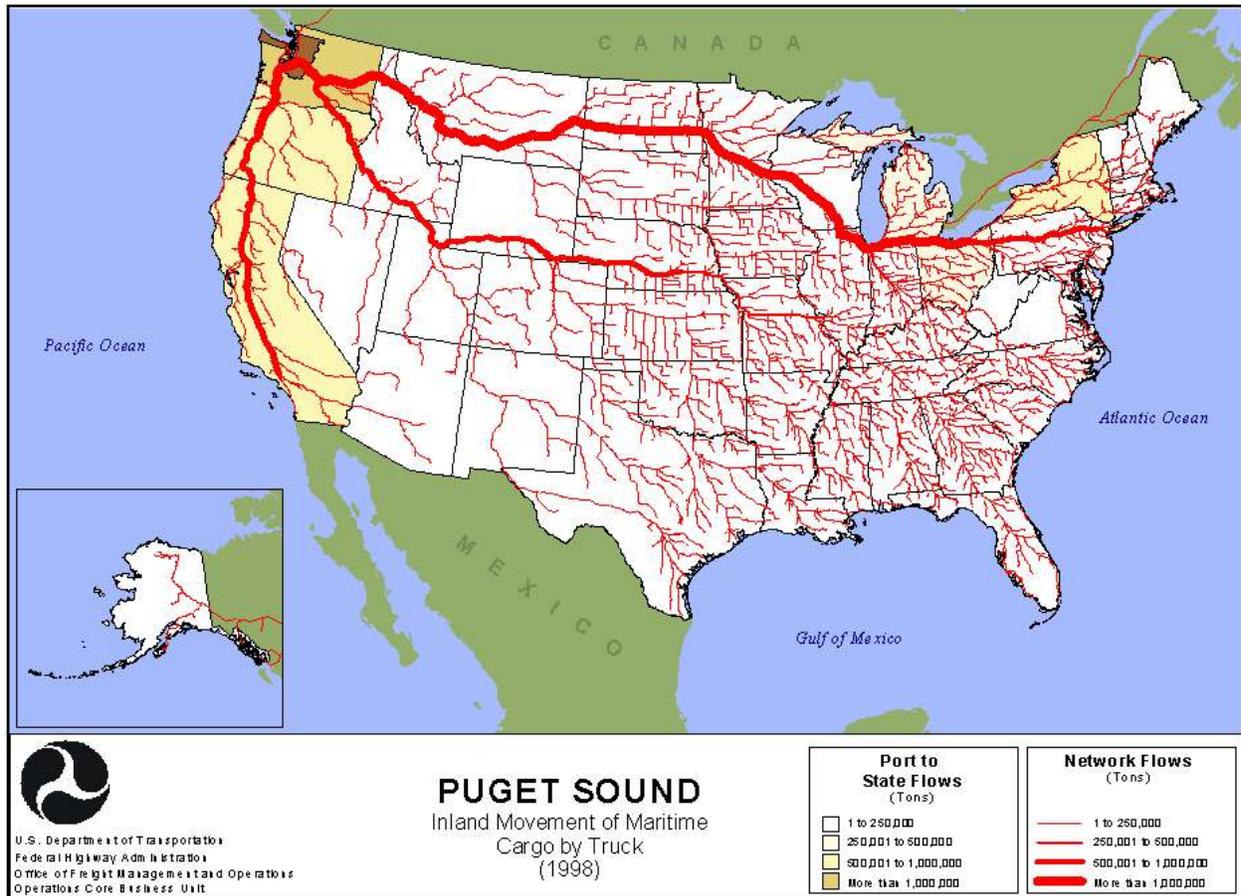


Figure 1: Example of a first generation FAF flow map

Overview of the Commodity Flow Survey

This research originally set out to describe how state departments of transportation and metropolitan planning organizations used the Commodity Flow Survey (CFS) in their planning and programming activities. The CFS is the result of long-standing collaboration between the U.S. Department of Transportation and the Bureau of the Census. Funding for the survey has traditionally been contributed by both agencies, with the majority coming from the USDOT. Since its resurrection in 1993 the Bureau of Transportation Statistics (BTS) has directed the USDOT involvement in the program.

The CFS is conducted in conjunction with the U.S. Economic Census, which is conducted twice each decade in years ending with 2 or 7. While the Economic Census dates back to 1954, and migrated to its current five-year schedule in 1967, the CFS has only been regularly conducted since 1993. Data for each survey generally become available two years after completion of each CFS. The results of the program are used in transportation and energy planning analyses by a variety of federal, state, and local government agencies, as well as the private sector. It is important to note that the primary goal of the CFS is to supply the federal government with aggregate statistics about the movement of goods across the nation. Those needs do not require a great deal of spatial or temporal detail, making the CFS ideal for its intended purpose. That expediency, however, greatly reduces the utility of the CFS as a standalone product for state and urban transportation planners and modelers.

A brief description of the CFS is presented in the following section. It is not a detailed description, as those details are adequately described in the CFS documentation and in the recent CFS conference report (Hancock, 2006). The CFS is also described at length in several recent NAS publications and projects, to include:

- TRB Special Report 276, “A concept for a national freight data program” (2003)
- NCHRP Synthesis 358, “Statewide travel forecasting models” (2006)
- NCHRP Synthesis 384, “Forecasting metropolitan commercial and freight travel” (2008)
- NCHRP Report 606, “Forecasting statewide freight toolkit” (2008)
- NCFRP 12, “Specifications for freight transportation data architecture” (in progress)
- NCFRP 20, “Guidebook for developing sub-national commodity flow data” (in progress)

This research was not intended to recount the work or findings presented in those reports. Instead, the brief summary is followed by a synopsis of its applications reported by those interviewed and the literature. Several recurrent findings and opinions about the CFS are presented in the following section. Several of the more sophisticated modeling applications of the CFS are presented in the next chapter.

Survey Structure and Content

The Commodity Flow Survey (CFS) is normally conducted every seven years as part of the U.S. Economic Census. The first CFS was carried out in 1977. An effort was made to reorient the survey from a shipper to a carrier focus in 1983, but the survey was abandoned while in progress due to

unacceptably high error rates and presumably uncorrectable survey design flaws.³ The CFS resumed with a shipper focus in 1993, with subsequent surveys in 1997, 2002, and 2007. Data from the 2002 survey were the most recent available at the time of this research.

The CFS collects detailed data on shipments from selected establishments nationwide. They are principally manufacturing firms, although certain mining, wholesale, and a limited number of retail sectors are included as well. Because the CFS is part of the Economic Census participation is mandatory. The firms are selected from a statistical register of all such firms in the USA maintained by the U.S. Bureau of the Census. In 2002 approximately 760,000 such firms were included in the registry. A three-stage stratified sampling process was used to select the 50,000 firms included in the survey. Unfortunately, the sample size has been decreasing steadily over time. This is likely at least partially the result of budgetary constraints, although it has been asserted that the smaller 1997 sample size was more statistically efficient than the twice large 1993 sample. In 2002 the 50,000 sampled firms represented only 6.6 percent of the estimated population. In the 1997 survey 100,000 firms were sampled, or 13 percent of the 770,000 firms in the registry. In 1993 about 200,000 firms were surveyed, representing 25 percent of the 790,000 firms in the sample frame. It is encouraging to note that the 2007 survey sampled about 100,000 establishments, back to the level in 1997.

A mail-out, mail-back survey has been used to collect the data since 1993. Respondents are asked to record data for one full week in each of the four calendar quarters of the survey year. All outbound shipments from the surveyed firm by all modes of transportation, including multiple modes, are included in the survey. It was estimated that approximately three million individual shipments were captured by the survey in 2002.

The results of the CFS are available as a series of printed reports and limited data, available for download from the BTS website and on CDROM. In the past the latter has included some summaries not available on the web, such as origin-destination tables. The reports contain the tabulated summaries at the state and national levels shown in Table 1. Tables summarizing the reliability of the estimates are also included in the reports. Some limited data are also published for larger metropolitan areas, but a substantial amount of those data are masked. It is not easy, and in some cases not possible, to determine whether no firms producing a given commodity existed within the metropolitan area or that none that did were surveyed. Federal researchers have access to the survey data, but they are unavailable to outside researchers.

Table 1: Summary tables included in the 2002 CFS state reports

Table	Shipment characteristics by...
1	Mode of transportation for state of origin
2	Total modal activity for state of origin
3	Mode of transportation and distance shipped for state of origin
4	Mode of transportation and shipment weight for state of origin
5	Two-digit commodity for state of origin
6	Two-digit commodity and mode of transportation for state of origin
7	Outbound by state of destination for state of origin
8	Inbound by state of origin for state of destination
9	Mode of transportation for state of origin (compared to previous survey)
10	Commodity group for state of origin (compared to previous survey)

³ Limited data summaries and a brief description of the survey and its problems were available as recently as 2005 on the CFS website maintained by the U.S. Bureau of the Census, but are no longer found.

Table 2: Modes of transportation coded in the CFS

Single modes	Multiple modes
Truck (total, for-hire, and private)	Parcel, U.S. Postal Service, or courier
Rail	Truck and rail
Water (shallow draft, deep draft, and Great Lakes)	Truck and water
Air (including truck and air)	Rail and water
Pipeline	Other multiple modes

The original 1977 and 1993 surveys classified commodities using the Standard Transportation Commodity Codes (STCC), a system originally developed by the American Association of Railroads. It was found to not work as well for commodities mostly transported by truck or air, which led to the creation of the Standard Classification of Transportable Goods (SCTG). This classification scheme has been size in the CFS since 1997. A description of the two-digit SCTG codes is shown in Appendix A. The modes of transport are divided into single and multiple modes, with several separate modes in each category, as shown in Table 2.

Data were reported for National Transportation Analysis Regions (NTARs) in the 1993 survey. These regions were aggregations of Bureau of Economic Analysis (BEA) economic areas. A stated preference by users for data the state and metropolitan area level resulted in reporting at those levels in subsequent surveys.

State and Urban Applications Involving the CFS

There have undoubtedly been many attributions or direct use of the CFS tabular summaries in policy analyses, documents, and briefings. It is likely that most, if not all, state DOTs have used the CFS in this manner. For example, Chin & Hwang (2001) compared the tabular summaries from the 1993 and 1997 surveys to document an increase in the share of goods carried by trucking nationwide. Identifying such uses of the CFS was not part of the goals of this research, as it would have required interviewing thousands of professionals in hundreds of transportation agencies across the country.

It was assumed at the outset of this research, based on the adaptations carried out by the author, that many agencies had made more sophisticated use of the CFS. Examples of such uses include using it to build models and other analytical toolkits, direct use in major studies, or as validation targets for locally collected or third-party freight data. However, interviews conducted in 2006 to identify such uses of the CFS identified very few such applications, although some not reported by agency interviews were found in the literature. Most respondents cited the limitations of the CFS for use at the state and local level as significant impediments, although several respondents also cited lack of familiarity with the product, concerns about high sampling variability reported for the survey, and lack of staff capacity.

The limitations of the CFS failed to deter a few determined researchers and practitioners, however. Hancock & Sreekanth (2001) developed factors for converting total tonnage to railcar equivalents using the STB rail waybill sample, and applied them to the 1993 CFS. Krishnan & Hancock (1998) also used the CFS in conjunction with county-level employment estimates in Massachusetts counties to estimate truck flows on state highways. They also devised a process to convert tonnage estimates to truck flows, and allocated origins and destinations attributed to the state by the CFS to areas within it. The resulting origin-destination matrices were assigned and found to be within tolerable limits. Their use of employment within each area as the means for allocating statewide estimates to substate

areas is noteworthy, as it represents the simpler and more straight-forward of the two approaches used to synthetically disaggregate the CFS data. Sivakumar & Bhat (2002) used a similar process to allocate state-level CFS data from Texas into substate areas. The resulting allocations were then used with an intercity traffic assignment model to study interregional flows in Texas.

The second method used to disaggregate the state-level flows involves the use of input-output (IO) accounts or coefficients. This overcomes a key limitation of using employment data alone. While it is sometimes possible to unambiguously associate employment categories with specific commodities on the production end, it is often difficult to differentiate their consumers. For example, basic chemicals (SCTG 20) are often produced by chemical manufacturing (SIC 28 or NAICS 325). However, they are consumed by a wide variety of different industries and households. An IO matrix depicts the level of transactions industries, and between industries and commodities. These are most commonly reported in annual dollar terms, and are either aspatial or for large regions. IO accounts for the U.S. are available through the Bureau of Economic Analysis, U.S. Department of Commerce, and many states have accounts specific to their industrial structure. It is very rare to find IO matrices at the urban or metropolitan level.

Several attempts to construct statewide models using 1993 CFS data coupled with IO accounts demonstrated their relative superiority of using the latter in lieu of employment data. Donnelly (1997) described how using IO make and use coefficients improved the performance of Michigan's statewide freight model. A first prototype simply used employment broken down into 10 categories as the attractor in destination choice, with commodities arbitrarily assigned to each category. The final model used a two-stage allocation based on IO make and use coefficients⁴, with allocation to freight analysis zones based on its share of employment within each sector defined in the IO table. The resulting model, described further in the next chapter, correlated much better with observed flow data. Vilain et al. (1999) described a similar methodology, with a focus upon estimating the potential for new modes of transportation. They used the IO data primarily to define the attraction end of freight trips, with the production end more easily associated with specific industry groupings and related employment estimates.

Similar work was undertaken in Wisconsin during the same period. Huang & Smith (1999) provides an overview of the early work undertaken there. They tested a number of different modeling approaches, to include direct imputation of trip matrices from the CFS database, estimation of a series of gravity models constrained to trip length frequency distributions in the database, and a synthetic matrix estimation procedure that used a limited sample of observed truck counts to adjust zonal productions and attractions. As in Michigan, the derived models required substantial calibration in order to match observed data. Sorratini (2000) describes the first process in greater detail, including how Transearch data were used to adjust the production and attraction rates initially estimated from the 1993 CFS data. Freight attractors were developed using IO coefficients derived from the IMPLAN model, a widely used economic forecasting package. Interestingly, this work does not appear to have been folded into parallel development of the statewide model used by the Wisconsin DOT (Giaino & Schiffer, 2005), which was to be based wholly on the Transearch data.

At the time of this research none of these models were still in use.

⁴ Make and use coefficients that reveal the mix of goods required to produce \$1 of output or consumption, respectively, can be derived from the IO flows. These coefficients are typically used in lieu of the actual flows, as they scale to any level of production and consumption. Hewings (1985) and de la Barra (1989) both provide an excellent description of their typical derivation.

At least one case was reported in the literature of an aborted attempt to use the CFS for statewide planning and modeling. Aultman-Hall et al. (2000) described the analysis of several statewide freight issues in Kentucky, as well as other projects within the state that required data on freight movements. Information about the potential for modal substitution was particularly sought. After evaluating the CFS they decided to use Reebie Transearch data instead, in part because of its ability to depict trip ends by commodity and mode of transport. The CFS, by contrast, reports commodity and mode of transport separately, but rarely together.⁵

No evidence could be found of any attempt to use the CFS at the urban level.

Assessments of the CFS

The CFS has been extensively reviewed by several TRB and NAS panels over the past decade, and has been a key topic at several conferences. These have included:

- A one-day BTS Roundtable on the CFS was held in November 2000 in Washington, D.C. The high cost of the CFS motivated interest in finding alternative methods for meeting user needs. The workshop examined several options for redirecting the CFS.
- TRB and NYSDOT co-sponsored “Data needs in a changing world of logistics and freight transportation,” a two-day conference held in Saratoga Springs (Meyburg & Mbwana, 2002). The conference sought to prompt a first principles examination of what information was required for effective planning and decision-making, and action plans for moving forward on them.
- The NAS convened a committee to maintain the momentum from the Saratoga Springs conference, and to specifically recommend a national agenda and framework for overhauling freight data programs (TRB, 2003a).
- In same year a NAS committee charged with reviewing the BTS data programs produced a set of recommendations for the CFS, based upon interviews with a variety of government and private sector users of the data (TRB, 2003b).
- A Commodity Flow Survey Conference was held in July 2005 in conjunction with the TRB Midyear Meeting. The Proceedings of the conference are perhaps the most comprehensive description of CFS activities published to date (Hancock, 2006).
- TRB convened a conference, “Freight demand modeling: tools for public sector decision-making,” in September 2006 in Washington, D.C. Although the CFS was not the focus of the meeting it was mentioned in several sessions, and its pivotal role in statewide modeling was emphasized (Hancock, 2008).

Several recurring themes have emerged from these meetings. It has been widely acknowledged that the CFS is a national data program, designed to fulfill specific USDOT surveillance and reporting requirements. It was never intended to be used as a resource for statewide and urban planning and modeling, or at least not below the level of individual states. However, with a dearth of reliable public data on commodity flow movements at any scale of geography several attempts have been made to use the CFS to fill knowledge and information gaps it was never intended to. Perhaps not surprisingly the results have been unsatisfactory in many cases.

⁵ Tables of total originating shipments by two-digit commodity and mode of transport are available from the CFS summaries (Table 6 in the 2002 publications), but many values are either masked or too small for reporting. Moreover, the destination of shipments is not reported.

The inability to obtain data from the CFS at the substate level has been acutely felt by all users of the data. Data are available in many states for the metropolitan area(s) and the remainder of the state. Despite this improvement, several respondents reported that the survey was not useful to them in any way because of this limitation. Several factors undoubtedly influence the scale at which the data are reported, to include limitations of sample sizes, incomplete industry coverage, a dwindling response rate⁶, and federal requirements for reporting them at a level that preclude the identification of individual firms. With small samples it is difficult to differentiate sampling error from the inherent variability in the data, and only the incidence (but not magnitude) of small levels of activity⁷ is reported. This is viewed as an insurmountable obstacle for using the CFS at the metropolitan level, and for anything other than statewide reporting at the state level.

There has been widespread concern voiced about the continually decreasing sample size used in the survey. The 2002 survey included half as many firms as the 1997 survey, which in turn was half the size of the 1993 survey. The perceived loss of accuracy and coverage aside, there are some subtle implications of an ever-smaller sampling frame. The number of masked entries or those below one unit of measurement will likely increase as the number of surveyed firms decreases. It is likely that the effect is not linear; a 50 percent reduction in surveyed firms is likely to lead to an even larger number of masked or indeterminate data points. No systematic study of this effect of smaller sample sizes has been published, but is of great concern to users of the data.

A number of users have advanced the case for release of microdata, comparable to those reported for the Vehicle Inventory and Use Survey⁸ and Public Use Microdata Sample (PUMS). For the most part federal officials have resisted the idea, citing the high likelihood that state of origin or destination would need to be masked in such data. Even an inadvertent breach of respondent confidentiality might threaten the entire program, a risk that all concerned are understandably keen to avoid. The cost of preparing such a product, to include the lengthy negotiations likely over disclosure avoidance, has been cited as a concern as well.

From a pragmatic standpoint any remaining limitations of the data are somewhat academic, as the aggregate nature of the data and the small sample size preclude directly using the data for most analyses other than comparing and contrasting the tabular summaries (e.g., showing the growth in mode shares or ton-miles across successive surveys, or showing changes in overall commodity mix). These are unquestionably important uses of the data, and suffice for some enquiries. However, a large number of users were found to be eager to obtain origin-destination data by mode and commodity, and to be able to generate customized summaries. The solution in all cases has been to synthesize results by using other data in conjunction with the CFS. These range from simple allocation of statewide origins to counties or traffic analysis zones as a function of total or related⁹ employment to indirect use by teasing out relationships that could be used in a larger modeling framework. These indirect uses are the subject of the following chapter, which profiles two innovative uses of the CFS by state departments of transportation.

⁶ The CFS is conducted as part of the Economic Census, and its completion is therefore mandatory. The response rate in 1997 was 75 percent, and fell to 70 percent in 2003. The rate is expected to be even lower in 2007 due to only using mail-out, mail-back survey forms.

⁷ The incidence of less than one (including zero) unit of measurement is reported as a dash in the CFS reports. One unit of measurement is typically \$1 million in value, one million ton-miles, or 1,000 tons.

⁸ The VIUS was discontinued in late 2005 due to budgetary cutbacks at the Bureau of the Census. Data from the two most recent surveys (1997 and 2002) are still available on the Census Bureau website.

⁹ Some authors reported using employment in the same industries as the CFS covers. This includes the manufacturing sector as well as selected agricultural, mining, wholesale, and retail industries.

Advanced CFS Applications

The most sophisticated uses of the CFS outside of the USDOT were in the development of statewide freight models. Such work in Michigan and Wisconsin was based upon data from the 1993 CFS, which became available to users in late 1996. In both cases the CFS data were used to estimate functions that described the generation and attraction of commodities, allocation to destinations, and conversion to vehicle equivalents. Both built upon methods largely pioneered by Polenske (1974) and Roberts et al. (1977). While similar in some respects, the two applications – considered to be the most advanced uses of the CFS in statewide modeling – are informative in many aspects. More importantly, both were developed for use by their respective state departments of transportation, embodying the functionality required to address current policy and planning topics. The Michigan model is discussed first, followed by sections describing models in Wisconsin and Indiana. The models are then contrasted, and lessons learned from their development and applications are presented.

Michigan Statewide Truck Model

The Michigan DOT was one of the first states to undertake statewide travel modeling, with several reports published in the late 1960s found in the NTIS database. Work started on their second generation model in 1993, and substantially completed by the end of 1998. The multi-phase development of a freight model was specified in 1994 as part of that work. It recommended that a truck model be built initially, based in part on the 1993 CFS, with later versions expanding to include other modes of transport, an explicit mode choice model, and data from the 1997 CFS. The project illustrated the potential for using the CFS in statewide modeling, as well as the limitations imposed by the data and possibilities for using it in conjunction with other freight and economic data.

A bi-level representation of geography was used in the model. The person statewide travel model was built on a system of 2,391 traffic analysis zones. Freight data were not available at that level, so a more aggregate commodity analysis zone (CAZ) structure was defined that included counties within Michigan and in the adjacent part of surrounding states, then Bureau of Economic Analysis (BEA) areas, and finally Census regions towards the west coast. Canadian provinces were included, as was Mexico. The 1993 CFS was directly used in several ways:

1. Trip generation rates, measured in annual dollars terms, were derived for each commodity as a function of the gross state product by industry. Input-output make and use coefficients were used to map commodities to industries and vice-versa. The derivation was complex, attempting to equate millions of dollars of goods shipped per thousand employees in each of the generating sectors.
2. Average trip lengths for for-hire and private trucks were used in conjunction with similar data gleaned from a series of truck intercept surveys conducted at several locations across the state in 1994-95.
3. The ratios of dollars to tons reported in the data were used to convert the generated annual dollar flows into tonnage equivalents. A separate step (and data sources) was used to convert tonnage to daily truckload equivalents.

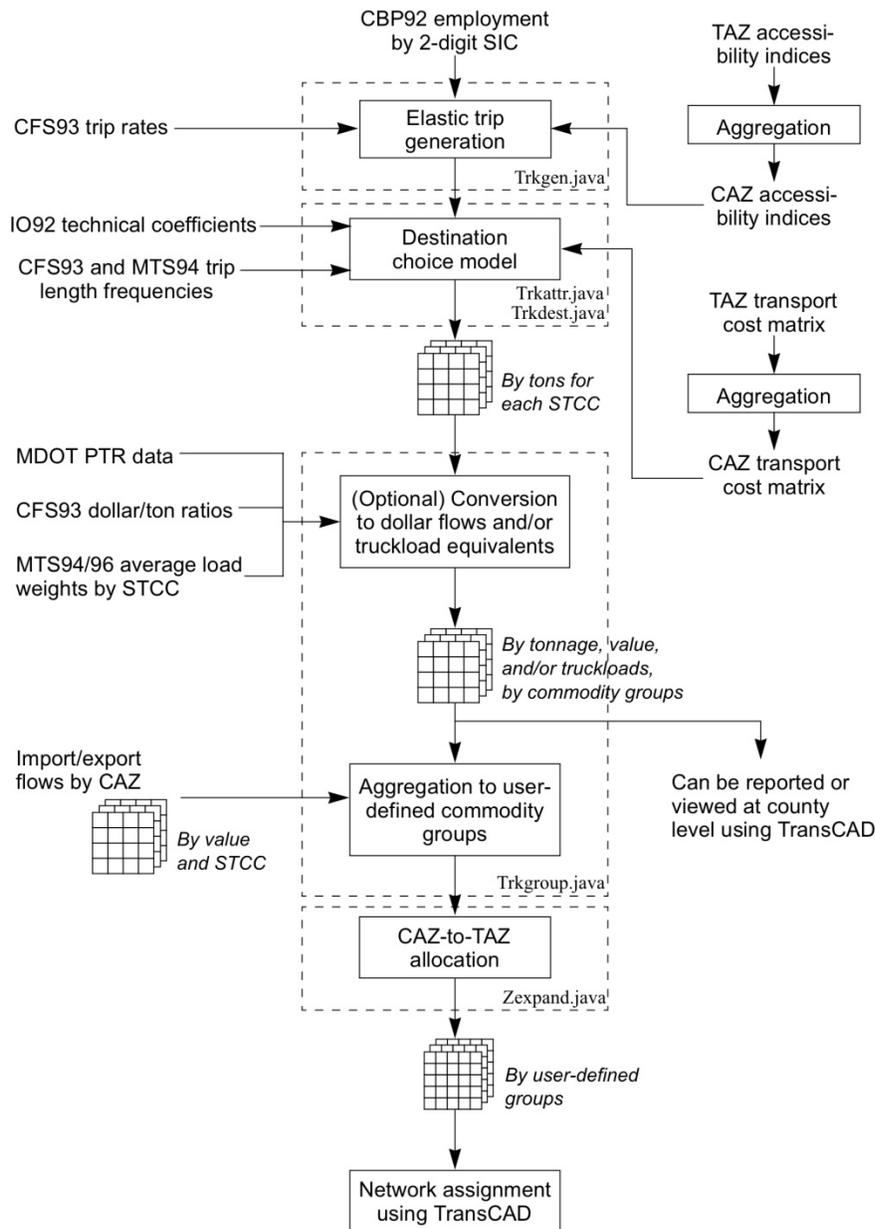


Figure 2: Schematic of the second generation Michigan statewide model

The overall flow of the model is shown in Figure 2. Trip generation and destination choice was carried out at the CAZ level. Trip generation was carried out for the entire state as a whole, and then allocated to counties within Michigan on the basis of each county's share of overall statewide employment in each sector and the percentage of ton-miles carried by truck reported in the CFS.¹⁰ A

¹⁰ These static mode shares were used for both the base and future years, as the 1993 survey was the first one completed since 1977, and no reliable information about how mode share was changing over time were available. Today trends can be spotted across the several surveys that have been completed in the intervening years since this model was completed.

similar process was used for CAZs outside of Michigan. Because the CAZ system covered all of North America a separate external trip model was unnecessary.

A destination choice model was used to allocate the generated trip origins. Most destination choice models are mathematically equivalent to a singly-constrained gravity model. This was an appropriate formulation, as the CFS was a shipper (origin) survey. Summaries of trip destinations were not published, and attempts to obtain custom summaries of such were not successful. Trip length frequency distributions were available for some commodities that were well represented in the truck intercept surveys. However, the average trip length reported in the CFS and idealized distributions were used for infrequently observed commodities. An additional term was included in the utility expression to bias the number of trips kept within the state to roughly match truck counts at the state border. The output of this procedure was a set of CAZ-to-CAZ truck trip matrices by commodity.

The final step at the CAZ level was its synthetic allocation to the much finer TAZ system, a step that also involved converting the flows from annual dollars to daily truckload equivalents. The reported ratio of value to tonnage by commodity reported for Michigan was used when such data were reported in the data. However, the Michigan value or tonnage (or both) were masked for many commodities carried by truck. Comparable data from adjacent states were used if reported in the data, but in some cases national averages had to be computed. These factors, along with average load weights by commodity (obtained from the 1994-95 truck intercept surveys) and ratios of daily to annual flows from weigh-in-motion and permanent detectors were, were applied separately for each commodity. The commodities were grouped into fewer categories to mimic the commonly observed combinations of goods in the surveyed trucks.

Several methods were devised to allocate the flows from the CAZ to the TAZ level, which was carried out so that the trucks could be assigned alongside autos in a multiclass traffic assignment. Because some commodities were shipped relatively infrequently the trucks were typically randomly assigned to a TAZ within the origin and destination TAZ. Not shown in Figure 2 is the treatment of international trips. Detailed foreign trade statistics were used to generate flows between Ontario and Michigan. These trips were coded at the state level, and so were allocated first to CAZ and then TAZ using the same method described above for domestic trips.

The flows from the model were compared to truck counts on federal and state highways within Michigan. Flows on Interstate and other major highways compared well with observed data. However, flows on lower functional class roadways did not match as well, casting doubt on whether the model could be used for project planning or to support engineering analyses. A direct comparison between the trips generated using this method versus more traditional truck trip generation models¹¹ was not reported, but it seems quite conceivable that the two would diverge considerably given the differences in methodology. The CFS, and therefore models like this one derived from it, did not include all sectors of the economy. Trips made from wholesale to retail establishments, and from there to households, were not included in the CFS. However, they constitute a sizeable portion of trips made in smaller communities, further reducing the accuracy of the model. In some respects it is surprising that the model replicated reality at all at fine levels of detail, given that it was applied at the county level, and was based on trips generated at the statewide level and then allocated to individual counties.

¹¹ Such models typically generate truck trips (often by truck type) as a function of the number of employees per firm or the square footage of their facilities. Fischer and Han (2001) provide an excellent summary of such models.

Wisconsin Statewide Model

A similar approach was used in the development of one of several statewide freight models in Wisconsin over the past two decades. Cambridge Systematics developed their more recent freight model based primarily on the Transearch data (Proussaloglou et al., 2007). However, earlier work reported by Huang & Smith (1999) describes a number of attempts to develop truck demand estimates from the CFS. Their paper gives a high-level detail of the entire modeling process. A more detailed description of how the CFS was used was reported by Sorratini (2000). He went further than the developer of the Michigan model, particularly with respect to fusing the CFS with other data sources. Freight generation rates by commodity derived from the 1993 CFS were used to estimate total productions in tons per employee at the state level. Total tonnage by heavy truck was derived by multiplying total production by the share of tons moved by truck reported in the CFS. The overall process is shown in Figure 3. The commodities were classified using the Standard Transportation Commodity Codes and mapped to 28 two-digit SIC sectors using the IMPLAN model (Deller et al., 1993; Lindell et al., 2006). The total productions were allocated to counties based upon employment by sector from the County Business Patterns. In these aspects the Wisconsin approach mirrors Michigan. However, the similarities stopped there.

The productions were allocated to traffic analysis zones within each county based on population, as the researchers had doubts about the accuracy of employment data by sector at that level. They then used Global Insight (then Reebie Associates) Transearch data on average tons per truck to derive annual truckload equivalents, which were then converted into daily estimates by dividing by 312. The latter was derived by assuming that trucks operate six full days per week.

The freight attraction model was, in the words of the author, “not as straight-forward as deriving freight productions.” Input-output accounts from an IMPLAN model implemented in Wisconsin were used to derive attractions for businesses and households as well as freight distribution centers. Input-output make coefficients (the inputs required to produce a given unit of output of a given industry) were derived for each of the sectors used in trip production. Each industry consumes a different mix of goods (commodities by two-digit STCC code in this case) in order to generate its output. The total productions by commodity generated in the previous step were added to imports to generate the total attractions at the state level. The total attractions were balanced to total productions by commodity, transforming them from annual dollars terms (in the input-output matrix) to total tons.

The statewide attractions by commodity were allocated to the county level using the CBP data. The share of total employment in the consuming industries in each county drove the allocation process, the same process in used in trip production. And as in trip production, population within each zone was used to allocate the county-level attractions to TAZ. The tons per truck by commodity from the Transearch data was used in the final step of converting annual tons to daily truckloads using the same process described earlier.

A gravity model was used to distribute trips from three purposes (internal, internal-external, and external-internal) among competing destinations. The resulting trip matrices were assigned and compared against observed counts on five screenlines and at 104 individual locations on highways across Wisconsin. The model performed slightly better on Interstate highways than federal or state facilities, and reached 42 percent root mean squared error (RMSE) overall. As in Michigan it generally replicated intercity flow patterns well, although different patterns at different screenlines were observed. In overall terms the two models performed equally well despite their differences.

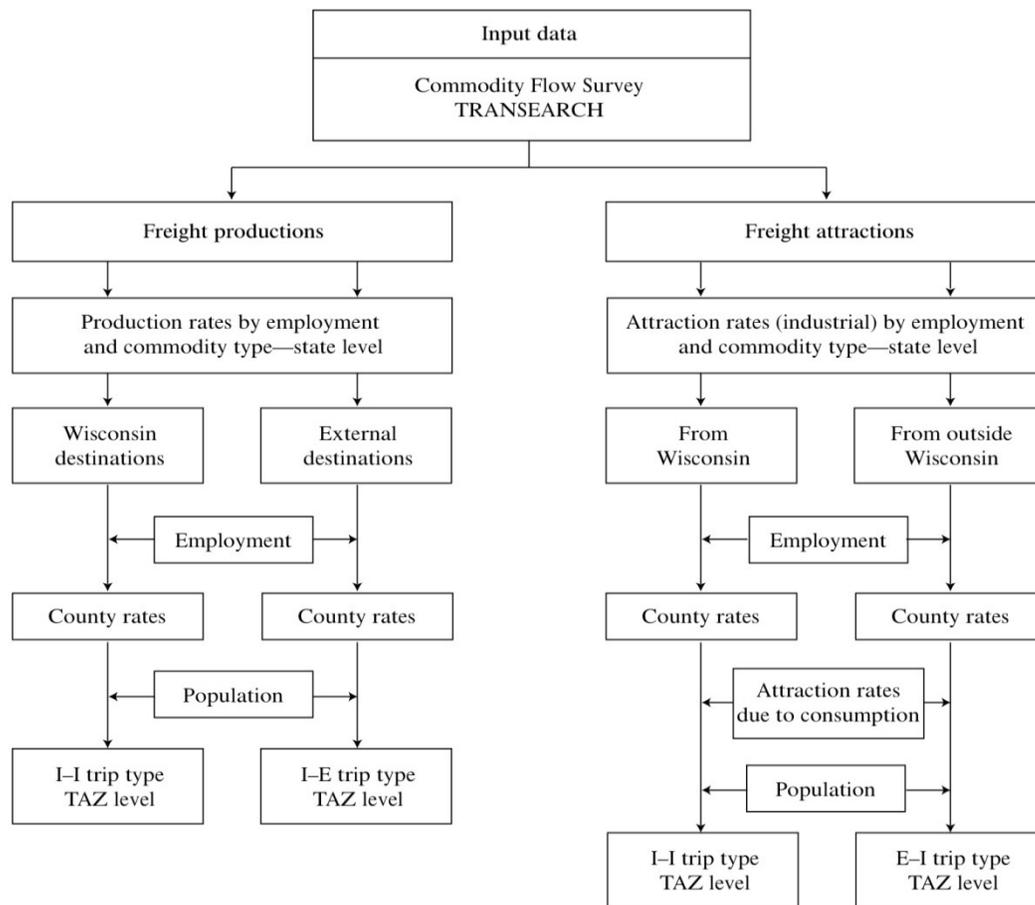


Figure 3: Truck trip generation in Wisconsin (Source: Sorratini, 2000)

Freight Flows of Indiana

At least two generations of freight models built partially with the CFS have been undertaken in Indiana. The first sought to build a commodity flow database for Indiana based upon patterns and relationships found in the 1993 CFS (Black, 1999). Their work on a second generation model is extensively described by Black (2006), which describes how he used the 1997 CFS and other sources to develop a four-step statewide freight travel demand model. The report is as much a diary of the various obstacles and ineffective strategies he overcame as it is a comprehensive model description. While the approach described is broadly similar to that developed in Michigan and Wisconsin it is different in many key respects worth recounting here.

The CFS was used in the trip generation, distribution, and mode split parts of the model. Linear regression models were developed for each two-digit SCTG commodity. The models sought to explain annual trip productions at the state level as a function of statewide employment in one or more sectors, and in some instances, population. Each observation represented the state-level total annual tons produced in each state and statewide employment in the industries thought to correspond to the commodity in question. Data for some states were masked in the CFS because of sampling variability, too few shipments (less than 1,000 tons/year) for that commodity, or protect respondent confidentiality. Thus, in many instances less than 51 observations were included in the regression.

The author laments the lack of accuracy in the fitted slope parameters (the intercept was purposely set to zero to avoid negative trips or trips from zones with no employment). Yet two-thirds of the models yielded an adjusted r^2 greater than 0.8, and only for five commodities was r^2 less than 0.5. At first glance these results appear encouraging. Even small errors in this case are significant, however, as the dependent variable (annual tons shipped) is measured in thousands of tons. In some cases a simple ratio of total tons shipped, as reported in CFS, to total employment was used where the regression provided unacceptable or illogical results.

A similar approach was used to develop linear regression models for trip attractions. As with the productions the goodness of fit varied by commodity, but in general the results were not as good as for trip productions. This was attributed in part to the change in industrial and commodity classifications. Black asserts that the correspondence between the older SIC and STCC codes (used in the 1993 survey) were closer than the NAICS and SCTG codes used starting in 1997. Note that he is using a direct mapping of commodities and industries based on their description rather than the IO approach used by others, although he included (presumably using stepwise regression) several industries in the attraction models to indirectly arrive at the mappings. For example, the attraction model for SCTG 40 (miscellaneous manufactured products) was defined as:

$$\text{Attr[SCTG 30]} = (0.000235 * \text{Pop}) + (0.031 * \text{Empl[SIC 321]}) + (0.014 * \text{Empl[SIC 313]})$$

Employment in two sectors – glass production (SIC 321) and footwear manufacturing (SIC 313) were found to be significant factors of attraction.

The production and attraction equations were applied at the county level within Wisconsin, which was a deviation from the use of these relationships at the state level (because the data were reported at that level) in other models.

The CFS reports average trip lengths by commodity and mode, although in many instances these values are masked because reporting the combination of those two attributes, coupled with the state of origin, would reveal shipments from individual firms. However, they are rarely masked for all states, which allowed such values to be obtained for all commodities. The CFS also reports, in a different summary table, the percentage of shipped tonnage by coarse distance intervals. Attempts to calibrate a doubly-constrained gravity model for trip distribution to these distributions were not successful. After trying a variant constraining only the productions better results were achieved by going back to the doubly-constrained model but focusing on data specific to Indiana rather than national summaries.

Data on mode shares by commodity and distance were used to allocate the distributed trip to six modes of transport. The resulting flows by commodity were assigned to a multimodal network, with an emphasis on flows on highways and railroads. They are unique among the applications of the CFS found in their modeling of rail traffic, as all other states focused almost solely on truck traffic.

The Freight Analysis Framework

The Freight Analysis Framework (FAF) was originally designed as a tool for federal policy analyses. The first version was created in 1997-99 using existing “off the shelf” data from both the public and private sectors. Included in the latter were economic forecasts for each state, as well as some third-party products such as the Reebie Transearch data. It represented a proof of concept for a national freight model, as well as providing immediate estimates that could be used by the USDOT.

The second generation of the FAF framework (FAF2) began in 2003, and was initially completed in 2005. It represented a substantial overhaul of the original framework, with the goal of making the model more widely useable and transparent. The CFS became the backplane of the FAF2, in contrast to being hardly used in the FAF1 work. Transparency and extensibility became key goals of the program, which influenced subsequent design and dissemination strategies. One important benefit of this emphasis is that all of the FAF2 products and documentation are available online. The latter includes over 300 pages of comprehensive information about all aspects of the framework, with the exception of the underlying economic forecasts. Data dictionaries and supporting materials are provided as well. The site layout and content are exemplary, and should serve as a model for other USDOT and federal programs.

The FAF is not so much an analytical framework – a phrase that suggests underlying models – as it is a framework *for* analyses. With the exception of the forecasting process there are no analytical methods or products embedded within the FAF. There is no mode choice model included, nor any other policy-sensitive models. As a consequence, existing patterns and trends are extrapolated into the future. This can hardly be construed in a negative light, considering that existing choices and preferences remain invariant in most travel demand forecasts. Moreover, the amount of progress made over the past several years is remarkable nonetheless. Perhaps most importantly, the FAF can be used with statewide or national freight models in a far easier manner than the CFS alone.

Unlike the CFS upon which it is built, the FAF2 provides a comprehensive database of freight flows rather than limited static tabulations. Two principal databases are produced:

- *Origin-destination flows of freight by commodity and mode of transport*, measured in tons. These are provided at the county level or international gateway for users within the USDOT, and at the zonal level or international gateway for all other users.
- *Estimates of flows on major routes and segments of highways*, by mode of transportation. A range of tonnage is provided at the route level, while truckload equivalents and tons of truck-carried freight are available for segments. Additional attributes are available to USDOT users, such as the commodity and origin-destination data.

These are available for the base year (2002 and provisionally for 2008) and for forecasts in five year intervals from 2010 through 2035. All modes of transport are included except pipeline, with trucking divided into three groups (undifferentiated except by carrier type in the CFS). Like the CFS, commodities are classified using the Standard Classification of Transportable Goods (SCTG) system. Some data from other sources, expressed in either STCC or HS codes, are transformed into SCTG for reporting in the FAF.

One important aspect of the FAF2 is its more detailed level of spatial resolution. For internal use within the USDOT data are available at the county level. The data is substantially more aggregated when released to all other users, who can only obtain information at the scale of 138 freight analysis zones covering the USA and major ports¹². These CFS regions included in these zones are shown in Figure 4. While still more aggregate in nature than many state and local planners need the data, it represents a significant improvement over the CFS. It can be seen that several major metropolitan areas are defined in the zone system. This allows these areas to more accurately deduce the flows entering, leaving, and going through their respective areas. As will be discussed in detail later, this feature has been successfully used by several MPOs.

The FAF2 provides a far more holistic picture of freight flows than the CFS alone. The latter does not include foreign trade, whereas the latter does. An estimate of exports has been published by the CFS team, but the total volume appears quite low when compared to foreign trade statistics. Because the CFS only captures outbound shipments from domestic firms it does not include any movements of foreign imports. As a consequence, data on foreign trade, to include land crossings by truck and rail between the three countries in North America, must be added to the CFS in order to understand the full range of good flowing through the transportation network.

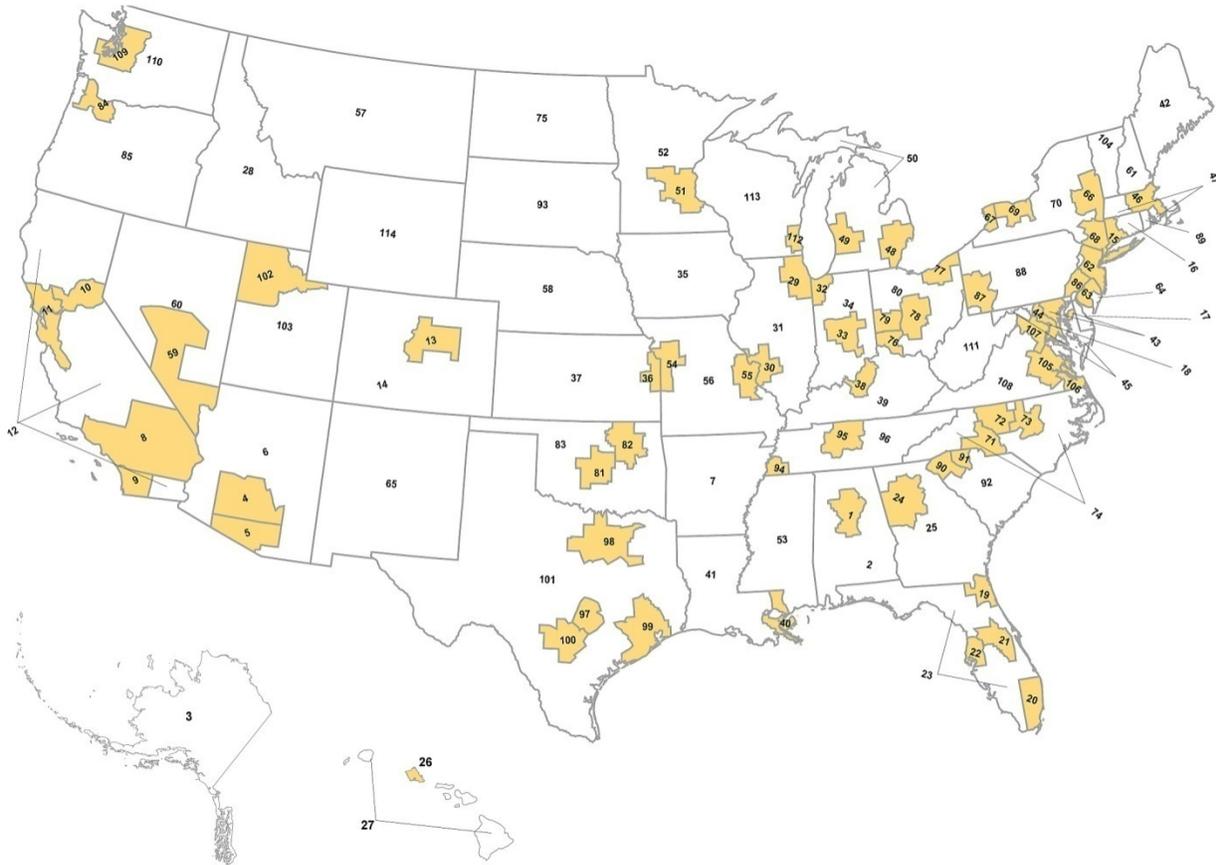


Figure 4: FAF2 domestic zones (Source: FHWA, 2009)

¹² Included in the FAF2 coverage are 114 CFS regions (shown in Figure 4), 17 metropolitan areas that represent major international gateways (land crossings or marine ports), and seven international trading regions.

The FAF2 overcomes this problem by including foreign trade flows reported by the U.S. Department of Commerce. Their foreign trade statistics are a complete recording of all imports and exports, eliminating the uncertainty and noise associated with small sample sizes. These data are currently extracted by the BTS into a monthly series known as the North American Transborder Freight Data.¹³ In the near future it is expected that these data will be replaced with comparable information from the more recent International Trade Data System.

The FAF uses data from a variety of sources to construct its estimates:

- Commodity Flow Survey (both publicly available and microdata)
- Vehicle Inventory and Use Survey (VIUS; discontinued after 2002 survey)
- Carload Waybill Sample data collected by the Surface Transportation Board (both spatially aggregated public use and detailed data available only to government agencies)
- Domestic and international Waterborne Commerce data collected by the U.S. Army Corps of Engineers
- U.S. Air Freight Movements data compiled by the Bureau of Transportation Statistics

The CFS is further divided into “within scope” and “out of scope” data. The former are those data reported by the Census Bureau, as described in Chapter 2. The latter is a series of custom analyses by the researchers working under the direction of the BTS to fill gaps in the CFS. These include sectors of the economy not covered in the CFS, such as construction and services, as well as exports and imports. Collectively these data are used to provide a more accurate and precise picture of freight flows than can be gleaned from the CFS alone.

Both current and forecasted freight flows are available from the FAF. The original FAF2 estimates were based upon the 2002 CFS and associated data. Forecasts based on them are available from 2010 through 2035 in five year increments. A provisional estimate for 2008 is also available. These forecasts are driven by a series of exogenous state-level economic forecasts prepared by Global Insight. Information about the assumptions and methodology underlying these forecasts are not documented, and the economic forecasts themselves are proprietary. Thus, this part of the FAF2 is not as transparent as most of the remaining parts. This has become an issue in recent years, especially after the sharp economic downturn in late 2008. However, many states noticed a decline in economic activity prior to the market crash, and had begun to question the reasonableness of a forecasted doubling of freight traffic between now and 2035. The inability to adjust these forecasts independently of the commodity flows influenced by them has made it all but impossible to lower the forecasts in an informed and rigorous manner.

Reported Uses of the FAF

Considerably more direct uses of the FAF were identified than for the CFS. In some instances these included the extraction of customized summaries of the data, while in other instances the data have been used as an adjunct to or integral part of metropolitan and statewide freight models. Some of the more recent applications include:

Use in Planning and Policy Studies

- The Binghamton (NY) Metropolitan Transportation Study used the FAF2 regional origin-destination data to create summaries of trading partners, imports and exports, and domestic

¹³ Further information about this product can be found at <http://www.bts.gov/programs/international/transborder/>.

flows by commodity and mode of transport for their portion of the NY remainder zone (zone 70 in Figure 4).

- The Maine DOT created summaries of base and forecast year inbound, outbound, and internal flows for the state (zone 42 in Figure 4) by commodity and mode of transport. Several charts and tables of these data were included in their Integrated Freight Plan Update.
- The West Coast Corridor Coalition generated summaries of trade flows through west coast seaports in order to better understand trade and traffic patterns associated with trade with the Pacific Rim over the next 20 years.
- Fresno County (CA) used the FAF2 to validate 2002 truck traffic estimates produced by the Intermodal Transportation Management System. Because Fresno sits in “the remainder of California” (zone 12 in Figure 4) they used both databases at that level. Data from the FAF2 were used to convert value to tons for comparison, and the results were used in Phase III the San Joaquin Valley Goods Movement Study.
- The Texas DOT and Baltimore Metropolitan Council used the FAF2 origin-destination trip matrices as a comparison to Reebie Transearch data used in various studies. In Texas the comparison was carried out as part of a study of the impacts of the North American Free Trade Agreement (NAFTA). In Baltimore it was used to validate Transearch estimates of county-level flows by county, direction, and composition.
- Broward County (FL) identified routes into and from Port Everglades in Fort Lauderdale using the FAF2 origin-destination database.
- The Delaware Valley Regional Planning Commission developed tables from the FAF2 origin-destination database to illustrate domestic and foreign freight flows into the Philadelphia region. The latter included summaries of exports, imports, and trading partners. This information was included in their Freight Facts, used by policy makers and the public to better understand regional freight trends.

Use in Freight Modeling

- The Appalachian Region Commission used the FAF2 regional origin-destination matrices to factor a county-to-county trip matrix synthesized from the FAF1 data. Twelve commodity groups were defined, and flows were converted from annual tons to daily trucks. The resulting demand was assigned to a multimodal network, and further adjusted to match observed counts using synthetic matrix estimation. Forecasts for 20230 and 2035 were also developed.
- The American Association of Railroads developed growth factors from 2002 to 2035 from the FAF2 and then applied them to existing county-to-county trip matrices derived from the STB Carload Waybill Sample. The results were assigned to the ORNL rail network to estimate flows and levels of service in primary rail corridors.
- The Atlanta Regional Commission developed an external trip model for their freight model using the FAF2 origin-destination database. Their modeling region was expanded to match the boundaries of the FAF2 zone it resides in (zone 24 in Figure 4), and flows crossing this cordon were allocated to counties within the Atlanta region.
- FAF2 estimates of international flows and through trips passing through Indiana were included in a post-processor for their statewide model truck component. The freight flows generated complemented the internal freight and non-freight truck flows generated by the statewide model.
- The Florida DOT is developing a methodology for allocating the FAF2 origin-destination flows to regions within the state, and eventually to the county level. The FAF2 will be used

in conjunction with other data to generate the county-level estimates. These will be used to update and validate their statewide freight model, to include derivation of generation and production parameters for internal and special generator trips. Mode split factors will be developed as well.

- The FAF2 origin-destination database was used to develop estimates of internal-external and through truck trips for a truck model of the San Diego region. Particular attention was paid to modeling of truck trips crossing the U.S.-Mexico border.
- A multi-level statewide model was developed in Maryland, where the FAF2 origin-destination database and forecasts are used to model truck flows through, into, and out of the state. As part of this process the flows are allocated to the county level across the country and then assigned to the ORNL freight network to define the entry and exit points of interstate flows entering the modeling area.

Several of these applications are described in the next chapter. In addition to the uses listed, several agencies are in the process of developing FAF2 applications. The Colorado DOT is developing methodologies for using the FAF2 for project-level and regional programming and planning studies. The New Jersey DOT is studying ways to use the FAF2 to update their statewide freight model as well. It is highly likely that additional uses of the FAF2, not yet reported in the literature or known in practice, have occurred as well.

Users Perception and Use of the FAF

Most users of the FAF2 were excited about it and the applications they were able to use it for. Former users of the CFS alone or FAF1 were particularly appreciative, citing improved transparency, greater level of detail, availability of forecasts, and online documentation as significant benefits of the FAF2. Most of those interviewed understood the amount of work that went into the product “behind the scenes,” particularly with respect to reconciling the various data used to construct the databases. Several mentioned their frustrating experiences trying to align the CFS with other freight data sources, particularly the carload waybill data (for rail) and land trade statistics between the three North American countries.

The FAF2 website was frequently cited for its ease of use, completeness, and openness. No negative comments about it were received, although several offered suggestions for improvement. The most commonly mentioned issue revolved around its “one way” flow of information. That is, it functions more as a repository than vehicle for exchange of ideas, user support (of questions and issues), or sharing user-generated products. Some felt that the documentation, while complete, was difficult to approach for new users. A roadmap, tutorial, wiki, or similar mechanism for presenting a high-level summary of the framework would appear useful. The ability to click into or download more detailed documents from this high-level overview seemed widely desired, as was a search function across all documentation. All of this said, it should be emphasized that all respondents viewed these as ways of improving an already useful product rather than deficiencies that distract from it.

The level of spatial resolution of the origin-destination database was frequently mentioned. Users of the CFS alone were particularly critical of this limitation (Meyburg & Mbwana, 2002; Hancock, 2006), as noted in Chapter 2. The level of frustration was not as great with the FAF2, in part because of the increased level of geographic detail that tended to correspond to the counties modeled in major metropolitan areas. Despite this there was widespread desire to obtain OD data at the county level. Knowing that these data were available to USDOT planners increased the disappointment felt by state and local FAF2 users. In fact, several users allocated the data themselves to the county level

using techniques of varying sophistication, a topic addressed in the next chapter. Most acknowledged that the data probably had unacceptable inaccuracy at the county level, but curiously felt they would rather have questionable data than none at all.

While it is thought that long distance freight trips are well represented in the CFS local trips have historically been under-estimated, although the extent of such has never been documented. It has been hypothesized that most of the travel within cities is related to the delivery of goods to retail establishments and households, movement of construction materials, and transport of waste materials. These activities are largely excluded from the sampling frame of the CFS. Because the FAF2 is built upon the CFS the same omissions follow for it. Moeckel & Donnelly (2008) estimated that the FAF2 underestimates local urban trips by a factor of ten. This limits the use of the FAF2 for estimating total urban truck movements or flows within urban areas.

If there was a criticism voiced about the FAF2 it generally surrounded the methods and assumptions used to create its forecasts. Most users did not understand how the forecasts were created and reported little success in locating documentation on the issue.¹⁴ This is particularly an issue at the time of this writing, as many users openly questioned the validity of forecasts developed before the economic downturn in October, 2008. Many felt that a slowdown in the economy was evident even before then, and that such trends invalidated the optimistic growth apparently assumed in the FAF2 forecasting process. Several users reported “just blindly reducing the numbers,” in effect lowering the forecasts by the reduction in growth implied by the forecasts. Others stated that they planned to continue with the FAF2 forecasts, but heavily laden with caveats about expectations that they would likely be revised downward when the third generation FAF was unveiled. A robust methodology for adjusting the forecasts, at least by users outside of the USDOT, does not appear possible. Thus, this shortcoming of the FAF2 probably cannot be overcome until the FAF3 becomes available.¹⁵

¹⁴ Few of the details of the forecasting process are published, as the process is heavily based on proprietary economic forecasts supplied exogenously. Thus, the inability to find documentation about it resulted not from poor website design or incomplete documentation, but rather because such details cannot be divulged.

¹⁵ Version 2.3 of the FAF will be released in the fall of 2009. While ideas and plans for FAF3 were discussed in a FAF Users Meeting earlier this year no official release schedule has been published.

FAF2 Adaptation Success Stories

With over two dozen known uses of the FAF2 outside of the USDOT it is difficult to describe and learn from each in detail. Several completed and operational examples are presented to illustrate how the FAF2 products are being used. These focus upon modeling applications, as the methods and outcomes of creating customized summaries from the FAF2 are apparent to readers.

FAF2 in Urban Freight Modeling

The levels of spatial and temporal detail in the FAF2 are not sufficient unto themselves for use in urban freight models.¹⁶ Even at the county level such estimates would be too coarse for most analyses. Moreover, local flows are absent from the FAF2 database. These are defined as trips of 50 miles or less. This distance is considerably longer than average urban truck trip lengths reported in the literature (Beagan et al. 2007, Holguin-Veras & Thorson 2001, Woudsma 2001). There do not appear to be options for overcoming this omission. While this does reduce some of the functionality of the FAF2 in urban areas it does not preclude it. One very promising area of research and development has focused on using the FAF2 to define external trips for urban truck models. Two examples of such applications, from San Diego and Atlanta, are presented.

Atlanta

Like most other major metropolitan areas, external truck trips in the Atlanta region were generated by matching observed truck counts at the periphery of the modeled area to employment within it. Lacking local data, assumptions were made about the split between internal-external (and vice versa) and through trips, as well as the allocation of through trips entering the region to exit points at other external stations. Trend extrapolation of the count history was used to estimate future volumes at each of the external stations. The base year external truck trip matrix was simply increased by these growth factors to arrive at estimates for the forecast years. Because the external stations served as surrogates for trading partners outside of the Atlanta region, and nothing was known about them other than the volumes crossing them, the resulting external truck model was incapable of informing planners and decision-makers about trading partners, the commodities carried, and other characteristics of external trips. Because external trips comprise a large percentage of the total trucks operating on Atlanta freeways and highways this gap in knowledge precluded a holistic understanding of regional freight patterns or policies based on such.

The FAF2 origin-destination database and factors for converting tons to truckload equivalents were used to develop a more robust and informative external truck trip model. The goal was to produce an external truck trip matrix that matched observed counts at the external stations, which in turn required placing Atlanta within the context of a major city in the Southeast U.S. Zone 24 in Figure 4 was slightly larger than the twenty-county area covered by the ARC travel demand model, but the two areas were close enough in definition to assume they were equivalent. Thus, there was near-perfect correspondence between the FAF2 and ARC definitions of the Atlanta region.

¹⁶ A possible exception might be for modeling flows to and from major marine ports, some of which are represented as a single zone in the FAF2 zone system. If assumptions were made about the precise trip end outside of the port these flows could be assigned to a network and added to automobile and non-port-related truck traffic.

The first step in the construction of the model was to generate flows on the highway within, through, and near the Atlanta region. This required a fairly detailed network, at least within Georgia. The network supplied with the FAF2 was ideal for the purpose, but included far more network detail than the corresponding FAF2 zones. A simple process was devised for allocating the flows to individual counties within the Atlanta region. Employment by county was used to divide the flows destined to or from FAF2 zone 24 to the counties. The resulting assignment results on a national level and into the Atlanta region are shown in Figure 5 and Figure 6, respectively.

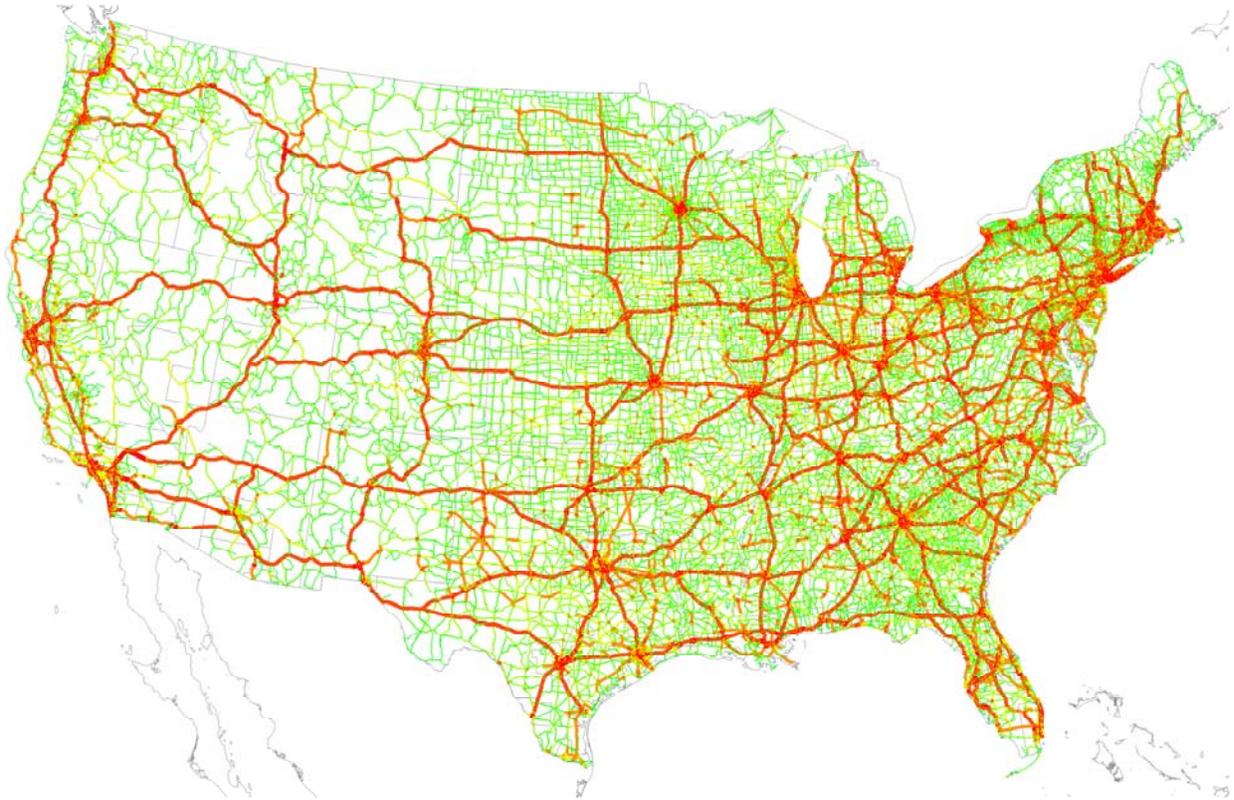


Figure 5: Assignment of synthetic county-to-county truck flows to the national network

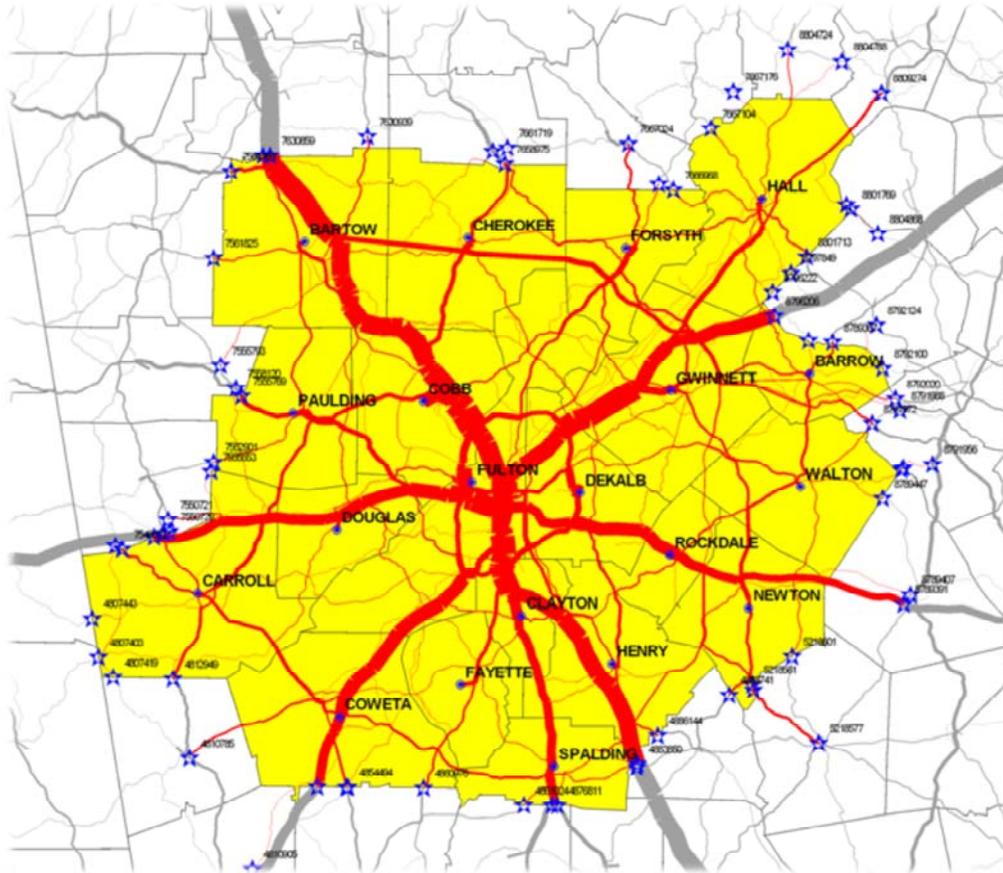


Figure 6: FAF2 assignment flows in the Atlanta region

The shaded counties in Figure 6 correspond to FAF zone 24 in Figure 4. In the external trip model the zone is further broken down in the 24 counties shown. Centroid connectors from the center of commercial activity in each county connect each county to the highway network. This is still too coarse for modeling flows within the Atlanta region. For example, a large number of trucks use I-285 (the circumferential freeway in the center of the region) rather than traveling through downtown Atlanta as shown in Figure 6. However, for the purposes of external trip modeling the distribution of flows most important are those crossing the external cordon, which in this case corresponded to observed data as well as perceptions of local planners about freight patterns in the region.

The flows at this point were still expressed in annual tonnage, the unit of measure in the FAF2 database. Average weekday truck counts at the external stations provided the calibration targets. Factors for converting tonnage to truckload equivalents published by the FAF2 team were used to generate estimates of annual trucks, and peaking characteristics from the traffic recorders were used to convert annual to average daily flows. The initial values of the tonnage-to-truckload conversion factors worked surprisingly well. Only minor adjustments were required to obtain the fit shown in Figure 7. While a closer overall fit might have been possible by further tuning conversion factors for individual commodities no data were available to guide the process, and there appeared to be little to gain from the exercise. Some external stations did not fit well, which was more likely due to imprecision in the allocation of flows to counties or in the origin-destination data themselves.

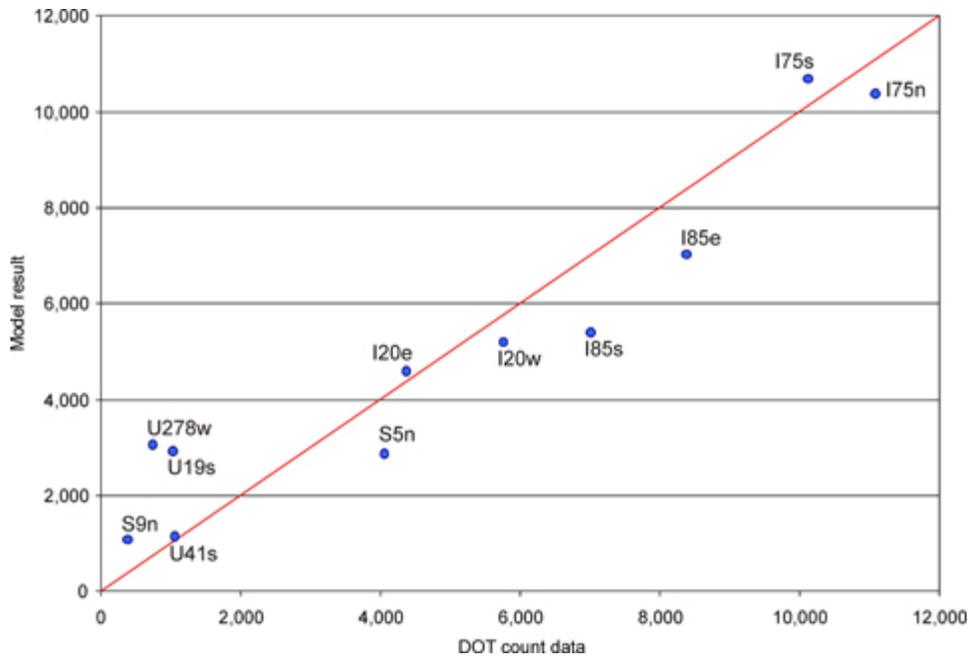


Figure 7: Comparison of observed versus estimated external station flows

Once the flows were converted into daily truckload equivalents the final step of constructing the external trip matrix for the Atlanta regional model was undertaken. Selected link analyses of the external stations were carried out to build this matrix, a subset of which is shown in Figure 8. The trip end within the Atlanta region is still at the county level, which is still too coarse to be directly used in the Atlanta model. This was overcome by reconciling estimates of internal-external trips (and vice-versa) from Atlanta's truck model with the total flows to and from the external stations. The two trip matrices were merged so that the origin or destination within the Atlanta region (at the traffic analysis zone level) were coupled with external destinations instead of external stations (from the FAF2 origin-destination matrices).

Forecasts are constructed in same manner, substituting the FAF2 2035 origin-destination flows for current ones. The same tonnage-to-truckload ratios used in the base year were assumed to apply in the future, although nothing in the process precludes the analyst from making different assumptions. They might adjust them downward, for example, to reflect anticipated shifts to smaller vehicles or payloads, or upwards to capture the effects of more consolidation and empty backhauls¹⁷ in intercity trips.

¹⁷ Empty trip are not explicitly represented in the FAF2. Because the CFS focuses upon shipments made by the surveyed firms the incidence of an empty truck arriving at or leaving the premises would go unreported. Truck intercept surveys would capture the incidence of such activity, but no such source of data is known to be used in the FAF2.

Truck Flow Matrix
Daily Flows in 2002

From	Exit points										Counties						EE Sum	EI Sum
	4807403 STATE ROUTE 5	4807419 STATE ROUTE 100	4807443 STATE ROUTE 166	4812949 US ROUTE 27	8807730 US ROUTE 27	8810626	12770839 FOCKDALE	12770138 NEWTON	12770139 WALTON	12770140 BARTOW	12770881 BARTOW	12770834 BARTOW	12770881 CARROLL	EE Sum	EI Sum			
Entry points	-	-	46.26	-	-	-	-	-	-	-	-	-	-	42.09	23.61			
4807419	-	54.19	-	-	-	-	-	-	-	-	-	-	-	46.26	0			
4807443	-	-	-	-	-	-	-	-	-	-	-	-	-	54.19	0			
4812949	-	-	-	-	-	-	-	-	-	-	-	-	-	703.36	254.65			
4810785	-	-	-	-	-	-	12.14	6.75	6.12	5.58	-	-	-	113.86	559.23			
4810905	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0			
...	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0			
8807730	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0			
8810626	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0			
12770839	0.25	-	-	-	-	-	15.2	14.63	14.57	20.5	19.8	-	-	-	-			
12770138	0.15	-	-	-	-	-	15.2	9.71	9.66	15.59	14.89	-	-	-	-			
12770139	0.14	-	-	-	-	-	14.63	9.71	9.08	15.01	14.31	-	-	-	-			
12770140	0.13	-	-	-	-	-	14.57	9.66	9.08	14.96	14.26	-	-	-	-			
12770842	0.25	-	-	-	-	-	20.28	15.37	14.79	20.66	19.97	-	-	-	-			
12770144	0.44	-	-	-	-	-	29.28	24.36	23.79	29.66	28.96	-	-	-	-			
...	0.26	-	-	-	-	-	20.5	15.59	15.01	14.96	20.19	-	-	-	-			
12770834	1.3	-	-	-	-	-	19.8	14.89	14.31	14.26	20.19	-	-	-	-			
12770881	39.71	54.19	46.26	921.91	0	0	0	0	0	0	0	0	0	0	0			
EE Sum	15.58	0	0	267	0	0	0	0	0	0	0	0	0	0	0			
IE Sum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			

Figure 8: 2002 external trip matrix for the Atlanta region

San Diego

A similar process was applied in San Diego to generate external trip matrices, a task made more difficult because of the immediate proximity of Tijuana to the south and the Los Angeles region to the north. The resulting model also needed to differentiate between several classes of trucks, and to adjust the FAF2 forecasts downward to reflect current economic trends. This work, undertaken for the San Diego Association of Governments (SANDAG), was an adaptation of work recently completed by Cambridge Systematics for the Southern California Association of Governments (SCAG). The decision to build upon the SCAG work was simple, as it involved locally collected survey data and validation to Southern California conditions.

A bi-level model was employed in San Diego, as shown in Figure 9. The traditional approach to modeling urban trucks is shown on the left. As noted, the structure of this model was adapted from the SCAG heavy duty truck model. Daily internal trips made by medium and heavy duty trucks¹⁸ are represented in this level of the model. The FAF2 was used to estimate the characteristics of external trips (those with one or both trip ends outside of San Diego County), as shown on the left-hand side of the Figure. These are converted from annual tons to daily truckload equivalents before being combined with internal (local) trips. The combined trips are allocated to period of the day and included in a multiclass network assignment. The results are fed to EMFAC2007, the emissions model provided by the California Air Resources Board.

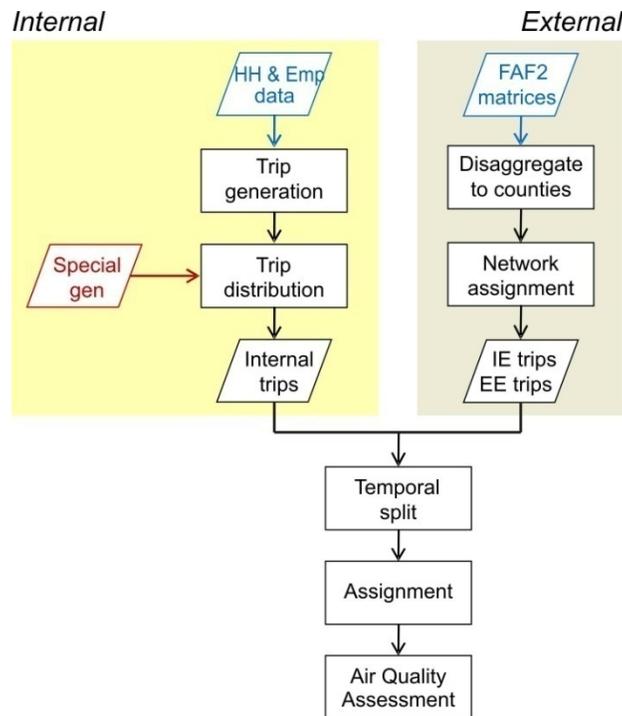


Figure 9: Structure of the SANDAG truck model

¹⁸ Light duty trucks, also known as light goods vehicles (LGVs), were not included in this model. They were also not included in the SCAG model, and no local data were available from which to estimate such trips. Only a quarter to one half of all urban LGV movements involve freight (Holguin-Veras & Patil 2005), and the count data available to SANDAG did not differentiate between freight and other vehicle uses.

San Diego had the same good fortune as Atlanta, in that their model area corresponded to a single FAF2 zone. In addition, the land border crossing at Otay Mesa is another FAF2 zone, as is the Los Angeles region to the north (zone 8 in Figure 4). Because these zones are too coarse relative to the network detail, and specifically routes crossing the San Diego county border, a national assignment of the entire FAF2 origin-destination database at the county level was carried out. Flows were allocated to counties within states on the basis of their share of statewide employment.¹⁹ The flows into and out of San Diego County were translated from annual tons into daily truckloads using a process developed by Battelle (2002).

The resulting assignment produced reasonable flow patterns, but lower than observed. It was determined that the estimates of flows through Otay Mesa were significantly under-estimated in the process. Further investigation revealed that a up to one-half of the trucks crossing the border there are empty, and that others are likely to have smaller loads than domestic intercity truck trips. The latter is due in part to the just-in-time manufacturing of the maquiladora industries and relatively low cost of labor in trucking. The factors for converting tonnage to truckloads were adjusted to match counts at Otay Mesa, which raised the overall flows within the county to acceptable levels of accuracy.

In most other aspects the San Diego adaptation of the FAF2 was carried out the same way as in Atlanta. One significant difference was its use in forecasting. The FAF2 forecasts extend in five year increments to 2035. SANDAG develops forecasts to 2050, requiring extrapolation of the FAF2 forecasts another 15 years beyond its horizon. Plotting the total FAF2 flows between 2002 and 2035 revealed that the forecast assumed an exponential rate of growth of commodity flows (thin grey line in Figure 10). Extrapolating that curve to 2050 would have produced clearly unlikely rates of growth. This growth rate appeared optimistic at the time the model was developed in San Diego, and even more so today. Moreover, it appears to be influenced by an assumption that fuel prices will remain constant beyond 2006 (Battelle 2007), an assumption that appears increasingly unlikely.

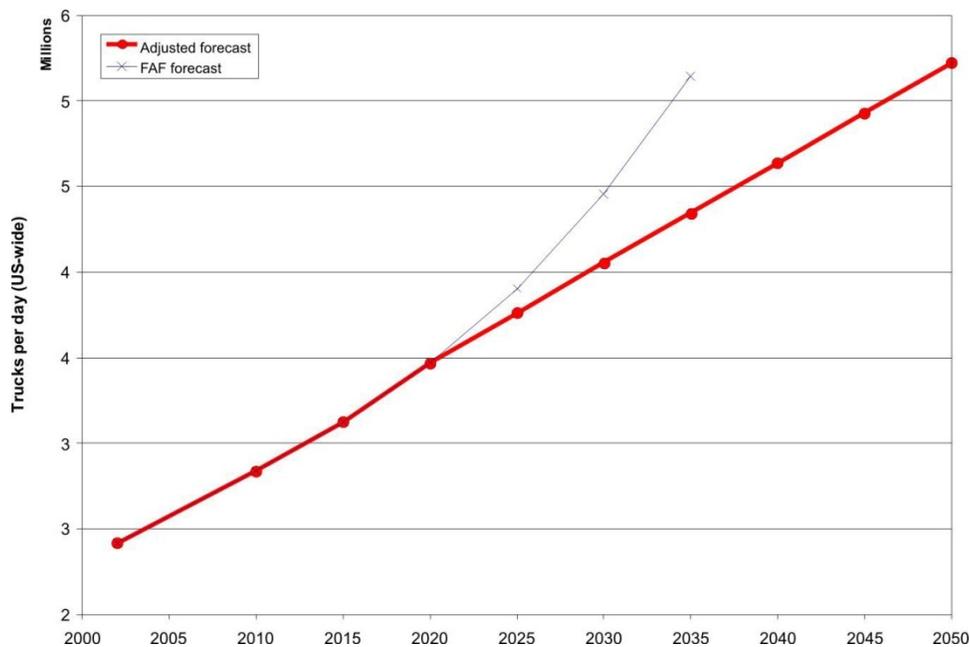


Figure 10: Original and modified overall growth rates

¹⁹ A more sophisticated approach using input-output coefficients to link specific industries and commodities is being developed at the present time.

Because the rate of growth between 2002 and 2020 appears relatively constant the same rate of growth in overall economic growth was assumed. This corresponds to the solid red line in Figure 10. Unfortunately, the assumed rate of growth cannot be adjusted in the FAF2, nor are there several forecast series (e.g., low, medium, and high growth rates) published on the FAF2 website. Lacking insight into the growth rates assumed for each specific region and economic sector the developer simply increased all components of the forecast in the same manner.

Statewide Modeling Uses of the FAF2

Many state DOT planners reported using data from the FAF2 to create charts and tables of current and future freight demand levels, origin-destination patterns, identification of key trading partners, and flows through major corridors. Most of these analyses involved summarizing these values from the origin-destination databases and flow estimates published on the FAF2 website. The resulting summaries were used in briefing papers, corridor and major investment studies, modal analyses (particularly truck-rail diversion), statewide freight plans, and overall statewide transportation plans.

The Maryland Statewide Travel Model (MSTM) was constructed with the FAF2 databases as an integral part. This was motivated in part by the desire to be able to quickly update the model as provisional FAF2 forecasts became available. The overall structure of the model is shown in Figure 11. The overall model is implemented in three geographic levels:

- A regional level covers the entire USA, with county and state level detail in and near Maryland, and progressively more aggregation as distance from the state increases. Long distance freight and person trips are modeled at this level. The FAF2 is used for the former, estimating multimodal freight trips into, out of, and through Maryland.
- A statewide level covers Maryland and parts of adjoining states with 1,607 traffic analysis zones, and is the level at which internal trips are modeled within Maryland. Person and truck travel models based upon comparable structures from the Baltimore and Washington (D.C.) urban travel models were developed at this level. Of particular note in this case is the fact that the truck model at this level generates and distributes internal, internal-external, and external-internal trips.
- An urban level includes the traffic analysis zones, trip matrices, and network data from existing urban and metropolitan travel models within the statewide level. Data generated by those models are available to the statewide model, but they are not run as part of the statewide model.

A first generation of the model was recently completed. The implementation of the FAF2 is similar to the urban models already described, although with some unique twists. Annual tonnage and value by mode of transport and commodity are derived for the FAF2 zone system, and then allocated to counties across the country using a combination of proportional allocation based on employment in sectors included in the CFS and input-output coefficients. Flows by non-trucking modes are included initially, but because there is currently no mechanism for further representing them in the model at this point they are omitted from further analyses. The remaining annual tonnage flows by truck are converted to truckloads by truck type using the procedure described by Battelle (2002), which is the same method used by the FAF2 team in preparation of their database of network flows. The resulting county-to-county trip matrix is then assigned to the FAF2 network to obtain the annual flows shown in Figure 12. These are then scaled to average daily flows using permanent traffic recorder data. The Battelle factors were adjusted, to a maximum of 10 percent from their original values, to bring the assigned FAF2 flows into alignment with observed truck counts at the Maryland border.

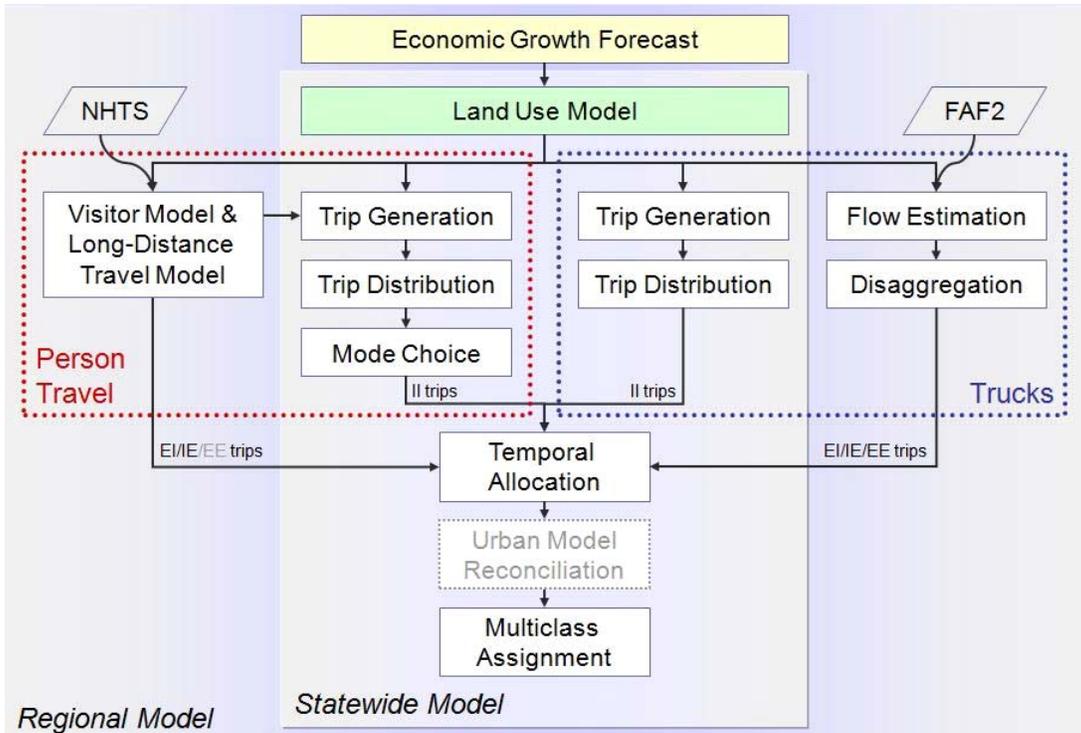


Figure 11: Structure of the Maryland Statewide Travel Model

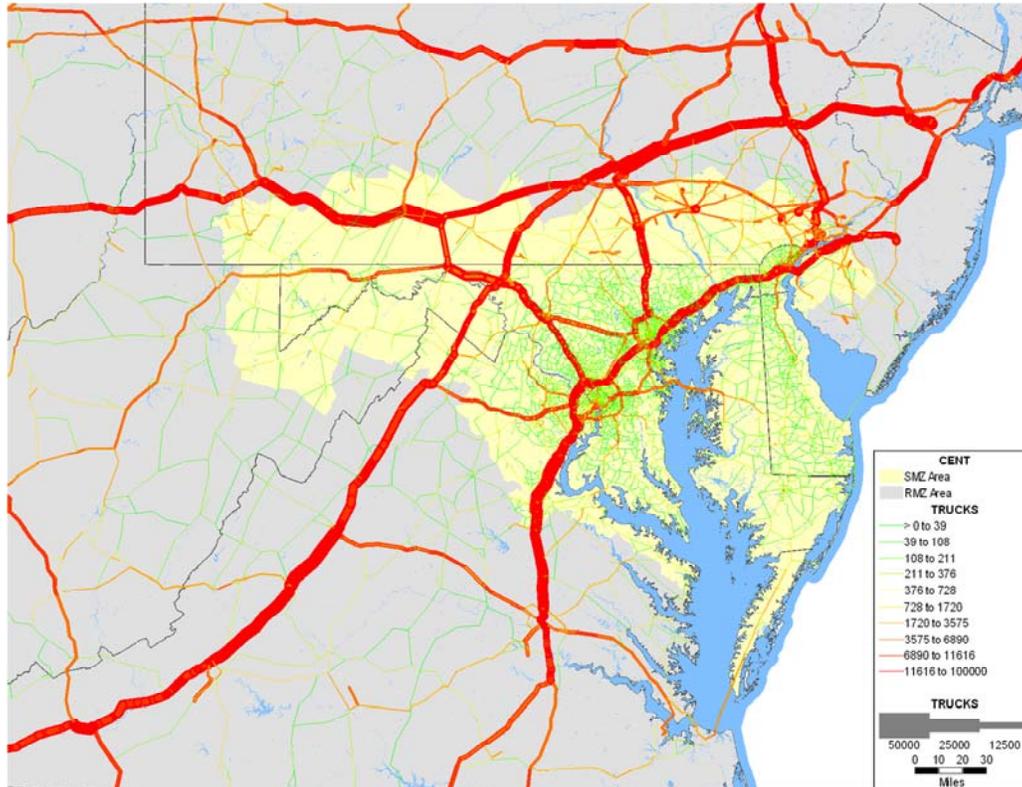


Figure 12: Assignment of modified FAF2 origin-destination truck data to Maryland roadways

At this point the internal-external and external-internal trips generated at the statewide level must be reconciled with comparable flows estimated using the FAF2 adaptation described above. The former have the advantage of being generated at the traffic analysis zone level and being constrained to counts at the edge of the state. The disadvantage, of course, is that nothing is known about where outside of Maryland they are bound to or from, nor is anything known about what they carry. The regional model (FAF2 adaptation) overcomes these shortcomings typically associated with urban external truck models, but at the expense of only being able to identify the county of origin or destination, which in most parts of Maryland are coarser than traffic analysis zones. The reconciliation step attempts to merge the best parts of these complimentary but separate models. In cases where the two models overlap (i.e., flows to or from the same Maryland county to the same entry or exit point at the Maryland border) the traffic analysis zone inside Maryland is simply appended to the origin-destination flows. In some instances several traffic analysis zones within each county are used. In cases where the two models do not overlap the statewide flows are replaced by those from the regional model, and randomly assigned to a traffic analysis zone within the Maryland county end of the trip. In the majority of cases the two models overlap, making the reconciliation straight-forward and consistent.

Once the reconciliation step is completed the flows are allocated to periods of the day and then combined with auto and internal truck flows in a multiclass auto assignment. The validation of the model was carried out using 50 medium and heavy duty truck counts and HERS vehicle miles of travel (VMT) estimates for the modeled base year. While the model validates well across four screenlines across the state its performance at the link level is only moderately satisfactory. Counts at five of the 50 count locations are off by a factor of 4, although most are less than that. Moreover, many come from roadways over which less than 5,000 trucks pass daily, many of which exhibit high variation in the number of daily trucks counted. Part of this might be due to the only cursory level of calibration and validation carried out in the first year of the model, but part of it is due to the inherent variability of freight flows. The former will likely be addressed in the second year of model development, while the latter is unavoidable.

Key Research Findings

The lack of data about freight demand and flows at almost all levels of geography has historically hindered their inclusion in state and local transportation plans and investments. The desire to use federal data sources to fill some of these data gaps has been equally long-standing. However, success stories have been few and far in between. In the case of the CFS the limitations of the data are substantial below the national level. As a consequence, most users abandoned attempts to use it. Only a small number of very determined researchers overcame these limitations, and doing so still involved several assumptions and compromises. As a consequence freight was often neglected in transportation planning, providing a disservice to its users and perhaps limiting the economic potential and competitiveness of cities and regions in the process.

For state and local planners and engineers the FAF2 represents a major improvement in usability over the CFS. This is in part attributable to subtle but significant differences in the two programs. The CFS is a survey program. While as much as 80 percent of the flows represented in the FAF are thought to emanate from the CFS it incorporates and reconciles several other important data sources, includes analytical outcomes such as forecasts, and includes data in trip matrix and network formats that are more immediately useful to analysts. The highly aggregate nature of the CFS, coupled with its availability only in summary table form – and often with large numbers of masked values – rendered it largely useless for urban and metropolitan planning. Its utility for statewide planning was limited to those cases where summaries at only the statewide level were required. Users contacted as part of this research felt that such uses represented only a small percentage of their data needs. Before the FAF those needs went largely unmet, or were supplied by commercial vendors of comparable data. Roughly two-thirds of the users contacted reported using commercial data, particularly the Reebie Transearch data, in place of or in conjunction with the CFS. While its internal workings were not known, the Transearch data were largely perceived as having overcome the limitations imposed by the CFS.

None of the limitations of the usage of the CFS encountered are unique to this research. Most have been reported before in several venues. However, it must be kept in mind that that purpose of the CFS was never to support state or local transportation planning, but to merely inform it in a general fashion. To that end it arguably succeeded. Users can compare data from their states to others, and in some cases learn the about key trading partners. Using the CFS for more advanced or sophisticated applications was rare, and involved synthesizing or inferring significant portions of the data required for more detailed analyses. Moreover, it was clear that the CFS alone could not be altered or extended to provide the freight data so eagerly sought by state, metropolitan, and urban transportation planners.

The FAF2 has been a welcome enhancement of the CFS for understanding and forecasting freight at the state and local levels. It overcomes many of the limitations of the CFS while building upon its strengths, albeit not yet at the level of spatial resolution required by most users. However, by defining regions within most states that correspond to metropolitan areas the data are much more usable, especially for illustrating their linkages with other states and cities. The openness of the FAF2 process and online availability of documentation and data is quite remarkable in quality and extent, and is a model for how federal programs and information could be made readily available to a broad

spectrum of users. In stark contrast to the CFS almost all aspects of the FAF2 can be discovered by examining the documentation, conferring with the FAF2 staff, and through contact with others users. Moreover, the sponsors of the FAF2 program are actively seeking to extend it in several useful ways, and appear to consider user requirements and desires when deciding upon next steps.

An equally important finding is that many MPO and state DOT planners reported using the FAF2 themselves for their analyses, citing the ease of access of the data, excellent documentation, and ability to quickly tabulate the data and express it in network flow terms. This is in stark contrast to the CFS, where consultants and academics completed virtually all of the model development and original analyses associated with its use.

The highly commendable work to date is incomplete, in that several important issues remain unresolved. The lack of documentation and transparency of the economic forecasts used in the FAF was cited by several users as a serious limitation of the framework. While some users reported being worried about the optimistic growth included in the FAF2 forecasts before the economic downturn in October 2008, almost all expressed reservations about their validity afterwards. The forecasts themselves were not necessarily suspect, but the assumptions underlying them were thought to be based upon rosier economic trends from the late 1990s and early years of this decade. It is not clear how this issue can be resolved in the short term. Suggestions included the development of forecasts based upon several growth scenarios (e.g., low, medium, and high rates), the use of an open source forecasting methodology or use of government forecasts such as the Annual Energy Outlook²⁰, and the use of Delphi panels.

Several users noted ways in which the FAF2 website could be improved in the near term. For many the thorough documentation was highly welcome and very helpful. For others the sheer volume was daunting. The lack of a documentation roadmap and hierarchical levels of documentation proved frustrating to some users, especially those who had not used the CFS or FAF1 previously. High-level documentation of the model was considered inadequate by several survey respondents, with some lamenting that it went immediately into great detail. An overview document that explains the derivation, structure, and capabilities of the FAF2 was highly desired by users. Extensive uses of hyperlinks in such a document might enable a user to drill down quickly to the level of detail required. Many also desire a searchable set of documentation, capable of searching the entire site, such as the interface provided by Google search or comparable search engines. A wiki interface for documentation might provide most of the desired functionality. However, most respondents were very quick to admit these were minor issues, and were highly appreciative of the documentation provided.

A final observation concerns the flow of information about the FAF2. Some users expressed disappointment that information only flowed from FHWA to users on the website. There is no provision for user interaction other than ability to find email addresses of FHWA staff members. Some expressed a desire for forums where specific requests for enhancements could be posted, as well as questions and bug reports. A section for reports and products generated was highly desired by users as well. Whether open forums can be accommodated on a federal website is an open question. However, a parallel website (<http://fafusers.org>) could easily be operated by users, and perhaps maintained under small contract with the FAF2 team.

²⁰ The AEO was developed by and is maintained by the Energy Information Administration, U.S. Department of Energy. Further information about it can be found at <http://www.eia.doe.gov/oiaf/aeo/economic.html>.

Despite these current limitations – some of which are being addressed by FHWA as this report is written – the FAF2 remains a remarkable asset for state and local freight transportation planning. It has filled a significant gap for planners and engineers in those organizations, as well as the public policymakers they advise.

References

- Aultman-Hall, L., Johnson, B. and Aldridge, B. (2000), "Assessing potential for modal substitution from statewide freight commodity flow data", *Transportation Research Record*, Vol. 1719, pp. 10-16.
- Battelle Institute (2007), "Forecasts of economic variables that impact passenger and freight demand and the implications of alternative economic assumptions on modal travel demand", Commission Briefing Paper 4B-06, viewed December 18, 2008, from http://transportationfortomorrow.org/final_report/pdf/volume_issue_papers/paper4b_06.pdf.
- Battelle Institute (2002), "Freight Analysis Framework highway capacity analysis: methodology report", Office of Freight Management and Operations, U.S. Department of Transportation.
- Beagan, D., Fischer, M. and Kuppam, A. (2007), Quick response freight manual II, Report No. FHWA-HOP-08-010, U.S. Department of Transportation, Washington, D.C.
- Black, W.R. (1999), "Commodity flow modeling", in "Statewide travel demand forecasting conference proceedings", Transportation Research Circular E-C011, Transportation Research Board, Washington, D.C., pp. 136-154.
- Black, W.R. (2006), "Freight flows of Indiana", Report FHWA/IN/JTRP-2006/27, viewed February 2, 2010, from <http://docs.lib.purdue.edu/jtrp/276/>.
- Chin, S-M and Hwang, H-L. (2001), "In, out, within, and through: geography of truck freight in the lower 48", *Transportation Research Record*, Vol. 1768, pp. 18-25.
- Costinett, P.J. (1995), "Michigan statewide travel demand forecasting model", in Proceedings of the Fifth National Conference on Transportation Planning Methods Applications, Vol. 2, pp. 29-41.
- De la Barra, T. (1989). *Integrated land use and transportation modelling: decision chains and hierarchies*, Cambridge University Press, Cambridge.
- Deller, S.C., Sumathi, N.R. and Marcouiller, D. (1993), "Regional economic models for the State of Wisconsin: an application of the Micro-IMPLAN modeling system", Staff paper 93.6, Department of Agricultural Economics, University of Wisconsin-Madison.
- Donnelly, R. (1997), "Michigan's statewide truck model: revision and preliminary results", in Proceedings of the Sixth TRB Conference on Transportation Planning Methods Applications, pp. 107-116.
- Federal Highway Administration (2009), FAF2 data and documentation: 2002-2035, viewed August 5, 2009, from http://ops.fhwa.dot.gov/freight/freight_analysis/faf/cfs_faf_zones.pdf.
- Fischer, M. and Han, M. (2001), "Truck trip generation data", Synthesis 298, National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C.
- Gaiimo, G. and Schiffer, R. (2005), "Statewide travel demand modeling: a peer exchange", Transportation Research Circular E-C075, Transportation Research Board, Washington, D.C.
- Hancock, K.L. (2006), "Commodity flow survey conference", Transportation Research Circular E-C088, Transportation Research Board, Washington, D.C.
- Hancock, K.L. (2008), "Freight demand modeling: tools for public-sector decision making", Conference Proceedings 40, Transportation Research Board, Washington, D.C.
- Hancock, K. L. and Sreekanth, A. (2001), "Conversion of weight of freight to number of rail cars", *Transportation Research Record*, Vol. 1768, pp. 1-10.
- Hewings, J.G.D. (1985), *Regional input-output analysis*, Sage Publications, Beverly Hills, CA.
- Holguin-Veras, J. and Patil, G. (2005), "Observed trip chain behavior of commercial vehicles", *Transportation Research Record*, Vol. 1906, pp. 74-80.

- Holguin-Veras, J. and Thorson, E. (2001), "Trip length distributions in commodity-based and trip-based freight demand modeling: investigation of relationships", *Transportation Research Record*, Vol. 1707, pp. 37-48.
- Huang, W-J and Smith, R.L. (1999), "Using commodity flow survey data to develop a truck travel-demand model for Wisconsin", *Transportation Research Record*, Vol. 1685, pp. 1-6.
- Krishnan, V. and Hancock, K. (1998), "Highway freight flow assignment in Massachusetts using geographic information systems", *Transportation Research Record*, Vol. 1625, pp. 156-164.
- Lindell, S., Olson, D. and Alward, G. (2006), "Deriving multi-regional models using the IMPLAN national trade flows model", *The Journal of Regional Analysis & Policy*, Vol. 36, No. 1, pp. 76-83.
- Meyburg, A. and Mbwana, J. (2002), "Data needs in the changing world of logistics and freight transportation: conference synthesis", New York State Department of Transportation, Albany, NY.
- Moeckel, R. and Donnelly, R. (2008), "Simulation of intrazonal traffic flows: the end of lost trips", paper presented at the Computers in Urban Planning and Management Conference, Hong Kong.
- Nellett, R., Banninga, G., Johnson, C., Witherspoon, L. and Whiteside, L. (1999), "Organizational and institutional issues in statewide forecasting", in "Statewide travel demand forecasting conference proceedings", Transportation Research Circular E-C011, Transportation Research Board, Washington, D.C., pp. 100-115.
- Parsons Brinckerhoff Quade & Douglas, Inc. (1999), "The second generation Michigan statewide truck travel demand forecasting model", Michigan Department of Transportation.
- Polenske, K. (1974), "Interregional analysis of U.S. commodity freight shipments", Report SP-389, Society of Automotive Engineers, New York.
- Prousaloglou, K., Popuri, Y., Tempesta, D., Kasturirangan, K. and Cipra, D. (2007), "Wisconsin passenger and freight statewide model: case study in statewide model validation", *Transportation Research Record*, Vol. 2003, pp. 120-129.
- Roberts, P.O., Ben-Akiva, M., Terziev, M.N. and Chiang, Y.S. (1977), "Development of a policy sensitive model of freight demand, Phase 1 report," CTS-77-11, Massachusetts Institute of Technology.
- Sivakumar, A. and Bhat, C. (2002), "Fractional split-distribution model for statewide commodity-flow analysis", *Transportation Research Record*, Vol. 1790, pp. 80-88.
- Sorratini, J.A. (2000), "Estimating statewide truck trips using commodity flows and input-output coefficients", *Journal of Transportation and Statistics*, Vol. 3, No. 1, pp. 53-67.
- Transportation Research Board, 2003a, "A concept for a national freight data program", Special Report 276, Transportation Research Board, Washington, D.C.
- Transportation Research Board, 2003b, "Third letter report of the Committee to Review the Bureau of Transportation Statistics' (BTS) survey programs", viewed April 4, 2005, from <http://www.nas.edu/trb/>.
- Woudsma, C. (2001), "Understanding the movement of goods, not people: issues, evidence and potential", *Urban Studies*, Vol. 38, No. 13, pp. 2439-2455.

Appendix A: Standard Classification of Transportable Goods (SCTG) Codes

Group	SCTG	Description
01-05 Agricultural products and fish	01	Live animals and fish
	02	Cereal grains (including seed)
	03	Other agricultural products except for animal feed
	04	Animal feed and products of animal origin, N.E.C
	05	Meat/fish/seafood and their preparations
06-09 Grains, alcohol, and tobacco products	06	Milled grain products and preparations, and baked goods
	07	Other prepared foodstuffs, and fats and oils
	08	Alcoholic beverages
	09	Tobacco products
10-14 Stone, non-metallic minerals, and metallic ores	10	Monumental or building stone
	11	Silica and quartz sands
	12	Limestone, gravel, and crushed stone
	13	Non-metallic minerals, N.E.C.
15-19 Coal and petroleum products	14	Metallic ores and concentrates
	15	Coal
	16	Crude petroleum and bituminous oils
	17	Gasoline and aviation fuels
20-24 Basic chemicals and chemical and pharmaceutical products	18	Fuel oils including diesel and Bunker C
	19	Coal and petroleum products, N.E.C.
	20	Basic chemicals
	21	Pharmaceutical products
25-30 Logs, wood products, and textiles and leather	22	Fertilizers
	23	Chemical products and preparations, N.E.C.
	24	Plastics and rubber
	25	Logs and other wood in the rough
31-34 Base metals and machinery	26	Wood products
	27	Pulp, newsprint, paper, and paperboard
	28	Paper or paperboard articles
	29	Printed products
	30	Textiles, leather, and articles of textiles or leather
	31	Non-metallic mineral products
35-38 Electronics, motorized vehicles, and precision instruments	32	Base metal in primary and semi-finished forms and finished basic shapes
	33	Articles of base metal
	34	Machinery
	35	Electronic and other electrical equip. and components, and office equip.
39-43 Furniture, mixed freight, and miscellaneous manufactured products	36	Motorized and other vehicles (including parts)
	37	Transportation equipment, N.E.C.
	38	Precision instruments and apparatus
	39	Furniture, mattresses and mattress supports, and lamps and signs
	40	Miscellaneous manufactured products
	41	Waste and scrap (except of agriculture or food)
	43	Mixed freight

N.E.C. = not elsewhere classified

Appendix B: Comparison of CFS, FAF, and Transearch Data Products

Most of the states interviewed reported using the Global Insight (formerly Reebie Associates) Transearch data, a proprietary freight database whose widespread use dates back to the early 1980s. An evaluation of the Transearch data was beyond the scope of this research, but its apparently widespread use in place of or in conjunction with the CFS and FAF suggests the need for at least a cursory review of the key differences between these products. Table B1 summarizes the key differences between these products. Neither the author nor the Transportation Research Board and its sponsoring agencies endorse these products, although the author assumes all responsibility for errors in the information presented. Readers are strongly encouraged to consult the developers of these data prior to deciding on the suitability of one or more of them for their intended applications.

Table B1: Ad hoc comparison of three multimodal freight data sources

Attribute	Commodity Flow Survey (CFS)	Freight Analysis Framework (FAF)	Transearch
Brief description	Diary-based survey of shipments from manufacturers, as well as certain agricultural, mining, and wholesale sectors conducted as part of the Economic Census since 1993. Firms are selected from a national register, with sample sizes declining in recent surveys due to funding constraints.	Synthetic database of freight flows by commodity and mode between broad geographic regions. Data are available for 2002 base year, as well as forecasts in five-year intervals from 2010 to 2035. The FAF was constructed using a variety of public data sources, to include the CFS, (rail) carload waybill sample, USCOE waterborne commerce data, and USDOC foreign trade data.	Synthetic database of freight flows by commodity and mode between user-defined units of geography at the county or higher level, available for a variety of historical and forecast years.
Products	Summary tables of domestic and national and state-level freight activity and associated confidence intervals. In some years partial origin-destination matrices are available.	Origin-destination matrices of flows by mode and commodity are available for base and forecast years, as well as a large amount of documentation online. Network data are available, as well as estimates of network flows by mode. Some supporting data, such as value-ton ratios, are also available.	Origin-destination matrices of flows and summary tables by mode and commodity are available for user-specified years.
Developer	U.S. Census Bureau	Federal Highway Administration	Global Insight (formerly Reebie Associates)
Transparency	Moderate: Methodology and assumptions well documented, but data are mostly available only in highly aggregated summaries.	High: Methodology and assumptions very well documented, and data are available at sub-state level in most cases.	Low: It is thought that the database is based upon a fusion of public databases (including the CFS) and private surveys, but the methodology and assumptions are not documented.
Commodity detail	2-digit Standard Classification of Transportable Goods (SCTG) since 1997	2-digit Standard Classification of Transportable Goods (SCTG)	2-digit Standard Transportation Commodity Codes (STCC), with 3-digit detail available in some cases.
Modal detail	17 single and multiple modes	Seven modes (truck, rail, air, water, intermodal, pipeline, and other)	Typically available for major modes (truck, rail, air, water, pipeline), with subcategories of trucks
Time series	1977, 1993, 1997, 2002, and 2007	2002 base year, and forecasts in five-year intervals from 2010 to 2035	Typically available for a variety of user-defined years, which are likely to be interpolated in some cases.

Table B1: Ad hoc comparison of three multimodal freight data sources (Continued)

Attribute	Commodity Flow Survey (CFS)	Freight Analysis Framework (FAF)	Transearch
Spatial detail	Most reports are available at the state level, with some reports available (with considerable masking of observations) for metropolitan areas.	138 zones defining major metropolitan areas, major ports, and the remainder of the state they are included within. Some states represented by a single zone. Also includes seven international regions.	Data are available for user-defined zones consisting of one or more counties. The flexibility of zonal definition is a particularly strong point of the Transearch data, although the methods used to allocate flows to the county level are not disclosed.
Cost	Free	Free	Pricing determined on a case-by-case basis, depending upon level of detail required, number of years included, etc.
Support	Low: Limited support available from Census Bureau and BTS staff via email.	High: Full-time FHWA staff generally provide quick phone and email support, and can obtain help from contractors when necessary.	High: Fully supported by vendor, who typically provides quick support, and can create custom modifications and revisions when required.
Accuracy	Almost impossible to assess aside from what is known about survey size and published statistical accuracy.	In several states flow matrices assigned to the ORNL or national TransCAD networks the estimated daily flows look reasonable and usually follow volume patterns reported in traffic monitoring data. However, in a few cases there have been notable exceptions, where it appears the statewide totals are correct but the directionality is wrong. It is not possible to know whether this is a result of sampling error of the primary data, artifact of spatial aggregation, or the result of one or more models used to synthesize the data. Some users reported that the FAF understates local truck trips by a large amount, perhaps as much as a factor of 10. Thus, the assigned flows are thought to only represent intercity truck trips. No comparable analyses of the accuracy of flows by other modes have been published.	Matrices derived from the Transearch data assigned to the TransCAD networks typically appear as or more accurate than the assigned FAF flows, but in other instances they diverged considerably more. In some cases providing this feedback to the vendor resulted in updates to the data that improved its correspondence to counts. However, because their methodology and assumptions are proprietary it is not possible to know how the flows were adjusted, or if the process was reasonable. Like the FAF data, it appears that the Transearch data do not include most urban truck trip generation, although its inclusion of secondary freight generators (unique to Transearch) appears to result in a smaller error than the FAF. Published analyses of the accuracy of flow estimates for other modes are not known to exist.

Table B1: Ad hoc comparison of three multimodal freight data sources (Continued)

Attribute	Commodity Flow Survey (CFS)	Freight Analysis Framework (FAF)	Transearch
Potential shortcomings	<p>Large numbers of observations are masked due to confidentiality or data quality concerns. Most of the data are only available as summary statistics at the state level. The resulting summaries for truck flows suggest that only intercity trips are captured, which is perhaps not unreasonable given that the CFS primarily targets the manufacturing sector. Flows from wholesale and retail activities, which likely dominate urban flows, are not included in the survey.</p>	<p>The CFS is the only known source of trucking data, which brings their shortcomings into the FAF with them. In addition, the forecasts are based on a proprietary external economic forecast that appeared optimistic even before the economic downturn in late 2008. There is no satisfactory method for adjusting the economic assumptions or removing their effects from the commodity forecasts. Thus, only a brute-force reduction of the flow matrices can be carried out to dampen the assumed rates of economic growth included in the forecasts. Unfortunately, the economic forecast methodology and assumptions are the only aspect of the FAF not well documented.</p>	<p>The lack of transparency in methodology and assumptions precludes understanding how these data differ from the FAF, or what advantages it offers. The use of STCC codes frustrates efforts to directly compare the two data sources. It's quite possible that these data suffer from most or all of the shortcomings of the CFS and FAF, but aside from the same under-representation of local urban truck trips it is difficult to know.</p>