

INITIAL THOUGHTS ON FAF2 EXPERIENCE AND FAF3 DESIGN

Rolf R. Schmitt

November 4, 2008

The Freight Analysis Framework (FAF) is one of the most popular statistical products of the Federal Highway Administration (FHWA). The FAF includes estimates of the weight and value of commodity movements by origin, destination, commodity, and mode for the most recent Economic Census year, 30-year forecasts, and a network database in which tons are converted to truck payloads and assigned to specific routes on the highway network. FAF maps and tables are featured in publications such as the annual *Freight Facts and Figures*, as well as in congressional testimony and in policy studies such as the biennial Status of the *Nation's Highways, Bridges, and Transit: Conditions and Performance*.

The final release of the second generation FAF, based on the 2002 Commodity Flow Survey (CFS), is scheduled for early 2009. This release, version 2.3, will include minor adjustments to the 2002 origin-destination database, which will then be applied to the 2010-2035 forecasts and the 2002 and 2035 network databases. Version 2.3 will also include a distance matrix for estimating ton-miles, major corrections to the 1997 historical origin-destination file, and a web-based tool for creating tables and extracting portions of the origin-destination database.

The third generation FAF will be benchmarked on the 2007 CFS with forecasts from 2015 through 2040. FHWA will compile user feedback on FAF2 and design FAF3 throughout the fiscal year ending in September 2009. FAF3 will be implemented as data for the 2007 base year become available, particularly the CFS in December 2009.

This paper is an initial summary of goals, successes, and shortcomings of FAF2, and an initial outline of goals and strategies for development of FAF3. FAF3 is expected to be a modest improvement over FAF2, rather than the significant evolution between FAF2 and FAF1.

FROM FAF1 TO FAF2

While FAF1 produced very popular and useful maps and statistics, it had major shortcomings. FAF1 was based in part on proprietary data and made relatively little use of data from the CFS. As a consequence, statistics between FAF1 and the CFS were inconsistent, and FAF estimates of commodity flows among areas smaller than states could not be shared with the public. FAF2 was built primarily on the CFS, and all FAF2 estimates were made available for download from the web without charge.

Freight Analysis Framework: Issues and Plans, dated 24 September 2004, guided the development of FAF2. The paper identified six general goals.

1. Integrate more completely the FAF with data from the Economic Census. FAF2 made far greater use of the CFS and the Vehicle Inventory and Use Survey (VIUS) than FAF1. FAF2 estimates of domestic commodity flows required extra steps to extract exports from the CFS file. The Census Bureau has agreed to provide a 2007 CFS domestic file without exports to simplify processing of FAF3. Since the VIUS was discontinued after 2002, FHWA is estimating key VIUS tables for 2007 from trend analyses and correlations to data from related sources.

2. Assure quality of FAF data for the benchmark years. FHWA analyzed several quality questions raised by FAF users and determined that most questionable values could be explained. Questionable values that could not be explained have been adjusted through releases 2.1 and 2.2 of FAF origin-destination data for the base year of 2002, and in forthcoming release 2.3 for 2002 and forecast years. FHWA needs to compile the questionable values and their resolution to assist design of a timely FAF3 quality assurance process.

3. Provide timely updates to FAF data products. FHWA initiated an annual provisional update program to provide previous year estimates of the origin-destination database by March 31 of the current year. Provisional estimates provide value and tons by mode, origin, and destination, but not by commodity. The network database is updated for internal analyses, but the updates are not published. FHWA has been less timely with the speed at which summary tables and maps have been published from FAF2. While three releases of very large FAF2 databases have been posted to the web, the FAF website was still filled with FAF1 maps long after the release of version 2.2. Many useful summary tables developed for internal analyses and in response to special requests have not been posted. The only FAF2 tables and maps published on the web at the end of FY 2008 appear in *Freight Facts and Figures*.

4. Assure that FAF methods and products are transparent and can be reproduced. Transparency of the base case and provisional estimates is supported by thorough documentation and by use of public data sources. FAF2 documentation is comprehensive but complicated and somewhat disjointed. More carefully constructed user guides and technical documentation would be helpful. Forecasts are not transparent, even in the basic assumptions used, because a proprietary system was employed.

5. Help state and local governments make effective use of FAF products in conjunction with local understanding of freight activity. FHWA developed a FAF-based update to the *Quick Response Freight Manual*, sponsored applications of FAF data in several states, supported several modeling and data initiatives through the National Cooperative Freight Research Program (NCFRP), established a web-based clearinghouse for FAF and other freight analysis tools under the Freight Model Improvement Program, and is creating special tabulations for use in California. A web-based tool for extracting FAF origin-destination data without having to download and manipulate the entire file is under

development. Nearly all FAF user support is provided by two individuals at FHWA headquarters who have many other duties, and through FHWA-sponsored courses that are currently out of date. Effective technical assistance requires better use of FHWA field resources, updated course material, and development of a cadre of FAF experts in public agencies and academia.

6. Continue to work with customers to improve the usefulness of FAF products. Key elements of the FAF2 design responded to issues raised by customers through a facilitated workshop of invited power users, a special panel of the Transportation Research Board (TRB), and public discussions such as the Freight Data User Forum at the TRB annual meetings. Ongoing customer questions and suggestions are tracked informally, but have not been catalogued and summarized.

The 2004 document raised a number of specific design issues for FAF2. Most issues are relevant to FAF3.

Geographic coverage and detail. FAF2 covers all flows to, from, and within the US for 114 CFS regions (major metropolitan areas and balances of states), 17 additional metropolitan areas that serve as major international gateways, and 7 international regions. FAF2 does not include through shipments (a.k.a. in-transits) from foreign origin to foreign destination via the US, and the origin-destination database does not provide flows among individual counties. FHWA is initiating work with Canada and Mexico to estimate in-transit commodity flows, and increases in the number of CFS regions in 2007 thanks to doubling of the CFS sample may eliminate the need for identifying additional international gateways. While customers frequently request county-to-county flows, FAF2 estimates and the CFS are limited flows among multi-county regions. FAF2 disaggregates region-to-region flows to the county level as an intermediate step in creating the network database, but the temporary file is not published because flows among individual county pairs have significant errors that generally offset one another once loaded onto the published network database. Increased statistical reliability from the greatly expanded CFS sample in 2007 reopens the question of whether a standard disaggregation method for creating county-to-county flows with a prescribed set of locally collected supplemental data should be developed as an extension of FAF3.

Modal coverage and detail. FAF2 was expanded to include all modes (truck, rail, water, air, and pipeline), as well as two categories of intermodal (truck-rail and other). Tonnage by truck in the origin-destination database includes long-distance and local, but only tonnage moving at least 50 miles is converted into freight-hauling trucks in the network database. "Intermodal" in the FAF is based on CFS definitions, which include shipments by postal and courier services and any shipment using more than one mode. This categorization of "intermodal" is much broader trailer-on-flatcar or containerized service, sometimes leading to confusion. Can better modal definitions be developed within the confines of the 2007 CFS?

Commodity coverage and detail. FAF2 is based on the Standard Classification of Transported Goods (SCTG) at the 2-digit level. The SCTG is used in the US and Canada,

and is based on the Harmonized System that underlies trade statistics throughout the world. SCTG has some comparability issues with commodity classifications based on industry of origin, such as the Standard Transportation Commodity Classification (STCC) system used primarily by railroads; however, SCTG provides a critical link between transportation and trade data, and has more useful distinctions of commodities carried by all modes (compared to the STCC's emphasis on bulk goods that move by rail). The 2-digit level of the SCTG does not break out ethanol and other commodities that have grown in importance since mid 1990's, and it does not provide a direct way to classify flows between hazardous and non-hazardous cargo. Many FAF customers would prize an origin-destination matrix for hazardous cargo, especially by hazard class.

Network coverage and detail. In the FAF network database, commodities are routed over the entire National Highway System, the entire National Network designated for conventional combination trucks, and additional highway mileage connecting freight activity centers. The network database does not identify freight-hauling truck moves between places less than 50 miles apart, nor does it relate commodity flows among FAF regions to individual rail lines, waterways, or pipelines. Early plans for FAF2 to create probability-based assignments of traffic within FAF regions and between adjacent regions to handle local traffic were not realized. FAF2 used the FAF1 strategy of disaggregating flows to counties and selected sub-county generators such as major ports and assigning the flows to individual routes. FAF assignments were matched to truck volume estimates for individual highway segments from the Highway Performance Monitoring System (HPMS), and revealed several quality problems with HPMS data.

Timeliness. In response to customer requests for data more up-to-date than once every five years, FHWA initiated a program of FAF2 provisional annual estimates. A provisional estimate of FAF origin-destination database is created by March 31 for the previous year, adjusting the base year estimates of weight and value by origin, destination, and mode with economic and modal growth factors. Estimates by commodity type are not made. Since methods for the provisional annual estimates are imprecise and evolving, each year's estimates supersede rather than supplement the previous year's estimates. In addition to the annual provisional estimates, a comprehensive and expedited set of FAF summary tables and maps would enhance customer satisfaction.

Temporal variation. FAF2 estimates annual averages, and has no seasonal or daily variation in flows. Network assignment is based on peak period congestion. Routing changes in response to bad weather are simulated through exogenous adjustments to network impedances in applications of FAF data. Temporal variation could be tested with observed data from FHWA's Freight Performance Measurement program.

Shipper and carrier cost as a FAF input and output. Cost of transport is a key variable in evaluating possible shipper and carrier responses to public policies and to forecasts based on changing fuel, labor, and other costs. Collection of requisite public data was discontinued in the aftermath of deregulation, and no strategy has been articulated for obtaining cost data in the future.

Analytical tools. FAF1 was initially intended to be a policy analysis tool. FAF2 has been used to estimate origin-destination patterns of flows through portions of the highway network, particularly for analyses of network disruptions. This internal FHWA capability is similar to the public *Geo-freight* application for CFS data. FAF2 data have been used in policy analyses such as the Highway Economic Requirements System and freight bottleneck studies, providing base conditions and trend forecasts without policy interventions. FAF2 data can also be an input to cost allocation and vehicle size and weight analyses. Special tabulations of FAF2 are under development for California as a key input to air quality and greenhouse gas emission studies. Other applications have been documented in FHWA's update to the *Quick Response Freight Manual*. Additional applications may evolve through the NCFRP.

Passenger travel and general traffic forecasts. The success of FAF2 has inspired FHWA plans for the development of a passenger travel equivalent of the FAF to forecast external and through trips for statewide and corridor planning. The proposed model could also be used to replace or supplement traffic growth forecasts in the HPMS with nationally consistent forecasts.

Public versus commercial data. The shift from commercial to public data for FAF2 base year and provisional estimates resulted in greatly enhanced transparency, credibility, and public access. Transparency will become even more important if Congressional interest in FAF estimates as apportionment factors becomes law. Some customers have recommended that FAF forecasts also be made with open-source, public methods to support sensitivity and what-if analyses. Can open-source methods achieve adequate robustness and sophistication within available time and budget for FAF3 compared to commercial sources? Would development of open-source methods cause unfair competition with the private sector or create needed competition in an arena with few vendors? Do third-party forecasts provide more or less credibility than open-source forecasts controlled by government agencies? Can other federal forecasting systems, such as the Macroeconomic Activity Module of the National Energy Modeling System, provide an effective substitute for the current FAF forecasting procedures?

Feedback. The FAF2 team promised to seek and incorporate feedback from customers. While limited by the small size of the FAF management team and personnel turnover, FAF2 responded to customer feedback with improvements in each of its numbered versions. FAF3 will include outreach to power and novice users to improve the quality and ease of use of FAF products.

PROPOSED FAF3 GOALS

- Produce the FAF origin-destination commodity flow and highway network databases for 2007, forecasts to 2040, and provisional annual updates with equal or improved transparency and quality, maintaining comparability of FAF origin-destination commodity flow estimates for the base years of 1997, 2002, and 2007.

- Enhance access to FAF data through an extensive set of maps and summary tables, and through web-based, interactive methods to create tables and maps on demand.
- Establish statistically reliable, transparent methods for estimating ton-miles by mode for the nation and by state.
- Add local freight and intermodal trucking to the highway network database.
- Expand the network database to include rail and water.
- Establish a hazardous materials origin-destination database.
- Establish procedures for disaggregating FAF commodity flows into statistically reliable estimates of flows among counties based on a prescribed set of locally collected supplemental data.
- Improve linkages between FAF data and freight performance measures.

PROPOSED FAF3 PRODUCTS

- An origin-destination commodity flow database of tons and value (in 2007 chained dollars) by origin, destination, international gateway (if import or export) or gateways (if in-transit), mode, and commodity for 2007, with forecasts for 2015, 2020, 2025, 2030, 2035, and 2040 (or 2040 with no, one, or two other years).
- An origin-destination hazardous materials flow database of tons and value (in 2007 chained dollars) by origin, destination, international gateway (if import or export) or gateways (if in-transit), mode, and hazard class for 2007, with forecasts for 2015, 2020, 2025, 2030, 2035, and 2040 (or 2040 with no, one, or two other years).
- An origin-destination average distance database for estimation by origin, destination, international gateway (if import or export) or gateways (if in-transit), and mode for 2007.
- A highway database with capacity, impedances, tonnage carried, and annual average daily traffic by highly aggregated vehicle types for 2007 and 2040.
- A railroad database with capacity, impedances, tonnage carried, and annual average daily carloads by intermodal, unit, and other trains for 2007 and 2040.
- A waterway database with capacity, impedances, tonnage carried, and annual average daily vessels by highly aggregated vessel types for 2007 and 2040.
- A provisional annual estimate of the origin-destination commodity flow, origin-destination hazardous materials flow, and network databases, with values in both 20007 chained and current dollars).
- A series of national and state tables for FHWA publications and the FAF website based on tables developed for FAF2, and maps for *Freight Facts and Figures* and the *Atlas of Freight Transportation in the United States*.
- A comprehensive set of user guides, training material, and technical documentation.
- An estimate of ton-miles by mode for movements to, from, within, and through each state.

- Procedures for disaggregating FAF flows to the county level with a prescribed set of locally collected supplemental data.

PROPOSED FAF3 STRATEGY AND TIMELINE

FHWA is reassembling the FAF2 team to develop FAF3. FHWA proposes to solicit customer feedback in early 2009, and have draft specifications for FAF3 products and procedures in summer 2009. Procedures not dependent on the CFS will be initiated as plans are finalized, and remaining procedures will be initiated as soon as CFS data are available in December 2009. FAF3.0 is proposed for release in summer 2010 with the 2007 origin-destination commodity flow database, the 2007 highway network database, and initial ton-mile estimates by state. FAF3.1 is proposed for release by December 2010 with forecasts, the rail and waterway network databases, and detailed ton-mile estimates.

Provisional annual estimates will continue to be released by March 31 for the previous year. Provisional estimates for 2008 (released in 2009) and 2009 (released in 2010) will be based on FAF2. Provisional estimates for 2010 (released in 2011) and beyond will be based on FAF3.

As was done for FAF2, FHWA plans to use Oak Ridge National Laboratory and its partners to develop specifications for FAF products and procedures, estimate the 2007 origin-destination database, and assure that other FAF products meet specifications and are integrated effectively. The forecasts, network databases, ton-mile estimates, provisional estimates, web-based data access and mapping tools, and user guides will be developed mostly through contractors who have indefinite delivery/indefinite quantity agreements with FHWA, though some tasks closely aligned with the 2007 origin-destination database may remain with Oak Ridge.

Procedures for disaggregating FAF flows to the county level with a prescribed set of locally collected supplemental data will be pursued on a parallel track to FAF3. The NCFRP solicitation for 2009 resulted in two proposals for closely related work.

SPECIFIC FAF3 BUILDING BLOCKS

Projects initiated in FY 2007-2008 related to FAF3 include:

- TransBorder Freight Data. FHWA provides financial support to the Bureau of Transportation Statistics to maintain the preparation and publication of TransBorder data from the Census Bureau.
- Port drayage estimates. Cambridge Systematics is compiling data from a range of individual ports and will estimate modal and destination shares of inland moves of international trade in 2007 for the FHWA freight office, due in early 2010. The Tioga Group is conducting a related study of truck drayage practices for the NHFRP, due in summer 2010.

- Synthetic VIUS. Battelle is estimating key VIUS tables for 2007 from trend analysis and from correlations with related databases for the FHWA freight office, due in early 2010.
- FAF network improvements. Battelle is linking the National Bridge Inventory and adding other attributes to the FAF highway network database, due in fall 2009.
- Distribution centers database. Battelle is compiling a geographic database of distribution centers for the FHWA freight office, due in fall 2008.
- FAF disaggregation method. Cambridge Systematics is documenting its procedures for the FHWA freight office, due in fall 2008.
- Specifications for a national freight data architecture. Texas Transportation Institute is reviewing definitions and classifications used in the FAF, the International Trade Data System, and other freight data programs to identify potential contents and utility of a national architecture for the NCFRP, due in early 2010.

Projects planned for FY 2009 related to FAF3 include:

- FAF3 design and initial implementation, including FAF2 customer feedback sessions at the TRB annual meetings in January and a facilitated meeting of invited power users in spring of 2009.
- Estimation of in-transit flows for the FAF and the Surface Transportation Working Group of the North American Interchange on Transportation Statistics.

Additional activities needed for FAF3 planning include:

- Compile and evaluate user comments and questions related to FAF2.
- Compile and evaluate all FAF2 data quality issues and resolution of those issues.
- Review FAF2 documentation and propose improvements in completeness and usefulness.
- Identify issues related to provisional updates of the FAF network database, such as availability of timely wayside count data for traffic assignment.
- Identify issues related to establishment of rail and waterway network databases parallel to the highway network database.
- Estimate costs and evaluate the pros and cons of establishing an open-source forecasting system, continuing the current FAF forecasting procedures, or adopting another federal forecasting system for some or all of the FAF and for FHWA's proposed passenger travel analysis framework.
- Determine whether user interest warrants intermediate 5-year forecasts of the origin-destination database.
- Identify methods and challenges for establishing an origin-destination database for hazardous materials flows.
- Review changes in CFS geography between 2002 and 2007 and evaluate the effectiveness of FAF2 foreign trade regions to improve FAF3 domestic and international geography and to develop a plan for maintaining comparability of origin-destination estimates for 2007, 2002, and 1997.

- Identify ways to simplify and improve the estimates of flows covered by the CFS using the larger CFS sample and the special domestic-only CFS origin-destination file.
- Identify potential improvements in methods used to estimate commodities that are out of scope for the CFS.
- Identify changes in FAF data sources and determine effects on their use in the FAF, especially as control totals.
- Review and either revise or eliminate methods used in FAF2 to establish control totals for value, tons, and ton-miles.
- Identify 3-digit commodities that have grown in importance since the 1990s and could be treated separately without compromising disclosure of the remaining 2-digit category.
- Identify potential improvements in definitions of modes and intermodal combinations.
- Identify potential improvements in methods used to estimate average region-to-region distances and ton-miles.
- Catalogue all FAF2 tables developed for publication in printed reports and on the web or for internal analyses, and indicate how those tables should be used as part of the quality assurance process.
- Inventory methods used to disaggregate FAF region-to-region flows and determine best practices for estimating the FAF network database and for estimating county-to-county flows with a prescribed set of locally collected supplemental data.
- Identify probability-based methods for distributing local freight movements within zones and between adjacent zones over the highway network.
- Evaluate the benefits and costs of providing web applications to the public for producing FHWA-published FAF origin-destination and network maps on demand, for mapping origin-destination patterns of freight flows through portions of the network, and for estimating possible patterns of freight rerouting following disruptions to the network.
- Identify useful linkages between proposed FAF products and freight performance measures.
- Develop a comprehensive FAF3 product plan.

COMMENTS AND SUGGESTIONS

Questions or comments regarding FAF may be referred to the FAF project manager, Michael Sprung, via email: faf@dot.gov or phone: 202-366-9047.