

Freight and Fuel Transportation Optimization Tool Technical Documentation

FTOT 2019.3 Public Release Version

DOT-VNTSC-FAA-19-06

DOT/FAA/AEE/2020-01

October 31, 2019

Prepared for:

Federal Aviation Administration
Office of Environment and Energy
Washington, DC



Notice

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for the contents or use thereof.

The United States Government does not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the objective of this report.

| REPORT DOCUMENTATION PAGE | | | <i>Form Approved</i> <i>OMB No. 0704-0188</i> | |
|--|--|---|--|---|
| Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503. | | | | |
| 1. AGENCY USE ONLY (Leave blank) | | 2. REPORT DATE October 31, 2019 | | 3. REPORT TYPE AND DATES COVERED Document, 9/2018-9/2019 |
| 4. TITLE AND SUBTITLE Freight and Fuel Transportation Optimization Tool Technical Documentation: FTOT 2019.3 Public Release Version | | | | 5a. FUNDING NUMBERS FB48CS/SG287 and SG291 |
| 6. AUTHOR(S) Kristin C. Lewis, Matthew Pearlson, Alexander Oberg, Olivia Gillham, Scott Smith, Gary Baker, Amy Vogel | | | | 5b. CONTRACT NUMBER 693KA9-18-N-00022 |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Volpe National Transportation Systems Center, 55 Broadway, Cambridge, MA 02142 | | | | 8. PERFORMING ORGANIZATION REPORT NUMBER DOT-VNTSC-FAA-19-06 |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Federal Aviation Administration, 800 Independence Ave, SW, Washington, DC 20591 | | | | 10. SPONSORING/MONITORING AGENCY REPORT NUMBER DOT/FAA/AEE/2020-01 |
| 11. SUPPLEMENTARY NOTES | | | | |
| 12a. DISTRIBUTION/AVAILABILITY STATEMENT | | | | 12b. DISTRIBUTION CODE |
| 13. ABSTRACT (Maximum 200 words) Volpe created the Freight and Fuel Transportation Optimization Tool (FTOT) to support the FAA, DOE, and the U.S. Navy's Office of Naval Research to assess optimal transport options for freight and fuel supply chains. FTOT is a flexible scenario-testing tool that optimizes the transportation of materials for future energy and freight scenarios. FTOT models and tracks commodity-specific information and can take into account conversion of raw materials to products (e.g., crude oil to jet fuel/diesel) and fulfillment of downstream demand. This report documents the components and functionality of the FTOT public release 2019.3. | | | | |
| 14. SUBJECT TERMS Freight, fuel, geospatial information systems, GIS, network analysis, transportation optimization. | | | | 15. NUMBER OF PAGES 42 |
| | | | | 16. PRICE CODE |
| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT | |

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18

298-102

Table of Contents

| | | |
|-------|---|----|
| 1 | CHANGE LOG IN VERSION 2019.3 | 6 |
| 2 | INTRODUCTION | 6 |
| 2.1 | FTOT Use Cases | 7 |
| 3 | FTOT STRUCTURE AND COMPONENTS | 9 |
| 3.1 | Key software modules in FTOT | 11 |
| 3.2 | FTOT GIS Network..... | 11 |
| 3.2.1 | Data Sources | 12 |
| 3.2.2 | Data Preparation..... | 13 |
| 3.2.3 | Configuration of Facilities | 15 |
| 3.2.4 | Exporting the Data | 16 |
| 3.3 | NetworkX | 16 |
| 3.4 | Optimization | 16 |
| 3.4.1 | Routing and Optimization Factors | 17 |
| 3.4.2 | Candidate Generation Option in Optimizer..... | 21 |
| 3.4.3 | End-to-end Source Tracking Option in Optimizer..... | 22 |
| | INSTALLING REQUIRED PROGRAMS..... | 22 |
| 4 | FTOT INPUTS | 24 |
| 4.1 | XML Schema..... | 24 |
| 4.2 | Configuring FTOT Scenarios in the XML File | 25 |
| 4.3 | Facility Data Input Files | 29 |
| 4.4 | Facility Geodatabase (GDB) | 30 |
| 4.5 | Default Data/Input Parameters | 30 |
| 5 | FTOT OUTPUTS..... | 33 |
| 5.1 | Reports..... | 33 |
| 5.2 | Graphics Dashboard..... | 34 |
| 5.3 | Maps..... | 35 |
| 6 | SETTING UP A SCENARIO..... | 37 |
| 7 | RUNNING A SCENARIO..... | 38 |
| 8 | SUPPLEMENTARY TOOLS | 38 |

| | | |
|-----|-----------------------------|----|
| 9 | TROUBLESHOOTING | 39 |
| 9.1 | Known Bugs and Issues | 40 |
| 10 | ACKNOWLEDGEMENTS | 40 |
| 11 | REFERENCES | 40 |

| | | |
|------------|---|----|
| Figure 1: | The four use cases for which FTOT is designed. | 7 |
| Figure 2: | Generalized supply chain structure in FTOT showing origins, processors, and destinations, with multiple commodities flowing along the various steps in the process. Commodities can be facility-type specific or can share facilities (e.g., all types of crude oil can flow into any refinery, versus jet fuel only permitted to flow to airports, diesel to other types of demand centers). The number of steps in the supply chain, commodities, processor types, and availability/schedule can all be expanded..... | 8 |
| Figure 3 - | Analytical tool data flow schematic showing the key components/roles of each component of the FTOT | 10 |
| Figure 4: | Default multimodal network feature classes included in base FTOT network. The user will add feature classes for origin, processor, and destination facilities. | 12 |
| Figure 5 - | FTOT multimodal transportation network..... | 15 |
| Figure 6: | Attributes of the facilities and network segments in FTOT that the user can adjust to build a scenario..... | 24 |
| Figure 7 - | Split screen practice for running FTOT scenario. The command line is on the left, and the scenario opened in the Windows Explorer is on the right..... | 37 |

1 CHANGE LOG IN VERSION 2019.3

- ADDED – issue #148 – processor facilities can now handle two input commodities.
- ADDED – issue #137 – commodity-mode specific restrictions are enabled.
- UPDATED – issue #108 – tableau dashboard: display all facility types at once by default.
- UPDATED – issue #124 and #46 - emission factors are updated to reflect 2019 values.
- DEPRECATED – issue #113 - getter and setter definitions in the ftot_scenario.py module.
- FIXED – bug #131 – commodity names are now case insensitive for input CSV files
- FIXED – bug #147 – optimizer passes back nearly zero values and causes bad maps
- FIXED – bug #152 – D Step throws an exception when no flow optimal solution
- FIXED – bug #149 – F step throws an exception if there is extra blank space in an input CSV file
- FIXED – bug #108 - tableau dashboard: units label for material moved by commodity and mode

2 INTRODUCTION

Volpe created the Freight and Fuel Transportation Optimization Tool (FTOT) to support the FAA, DOE, and the U.S. Navy’s Office of Naval Research to assess optimal transport options for freight and fuel supply chains. FTOT is a flexible scenario-testing tool that optimizes the transportation of materials for future energy and freight scenarios. FTOT models and tracks commodity-specific information and can take into account conversion of raw materials to products (e.g., crude oil to jet fuel and diesel) and the fulfillment of downstream demand.

FTOT’s network analysis is based on a unique, multimodal, flowable Geographic Information System (GIS) network that assigns costs, capacity and existing flow information, and weighting factors to each link. It takes origin and destination (OD) data, identifies best routes among the origins and destinations, and then optimizes flow patterns by minimizing overall scenario-wide costs. The costs include transport costs as well as user-defined weights and penalties (e.g., to prefer larger roadways over smaller), transloading (mode switching) costs, and, if necessary, capital costs to build new processing locations in between the origins and destinations to convert/refine raw material. FTOT encourages flows toward common freight paths while also providing the option of taking into account modal flow capacity and capacity availability (capacity minus existing flows) by mode, based on user needs. Outputs include commodity, facility, and link-level information on flows and demand fulfillment, maps, and summary graphics.

Based on variations in infrastructure (e.g., of the transportation network), cost elements, conversion factors, and origins, destinations and waypoints in the scenario, among other parameters, the user is able explore how variants of future scenarios will impact optimal multimodal freight transport costs, transport vehicle/vessel/pipeline demand, and emissions.

In summary, FTOT provides a “best case” transportation outcome of the scenario that can be used as a screening tool for feasible scenarios and as a benchmark for comparing against real world outcomes.¹

The first public release of FTOT (Beta Version 2019.1) was made available on the GitHub code hosting site at (<https://github.com/VolpeUSDOT/FTOT-Public>) on June 24, 2019. FTOT Beta v. 2019.1 is considered an “experts’ release,” requiring the user to have some knowledge of Python and ArcGIS to run. It enables the user to analyze scenarios using either the default network and parameters or user-defined details. The Beta software and documentation was sent to 12 initial testers outside of the Volpe organization. Feedback was obtained from the testers and used to improve the tool for a wider general public release. FTOT version 2019.3 incorporates numerous bug fixes, enhancements, and updates as noted in the change log above.

2.1 FTOT Use Cases

One of the unique aspects of FTOT is that it focuses on transport for supply chains as opposed to a single commodity. Generally these take the form of:

Origins (+ outgoing storage) → Processor facilities type 1 (pass through or conversion) (+ incoming/outgoing storage) → Processor facilities type...x → Destinations (+ incoming storage).

The four key use-cases are described in Figure 1.

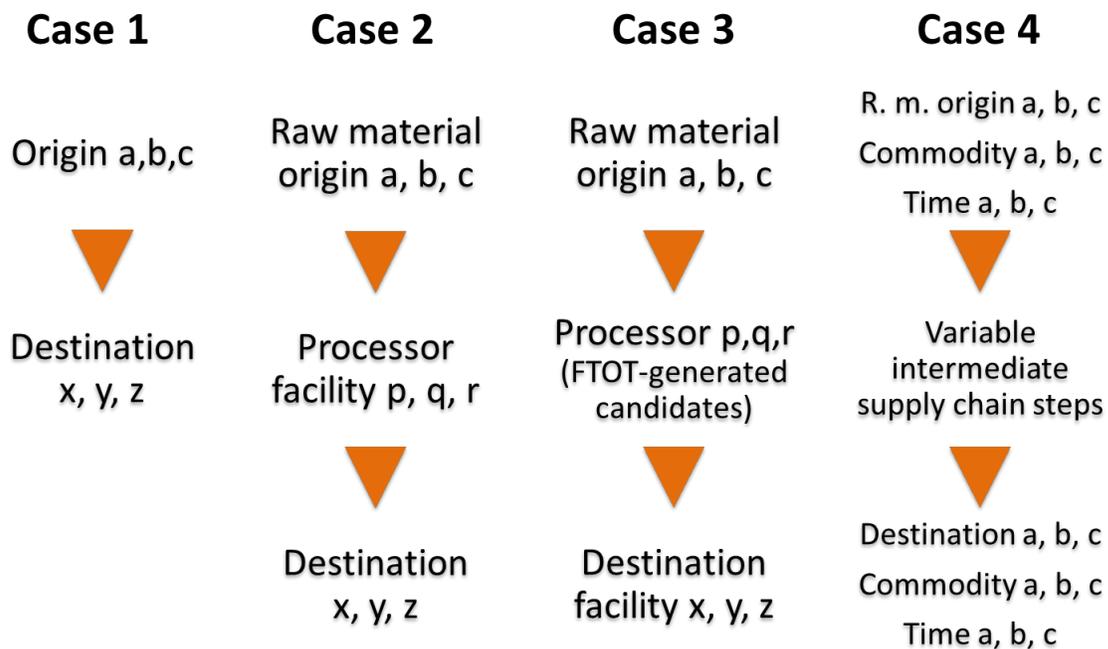


Figure 1: The four use cases for which FTOT is designed.

¹ Note: Optimal scenario results can differ from real-world patterns due to complexities in transport availability, contracting mechanisms, habits, and nuances in demand and pricing, among other system dynamics.

Multiple commodities can be flowed, converted, and tracked through the analysis to enable optimization of complex supply chains. These commodities can share origins, processors and destinations or can be canalized to separate processors or destination types depending on commodity and designated facility compatibilities. Furthermore, FTOT includes the option to schedule facility/route availability down to the daily level. The supply chain can be expanded by any number of steps, commodities, processor type, and time (e.g., facility availability, route availability). The generalized supply chain structure used in FTOT is shown in Figure 2.

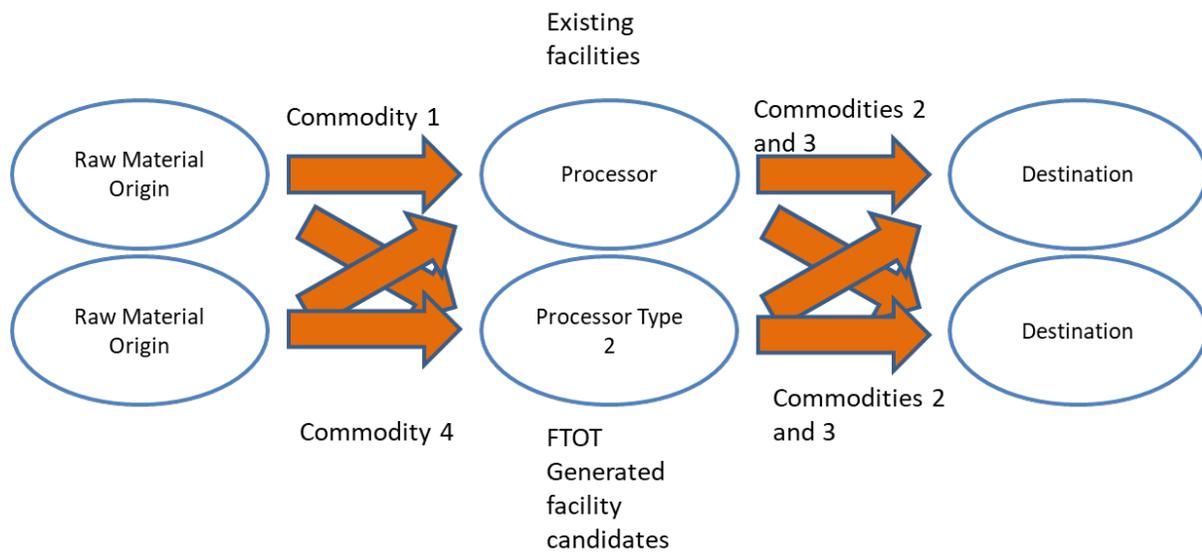


Figure 2: Generalized supply chain structure in FTOT showing origins, processors, and destinations, with multiple commodities flowing along the various steps in the process. Commodities can be facility-type specific or can share facilities (e.g., all types of crude oil can flow into any refinery, versus jet fuel only permitted to flow to airports, diesel to other types of demand centers). The number of steps in the supply chain, commodities, processor types, and availability/schedule can all be expanded.

3 FTOT STRUCTURE AND COMPONENTS

The FTOT model takes the GIS-based network, transforms it into a consumable multi-graph using an open-source network analysis tool, and feeds the multi-graph and scenario parameters into an optimization module. The optimizer selects among origins, processors and destinations and allocates flows among them to maximize commodity delivery and minimize overall cost, taking into account multiple cost elements, weightings, and constraints and described in Section 3. Figure 3 demonstrates the FTOT architecture and data flow. Installation instructions for these required software modeling tools are described in Section 0.

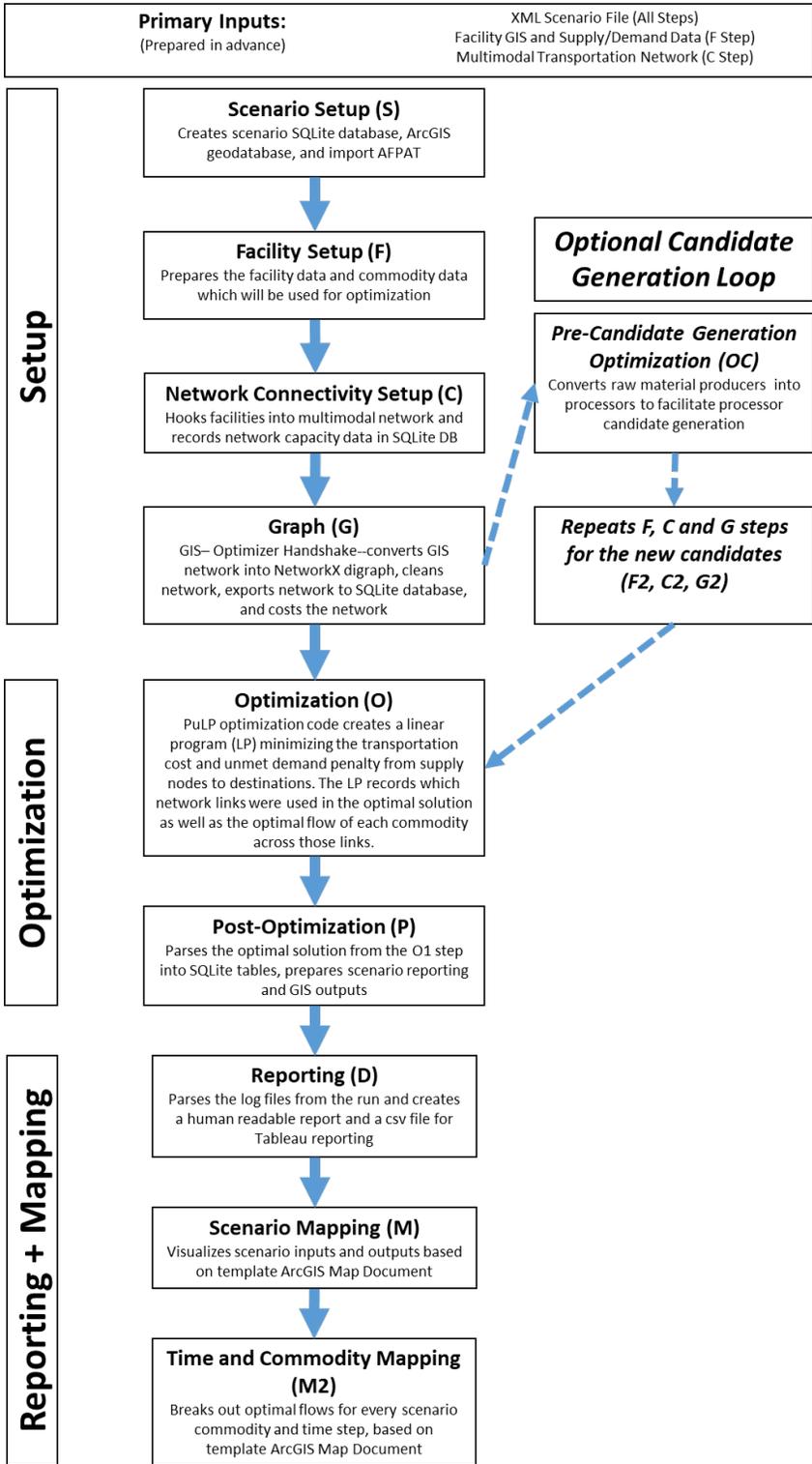


Figure 3 - Analytical tool data flow schematic showing the key components/roles of each component of the FTOT

3.1 Key software modules in FTOT

The following software tools are integral to the use of FTOT:

Python v2.7 is the scripting language supporting FTOT. FTOT relies specifically on the Esri ArcGIS installation of Python 2.7—which among many default modules, includes the Arcpy module for GIS functionality. Additional supporting Python modules are also installed on top of a base Python installation.

ESRI ArcGIS supports the GIS component of the tool. The Arcpy module is a python-based tool that allows access to ArcGIS functionality. FTOT uses arcpy to process the multi-modal network and facility locations defined for the scenario. Components of ArcGIS apply transport costs, impedances, capacity, and existing flows to the network. This network is then passed through NetworkX to the optimizer for analysis. After optimization, the outputs of the scenario are passed back to the GIS module to generate scenario maps.

NetworkX is used to interface between ArcGIS and the optimizer. NetworkX is a Python-based open source software package that enables network analyses. NetworkX is able to read in the scenario GIS data, and translate it into a format that is readable by the optimizer. NetworkX can be used for network flow optimization, but is not able to take into account some of the required elements in FTOT, specifically the ability to optimize using multiple paths between origin-destination pairs to allow for capacity-constrained analyses in which the optimizer will use a shortest path until capacity is reached, then route flows to the next shortest path, and so on.

The **PuLP** optimizer is a Python-based tool that uses standard operations research algorithms to solve optimization problems. Once a problem is described in mathematical terms, PuLP can be used to find the optimal values. In FTOT, the optimization objective is to satisfy maximum demand based on least cost. More detail on the optimizer and the COIN-OR solver can be found in the sections below.

A **Tableau dashboard** is used to generate a suite of graphical outputs based on the optimal scenario solution. To view the Tableau dashboard, one can use either the free Tableau Reader or a full version of Tableau Desktop v2019.2 or higher.

3.2 FTOT GIS Network

The elements of the FTOT multimodal network include feature classes for road, rail, water, and pipeline networks. Intermodal facilities where transfers between modes occur, waterway locks, and available tariff and capacity information for crude and product pipelines are also included in the network feature classes. Each pipeline category has three feature classes associated with them: 1) pipeline feature classes represent the segments associated with the pipeline network; 2) pipeline_trf_sgmts feature classes store the ordered segments associated with each tariff for which FTOT has reliable cost and origin/destination data; and 3) pipeline_trf_rts contains tariff-level information such as the cost of the movement and information on the pipeline and origin/destination pipeline stations. FTOT uses these feature classes together to limit pipeline movements to the predetermined routes that tariffs

determine. Default feature classes are shown in Figure 4. Full metadata for the multimodal network can be accessed in ArcMap or ArcCatalog, part of the ArcGIS suite of software that is required to run FTOT. A summary of the data sources behind the multimodal network is provided below.

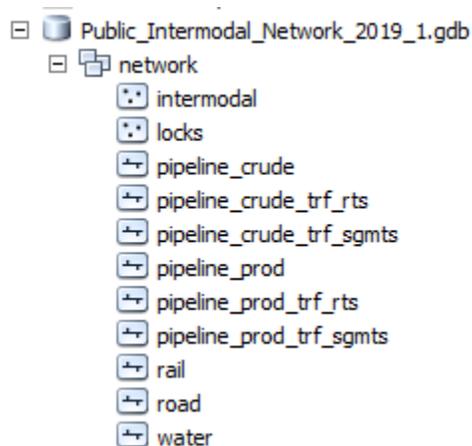


Figure 4: Default multimodal network feature classes included in base FTOT network. The user will add feature classes for origin, processor, and destination facilities.

3.2.1 Data Sources

The FTOT install includes a base multimodal transportation network composed of features and selected attributes from the following source datasets:

- Network source data / components
 - Road: FHWA Freight Analysis Framework (FAF) v. 4--2016) (accessed via the [Federal Highway Administration Office of Operations](#))
 - Rail: FRA North American Rail Network (NARN)—2018 (sent directly to Volpe by FRA). Only publicly available attributes of the NARN are included in the public release of the FTOT network.
 - Waterway: U.S. Army Corps of Engineers (USACE) Navigable Waterway Network--2018 (accessed via the [USACE Navigation Data Center](#)), modified based on additional research to exclude some non-navigable waterways.
 - Locks: USACE Locks Shapefile—2018 (accessed via the [USACE Navigation Data Center](#))
 - Crude and Petroleum Product Pipeline Network Data: Energy Information Administration (EIA) crude oil and petroleum product pipeline data- March, 2017 (accessed via [EIA Maps](#)), modified by Volpe based on station locations in a tariff dataset by LawIQ Inc. (February, 2018) and other publicly available pipeline network data.
 - Crude and Petroleum Product Pipeline Tariff Rate Data: Pipeline transportation rate and origin/destination information supplied by LawIQ Inc. For analysis purposes, all rates have been rounded (nearest \$1 for rates from \$0-\$20; nearest \$5 for rates from \$20-\$50, and nearest \$10 for rates over \$50). If further detail is needed, the user can contact info@lawiq.com for precise tariff-level rates for over 1,000

interstate and intrastate pipelines across 100 unique commodity types and 40 jurisdictions.

- Intermodal facilities: Volpe-developed list of intermodal facilities that handle energy products based on public data from:
 - EIA [Crude Oil by Rail Dataset](#) (November, 2014)
 - USACE [Master Docks Database](#) (February, 2016)
 - Corroboration from Surface Transportation Board Waybill Sample dataset (2013)
 - Pipeline station data from LawIQ Inc. (February, 2018)
 - Volpe research/review of facility locations using satellite imagery and other research of public data sources

3.2.2 Data Preparation

While the source data for road, rail, waterway and locks are provided in formats that FTOT can leverage with little modification, FTOT's pipeline network and intermodal facilities datasets were produced especially for FTOT and require standalone processes that integrate various datasets, described below.

To prepare the FTOT pipeline network, a series of programmatic methods are used to integrate tariff cost data and origin/destination station location data from LawIQ Inc. with their corresponding pipeline segments in the Volpe-modified EIA GIS pipeline network, providing a flowable route across the pipeline network for each tariff. The tariff costs are attached to each tariff's approximate route, and these sets of routes and costs become candidates for movements of relevant commodities in FTOT, often as part of larger multimodal movements. A product is considered to flow between existing origin-destination pairs within the pipeline network (i.e., no new injection or extraction points are allowed).

The intermodal facilities dataset was created by combining two separate efforts to identify 1) agriculture-related intermodal facilities, and 2) energy-related intermodal facilities. Intermodal facilities may be added or removed from time-to-time during periodic updates of the data sources.

The agriculture-related intermodal facilities dataset was created from the Bureau of Transportation Statistics' public intermodal facility list using specific criteria and visual review of satellite imagery to correct facility locations and eliminate excess facility points. Criteria for elimination included: lack of nearby significant intermodal facility in satellite imagery, proximity to other, more accurate intermodal points, points that did not have complete data and/or points that were not applicable to the FTOT network (e.g., made at least one linkage among road, rail, waterway, and pipeline). Rail intermodal points were added to the remaining list based on movements of specific, raw-material-relevant commodity types in the Railway Waybill Sample to identify potential railway entry points for agricultural commodities, which are likely to enter the system at smaller facilities than the main intermodal facilities identified by BTS. The commodity-specific origins were incorporated into the GIS layer and duplicates were eliminated. The final intermodal facilities layer for FTOT includes 383 unique intermodal facilities tagged specifically for the different modes that interface at each point.

The energy-related intermodal facilities were defined through a separate effort to integrate Army Corps of Engineers port facilities handling coal, crude oil and petroleum products; Energy Information Administration crude-by-rail facility data; and other facilities identified through a screening process involving Railway Waybill Sample data and manual review of satellite imagery.

FTOT includes a default national multimodal network including the road, rail, waterway, pipeline, and intermodal facility layers. The network is based on the authoritative sources and Volpe-produced datasets described above and is processed to simplify network segmentation as much as possible while retaining core required FTOT attributes—such as existing flows, capacity and other attributes helpful for determining network impedances.

For datasets not produced especially for FTOT, some additional filtering of the source network data is necessary. For example, in the FTOT network FRA's North American Rail Network is subset to the main subnetwork and major industrial leads within the continental United States. The USACE Navigable Waterway Network is subset to exclude waterways that are not navigable to a depth of nine feet (standard barge clearance). Note that water movements in FTOT are assumed to travel by barge, not ship.

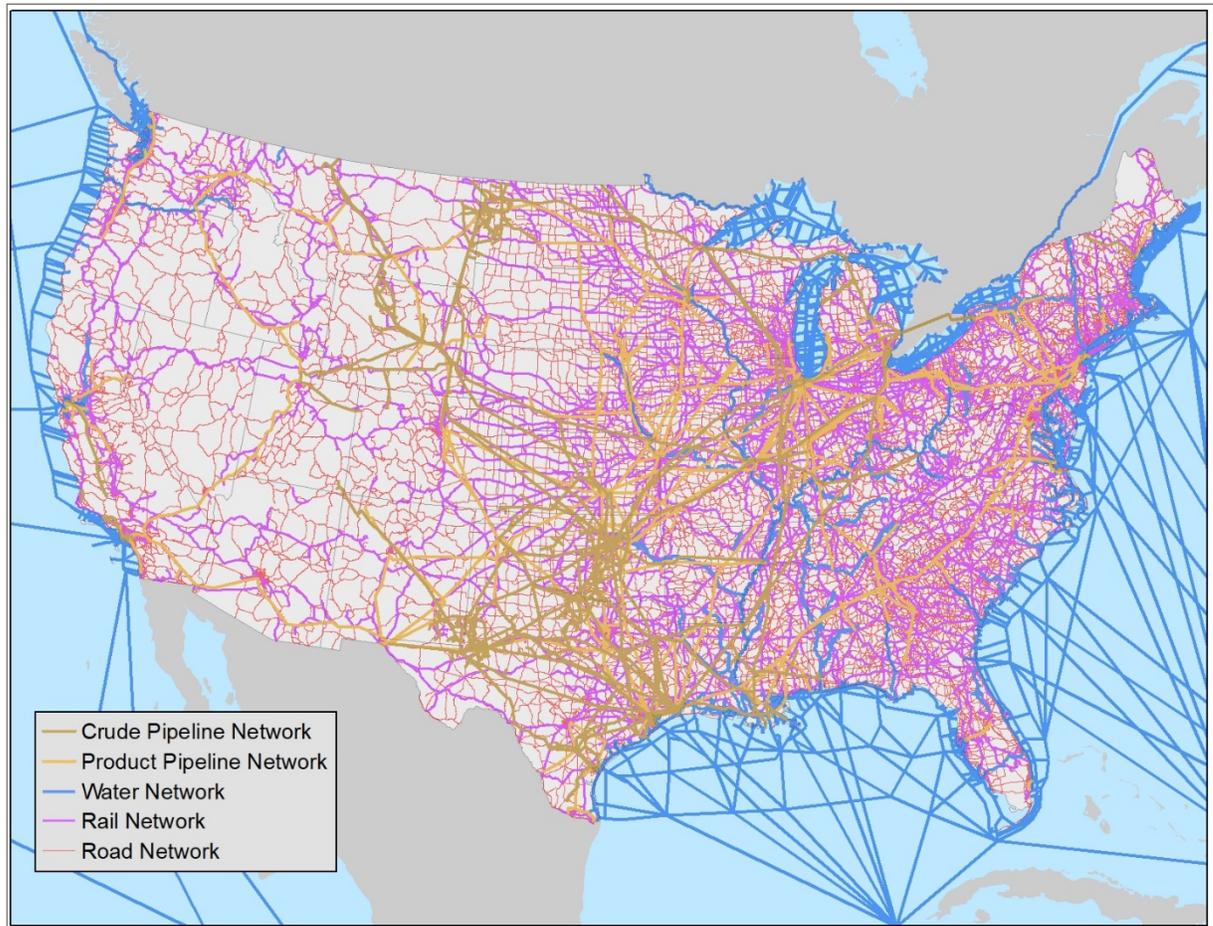


Figure 5 - FTOT multimodal transportation network

Figure 5 shows the resulting multimodal network that FTOT uses for routing analyses and display of results.

3.2.3 Configuration of Facilities

In addition to the multi-modal network, FTOT requires origination facilities to supply material in the scenario, and destinations with demand to terminate the flows. Origin facilities are by definition facilities that have an output but no input. Conversely, destinations are defined as facilities that have inputs but no outputs. Processors are an optional facility type that can be included in the supply chain. FTOT considers any facility that has inputs and outputs to be processors. Processors may process the material from one commodity to another, or simply act as a pass-through where material is aggregated or stored. During an FTOT scenario, FTOT copies user-specified facilities into the scenario network geodatabase and programmatically links them to the multimodal network. The process of connecting facilities to the network modes is known as adding artificial links. The fully connected network is stored in a feature dataset and includes the multimodal network, intermodal facilities, and origin/processor/destination facilities feature classes.

3.2.4 Exporting the Data

After the facilities are connected to the multimodal network with artificial links, the network feature classes are exported as shapefiles in preparation for the next step, converting the network into a graph. The NetworkX module, described in greater detail below, is used to create a graph of the connected network. NetworkX uses shapefiles as the geospatial input. Therefore, the feature classes for each network feature dataset are exported as shapefiles using Arcpy and stored in a temporary directory for NetworkX.

3.3 NetworkX

NetworkX is used to create a network graph from the geospatial network and facility data. Shapefiles are the required geospatial input. Each element of the network feature dataset (roads, rail, destinations, etc.) are exported as shapefiles and stored in a temporary directory. Next, NetworkX reads in each shapefile and adds it into a network graph.

Once the base graph is created, it is processed with the following procedure to ensure flows in both directions are enabled for certain modes (e.g. road, water, and rail) and restricted for other modes (e.g. pipeline). First, a copy of the network is duplicated. The duplicate is then reversed to enable flow in opposite directions. The duplicate-reverse graph and the original graph are then joined back together, after which the network edges are inspected individually and cleaned. Cleaning includes:

- Deleting duplicate artificial links:
 - Origin facilities should only have links flowing out.
 - Destination facilities should only have links flowing in.
- Deleting reversed pipeline links (pipeline only flows in one direction).
- Node IDs are converted from latitude longitude coordinates to integer values, which simplifies labeling and graph operations.
- Network costs are calculated based on link attributes and user-defined inputs. Two types of network costs are recorded. First, the dollar cost is based on the mileage and mode. Second, a routing cost is stored that includes impedances based on the network mode and link attributes. Impedances are used to encourage flows on main links of the network, and penalize minor link usage.

Once the network graph is cleaned, it is stored in a local SQLite database and is ready for the optimization steps. The following section describes the optimization: how the problem is defined and solved using the linear programming solver.

3.4 Optimization

The goal of the PuLP optimization process is to minimize the total cost of transporting enough material from the origin facilities, through the processor (if included in the supply chain), and on to the destinations to meet demand. This analysis includes factors for actual transportation costs that vary by

the mode and length of the routes, weightings for route links, capital costs for building refineries, and penalties for not meeting demand. The three components of the optimal cost are described below:

- 1) The actual dollar costs of transporting the material plus additional routing costs from impedances (weightings and penalties that force the tool to favor particular desirable characteristics of the routing).
- 2) The amortized annual capital cost of building any candidate facilities that are part of the solution. If no candidate facilities are included in the scenario, this cost is equal to zero.
- 3) Destination facilities with unmet demand are also treated as a “cost” in this system by applying a penalty per unit of undelivered commodity. Without this sort of penalty, the lowest cost solution would always be to transport nothing at all.

The FTOT network flow and facility location problem can be mathematically formulated as a linear programming problem (Bertsimas and Tsiklis 1997). Solving problems of this type optimally (or near-optimally) is one focus of operations research. Computational Infrastructure for Operations Research (COIN-OR) maintains a repository of open-source software for the operations research community. This repository is the source of the PuLP ("pulp-or" 2007, pulp-or 2014) and Coin-OR Linear Programming (CLP) (COIN-OR 2007) software used for the FTOT project. PuLP is a linear programming modeler, used to set up the linear problem and then call the CLP solver. The CLP solver does the work of solving the linear programming problem (Bertsimas and Tsiklis 1997) using the Simplex algorithm (Carreira-Perpiñán 2014).

The PuLP optimizer identifies a maximum or minimum value (in this case, total transportation, capital, and unmet demand costs) using a mathematical description of the supply chain scenario. In its application for FTOT, the PuLP optimizer takes the user specified facility information (geospatial and commodity supply/demand/processing data) and the transportation network and optimizes the paths and flows among all components. The goal of the optimization is to minimize the total cost of meeting as much destination demand as possible by utilizing all available supply and transportation modes.

3.4.1 Routing and Optimization Factors

This section describes each of the factors that are considered in the route optimization algorithm.

- 1) **Monetary costs** for each transportation option and for construction (if applicable)
 - For road, rail and waterway, this is applied as a cost per ton-miles for solids or per thousand gallon-mile for liquids.
 - Pipeline costs are based on tariff data described above.
 - Amortized annual capital cost for candidate processor construction is also considered if candidate processors are generated in the scenario.
- 2) **Modal flow capacity** and existing flows
 - Modes can be configured as permitted or not permitted in the scenario configuration file.

- Capacity constraints can be enabled by mode to reroute flows if and when available capacity on the network has been exceeded. Network capacity and background volumes are estimated from existing data sources which are further described in Section 4.5.

3) **Weightings and penalties**

- Unmet demand penalty:
- Capacity constraints can be enabled by mode to reroute flows if and when available capacity on the network has been exceeded. Network capacity and background volumes are estimated from existing data sources which are further described in Section 4.5.

4) **Unmet demand penalty** if a destination does not receive the desired quantity of a given commodity

- The optimizer adds a penalty (proportional to the amount of unmet demand) to the total cost of a potential solution.
- The penalty (per unit of unmet demand) is specified by the user in the scenario configuration file.
- If the magnitude of the penalty is low, the optimization will prioritize minimizing transportation costs; if it is high, the optimization will prioritize meeting all demand, with less sensitivity to the dollar cost of doing so.
- This penalty is required for the optimizer to function; if there is no penalty for not meeting demand, the lowest cost solution will always be to transport nothing at all (a “no-flow solution”). The penalty drives the model to move material rather than allow it to remain at the starting point; therefore, it represents the maximum allowable transport cost for a given unit of flow. However, the optimization will always choose the lowest cost path when possible for a given unit of material.
- The unmet demand penalty is also a common cause of errors for new scenarios. As a general guide, the unmet demand penalty will likely work best if set to be 10-50 times the average actual transportation cost. When limited geographic scope or time periods may result in a scenario that does not produce enough material to outweigh the cost of transporting material or building a candidate processing facility, a higher unmet demand penalty may be required to avoid a no-flow solution. In general, it may be necessary to raise this penalty when any other cost (e.g. rail transport) is raised, or else the optimizer will conclude that it is more optimal to transport less material. As a general guide, the unmet demand penalty will likely work best if set to be 10-50 times the average actual transportation cost. This ensures that raw material and products will be transported even over long routes. It should be noted that at very low flows, one may see non-intuitive results such as reduced utilization of raw material that is available. This is due to the optimizer electing the less expensive option of reducing flow overall instead of accumulating high transport and capital costs in addition to a still-large unmet demand penalty for a given route. Likewise, a very high unmet demand penalty may force the flow of materials in unanticipated ways. It is

advised to run scenarios with multiple unmet demand penalties to explore the sensitivity of a given analysis.

5) **Minimum flow requirements** applied to processing facilities

- For a processing facility to be used, it must process at least a certain amount of raw material – the user provides this minimum.
- Similarly, there is a maximum facility size restriction that the user also provides.

6) **Network impedances** to avoid unrealistic shortest path routes on road, rail, and water segments

- Impedances encourage flow over highly used paths and discourage flow over paths where freight traffic is less common.
- Like penalties, impedance weights are added to the routing costs and used for the optimization. They are distinct from the dollar cost that FTOT reports and summarizes.
- In the scenario configuration file, the user can adjust the impedance weights on each type of segment in the network.
- See Section 4.5 for a description of FTOT’s default impedances.

The number of impedance categories and their definitions depend on the data available for the particular mode—for example, in the road network impedances are based on roadway functional class and in the water network impedances are based on actual freight flows. For the rail network, while the internal (non-public) network relies on actual freight tonnage data from the railroads to set impedances, the public network relies on a simpler set of categories based whether or not the rail segment is part of the Strategic Rail Corridor Network (STRACNET) and whether or not it is owned or flowed on by Class 1 freight railroads.

7) **Restrictions on distance a commodity can travel** if a commodity has a maximum allowable transport distance

- Because some commodities have a maximum allowable transport distance, the optimization problem is set up so that potential routes are built link-by-link from each source facility; total distance travelled from the source facility to a given link is thus tracked.
- When the maximum allowable transport distance is reached, no further links can be added to the route on which that particular commodity travels from that particular source facility.

These costs and weightings are translated to mathematical decision variables and coefficients as follows:

| Variable | Explanation |
|-----------------|--|
| X_{iabj} | Flow, in units ² /time period, from origin facility i_{ab} to origin facility storage j_{ab} , where a is the commodity produced and b is the time period of production |
| $X_{jabj(b+1)}$ | Storage of commodity a at origin facility storage location j from time period b to time period $b+1$, in units/time period |

² “units” are defined by the user in the scenario xml file. FTOT is able to use a wide variety of units and convert between units. Supported units are detailed in the Pint python module documentation.

| Variable | Explanation |
|--------------------|--|
| X_{jabkab} | Flow, in units/time period from origin facility storage j_{ab} to processing facility k_{ab} where a is the commodity being transported and b is the current time period |
| $X_{kabmabc}$ | Flow, in thousand gallons /time period from processing facility k_{ab} to destination storage m_{abc} where a is the commodity that was processed, b is the current time period, c is the fuel being transported |
| $X_{mabcma(b+1)c}$ | Storage at destination storage location m from time period b to time period b+1 of commodity c, produced from commodity a, in thousand gallons/time period |
| $X_{mabcnabc}$ | Flow, in thousand gallons/time period from destination storage m_{abc} to destination n_{bc} where a is the commodity that was processed, b is the current time period, c is the fuel being transported |
| U_{nc} | Unmet demand of commodity c at destination n, in thousand gallons / year |
| E_{kabc} | Excess commodity remaining at processing facility k_{ab} of commodity c, where a is the commodity that was processed, b is the current time period. Non-negative. Counts toward upper and lower bounds on production, but is not used to fulfill demand. |
| y_k | 0 – 1 variable: 1 if processing facility k is used, 0 otherwise |

| Coefficient | Explanation |
|--------------------|--|
| U_{Piab} | Upper bound on flow of commodity a out of origin facility i_{ab} during time period b, in thousand gallons or tons / time period |
| C_{iabjab} | Transportation cost, in \$ / thousand gallons or ton, to flow commodity a during time period b from origin facility storage j_{ab} to processing facility k_{ab} |
| F_{Bk} | Fixed cost, converted to \$ / year, to build the processing facility k |
| U_{Bk} | Upper bound on flow out of processing facility k, in thousand gallons / time period |
| L_{Bk} | If processing facility j is used, lower bound on flow out of processing facility j, in thousand gallons / year |
| $C_{kabmabc}$ | Transportation cost, in \$ / thousand gallon, from processing facility k to destination storage m, of fuel c in time period b processed form commodity a |
| S_{Bac} | Conversion factor at the processing facility: inbound thousand gallons or tons of commodity a x S_{Bac} = outbound thousand gallons of commodity c |
| U_{nc} | Penalty (\$ / thousand gallon) for not meeting demand at destination n for fuel c |
| D_{nc} | Demand, in thousand gallons / year, at destination n of fuel c |
| M_x | Maximum flow (capacity) in units / time period on edge ³ r |

Then the problem for FTOT analysis is mathematically stated as follows:

$$\text{Minimize annual cost} = \sum_{jabkab} C_{jabkab} X_{jabkab} + \sum_k (F_{Bk} y_k) + \sum_{kabmabc} C_{kabmabc} X_{kabmabc} + \sum_{nc} (U_{nc} u_{nc})$$

Subject to the following:

³ A route between facilities can be composed of multiple edges, each of which may have a distinct maximum capacity

| Constraint | Explanation |
|--|---|
| (1) For each processing facility k_{ab} , $S_{Bac} \sum_{jabk} X_{jabk} - \sum_{kabm} X_{kabm} = E_{kabc}$ | Flow must be conserved at each stage relative to time and commodity (with the appropriate conversion factors); any excess is tracked. |
| (2) For each processing facility k_{ab} , $y_k U_{Bk} - \sum_{kabm} X_{kabm} - E_{kabc} \geq 0$ | If a processing facility is used, flow cannot exceed the upper bound during any time period, across all commodities, including excess. Note that if the processing facility is not used ($y_k = 0$), this constraint requires the flow out of the processing facility be 0. |
| (3) For each processing facility k_{ab} , $y_k L_{Bk} - \sum_{kabm} X_{kabm} - E_{kabc} \leq 0$ | If a processing facility is used, the flow out of it over a year must exceed the lower bound, summing all commodities. |
| (4) For each origin facility i_{ab} , $\sum_{iabj} X_{iabj} \leq U_{Piab}$ | Flow out of each origin facility does not exceed the origin facility upper bound, for each time period and commodity. |
| (5) For each destination n_c , $\sum_{mabc} X_{mabc} + U_{nc} = D_{nc}$ | Unmet demand plus flow into a destination is equal to that destination's demand, for each commodity. |
| (6) The y variables are binary (0 or 1) | A processing facility is used, or it isn't. |
| (7) The x , E , and u variables are non-negative | No negative flows are permitted |
| (8) For each time period b and edge r , $\sum x_r \leq M_r$ | For each edge and time period, total flow (sum of all x on that r) does not exceed capacity for that edge |

The optimizer then uses standard linear optimization techniques such as a revised simplex algorithm to solve the mathematical description of the problem to move material from origin to destination by selecting among paths and processor options for each unit of the commodities. The specific choice of algorithm is made by the COINMP_DLL solver, as implemented by PuLP. The allocation of commodities/materials among paths and facilities is based on meeting maximum demand while minimizing the total cost, without violating the constraints on minimum and maximum flow.

3.4.2 Candidate Generation Option in Optimizer

- FTOT can generate screening-level processor candidate locations to convert raw material to commodities demanded by the destinations. Candidate generation requires the user to specify the input and output commodity relationships, as well as the facility maximum and minimum size, minimum aggregation size for generating a candidate, and the amortized capital cost as a function of size.
- In order to get a solution, the user must add the `max_transport_distance` field in the RMP commodity input file. This field is mandatory for FTOT runs with candidate generation.
- Routes are built up starting from origin points out to maximum feedstock distance.
- Beyond the maximum raw material transport distance, FTOT converts raw materials to end products demanded by the final destinations using the most efficient conversion process specified in the candidate input file and flows these commodities to the final destinations as if

they were the commodity demanded. These routes are used to identify potential raw material aggregation along a lowest-cost path from the origin to the destination.

- FTOT identifies candidate nodes where sufficient raw material is aggregated on the network to support a potential facility. The minimum quantity of material can be set in the candidate processor commodity CSV file. Material that aggregates beyond the user defined raw material transport distance is ignored in the aggregation calculations.
- Candidate nodes are then assigned a process, and added to a candidate processor feature class.
- The facilities and connectivity steps are then run again to add the candidate processor locations into the scenario network as facilities and record the product slate and processor size.
- Optimization is then rerun and proceeds as usual.

3.4.3 End-to-end Source Tracking Option in Optimizer

- When detailed information is required about exactly where material flowed, an additional step can be run after the main optimization.
- This method provides the end-to-end path of every piece of demand met. It should generally only be run after “O”, the standard optimization. It will use that solution and define more specific flow paths. Running it from scratch, instead of “O”, is possible but exponentially more time and resource intensive, and yields the same information.

This is accomplished by creating flow variables for every edge and *source facility* combination; essentially, the same commodity from different sources is treated (and tracked) as distinct commodities.

INSTALLING REQUIRED PROGRAMS

Current FTOT install instructions are available here: <https://github.com/VolpeUSDOT/FTOT-Public/wiki>.

Before running a scenario in FTOT, the following programs must be installed:

- **ESRI ArcGIS** Version 10.6.1 (Geographic Information System (GIS) Program) performs the geospatial analysis elements of the tool. The Advanced License level of ArcGIS is required. FTOT may run on earlier versions of ArcGIS dating back to v 10.1, but for full support and functionality v 10.6.1 is highly recommended. Be sure to include the **Esri ArcGIS 10.6.1 base Python installation** (Version 2.7.14). This is installed alongside Esri ArcGIS and provides the majority of Python modules required for FTOT scripts to run. Special attention should be paid to this installation of Python. See the trouble-shooting section for issues with pre-installed Python environments.
- **ESRI ArcGIS 64-Bit Background Processing** is a separate installation of python and the arcpy modules. It is required for scenarios that exceed 2 GB of memory (RAM) during run time. It is highly recommended for installation and can be used for all scenarios, both large and small.
- **SQLite**—is used to store scenario variables and results. In order for the user to fully explore SQLite-generated outputs, users must install a SQLite database browser, such as DB Browser for SQLite.

- **NetworkX** Version 2.2 is an open source python-based suite of tools for manipulating and analyzing complex networks for which many publicly available algorithms and tools have been established.
 - **GDAL 2.2.4** is a required dependency for NetworkX to read in spatial input files for graph conversion.
 - **PuLP** Version 1.6.8 (Open Source Python Wrapper for Optimization Solvers) is a linear programming module written in Python. In FTOT, PuLP is used to link the solvers in the Computational Infrastructure for Operations Research project (COIN-OR). The **COIN-OR** project maintains an archive of open source software for operations research, including implementations of algorithmic solvers and modeling systems. This includes a number of open source optimization models, including a branch and cut solver (CBC) for mixed integer programming.
 - **lxml** Version 4.2.5 is a python package for handling xml and is used in FTOT to parse the scenario configuration XML file.
 - **Pint Version 0.8.1** is a python package for unit conversions and dimensionality checks. It is used throughout FTOT to harmonize user input data into a common set of units for solids and liquid freight movements.
 - **Imageio version 2.5.0** is a Python package required for creating animated graphics of FTOT mapping outputs.
- FTOT Version 2019_3** is the tool itself and can be downloaded from the Public GitHub repository [<https://github.com/VolpeUSDOT/FTOT-Public>].

4 FTOT INPUTS

FTOT scenarios require geospatial information for the network itself as well as associated attributes such as costs, impedances and weightings, capacity, movement restrictions, and schedule. The FTOT network must also include the facilities (origins, processing/waypoints, destinations) associated with the supply chain being analyzed, as well as their associated attributes such as facility minimum and maximum size, available supply/demand of input and output commodities and associated efficiency/conversion to products. These inputs are defined in a few input files:

- 1) a set of geospatial layers defining the network and facility locations
- 2) a scenario configuration file defines key input parameters applied to the network and optimization
- 3) a set of comma-delimited files to define the facility data

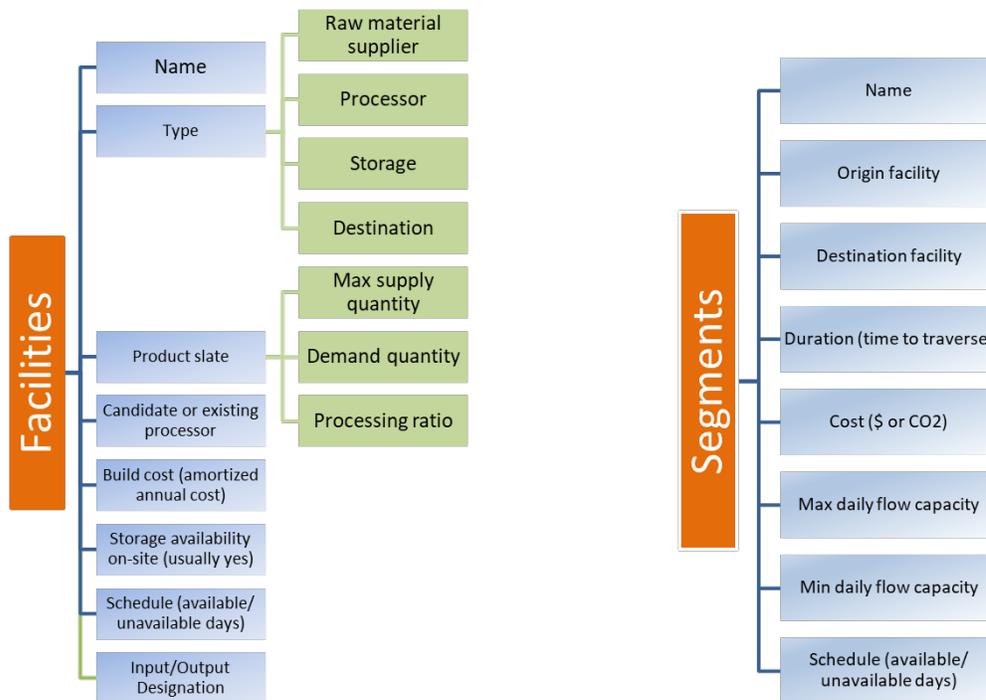


Figure 6: Attributes of the facilities and network segments in FTOT that the user can adjust to build a scenario.

4.1 XML Schema

Key parameters and the location of input data are defined with in a scenario configuration file. The scenario configuration file uses the Extensible Markup Language (XML) file format. An XML file is a structured text file that is formatted with markup language making it readable by both humans and machines.

To facilitate the input of many different variables and running of multiple scenarios, an XML-based “scenario file” approach supplies the inputs to the FTOT. This XML Scenario file tags each potential data source as a file source, a function, or a specific value. The XML Scenario file is a text-based file that can be modified in a text editor or integrated development environment for software development. The XML file uses tags to indicate the presences of variables that should be used in a scenario. FTOT can generate a new scenario xml file populated with certain default values for various parameters using the FTOT Tools suite. The user must update the placeholder values in the XML file with scenario specific information, and the default values should be reviewed and adjusted by the user.

The XML file is validated against an XML schema file. The ‘Master_FTOT_Schema.xsd’ file is located in the ‘lib’ folder of the project code repository (e.g. C:\FTOT\Program\lib). This schema ensures that the parameters input into the XML are in the format expected by FTOT. If the XML scenario file does not meet the requirements in the schema, FTOT raises an error and report the offending line and element.

4.2 Configuring FTOT Scenarios in the XML File

All scenario configurations are through the ‘scenario.xml’ file. The FTOT XML specifies input data file locations and user variables for the scenario.

The xml is divided into several sections:

- Scenario
 - Name, description, FTOT version
- Scenario Inputs
 - Base_Network_GDB – location of the multimodal network
 - Base_*_Layer – location of the facility locations for RMP, Destination, and Processors
 - *_Commodity_Data – location of the facility commodity csv files for RMP, Destination, Processors, and Candidate Processors
 - Default_Units_*_Phase – user specified units for the scenario. Mixed units are acceptable in input CSV files. FTOT uses the Pint module in Python to track and convert quantities to the default units specific in the scenario configuration file.
- Assumptions
 - Vehicle load capacity, fuel efficiency, and emissions are specified here. Users are free to modify values as necessary.
- Script Parameters
 - Network_costs
 - Modal costs: railroad, road, barge
 - note: pipeline costs are specified by tariffs values that have been rounded for public release.
 - Intermodal transloading costs
 - Artificial link distances to connect facility to the network
- Route Optimization
 - Permitted modes: road, rail, water, pipeline_crude, pipeline_prod

- Allows the user to selectively exclude modes from the optimization analysis
- Capacity options: to toggle capacity and background flows for the network links on or off.
- Penalty for not fulfilling depot demand: a penalty used in the optimization to drive flow in the scenario.
- Pulp Solve Options:
 - user specified variables for changing default parameters passed to the optimization solver (max solve time, fractional gap)

The following is a description of each element and parameter:

1. <Scenario xmlns="FTOTv2019.3"> The namespace used by the XML schema. This should not be modified by the user.
2. <Scenario_Schema_Version> - the version of the schema. Used to verify against changes to the XML schema, and code changes. The convention is: version X.Y.Z, where X = major version of FTOT, Y = breaking change to XML/Schema/Code, Z = non-breaking change to XML/Schema/Code. If a breaking change is made to the schema or code, the user must upgrade their XML file to meet the new schema standard.
3. <Scenario_Name> - a string used to identify the scenario in the reports and logs. Scenario Names should not contain spaces and is good practice to match the name of the scenario folder; e.g. 'Scenario_003_WheatStraw'.
4. <Scenario_Description> - a string used to describe the scenario. This is printed out in the reports and useful to quickly determine what is different between otherwise similar scenarios. For example if the unmet demand penalty is increased in a sensitivity analysis, calling that out in the scenario description is a good idea.
5. <Scenario_Inputs> - contains the following elements:
 - a. <Common_Data_Folder> - contains several files that are required by FTOT to run. For example, the default national multimodal network. The ftot_maps.mxd, and tableau_dashboard.twb are used for generating maps and Tableau dashboards, respectively. The files listed in the common data folder are not modified by FTOT. Instead, FTOT copies a local copy to the scenario folder and modifies that version.
 - b. <Base_Network_Gdb> - a string specifying the full path to the network FTOT will setup and use for the scenario. FTOT will copy the base version to the local scenario. Therefore, a single base network can be used across multiple scenarios without impacting the base version.
 - c. <Base_RMP_Layer>, <Base_Destination_Layer>, <Base_Processors_Layer> - are defined by a string specifying the full path to the FC containing the respectively facilities that FTOT will use for routing. Again, FTOT will create a copy of the base FC to the local scenario geodatabase and leave the base version untouched. In the case of the <Base_Processors_Layer>, it should be noted that these are user specified processor locations, and not candidates generated by the tool.
 - d. <RMP_Commodity_Data>, <Destinations_Commodity_Data>, <Processors_Commodity_Data>, <Processors_Candidate_Commodity_Data> - are full paths

to the CSV files that are used to define the commodity-facility relationships. The RMPs by definition only have “output” commodities that define the available supply from the facility. Destinations only have “input” commodities that define the total demand of a commodity at the destination. Processors have both “input” and “output” commodities that define the facility size and product-slate/processing ratio. The Processor Candidates commodity data file defines the maximum and minimum facility size, processing size, and cost formula for the amortized capital cost of the facility. NOTE: Processors are currently limited to a single input commodity, but may contain multiple output commodities.

- e. <Default_Units_Solid_Phase> and <Default_Units_Liquid_Phase> are used to specify the units FTOT will report the results in. The Pint python module is used to convert the quantities and units specified in the commodity input CSV files into the default units specified in these tags. Pint is able to handle a wide range of units, prefixes, and abbreviations. At run time, FTOT will process the default units with a simple conversion test and raise an error if it cannot use the user specified values because of a typo.
6. <Assumptions> - contains the vehicle load and CO2 emission factors for each mode and phase of matter. <Create_Network_Layer_Script> - Updates each network link with the cost functions defined for each mode and link type by the user. <Network_Costs> - The GIS module assigns costs to each link in the multimodal network based on the transport costs specified in the XML Scenario file. The dollar costs on the GIS network are the dollar amounts required to transport material over each particular link. The costs are defined as USD/1000 gallons for liquids, and USD/metric tonne for solid materials. The default cost values come from the [US BTS](#) average freight revenue per Ton-Mile. The 2004 data was used because it is the latest complete year reported in the table. The 2004 values were scaled to 2017 using the Producer Price Index. The liquid costs were estimated by multiplying the solid values by 2.87 to convert between thousand gallons and tons. The impedances apply a scaling factor to the base dollar cost. For example, FTOT is designed to prefer faster/larger links on the road network. Therefore, a multiplier of 1 is used for interstate highways, versus a multiplier of 1.3 for local roadways, resulting in a price of \$0.00045/gallon·mile for interstate highways and \$0.000585/gallon·mile for local roadways.⁴ Each mode is defined in the <Intermodal_Costs >
- a. <Railroad> - In Volpe’s internal FTOT network, rail impedance categories are based on FRA Rail Density Codes, which are sensitive, non-public data. In the public FTOT network, rail impedance categories are based on publicly available rail ownership data and whether or not the rail segment is part of STRACNET. In the public FTOT network, the final three impedance categories (Rail_Density_Code_2_Weight, Rail_Density_Code_1_Weight and Rail_Density_Code_0_Weight) are NOT utilized. These weights are used to encourage flows on portions of the rail network that see higher freight use (internal network) or part of STRACNET and associated with a Class 1 railroad (public network).

⁴ Assuming that a large tanker truck holds 8000 gallons, this corresponds to a cost of \$3.60 to \$4.68 per truck-mile, depending on functional class. Truck travel speed was calculated based on the 2007 estimated peak period speed estimated in the FAF for each link, assuming a minimum floor of 15 mph.

- b. <Truck> Weights are used to encourage flows on lower functional class roadways (e.g. interstates and highways). Truck_Interstate includes FAF Function Class 1. Truck_Principal_Arterial includes FAF Function Classes 2 and 3. Truck_Minor_Arterial includes FAF Function Class 4. Truck_Local includes all other FAF Functions Classes (excluding those above). More information on functional classes and the FAF road network is available at the following URL:
https://ops.fhwa.dot.gov/freight/freight_analysis/faf/faf4/netwkdbflow/index.htm
 - c. <Barge> Water impedances based on categorizing publicly available USACE freight volumes over the national waterway network. These weights are used to encourage flows portions of the waterway network that see higher freight use. The default value of 10.0 for waterways with no volume is meant to discourage flows on portions of the rail network with little to no documented freight use. Water_High_Volume includes waterways with volumes >= 10,000,000 tons. Water_Medium_Volume includes waterways with tonnage volumes < 10,000,000 and >= 1,000,000 tons. Water_Low_Volume includes waterways with tonnage volumes < 1,000,000 tons. Water_No_Volume includes waterways with no tonnage volumes documented. The source for the National Waterway Network and waterway network freight volume data is here: <https://usace.contentdm.oclc.org/digital/collection/p16021coll2/id/1450>
 - d. <Intermodal_Transloading_Costs>- intermodal costs specify the cost of switching from one mode on the network to another (e.g. rail -> road). Intermodal transfers can only take place as specified intermodal facilities on the network. A cost for solid material and a cost for liquid material transloading is specified by the user.
 - e. <Artificial_Links> Since facilities are not always located directly on the multimodal network, artificial links must be used to connect facilities to the network. A reasonable default is 5 miles, though longer artificial link distances can be conceivably used, particularly for the pipeline network where there is the least amount of detail. Shorter artificial link distances can be used when the facilities are known to be located directly on or adjacent to the relevant network modes
7. <Route_Optimization_Script>
- a. <Permitted_Modes> The following True/False flags determine whether or not a particular mode should be allowed for routing any flows in the scenario. The default is for all modes to be on.
 - b. <Capacity_Options> The following True/False flag determines whether network capacity should be considered as a constraint for flowing in the scenario. The default is for capacity to be off
 - c. <Background_Flows> The following True/False flags determine whether or not a particular mode should have its existing (background) flows considered. If True, then background flows (e.g. existing movements of freight on the network) will be considered. The default is for all modes to NOT have their background flows considered
 - d. <Minimum_Capacity_Level> The following setting (0-1) determines the minimum fraction of capacity that must be available for each network segment in the scenario. Setting

a value above 0 allows network segments that are already at capacity due to existing flows, to be traversible in this scenario up to the minimum capacity level

8. <Pulp_Solve_Options> Settings used to adjust the optimization parameters. Penalty_For_Not_Fulfilling_Depot_Demand is the also called the unmet demand penalty (UDP). This parameter applies a cost to each unit of demand at a destination that is not fulfilled. It is used to drive flow from in the scenario. Pulp_Solve_Options contains two additional parameters MaxSeconds and fracGap. These will be deprecated in future versions of FTOT because they are not supported by the PuLP developer.

4.3 Facility Data Input Files

In addition to the XML scenario file, FTOT requires a set of comma-separated-values (csv) files, which define facility-commodity information (e.g., supply/demand amounts, product slates and conversion efficiency, and minimum/maximum facility size, among other elements).

There are three comma-separated-variable input files available for facility definitions in FTOT. These files contains the facility name, facility type, commodity, quantity, units, and input/output flag for raw material producers (RMPs, which are the supply origin locations in the scenario), processors (which are way points or processing locations in the supply chain), and destinations (locations of demand for the final commodities).

The raw material producer, or origin, facilities file includes the following fields:

- facility_name; used to match the locations in the scenario gdb. NOTE: the facility_name in the input data CSV file must match the records in the facility_name field of the GIS FC. Facility_names should be unique and not duplicated. Care should be taken when using county or city names as these are often repeated across states. Facility_type, specifies the type of facility.
- Commodity; a unique name for the commodity. FTOT will create origin-destination pairs by matching the commodity names. For example, if an RMP specified commodity_A as an output, and a destination specified commodity_A as an input, FTOT would try and flow material from the RMP to the Destination.
- Value; the quantity of the commodity
- Units; the units of the quantity for the commodity
- Phase of matter; solid, liquid, (gas not currently supported)
- IO; specifies in the commodity is an input (i) into the facility, or an output (o) out of the facility. RMPs by definition ONLY have outputs, and destinations by definition ONLY have inputs. Processors must have at least one input AND at least one output commodity.
- Max_transport_distance (optional); a user specified maximum transport distance (miles) that this commodity can travel. This distance should be excluded from scenario runs in which all processors are fully defined for performance reasons. **However, note that this field is required in the Raw Material Producers input CSV file when FTOT is generating candidate processors.**

4.4 Facility Geodatabase (GDB)

Contains facility names and locations of destinations (dest), processors (proc), and raw material producers (rmps). Feature classes (FC) contain facility names and point locations of facilities. **Note: facility names must be unique** e.g. Middlesex County CT must be differentiated from Middlesex County MA. FTOT will match the geospatial location specified in the FC to the facility-commodity relationships specified in the CSV files defined above using the facility_name fields.

4.5 Default Data/Input Parameters

During FTOT development, certain costs, impedances and emissions factors were identified for scenario analyses and testing. These values are included in the FTOT default xml and network as a starting point for new users, but can be modified as desired. A few key default parameters and their sources are described below.

Table 1: Modal cost units and default values in FTOT.

| Transport mode | Cost (solids, per metric ton-mile) | Cost (liquids, per kgal-mile) |
|-------------------|------------------------------------|--------------------------------------|
| Roadway/Trucking | 0.19 | 0.54 |
| Railway/Rail car | 0.03 | 0.09 |
| Waterway/Barge | 0.02 | 0.07 |
| Pipeline | N/A | Actual tariff cost (varies by route) |
| Transloading cost | 12.35 | 40.0 |

Source for costs—[Bureau of Transportation Statistics Average Freight Revenue per Ton-Mile](#)—2004 data annualized to 2017 using the Producer Price Index

The default impedances are listed in Table 2.

Table 2: Modal impedance factors applied to the base modal cost based on modal link category (road) or existing flow bin (rail, waterway).

| Mode | Category / Bin | Impedance Factor |
|-----------------------|----------------------------------|------------------|
| Road | interstate | X 1.0 |
| | principal arterials | X 1.1 |
| | minor arterials | X 1.2 |
| | local roads | X 1.3 |
| Rail (public network) | STRACNET (Class 1 Ownership) | X 1.0 |
| | Other STRACNET | X 1.1 |
| | Class 1 Ownership (non-STRACNET) | X 1.2 |

| Mode | Category / Bin | Impedance Factor |
|----------|--|------------------|
| | Class 1 Rights (non-STRACNET) | X 1.3 |
| | Non-Class 2 (non-STRACNET) | X 1.4 |
| Waterway | | |
| | high volume links (> 10 million annual tons) | X 1.0 |
| | medium volume links (1 – 10 million annual tons) | X 1.3 |
| | low volume links (< 1 million annual tons) | X 1.6 |
| | 0 gross tons | X 10.0 |

Impedances listed in the table above are based on attributes contained within FTOT’s source network data. See Section 2.2.1 for more details on how to access these source datasets.

Network Capacity and Existing Flow

The capacity constraint can be enabled by mode to reroute flows once available capacity has been met. Network capacity and background volumes are collated or estimated from existing data sources and are included in the default multimodal network that is supplied with FTOT. Capacity is defined by the number of vehicles or vessels that can pass over a link in a 24 hour period.

- Road
 - Background flows and capacity data come from the Freight Analysis Framework Version 4 (FAF4) dataset
 - FTOT’s processed version of the FAF converts average annual daily traffic (AADT) to unidirectional daily flows
 - FTOT’s processed version of the FAF converts hourly capacity to daily unidirectional capacity
 - Available capacity = 1- (daily traffic / daily capacity)
- Rail
 - For Volpe’s internal network, background flows are based on FRA’s NARN, which includes segment-level estimated freight tonnage flows (NOTE: These background flow tonnages are non-public, sensitive data and neither the tonnages nor their binned counterparts are included in the public FTOT network).
 - The public FTOT network does not contain any background flow data for rail, as only publicly available attributes from FRA’S NARN are included in FTOT’s rail network.
 - For both the internal and public networks, daily unidirectional carload capacity is estimated based on track number and signaling type publicly available in NARN (and based on methodology developed by [Cambridge Systematics National Freight Infrastructure Capacity & Investment Study 2007](#))
- Waterway

- FTOT assumes waterway capacity only an issue at typical waterway chokepoints (e.g. locks)
- Background flow and capacity data based on US Army Corps Lock Characteristics and Lock Report data.
- FTOT network uses this data to estimate lock-specific barge equivalent flows and lock-specific barge equivalent capacity
- Pipeline
 - For crude oil pipelines, Oak Ridge National Laboratory's TRIM model estimates existing flows and capacity data for crude oil pipelines. This data is integrated into the default FTOT network. Crude pipelines not in TRIM are left without capacity constraint.
 - A source for product pipeline background flow and capacity data has not yet been identified, so these pipelines remain without capacity constraint in FTOT

5 FTOT OUTPUTS

FTOT generates four types of outputs. These are:

- 1) A human-readable text report containing the configuration history, results, elapsed run time, and any warnings given during the scenario.
- 2) A CSV file report that can be used to generate graphical dashboards showing summary map and statistics for the scenario, such as relative contributions of commodities and mode to cost, emissions, and VMT, amount of supply utilized and demand met in the scenario, and other elements. The CSV file can be used as an input to Excel, Tableau, or other data analysis programs.
- 3) A package tableau workbook (.twbx) that contains all the geospatial and summary result information required to display a Tableau dashboard in the free version of the Tableau Reader software.
- 4) Maps from each step of the analysis, including the network, facilities, candidate processing locations, and the optimal solution (optimal facilities, optimal flows, non-optimal facilities, etc.).

In addition to the four outputs listed above, the scenario results are stored in a sqlite database and in the scenario geodatabase. The user is free to create their own analysis or tools for reviewing this data.

5.1 Reports

Report outputs are generated as a csv file that is stored in the scenario folder. The csv report includes three groups of reporting results: the scenario configuration, the commodity-specific summary results and the facility summary results.

The scenario configuration details, reporting out the scenario name, description, and various input parameters from the xml, as well as the name and location of the GIS layers used to define the network and the facilities included in the scenario.

The commodity summary results include the following data by commodity and mode:

- Mass and volume movements
- Cost of movements
- Fuel burn
- Miles
- vehicles
- VMT
- CO2

The facility summary results include

- Facility utilization (amount of supply utilized or demand met)
- Facility utilization fraction (relative to total supply or demand)

5.2 Graphics Dashboard

A graphical dashboard output is created using Tableau (see example figures in accompanying Quick Start Documentation).

The Tableau dashboard is used to visualize the CSV report, and GIS facilities and routes from the analysis. It includes facility-level, commodity-level, and mode-level results. The results include quantities of material moved, utilization levels, CO₂, VMT, mileage, and costs.

To view the dashboard in Tableau:

- Double click on the tableau_dashboard.twbx file to open it in Tableau Reader.
- On the dashboard there are nine charts (NOTE: Charts are grouped by commodity type, and sorted alphabetically by mode).
 - From left-to-right, top-to-bottom:
 1. Key Input Parameters table
 2. Optimal Solution Routes by Mode (map)
 3. Size & Utilization of Facilities (map)
 4. Total vs. Optimal Supply & Demand Quantity
 5. Material Moved by Commodity & Mode
 6. Scenario Cost by Commodity & Mode
 7. Total vs. Optimal Supply & Demand Facilities
 8. VMT by Commodity & Mode
 9. CO₂ Emissions by Commodity & Mode
- Notes & Tips on Dashboard Symbology
 - Mode colors:
 - Pipe – brown
 - Rail – purple
 - Road – red
 - Water – blue
 - Facility Type colors:
 - RMP Optimal Supply – dark green
 - RMP Total Supply – light green
 - Destination Optimal Demand – dark orange
 - Destination Total Demand – light orange
- Notes & Tips on Dashboard Functionality

- In upper-right corner, the dropdown menu changes what facilities are viewed on the map “Size & Utilization of Facilities” map. The default view is “Demand”.
- Hover over an item on a chart (e.g. bar, icon, point, etc.) to display tooltip which contains additional data.

5.3 Maps

The FTOT map outputs are generated in the scenario’s Maps directory (see example figures in accompanying Quick Start Documentation) and include the following:

- The multimodal network used in the optimization (01_S_Step)
- Raw Material Producers (02a_F_Step_RMP, 02a_F_Step_With_Labels)
- User Defined Processors (02b_F_Step_User_DefinedPROC, 02b_F_STEP_User_Defined_PROC_With_Labels)
- Ultimate Destinations (02c_F_Step_DEST, 02c_F_Step_DEST_With_Labels)
- All Facilities (02d_F_Step, 02d_F_Step_With_Labels)
- Processor Candidates (03a_F2_Step_Processor_Candidates, 03a_F2_Step_Processor_Candidates_With_Labels)
- All Processors (03b_F2_Step_Processors_All, 03b_F2_Step_Processors_All_With_Labels)
- Final Optimal Routes (04a_O_Step_Final_Optimal_Routes_with_Commodity_Flow, 04b_O_Step_Final_Optimal_Routes_with_Commodity_Flow_NO_LABELS, 04c_O_Step_Final_Optimal_Routes_with_Commodity_Flow_JUST_FLOW)
- Aggregated Flows by State of Matter (04d_O_Step_Final_Optimal_Routes_With_Commodity_Flow_Aggregated_Solids, 04e_O_Step_Final_Optimal_Routes_With_Commodity_Flow_Aggregated_Liquids)
- Optimal and Non-Optimal Raw Material Producers (04f_O_Step_Optimal_and_Non_Optimal_RMP)
- Optimal and Non-Optimal Processors (04g_O_Step_Optimal_and_Non_Optimal_PROC)
- Optimal and Non-Optimal Ultimate Destinations (04h_o_Step_Optimal_and_Non_Optimal_DEST)

Note that many maps include variations that include/exclude labels. Since each scenario is different and all the maps are automated, some variations may work better than others depending on the scenario. The user is welcome to develop their own maps within ArcGIS leveraging the map document automatically generated in each scenario’s output.

In order to properly map aggregate flow volume, the map document (mxd) provided with FTOT must be modified by the user to reflect the desired classification bins. This involves modifying the symbology in the O_STEP_FINAL_OPTIMAL_ROUTE_SEGMENTS_AGGREGATED_LIQUID or O_STEP_FINAL_OPTIMAL_ROUTE_SEGMENTS_AGGREGATED_SOLID mxd group layers, depending on which one is of interest and/or relevant to the scenario.

- If the optional M2 Step of FTOT is run, additional time and commodity specific mapping is output into the scenario's Maps_Time_Commodity folder. These include: Maps showing location of flows for each commodity in the scenario
- Maps showing location of flows for each time step in the scenario
- Maps showing location of flows for each commodity/time step combination in the scenario
- An animation (.gif) representing flows for all time steps in the scenario
- Maps showing aggregate flows—both with and without facilities

Like the other aggregate flow maps, the mxd must be modified by the user to reflect the desired binning of the data.

6 SETTING UP A SCENARIO

As previously discussed, FTOT requires several input files to fully specify a scenario.

A scenario configuration file (scenario.xml) is used to define the location of key input data such as the multi-modal network, facility locations, and facility commodity information defined the user, as well as key input parameters and assumptions for the run.

A batch file (run.bat) is used to automatically run the sequence of steps necessary for an FTOT scenario to run. The batch file references the scenario configuration file for each step and must be modified for each scenario created.

Both of these files can be generated by the FTOT Tools suite.

Alternatively, the Quick Start scenarios and documentation are designed to familiarize the user with several scenarios of various complexities and numerous exercises for the user to explore. The quick start scenarios can also be used as a template for setting up new scenarios.

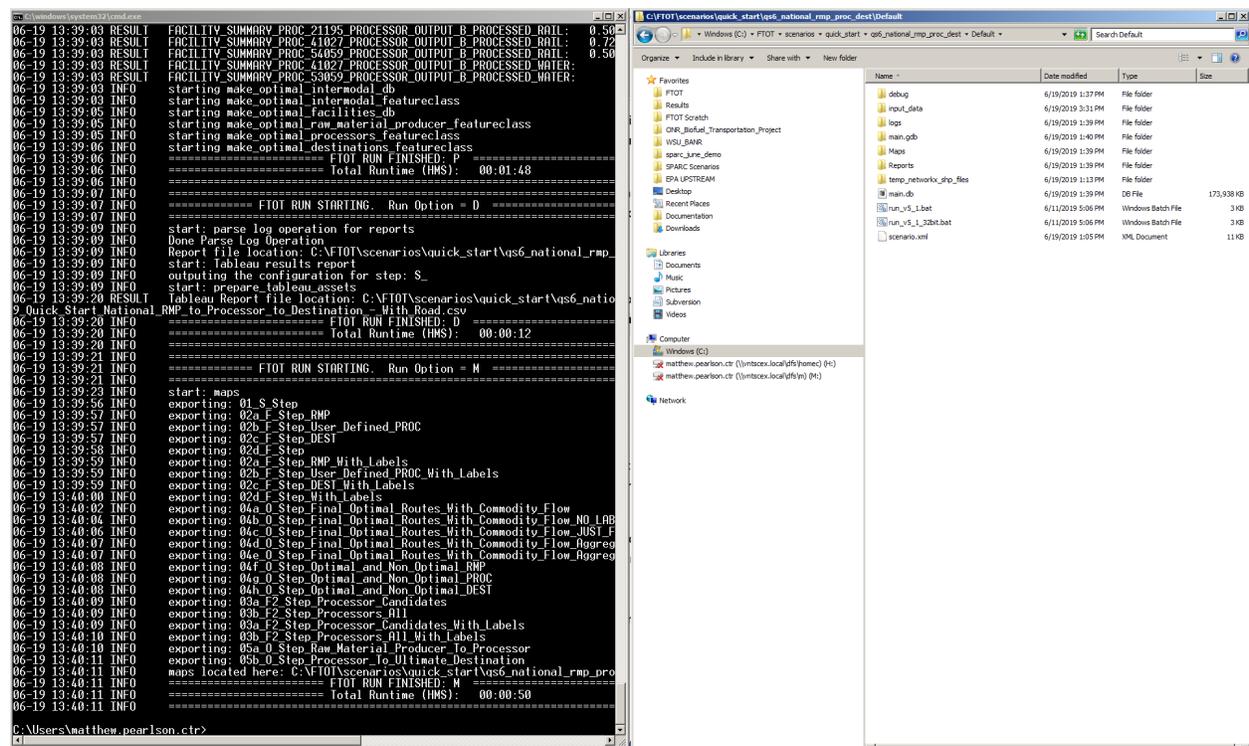


Figure 7 - Split screen practice for running FTOT scenario. The command line is on the left, and the scenario opened in the Windows Explorer is on the right.

7 RUNNING A SCENARIO

FTOT is a command line program that executes in the Windows shell. It is a recommended practice to run the scenarios with a command line window on the left, and the scenario folder opened in an explorer window on the right.

A scenario can be started by dragging the run.bat file located in the scenario folder into the command line and hitting [enter]. During the run, information is logged to the command line during the run to keep the user informed of the progress of the run and report results from each step. Detailed log files are created for each step and timestamped for each step in the .\logs folder. As new files are created in the various FTOT steps, they will become visible in the explorer window on the right.

8 SUPPLEMENTARY TOOLS

A set of supplementary FTOT tools is provided with the program to assist the user with a variety of common tasks. These tools include the following:

- **xml_tool:** Generates an xml scenario file based on either a 1) FTOT template with defaults or 2) existing, older user-generated xml file
- **bat_tool:** Generates a bat file for running a new FTOT scenario based on user-provided inputs.
- **compare_tool:** Generates a GIS-based feature class called 'optimized_route_segments_compare' containing a geometric comparison of two distinct scenarios of the user's choice. The user can add this feature class to an ArcGIS map document and symbolize based on the "Change_Commodity_Flow" field to visualize where flows either increased, decreased, or stayed the same when Scenario 2 is compared to Scenario 1. Later versions of FTOT will leverage this data for further visualization in Tableau.
- **Input_csv_templates_tool:** Generates the input data csv files for RMPs, processors, and destinations. The template for RMP input does not include the optional max_transport_distance field, so the user should add this if necessary (notably when running candidate generation scenarios).

The FTOT tools interface can be accessed by running the following in a command-line window:

32-Bit Python

```
C:\Python27\ArcGIS10.6\python.exe c:\ftot\program\tools\ftot_tools.py
```

64-Bit Python

```
C:\Python27\ArcGISx6410.6\python.exe c:\ftot\program\tools\ftot_tools.py
```

9 TROUBLESHOOTING

- Import errors. Python is missing a required module. See the installation guide [<https://github.com/VolpeUSDOT/FTOT-Public/wiki>], and verify that the module is installed in the Arcgis python environment.
- If a user is having difficulty running a batch file (BAT) or process, the user may need to include 'python' on the lines specifying the process. For example, the user would insert a 'python' before the syntax that runs the process. In pseudocode:

```
python PythonScriptFilePath XMLfilePath ProcessYouWantToRun  
USAGE EXAMPLE: python C:\FTOT\Program\python.py  
C:\FTOT\Scenarios\midwest1\scenario.xml
```

- If the user has a prior version of Python installed, there may be some conflicts with the new Python install included with ArcGIS. In that case, the user may need to add the following line to the BAT file:

```
path C:\Python27\ArcGIS10.6;C:\Program Files (x86)\ArcGIS\Desktop10.6\bin;%path%
```

- If the user is unable to install the ArcGIS 64-bit background geoprocessing, the user must run FTOT in 32-bit.
 - For Quick Start scenarios, run the batch files ending in run_32bit.bat.
 - For large or complex scenarios (e.g. Quick Start 6, a national scenario), FTOT may not run properly in 32-bit.
- If FTOT cannot find the GIS feature classes or facility commodity CSV files it will throw an exception error and exit the program. If this happens ensure that the appropriate scenario assets exists and the path is properly specified in the XML.
- Possible causes of no-flow solutions:
 - The unmet demand penalty is too low. Adjust the unmet demand penalty in the scenario configuration file. In general, it may be necessary to raise this penalty when any other cost (e.g. rail transport) is raised, or else the optimizer will conclude that it is more optimal to transport less material. As a general guide, the unmet demand penalty will likely work best if set to be 10-50 times the average actual transportation cost. It is advised to run scenarios with multiple unmet demand penalties to explore the sensitivity of a given analysis.
 - The candidate processing facility capital cost is too high. Increase the unmet demand penalty in the scenario configuration file, or decrease the capital cost in the processors candidate commodity CSV file.

- The processing facility lower bound are too high. Adjust the minimum processor size in the facility commodity CSV file. In scenarios involving candidate generation, a no-flow solution will occur if a value for maximum transport distance is not specified in the raw material producer commodity CSV file. Ensure that a maximum transport distance (in miles) is specified.
- No-flow solutions are possible if facilities are not located close enough to segments in the FTOT transportation network. To address, raise the artificial link distances specified in the scenario's XML to ensure facilities connect to the network (for most modes, the default artificial link distance is 5 miles). Lack of connectivity is unlikely unless facilities are located far away from populated areas or are located outside of the continental United States, which is the current extent of the FTOT transportation network.
- For capacity-constrained scenarios where existing background flows are being considered, no-flow solutions are possible when facilities are not accessible due to capacity constraints. Turn off the capacity constraint or increase the minimum available capacity in the scenario's XML to address this issue.

9.1 Known Bugs and Issues

- Errors with very large numbers:
 - If it seems like numbers are being capped or cut off in unexpected ways, the most likely cause is an improperly set constraint – double check the upper and lower bounds on processing facility flow, transportation and transloading costs, and demand vs production values. If none of those address the problem, it is possible that limitations of the programming language in handling very large numbers are having an impact. The COINMP_DLL solver is written in C++, which limits values depending on the data type to $\pm 2,147,483,647$ (Integer), 38 places (Float), 308 places (Double).

10 ACKNOWLEDGEMENTS

This research was funded by the U.S. Federal Aviation Administration Office of Environment and Energy through FAA contract 693KA9-18-N-00022 under the supervision of Nathan Brown. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the FAA.

11 REFERENCES

"pulp-or". 2007. COIN-OR.
 Bertsimas, D., and J. N. Tsitsiklis. 1997. Introduction to Linear Optimization. Athema Scientific, Belmont, MA.

Bureau of Transportation Statistics. 2011. Intermodal Terminal Facilities, available at http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/national_transportation_atlas_database/2011/index.html. National Transportation Atlas Database (NTAD)

Carreira-Perpiñán, M. Á. 2014. Simplex Method. MathWorld--A Wolfram Web Resource, created by Eric W. Weisstein.

COIN-OR. 2007. Clp.

Federal Highway Administration. 2013. Freight Analysis Framework (FAF) version 3.4.

Leonard B. Levine Associates. 2014. US Petroleum Product Pipeline Movements OD Pair and Base Rate Report. len.levine@lensonwashington.com.

McDuffie, W. 2013. Teleconference discussion of fuel transport costs. Volpe National Transportation Systems Center.

pulp-or. 2014. puLP: An LP modeler in Python.

Surface Transportation Board. 2011. Carload Waybill Sample.

United States Army Corps of Engineers, and Bureau of Transportation Statistics. 2011. U.S. Army Corps of Engineers Navigable Waterway Network, available at http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/national_transportation_atlas_database/2011/index.html. National Transportation Atlas Database (NTAD)

USDA. 2014. National Agricultural Statistics Service: Quick Stats.