Freight and Fuel Transportation Optimization Tool Technical Documentation

FTOT 2022.2 Public Release Version

July 1, 2022

DOT-VNTSC-FAA-22-06

DOT/FAA/AEE/2022-01

Prepared for:

Federal Aviation Administration
Office of Environment and Energy
Washington, DC

U.S. Department of Transportation
John A. Volpe National Transportation Systems Center

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			*Form Approved* *OMB No. 0704-0188*
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 July 1, 2022	3. REPORT TYPE AND DATES COVERED Document, 4/2022-6/2022	
4. TITLE AND SUBTITLE Freight and Fuel Transportation Optimization Tool Technical Documentation: FTOT 2022.2 Public Release Version			5a. FUNDING NUMBERS FB48CY21/_M359 and _M360
6. AUTHOR(S) Olivia Gillham, Kirby Ledvina, Kristin C. Lewis (PM), Mark Mockett, Alexander Oberg, Matthew N. Pearlson, Kevin Zhang			5b. CONTRACT NUMBER 693KA9-20-N-00013
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-22-06
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/2022-01
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 2022.2. This documentation is updated quarterly and available on the Public FTOT GitHub repository: https://github.com/VolpeUSDOT/FTOT-public			
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. 239-18

298-102

Table of Contents

Table of Figures and Tables

1 INTRODUCTION

Volpe created the Freight and Fuel Transportation Optimization Tool (FTOT) to support the Federal Aviation Administration (FAA), the Department of Energy (DOE), and the US Navy's Office of Naval Research in assessing 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 the 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 current version of FTOT is considered an "experts' release," requiring the user to have some knowledge of Python and ArcGIS Pro to run. It enables the user to analyze scenarios using either the default network and parameters or user-defined details. 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. 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. The first official full version of FTOT (2019.3) was released in October 2019. Since then, FTOT versions have been released quarterly. FTOT version 2021.1 was the first version of FTOT to be compatible with ArcGIS Pro and Python 3, ending compatibility with ArcGIS and Python 2—software and a programming language version which have reached end-of-life and are no longer being supported.

[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.

New FTOT functionalities and user materials are added every quarterly release. Specific details on updates associated with the latest release (and previous releases) can be found in the change log on the GitHub code hosting site at https://github.com/VolpeUSDOT/FTOT-Public/releases.

1.1 FTOT Use Cases

FTOT is a scenario testing tool for identifying the optimal solution for a supply chain scenario and enabling exploration of the scenario space (e.g., variations of one or many scenario elements).

One of the unique aspects of FTOT is that it can take into account not just multiple origins and destinations, but also intermediate waypoints ("processors") that the materials in the scenario must pass through (see Figure 1). Furthermore, the supply chain-oriented approach means that the materials that pass through those processors may experience losses and/or change in identity (for example, crude oil passes through a refinery, and the outputs are fuels such as gasoline, diesel, jet fuel, and naphtha). FTOT can take into account the change in identity and the relationships between inputs and outputs, as well as facility characteristics such as minimum and maximum input capacity, build cost (if not already built), and daily schedules to optimize the solution for the entire supply chain.

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 availability down to the daily level. The supply chain can be expanded by any number of steps, commodities, processor types, and times (e.g., facility availability).

Because of this unique approach, FTOT can be used to explore the effects of various aspects of supply chain conditions and characteristics. FTOT is particularly well suited for scenarios evaluating effects of:

- Changes in transportation infrastructure (e.g., gain or loss of transportation elements such as road links, bridges, etc.) or multimodal transfer points. This can include the assessment of hazard scenarios (e.g., flooding, etc.) based on loss of links in the network.
- Changes in supply chain / industry infrastructure (e.g., gain or loss of specific facilities or facility types within the supply chain like a new processing location, or separating processing of materials into one or more steps), or changes in capacity at facilities (e.g., expanding a processor to enable more throughput).
- Changes in supply of raw material from origins or demand at destinations (e.g., more agricultural production at origins, change in demand among destinations or overall).
- Effects of supply chain development options based on FTOT generation of potential processing facilities or inclusion of externally-generated candidate processing locations.
- Supply chain and optimal network resilience in the face of cumulative disruption.

FTOT enables the user to explore scenario variants and compare the optimal solutions with regard to routing and flow allocations and associated performance metrics including transport cost, emissions, vehicle miles traveled, and mileage on the network.

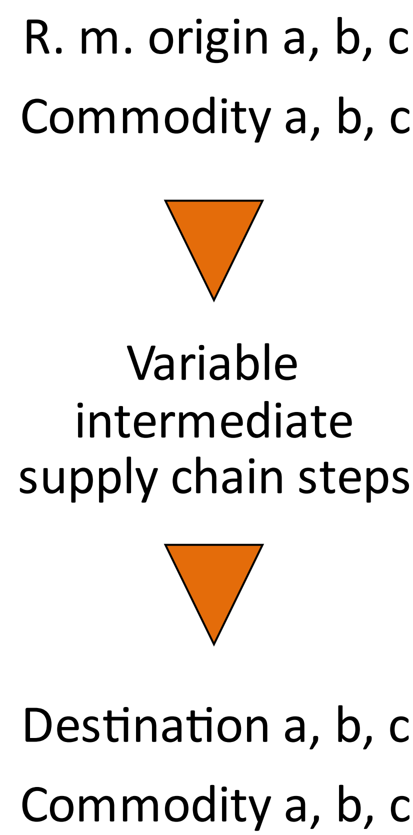

Figure 1: Generalized FTOT scenario structure incorporating supply chain elements.

2 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 as described in this section. Figure 2 demonstrates the FTOT architecture and data flow.

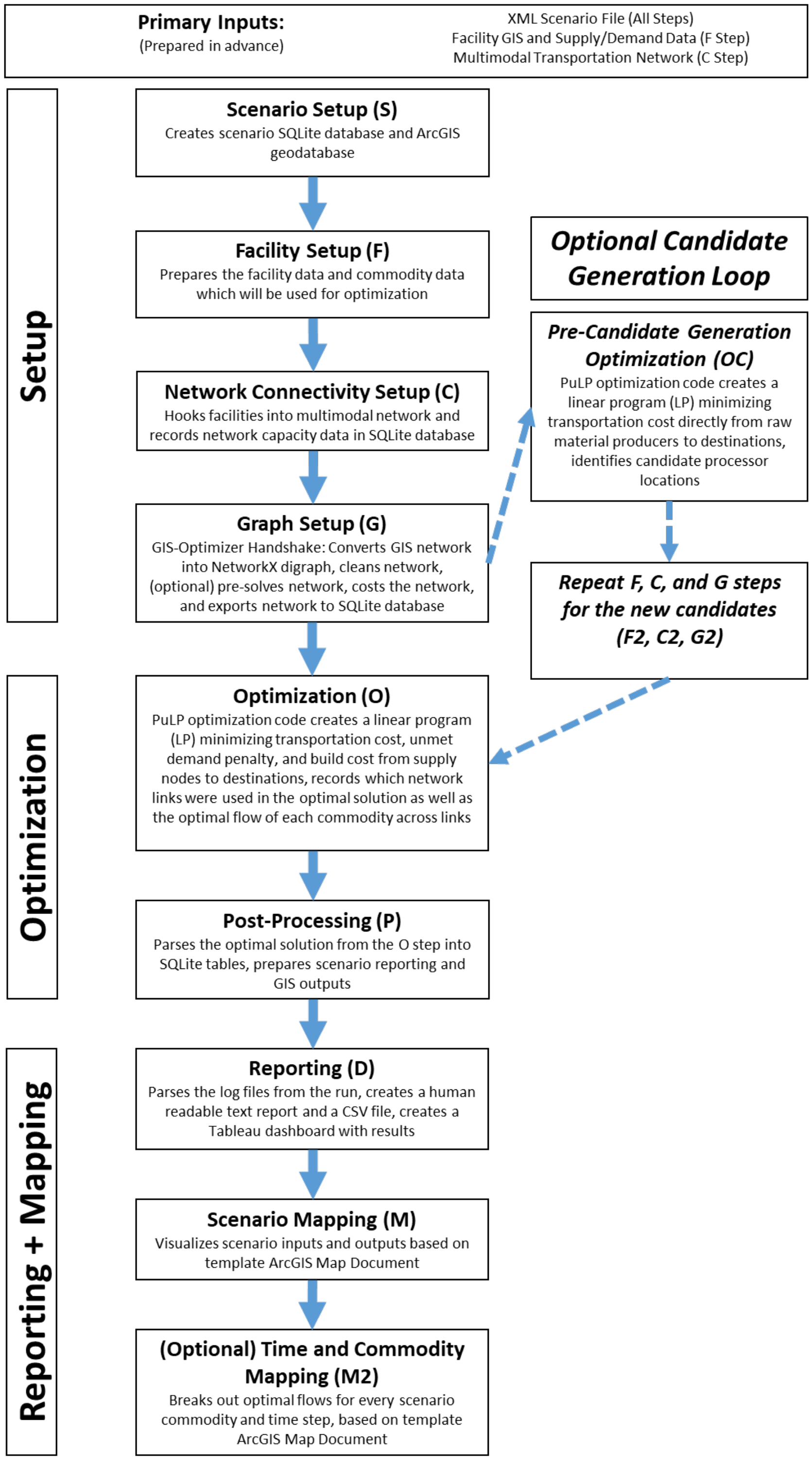

Figure 2: Analytical tool data flow schematic showing the key components/roles of each component of FTOT.

2.1 Key Software Modules in FTOT

The following software tools are integral to the use of FTOT. Installation instructions for these required software modeling tools are described in the standalone FTOT User Guide.

Python is the scripting language supporting FTOT. FTOT relies specifically on the Esri ArcGIS Pro installation of Python, which among many default modules includes the Arcpy module for GIS functionality. Additional supporting Python modules are also installed on top of the base Python installation that is stored within the FTOT installation directory.

ESRI ArcGIS Pro supports the GIS component of the tool. The Arcpy module is a Python-based tool that allows access to ArcGIS Pro functionality. FTOT uses Arcpy to process the multimodal network and facility locations defined for the scenario. Components of ArcGIS Pro 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 Pro 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 the FTOT optimization problem, 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. In addition, a network pre-solve functionality is included in FTOT, as specified in the scenario configuration file, for analyses that do not use capacity, candidate generation, nor maximum allowable transport distance. When this functionality is enabled, the NetworkX shortest_path algorithm is used to identify links for use in the optimization problem.

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.

2.2 FTOT GIS Network

Elements of the default FTOT multimodal network include feature classes for road, rail, water, and pipeline networks covering the continental United States. 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. The network is based on the authoritative sources and Volpe-produced datasets described below 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 the crude and product pipeline networks, each pipeline category has three feature classes associated with them: 1) pipeline feature classes representing the segments associated with the pipeline network; 2) pipeline_trf_sgmts feature classes storing 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 ensure that pipeline movements are restricted to the routes that tariffs allow. Unlike the other modes (road, rail, and water), pipeline movements can only enter and leave the pipeline network at locations where a tariff indicates it is possible.

2.2.1 Data sources

The base multimodal transportation network included with FTOT is limited to the continental United States and is 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). Note that the road network is not comprehensive and is limited to the National Highway System and other major freight corridors as defined by FHWA.
    - 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. The rail network included with FTOT is a comprehensive representation of all mainlines and major industrial leads, though it does not include minor industrial leads and yard trackage.
    - 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. The waterway network is considered comprehensive.
    - 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 compiled 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. Movements on the pipeline network are constrained to the origin-destination pairs designated in the dataset.

- 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

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. The pipeline network is limited to all known public pipeline tariffs compiled by LawIQ. Private pipelines and public pipeline tariffs that could not be assigned to a pipeline with known geospatial data are not included.

2.2.2 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 networks). 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.

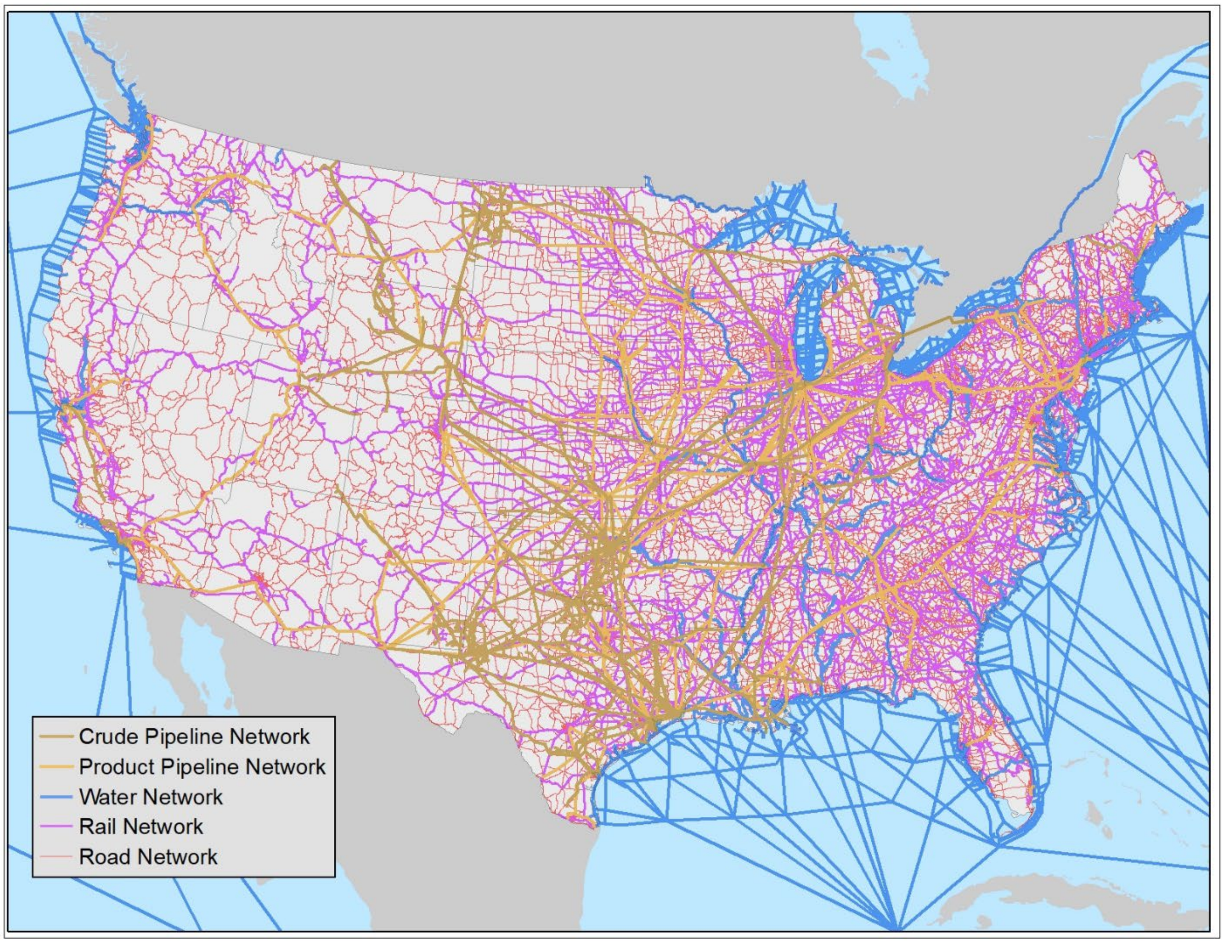

Figure 3: FTOT 2022.2 multimodal transportation network.

Figure 3 shows the resulting multimodal network that FTOT uses for routing analyses and display of results.

2.2.3 Network cost attributes

Network costs are calculated based on link attributes and user-defined inputs. Two types of network costs are recorded—the dollar cost and routing cost. For road, rail, and water networks, the dollar cost is a simple calculation based on the segment's length and a per-mile cost defined in the scenario configuration XML.

For pipelines, a rounded tariff cost is used based on the input data. Pipeline transportation rate and origin-destination information are 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. Private pipelines and some public pipeline tariffs that could not be assigned to pipelines with known geospatial data are not included.

In addition to a dollar cost, a routing cost is stored for road, rail, and water segments. This cost incorporates weights based on the network mode and link attributes. These weights are configurable and set in the scenario configuration XML. Weights are used to encourage flows on main links of the network and penalize minor link usage (e.g., to encourage use of the interstate highway network over more local roadways).

2.2.4 Customizing the FTOT GIS network

For most FTOT scenarios, users should utilize the default multimodal GIS network that is available for download with other supporting data through the FTOT GitHub site. Due to the complexity of the underlying multimodal GIS network structure, FTOT does not currently support using other networks. The network currently requires certain attributes (some of which are not easily available in other networks) and must be fully connected and flowable—including the existence of artificial links that connect the network segments to intermodal facilities.

In some circumstances, an FTOT user may wish to customize the existing default FTOT multimodal network. For example, a user may wish to add an additional network segment (or a handful of new network segments) to the network. This is achievable provided the FTOT user has some existing GIS experience and the patience to customize the GIS layers provided by FTOT, as the process is not automated. Adding network segments is not recommended if the user intends to add more than a dozen or so additional network segments, due to the complexity of integrating new data into an existing network. The FTOT User Guide walks through this process in detail.

In addition, segments can be automatically removed from the default FTOT network (i.e., 'disrupted') through the creation of an optional disruption data CSV. This is covered in Section 3.10 and in the User Guide.

2.2.5 Connecting facilities to the network with artificial links

In addition to the multimodal 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 both 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.

The process of connecting facilities to the modal networks is performed using artificial links. During an FTOT scenario run, FTOT copies user-specified facilities into the scenario network geodatabase and programmatically links them to the multimodal network by constructing artificial links. For each combination of facility location and modal network (e.g., road, rail, waterway, crude pipeline, product pipeline), FTOT uses Arcpy methods to determine the nearest point in the network to the facility. After comparing to user-specified maximum artificial link distance parameters, FTOT constructs new straight-line artificial links for each valid combination.

The fully connected network is stored in a feature dataset and includes the multimodal network, intermodal facilities, and origin/processor/destination facilities feature classes. More information on preparing facility location data for a scenario can be found in Section 3.4 and the User Guide.

2.2.6 Exporting the data

After the facilities are connected to the multimodal network with artificial links, the network feature classes are exported by FTOT 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.

2.3 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, 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 steps performed by FTOT during the finalization of the NetworkX graph include:

- 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 and 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, the routing cost is stored that includes weights based on the network mode and link attributes. Weights 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.

FTOT includes an optional feature to pre-solve the network using the NetworkX shortest path algorithm. The network density reduction (NDR) pre-solve step reduces run time for computationally intensive scenarios, but is currently not compatible with candidate generation, capacity, and maximum allowable transport distance. Reference Scenario 6 in the Reference Scenarios documentation provides an example of this functionality. Shortest paths are identified for each source and target location sharing a commodity. For example, RMP to Destination, RMP to Processor, and Processor to Destination source and target pairs are solved with the shortest path algorithm for the corresponding commodity. The routing cost is used as the weight for each link used by the shortest path algorithm.

The links identified in the shortest path solution are then stored in the SQLite database in the 'shortest_edges' table. The optimizer generates the linear program using only edges contained in the 'shortest_edges' table. This significantly reduces the complexity of building the problem in the O1 step and finding the solution in the O2 step. In addition, the links identified in each shortest path solution are stored in the SQLite database in the 'route_edges' table in order from source to target facility. Each route used in the optimal solution is also output in a timestamped optimal_routes_TIMESTAMP.csv file located in the Reports folder of the scenario, with descriptive statistics of the length, routing cost, dollar cost, and phase of matter of the commodity along each route.

The following section describes the optimization—how the problem is defined and solved using the linear programming solver.

2.4 Optimization

The goal of the PuLP optimization process is to minimize the total cost of transporting material from the origin facilities, through the processors (if included in the supply chain), and on to the destinations to maximize demand fulfillment and minimize scenario-wide cost. 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 processors (as either a fixed cost for processors with known locations or a unit cost for FTOT-generated candidate processors), and penalties for not meeting demand. The three components of the optimal cost are:

1) The actual dollar costs of transporting the material plus additional routing costs from impedances (weightings and penalties that force the tool to favor desirable characteristics of the routing, e.g., favoring larger roads over smaller).
2) The capital cost of building any candidate facilities that are part of the solution. This is typically an amortized cost based on the time step of the scenario to model realistic cost-benefit analyses

of transportation versus construction costs, but ultimately the cost is set by the user to best serve the scenario. Candidate processor facilities may be identified by FTOT in the candidate generation steps and/or supplied by the user if their potential locations are known; capital costs are specified per unit of commodity or as a fixed cost (e.g., from a siting tool), respectively. 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 Tsisiklis 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 Tsisiklis 1997) using the simplex algorithm (Carreira-Perpiñán 2014). Any users wishing to use an alternate solver can refer to the COIN-OR documentation for configuring a solver and replace any instance of “PULP_CBC_CMD” in the FTOT code (either in ftot_pulp.py or ftot_pulp_candidate_generation.py) with the name of the solver they wish to use.

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.

2.4.1 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 metric ton-mile for solids or per thousand gallon-mile for liquids by default. Users can specify units other than metric tons and thousand gallons.
   - Pipeline costs are based on tariff data described in Section 2.2.
   - Capital cost for candidate processor construction is also considered if candidate processors are provided by the user or 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. Average daily network capacity and background

volumes are estimated from existing data sources which are further described in Section 3.10.

3) **Weightings and penalties – 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. When limited geographic scope or time periods 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.

4) **Weightings and penalties – short haul penalty** for short movements on the water and rail network
   - A penalty for rail and water is added to the routing cost for all artificial links that connect facilities with the rail and water network. By default, this penalty is equal to the difference between the costs of routing 100 miles on local roads and 100 miles on rail (for rail) and the difference between the costs of routing 100 miles on local roads and 100 miles on water (for water). The magnitude of this penalty can be adjusted in the scenario configuration file.

- The penalty does not impact the actual dollar costs as reported in the results—it is leveraged as an additional impedance within the optimization to encourage the use of road for shorter commodity flows, reflecting a reality where rail and water are typically utilized for longer commodity flows.

5) **Minimum and maximum flow requirements** applied to processing facilities
    - A processing facility may have a defined capacity. The user provides this as a maximum input capacity for the processor. This is defined in the max_processor_input field of the proc.csv input file.
    - The user can also provide a minimum input capacity for the processor, which only applies if the processor is in use. That is, even though a minimum capacity is defined, the flow through a processor is allowed to be zero. This is defined in the min_processor_input field of the proc.csv input file; if not provided, the default minimum capacity is set to one-half the maximum capacity.
    - If no maximum input capacity is specified, then neither the maximum nor the minimum capacity is restricted.

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. The base costs represent the per metric ton-mile or per thousand gallon-mile dollar cost of traversing each mode. The impedances, or weights, act as multipliers on the dollar cost to produce a distinct routing cost, which helps encourage FTOT to route on portions of the network where one would expect to see more flow (for example, interstate highways over local roads, and Class 1 railways over Class 2 railways).
    - 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 3.10 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 on 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

- When 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:

Table 1: Decision variables used in the FTOT optimization problem.

Variable	Explanation
x_{iabjab}	Flow, in units / 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.[2]
$x_{jabja(b+1)}$	Storage of commodity a at origin facility storage location j from time period b to time period b+1, in units / time period.
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 units / 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 commodity 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 units / time period.
$x_{mabcnbc}$	Flow, in units / 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 commodity being transported.
u_{bnc}	Unmet demand of commodity c at destination n during time period b, in units / time period.
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.

Table 2: Fixed coefficients used in the FTOT optimization problem.

Coefficient	Explanation
U_{Piab}	Upper bound on flow of commodity a out of origin facility i_{ab} during time period b, in units/time period.
C_{jabkab}	Transportation cost, in \$ / unit, to flow commodity a during time period b from origin facility storage j_{ab} to processing facility k_{ab}.
F_k	Fixed cost in \$, to build the processing facility k. If processing facility k is used at all, this cost is paid once.

[2] "units" are defined by the user in the scenario configuration 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.

Coefficient	Explanation
U_{bk}	Upper bound on flow out of processing facility k, in units / time period b.
L_{bk}	If processing facility k is used, lower bound on flow out of processing facility k, in units / time period b.
$C_{kabmabc}$	Transportation cost, in $ / unit, from processing facility k to destination storage m, of commodity c in time period b processed from commodity a.
S_{kac}	Conversion factor at the processing facility k: inbound units of commodity a x S_{kac} = outbound units of commodity c.
P_{nc}	Penalty ($ / unit) for not meeting demand at destination n for commodity c for any time period.
D_{bnc}	Demand, in units / time period, at destination n of commodity c during time period b.
M_r	Maximum flow (capacity) in units / time period on edge[3] r. Using capacity in FTOT assumes time periods are daily.

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

Minimize **total cost** = $\sum_{jabkab} C_{jabkab}\, x_{jabkab} + \sum_{k} (F_k\, y_k) + \sum_{kabmabc} C_{kabmabc}\, x_{kabmabc} + \sum_{bnc} (P_{nc}\ u_{bnc})$

Subject to the following:

Table 3: Constraints used in the FTOT optimization problem.

Constraint	Explanation
(1) For each processing facility k_{ab}, $S_{kac} \sum_{jabkab} x_{jabkab} - \sum_{kabmabc} x_{kabmabc} = E_{kabc}$	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}, $y_k U_{bk} - \sum k_{abmabc}\ x_{kabmabc} - E_{kabc} \geq 0$	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 ($y_j = 0$), this constraint requires the flow out of the processing facility to be 0.
(3) For each processing facility k_{ab}, $y_k L_{bk} - \sum k_{abmabc}\ x_{kabmabc} - E_{kabc} \leq 0$	If a processing facility is used, the flow out of it over any time period must exceed the lower bound, summing all commodities.
(4) For each origin facility i_{ab}, $\sum_{iabjab} x_{iabjab} \leq U_{Piab}$	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, $\sum_{mabcnbc} x_{mabcnbc} + u_{bnc} = D_{bnc}$	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.

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

Constraint	Explanation
(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 formulation of the problem of moving 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.

2.4.2 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 relationship, as well as the facility maximum and minimum size, minimum aggregation size for generating a candidate processor, and the amortized capital cost as a function of size.

- In order to get a solution, the user must set a maximum transport distance for raw materials. This field is mandatory for FTOT runs with candidate generation.
- Potential paths are built up starting from origin points out to maximum raw material transport distance.
- Candidate generation requires two rounds of optimizations. In the first optimization, commodities flow from raw material producers towards ultimate destinations as the raw material commodity. Upon reaching the max transport distance, FTOT converts the raw material into the processed commodity demanded by the final destinations using the most efficient conversion process specified in the candidate processor input file. From this point, FTOT flows these processed commodities to the final destinations. These routes are used to identify potential raw material aggregation points along a lowest-cost path from the origin to the destination.
- After the first optimization, a post-processing step looks for points on the network where the raw material was aggregated on the network at an amount above the minimum aggregation size to support a processor facility. Raw material that aggregates beyond the user defined maximum raw material transport distance is ignored in the aggregation calculations.
- Candidate nodes are then assigned a processor ID 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 (the second round) is then rerun and proceeds as usual.

2.4.3 User-provided candidate processor option in optimizer

FTOT can also accommodate the use of candidate processors provided by the user from outside FTOT (e.g., identified by a siting tool). Like other facilities, FTOT expects two pieces of information about externally-generated candidate processor facilities: the geospatial location specified in the processor feature class, and a facility-commodity CSV file.

- In order to indicate that a processor is a candidate facility as opposed to an existing facility, the user must include the build_cost field in the processor commodity input file representing the amortized cost of building the facility across the given time period.
- The maximum raw material transport distance also applies to candidate processors provided by the user. While the maximum transport distance column is an optional input for the raw material producers input file when user-provided candidate processors are used, if it is provided then it will be applied in choosing optimal processors for routing.

FTOT reads in these facilities like any other input. The only difference is that the optimizer includes a cost penalty for building these facilities as specified by the user in the "build_cost" field. The user is able to incorporate both candidate processors that have been externally generated and those that are created during the FTOT candidate generation steps into their scenario.

2.4.4 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 the "O" step, the standard optimization. It will use that solution and define more specific flow paths. Running it from scratch, instead of after the "O" step, 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.

2.5 Post-Processing and Reporting

Once PuLP has found the optimal solution to the FTOT optimization problem, the decision variables are parsed into supply chain attributes and scenario metrics are calculated. Processors used in the scenario, optimal route flows, unmet demand at destinations, optimal storage flows, and excess material are identified from the optimal solution. Then, optimal route feature classes are constructed from the SQLite database. Network attributes like functional class and urban code are attached to the feature classes, which are then saved for mapping. Similarly, optimal RMP, processor, and destination feature classes are constructed for mapping. Finally, an optimal_scenario_results table is constructed in the SQLite database containing the following data by commodity and mode:

- **Commodity flow**
- **Miles of network** used
- **Liquid or solid unit-miles** traveled

- **Dollar cost**, based on flow, mileage, and base transport cost
- **Routing cost**, based on flow, mileage, base transport cost, and weights (impedances)
- **Vehicle loads**, based on flow and payload
- **Vehicle-miles traveled (VMT)**, based on flow, mileage, and payload
- **Emissions**, based on VMT and emission factors (for road mode) and flow, mileage, and emission factors (for non-road modes)
- **Fuel burn**, based on VMT and fuel efficiency
- **Facility input / output amounts** and utilization

The optimal scenario results table also includes **processor build costs** where relevant. Once the table is constructed, it is used to generate the FTOT output files described in Section 4.

3 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, processors/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 defining key input parameters applied to the network and optimization;
3) a set of comma-delimited files to define the facility and commodity data.

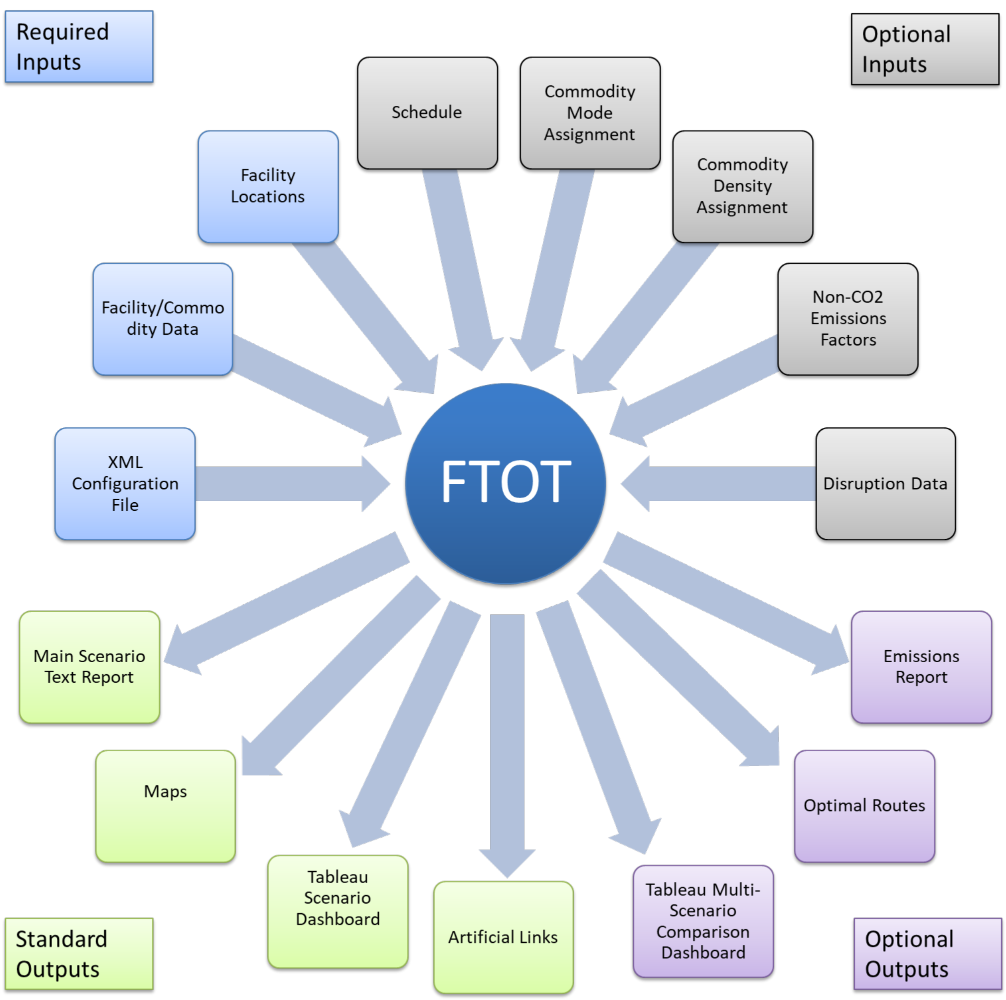

Figure 4: Attributes of the facilities in FTOT that the user can adjust to build a scenario.

3.1 XML Schema

Key parameters and the location of input data are defined within 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 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 reports the offending line and element.

3.2 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, XML schema version
- Scenario Inputs
    - `Base_Network_GDB` – location of the multimodal network.
    - `Disruption_Data` – (optional) location of the disruption CSV.
    - `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.
    - `Schedule_Data` – (optional) location of the schedule CSV file.
    - `Commodity_Mode_Data` – (optional) location of the commodity mode CSV file.
    - `Commodity_Density_Data` – (optional) location of the commodity density CSV file.
    - `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 specified in the scenario configuration file.
- Assumptions

        - Vehicle load capacity, fuel efficiency, and emission factors are specified here. Users are free to modify values as necessary.
        - `Detailed_Emissions_Reporting` flag – (optional) to toggle creation of emissions report on or off.
        - `Density_Conversion_Factor` – (optional) default density value used to calculate emissions for liquid commodities on rail, water, and pipeline modes.
    - Script Parameters
        - Network_costs
            - Modal costs: railroad, road, barge – base costs and impedances.
                - 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.
        - Short haul penalties for short movements on the water and rail network.
    - Route Optimization
        - Network pre-solve options: to toggle the network pre-solve using the NetworkX shortest path algorithm on or off.
        - 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.

3.3 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 contain 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).

Each facility input file includes the following fields:

- facility_name; used to match the locations in the scenario geodatabase. 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. Additionally, take care that extra trailing whitespace is not included, as “Whitewater” and “Whitewater ” will not match.

- 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. As noted above, care should be taken to avoid extra whitespace in the commodity names.
- value; the quantity of the commodity. Use of this field varies by facility type: for RMPs, it is the maximum available material; for destinations, it is the total demand; for processors, it defines how much output is created by the specified quantity of input. If a processor has more than one input commodity listed, it is assumed that all commodities must be used in the specified ratio.
- units; the units of the quantity for the commodity.
- phase_of_matter; solid or liquid, (gas is not currently supported).
- io; specifies if 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.
- schedule (optional); a user specifies the name of the availability schedule that each facility follows. The schedule name provided in the facility CSV should match one provided in the schedule CSV input file. The amount of product supplied or demanded at the facility on each day of the scenario is equal to the product of the commodity quantity and the fraction availability provided in the schedule.
- 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**.

For processors, there are three additional optional fields:

- max_processor_input (optional); here the user can specify the maximum amount of material a processor can take in. If this column exists, it should be the same for all rows of a single processor, not different by commodity. If this column does not exist, processor capacity is unbounded.
- min_processor_input (optional); here the user can specify the minimum amount of material a processor can take in if it is utilized. If this column exists, it should be the same for all rows of a single processor, not different by commodity. If this column does not exist, minimum processing capacity defaults to one-half the maximum.
- Build_cost (optional); to indicate that a processor is a candidate facility as opposed to an existing facility, the user must add the build_cost field in the processor commodity input file and specify a positive fixed build cost. Processors for which this field is 0 or missing are considered existing facilities. To use the optional build_cost field, the user should

enter numbers representing the amortized cost of building the facility across the given time period. If this column exists, it should be the same for all rows of a single processor.

3.4 Facility Location Data

FTOT requires GIS-based input datasets containing the facility names and locations of raw material producers (rmp), processors (proc), and destinations (dest). This information is stored in an ESRI geodatabase (GDB). These GIS feature classes (FC) must contain facility names and point locations for each facility. FTOT will automatically match the geospatial location to the facility-commodity data.

FTOT comes with one preexisting facility location geodatabase included in the common data subdirectory—point-based representations of every county in the United States.

In this dataset, county locations are placed at the geographic centroid of the most highly populated place within each county. While this is not always an ideal default location for some commodities, it ensures that each location can at least connect into the FTOT road network using the default artificial link distance. One unique feature class is provided for each facility type (raw material producers, processors, and destinations)—while this approach means each county has three identically-located facilities represented across these feature classes, FTOT requires distinct feature classes for each facility type, and this structure also ensures that each facility type and county combination has a unique facility name.

However, FTOT users may wish to use other preexisting datasets to represent FTOT facilities—or alternatively—create customized facility location data. For more in-depth information on leveraging other sources of facility location data or creating custom data, refer to Section 3.2.1 of the FTOT User Guide.

3.5 Schedule Input File (Optional)

Schedules can be used to run scenarios that span multiple days. This optional input file allows the user to define schedules of facility availability for each day in a scenario. An availability of 1 indicates that the facility produces or demands the same amount of the commodity as in the “quantity” column of the facility input file. An availability of 0 indicates the facility does not produce or demand any product on that day, and an availability of 1.5 indicates the facility produces or demands 1.5 times the amount in the “quantity” column in the facility input file.

The schedule input file includes three columns: schedule, day, and availability. Each schedule’s default value is indicated by day ‘0’ and any days with a different availability must be specified. All schedules for a scenario are the same length. The scenario schedule length is determined by the highest value in the ‘day’ column. FTOT forces all schedules for a scenario to be the same length to avoid mismatch issues between facilities with schedules of different lengths. An example of a schedule input is shown in Table 1.

Table 4: Example of schedule input file containing two 7-day schedules.

schedule	day	availability
weekdays	0	1
weekdays	1	0.5
weekdays	7	0.5
exceptDay3	0	1
exceptDay3	3	0

The table has two schedules. The 'weekdays' schedule has an availability of 0.5 on days 1 and 7 and availability of 1 on days 2 through 6. Since the largest value in the day column is 7, that is the scenario schedule length. As a result, the exceptDay3 schedule is also 7 days long even though only the default value, and day 3 value are specified. The 'exceptDay3' schedule has an availability of 0 on day 3 and an availability of 1 on days 1-2 and 4-7. An example schedule file is available in Reference Scenario 3.

The file name and file path should be added to `Schedule_Data` element in the scenario XML.

3.6 Commodity Mode Input File (Optional)

This optional input file allows the user to toggle different modes on and off for individual commodities. Users can also assign specific truck, railcar, or barge types to a commodity.

This CSV file must be included in order to include pipelines in the solution. Pipelines are included in the FTOT network but disabled for all commodities by default. This is to allow users to have flexibility in naming commodities while preventing commodities that are not supposed to flow on pipeline from utilizing it in the optimal scenario since it is usually the least expensive mode.

The commodity mode input file also allows for custom vehicle assignment by commodity and mode. Users can specify a vehicle label selected from the 'vehicle_types.csv' file to permit travel via a custom vehicle. Changing the vehicle type for a commodity updates the post-processing and reporting for that commodity and mode but does not affect the optimization.

3.7 Commodity Density Input File (Optional)

This optional input file allows the user to specify each commodity's density. Currently, density values are used only for calculating emissions from the transport of liquid commodities on rail, water, and pipeline modes.

3.8 Non-CO_2 Emission Factors Input File (Optional)

The detailed_emission_factors.csv file is located in the 'lib' folder of the project code repository (C:\FTOT\Program\lib) and is downloaded automatically with FTOT. FTOT uses the data in this file to calculate non-CO_2 emissions by commodity and mode. Users can toggle the `Detailed_Emissions_Factors` element in the scenario XML to output non-CO_2 emissions in a separate report file.

The detailed_emission_factors.csv file is prepopulated with emissions data for FTOT's default vehicles for transport on road, rail, and water modes as well as for the "small_truck" that is provided in vehicle_types.csv. The default emission factors will be used for all commodities on a mode *except* in the case that a custom vehicle is assigned to a commodity (see Section 5.6). If users create a new custom vehicle, they will need to (1) assign that vehicle to a commodity and (2) add custom emission factors to detailed_emission_factors.csv to include that vehicle in the detailed emissions report.

The detailed_emission_factors.csv file contains the following fields:

- vehicle_label – set to "Default" for the default vehicles assigned in the scenario XML. You can alternatively enter a vehicle label that matches the name of a custom vehicle in vehicle_types.csv. Note: The vehicle label is case-sensitive.
- mode – set to road, water, or rail
- road_type – set to "NA" for water or rail. For road, set to "Rural_Restricted", "Rural_Unrestricted", "Urban_Restricted", or "Urban_Unrestricted". Note: The road type is case-sensitive.
- Pollutant – can be set to one of the following values: CO, CO2e, CH4, N2O, PM10, PM2.5, VOC.
- Value – the emission factor for this entry. Include units of g/mi for road entries and g/ton/mile (or other unit of mass besides ton) for rail and water.

The vehicle emissions portion of Section 5.9 provides additional information on the prepopulated pollutant emissions factors.

3.9 Disruption Data (Optional)

The user can populate an optional disruption data CSV in order to disrupt (i.e., make unavailable) certain segments in the input FTOT multimodal network. At this time, only a link availability of 0 (fully disrupted) in the disruption data CSV is recognized by FTOT. In these cases, the links will be completely removed from the network and unavailable for any optimal solutions. This is useful in cases when the user would like to model a scenario in which a certain segment or corridor is unavailable due to some sort of hazard (e.g., flooding, earthquake) or other form of disruption (construction, motor vehicle crash).

3.10 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.

Modal Costs

Table 5: Modal cost units and default values in FTOT.

Transport mode	Cost (solids)	Cost (liquids)
Roadway/Trucking	0.22 USD/metric ton-mile	0.66 USD/kgal-mile
Railway/Rail car	0.047 USD/metric ton-mile	0.14 USD/kgal-mile

Transport mode	Cost (solids)	Cost (liquids)
Waterway/Barge	0.032 USD/metric ton-mile	0.097 USD/kgal-mile
Pipeline	N/A	Actual tariff cost (varies by route)
Transloading cost	12.35 USD/metric ton	40.0 USD/kgal

Source for costs—Bureau of Transportation Statistics Average Freight Revenue per Ton-Mile—2018 data.

Transloading costs based on development team experience.

Modal Impedance Factors (Weights)

The default impedances are listed in Table 6.

Table 6: 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 (FHWA Functional Class 1)	X 1.0
Road	Principal arterials (FHWA Functional Classes 2 and 3)	X 1.1
Road	Minor arterials (Functional Class 4)	X 1.2
Road	Local roads (Functional Class 5+)	X 1.3
Rail (public network)	STRACNET (Class 1 Ownership)/Density Code 7	X 1.0
Rail (public network)	Other STRACNET/Density Code 6	X 1.1
Rail (public network)	Class 1 Ownership (non-STRACNET)/Density Code 5	X 1.2
Rail (public network)	Class 1 Rights (non-STRACNET)/Density Code 4	X 1.3
Rail (public network)	Non-Class 1 (non-STRACNET)/Density Code 3	X 1.4
Waterway	High volume links (> 10 million annual tons)	X 1.0
Waterway	Medium volume links (1 – 10 million annual tons)	X 1.3
Waterway	Low volume links (< 1 million annual tons)	X 1.6
Waterway	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 3.2.1 for more details on how to access these source datasets.

Vehicle Emissions

FTOT estimates total emissions for various pollutant using unique emission factors for each mode. By default, FTOT generates CO_2 emissions as part of its main report. Optionally, users can toggle an option in the scenario XML to generate a separate emissions report for several non-CO_2 pollutants.

Road emissions. Truck emissions are calculated from emission factors *in units of g/mile* for four district road types: urban unrestricted, urban restricted, rural unrestricted, and rural restricted. Emission factor data sourced from EPA's Motor Vehicles Emission Simulator (MOVES) modeling system. Beginning with FTOT 2021.3, default truck emission factors are from the latest MOVES3 model version, released Fall 2020. FTOT's default fuel efficiency for road is then calculated as the weighted average of the CO_2 emission factors, based on each road type's VMT from the MOVES model, times 10.21 kg CO_2/gallon of diesel. The CO_2 intensity of diesel fuel is sourced from EPA's Emission Factors for Greenhouse Gas Inventories.

Non-road emissions. For rail, water, and pipeline, emission factors *in units of g/ton-mile* were sourced from Argonne National Laboratory's 2021 GREET Model.[4] Since non-road emission factors include a unit of mass (e.g., ton), FTOT applies a density conversion factor to calculate non-road emissions for liquid commodities. The default density for this conversion is 3.33 ton/kgal (the density of kerosene). To change this conversion factor, you will need to include a commodity density CSV input file as specified in Section 5.7 or update the `Density_Conversion_Factor` element in the scenario XML.

Table 7. Default emission factors by compound and mode.

Pollutant	**Description**	*Rural Restricted* **Road (g/mi)**	*Rural Unrestricted* **Road (g/mi)**	*Urban Restricted* **Road (g/mi)**	*Urban Unrestricted* **Road (g/mi)**	**Rail (g/ton/mi)**	**Water (g/ton/mi)**	**Pipeline (g/ton/mi)**
CO_2	Carbon dioxide	1338.31	1360.18	1343.74	1550.19	21.3	37.9	0.0
CO	Carbon monoxide	1.824	2.061	1.958	2.882	0.01350	0.08429	--
CH_4	Methane	0.039	0.083	0.062	0.155	0.001863	0.0003393	--
N_2O	Nitrogen dioxide	0.001	0.002	0.002	0.003	0.0005821	0.000892	--
NOx	Nitrogen oxides	3.925	4.114	3.998	5.683	0.1137	0.3015	--
PM10	Particulate matter (<10 microns)	0.131	0.178	0.159	0.370	0.001874	0.009820	--
PM2.5	Particulate matter (<2.5 microns)	0.088	0.099	0.094	0.153	0.001817	0.009525	--
VOC	Volatile organic compounds	0.181	0.194	0.188	0.259	0.007859	0.006925	--
CO_2e	CO_2-equivalents	1622.368	1630.850	1617.818	1829.149	21.553	38.192	--

Network Capacity and Existing Flow

[4] Emission factor data from GREET were provided by Dr. Michael Wang and Dr. Xinyu Liu at Argonne National Laboratory in January 2022.

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 as follows:

- 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 is only an issue at typical waterway chokepoints (e.g., locks)
  - Background flow and capacity data is 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.

Short Haul Penalties

The short haul penalties are fixed penalties per unit of flow that are used to prevent very short movements by rail and water, which are usually unrealistic (road is much more commonly used for short freight movements). Default values are calculated by mode and phase of matter using the default costs from Table 4. For rail, the default penalty is the difference between the default costs of routing 100 miles on local roads and 100 miles on rail. For water, the default penalty is the difference between the default costs of routing 100 miles on local roads and 100 miles on water. No penalty is applied to road and pipeline. Default penalties are presented in Table 6.

Table 8: Default short haul penalties in FTOT.

Transport mode	Cost (solids)	Cost (liquids)
Roadway/Trucking	No penalty	No penalty
Railway/Rail car	23.9 USD/metric ton	71.8 USD/kgal
Waterway/Barge	25.4 USD/metric ton	76.1 USD/kgal
Pipeline	N/A	No penalty

4 FTOT OUTPUTS

FTOT generates four main types of outputs (see Figure 4). 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.
3) A packaged Tableau workbook that contains all the geospatial and summary result information required to display a Tableau dashboard in the free Tableau Reader software or the licensed Tableau Desktop 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.).

Supplementary reports on the scenario results are also generated alongside the other main outputs for certain FTOT scenarios:

- Artificial links – summarizes the artificial links used to connect facilities to the multimodal network. Each row of the file specifies a facility in the scenario, and each column specifies a permitted mode in the scenario. Each entry in the table is the artificial link length (in miles) if the facility is successfully connected to that mode, or “NA” if the facility is unable to connect to that mode with the artificial link distance parameter specified in the scenario XML file. Note that stranded facilities (those that fail to connect to the multimodal network at all) are not included in the artificial links file.
- Optimal routes – when network density reduction is enabled to calculate shortest paths (see Section 2.3), an additional optimal routes CSV file is generated containing information on routes used in the optimal solution.
- Non-CO_2 emissions – when detailed emissions reporting is enabled, a detailed emissions CSV report is generated with total emissions for several non-CO_2 compounds. Emissions are reported by commodity, mode, and pollutant.
- In addition, the full scenario results are stored in a SQLite database and in the scenario geodatabase.

4.1 Graphics Dashboard

A graphical dashboard output is created using Tableau (see example figures in accompanying User Guide) and exported as tableau_dashboard.twbx file in the Reports outputs. After the scenario has finished running, the user can open this file in a Tableau application (e.g., Tableau Reader, Tableau Desktop).

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, CO_2, VMT, mileage, fuel burn, and costs.

4.1.1 Tableau Scenario Comparison Dashboard

The Scenario Compare tool is a supplemental FTOT Tool used to concatenate the results from one or more scenarios in a list.

The tool has two modes to generate the list of scenarios to concatenate: an automatic recursive search, and a user-specified search. In the recursive mode, the user provides a high-level directory and FTOT returns all sub-directories within it. In the manual mode, the user specifies which scenario folders should be added to the list until specifying the "done" keyword.

Once the data is concatenated, the tool zips the workbook, CSV report, geodatabase, and supplement files into a packaged workbook file (.twbx). The user can open this packaged workbook using Tableau Reader, which is available as a free download. The comparison dashboard allows the user to select one or more scenarios to compare against.

The comparison dashboard includes options to change the color of mapping route results by mode, commodity, and scenario name, as well as a Tableau Story format that allows users to step through each of the following dashboards:

- Supply Chain Summary
- Route Results by Commodity and Mode
- Facility Results by Supply and Demand
- Run Times
- Parameters

4.2 Maps

The FTOT map outputs are generated in the scenario's Maps directory (see example figures in the Map Appendix folders of the accompanying Quick Start and Reference Scenario Documentation) and include the following:

- The multimodal network used in the optimization
- Raw Material Producers
- User Defined Processors
- Ultimate Destinations
- All Facilities
- Processor Candidates
- All Processors
- Final Optimal Routes
- Optimal and Non-Optimal Raw Material Producers
- Optimal and Non-Optimal Processors
- Optimal and Non-Optimal Ultimate Destinations

All map names are suffixed with the following text—"default_basemap, "gray_basemap", "topo_basemap", or "streets_basemap"—based on the version of the mapping that is specified in the run.bat file for 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

Additional information on customizing FTOT map outputs is available in the FTOT User Guide.

5 SUPPLEMENTARY TOOLS

A set of supplementary FTOT tools is provided with the program to assist the user with automating 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. Note: Make sure to confirm the correct scenario values if updating an older XML file.
- **bat_tool:** Generates a bat file for running a new FTOT scenario based on user-provided inputs.
- **scenario_compare_tool:** Concatenates the results from one or more scenarios into a packaged Tableau workbook. The tool has two steps: (i) the user specifies an output directory where the concatenated results workbook will be stored, (ii) the user specifies the input directories which contain the results to be concatenated. This can be done recursively by supplying a top-level directory to automatically generate a list of sub-directories, or the user providing the scenario directories individually.
- **aggregate_raster_data**: Aggregates grid cell production data (e.g., USDA) by county.
- **generate_template_csv_files:** 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).
- **breakpoint:** Sets a python debugger breakpoint and enters a debugging REPL
- **replace_xml_text:** A tool to batch replace XML configuration elements recursively through a top-level directory. Useful for making changes to a number of scenario files, such as changing the base network geodatabase location.
- **network_disruption_tool:** Allows the user to automatically generate a network disruption CSV associated with a hazard scenario (e.g., NOAA sea level rise data, HAZUS data, etc.). In order to use this tool, you must have raster-based GIS data which identifies exposure levels due to some sort of hazard (e.g., flooding, earthquakes, etc.). Currently, this tool is only able to identify disruption of the road and rail portion of the default FTOT network.

For more details on running some of the key supplementary tools, consult Section 7 of the FTOT User Guide.

6 ACKNOWLEDGEMENTS

This research was funded by the U.S. Federal Aviation Administration Office of Environment and Energy through FAA contract 693KA9-20-N-00013 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.

7 REFERENCES

"pulp-or". 2007. COIN-OR.

Bertsimas, D., and J. N. Tsisiklis. 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.