Freight and Fuel Transportation Optimization Tool Technical Documentation

FTOT 2025.3 Public Release Version

September 30, 2025

DOT-VNTSC-FAA-22-06 Rev. 13

DOT/FAA/AEE/2022-01 Rev. 13

Prepared for:

Federal Aviation Administration
Office of Environment and Energy
Washington, DC

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

Volpe

Notice

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 September 30, 2025	3. REPORT TYPE AND DATES COVERED Document, 7/2022-9/2025

4. TITLE AND SUBTITLE Freight and Fuel Transportation Optimization Tool Technical Documentation: FTOT 2025.3 Public Release Version	5a. FUNDING NUMBERS FC14C222 FC14C323 FC14CA24
6. AUTHOR(S) Olivia Gillham, Kirby Ledvina, Kristin C. Lewis (PM), Mark Mockett, Matthew N. Pearlson, Peter Wilke, Jaewoong Yun, Kevin Zhang	5b. CONTRACT NUMBER 693KA9-22-T-00008 693KA9-23-T-00005 693KA9-24-T-00002
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Volpe National Transportation Systems Center, 220 Binney Street, Cambridge, MA 02142	8. PERFORMING ORGANIZATION REPORT NUMBER DOT-VNTSC-FAA-22-06 Rev. 13
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 Rev. 13
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 2025.3. 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 63
	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, or FTOT (https://volpeusdot.github.io/FTOT-Public) [1], 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 multimodal, flowable Geographic Information System (GIS) network aligned with the Generalized Modeling Network Specification (GMNS) [2] that assigns costs, capacity and existing flow information, and weighting factors to each link. It takes origin and destination 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, optional carbon dioxide (CO_2) emissions 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 to 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 built with Python and ArcGIS Pro and requires the user to have familiarity with manipulating data in Microsoft Excel and optionally ArcGIS Pro. 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

[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.

reached end-of-life and are no longer being supported. Version 2023.1 is the first version of FTOT to align the default network and codebase with the GMNS framework, allowing for the use of FTOT with a user-customized local, regional, or national network.

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 [3].

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 changes in 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 intermodal 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, differences in supply availability or cost among origins, change in demand amounts, or prioritizations among destinations or overall).
- 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.
- Consideration of the social cost of CO_2 emissions when selecting optimal routes.

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-distance traveled, and length of network used.

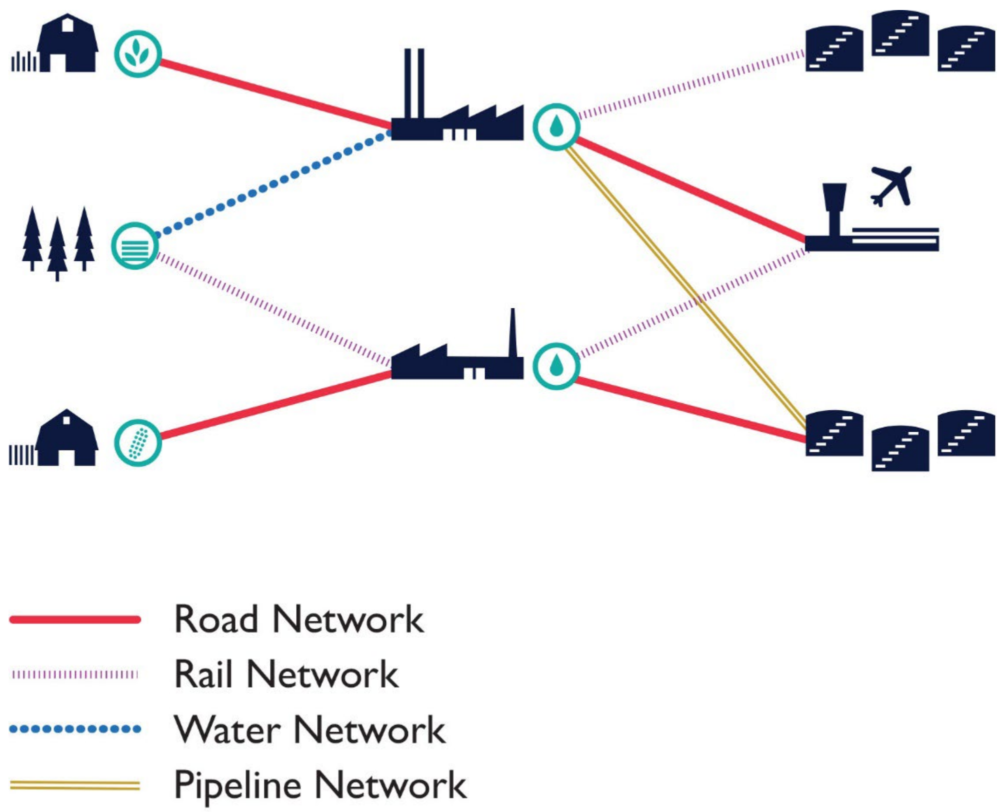

Figure 1: Generalized FTOT scenario structure incorporating supply chain elements.

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.

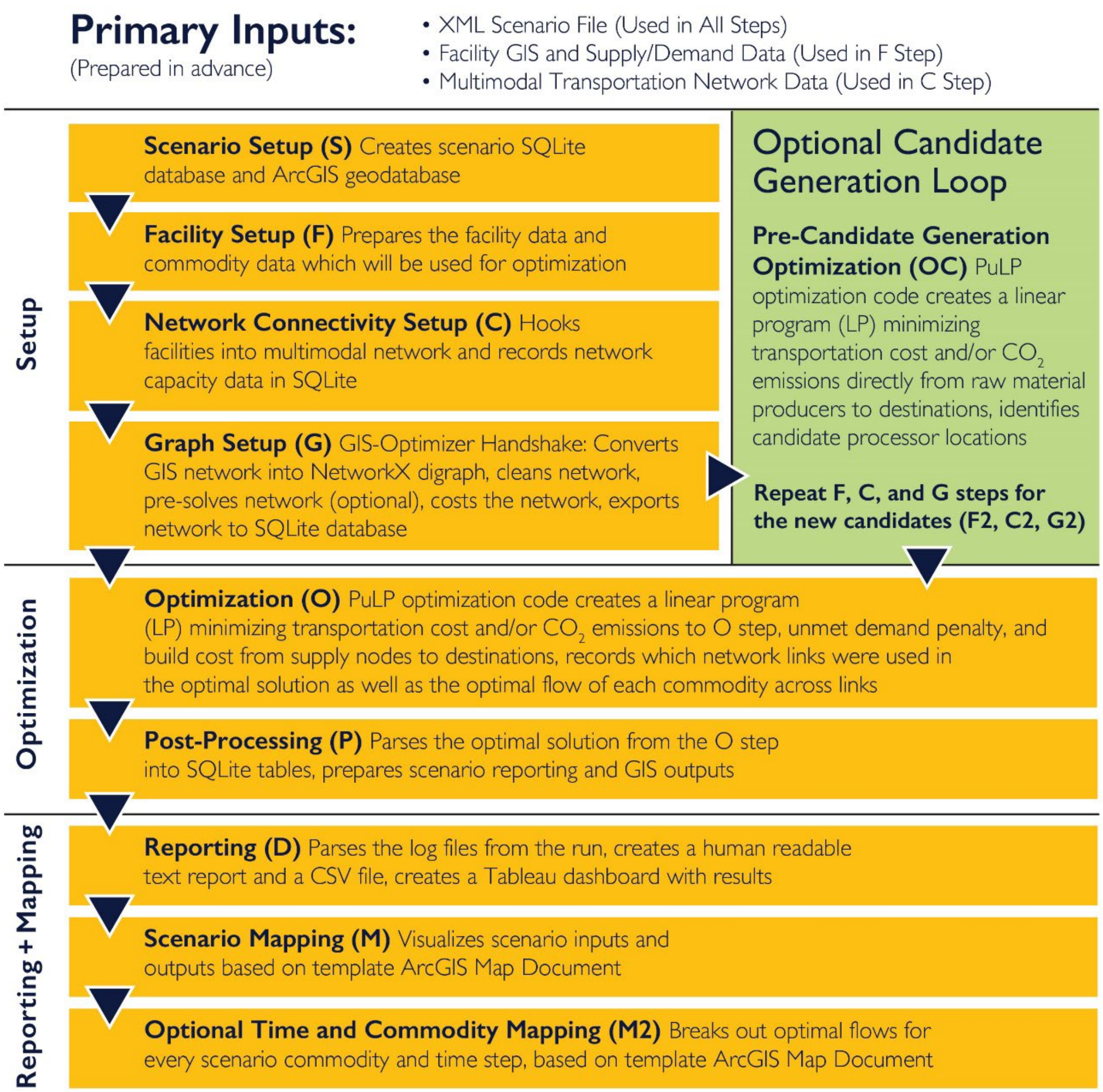

Figure 2: Analytical tool data flow schematic showing the key components/roles of each component of FTOT.

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 Section 2 of the 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 (Pint, PuLP, HiGHS, Imageio) 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 [4] to the optimizer for analysis. After optimization, the outputs of the scenario are passed back to the GIS module to generate scenario maps.

NetworkX [4] 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. When this functionality is enabled, NetworkX shortest path algorithms are used to identify links, and candidate processor locations during candidate generation (if used), for use in the optimization problem.

The **PuLP** [5] 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 solver options can be found in the sections below.

A **SQLite database** is used to store all data related to the FTOT scenario, from user inputs and network attribute data to the optimal solution and metrics calculated in post-processing.

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

Beginning with the 2023.1 release, FTOT has established a standardized network specification to allow the use of custom transportation networks within the tool. Prior to this release, FTOT was designed around the use of an FTOT-provided contiguous United States multimodal network. For new and existing

FTOT users who wish to run scenarios within the contiguous United States with the same network, the introduction of a network specification does not significantly change the program, as the default FTOT network has not substantially changed and still covers the same geographic extent. However, the introduction of a network specification does facilitate the use of customized local, regional, and national networks in substitution for the default FTOT network, making international and/or very local FTOT use cases possible. FTOT users who wish to use an alternative network rather than the default FTOT network can structure their network file to match the network specification schema and specify its file location in the scenario XML file.

2.2.1 FTOT network specification

The FTOT network specification is partially based upon a standardized network modeling format for sharing routable road networks called the Generalized Modeling Network Specification (GMNS). To align with GMNS, the FTOT network specification uses formal designation of field names for each mode and attribute to make these fields consistent and not tied to specific data sources. These field names provide a generalized representation of network components that can be populated using other data sources, such as the FHWA functional class system and FRA-specific rail corridor designations. Given the FTOT network complexity, there are some fields FTOT uses that were not already defined in the GMNS; a consistent approach has been used to make these fields as generalized as possible.

In addition to standardizing field names, the FTOT network specification provides flexibility with regard to the user's application of:

- Coordinate system: Any meters-based coordinate system is permissible as long as it is consistent among all feature classes making up the network.
- Impedance weight configuration: Impedance weights are based on the field "Link_Type", which allows various categorizations to be used to assign impedances to different classes of links within a mode. This allows users the flexibility to use any data to prioritize flows (e.g., number of lanes, speed buckets, traffic volumes, etc.).
- Distance unit: Distance is generalized to enable inputs in miles, kilometers, or other units, as long as they are used consistently throughout FTOT inputs. The distance unit used for the required length measurements in the network must match the default distance unit specified in the scenario XML file; results are output in the same units.
- Additional link attributes: Specific fields defined in the FTOT network specification are optional for users with limited access to data. FTOT will run with a basic network even when many optional attributes are missing, although certain optimization and post-processing considerations are limited by lack of these data. While FTOT does not utilize attributes beyond those included in the network specification, it does retain these attributes in the scenario network constructed by the program in case the user would like to access the additional attributes for user-generated reporting, mapping, or analysis. Note, however, that attributes not contained in the network specification are passed along to the modified scenario network 'as is' and are not re-calculated when links are split to allow scenario-specific facilities to be attached to the network in the C step. For this reason, users are discouraged from including non-network specification attributes that would need to be recalculated if network segments are split. For

example, if the user includes count or density link attributes within their network, these feature attributes may not be valid when the scenario network is processed to flow an FTOT scenario, as split segments will retain the same (and potentially inaccurate) counts and densities as the parent link.

2.2.2 FTOT network requirements

Overall requirements are described below:

- Each mode must be saved in a distinct feature class inside a feature dataset named 'network' located within a single ESRI File Geodatabase (GDB).
- Options for mode feature classes allowed in an FTOT network GDB are: road, rail, water, pipeline_crude_trf_rts, and pipeline_prod_trf_rts. A valid FTOT network includes a subset (or all) of these feature classes. For each mode feature class included, the network data fields/attributes should be named in accordance with the network specification schema.
    - Note 1: The pipeline modes require additional data layers if capacity and/or background flows are a consideration; see Section 2.2.3 below and Section 4.1.1 and Appendix B of the FTOT User Guide for more details.
    - Note 2: An optional locks data layer can be provided as an additional source of capacity and/or background flow data for the water mode; see Appendix B of the FTOT User Guide for more details.
- Each feature class must be in a consistent meters-based coordinate system appropriate to the region (e.g., USA Contiguous Lambert Conformal Conic is used for the default U.S. Network).
- All features should be digitized in a way that makes them fully connected and routable (e.g., grade separation should be honored, no gaps should exist between different features).
- Any intermodal facilities must be directly connected to the network associated with each relevant mode through a segment (artificial link) in the mode's feature class.
- The specific schema for each feature class is provided in Appendix B of the FTOT User Guide.

FTOT's supplementary network validation tool (see Section 5 below and Section 7.1.7 of the FTOT User Guide) can be used to confirm custom networks meet some of these requirements. In particular, the tool checks whether feature classes use a meters-based coordinate system, follow the correct naming conventions, and include the required data fields as described in the schema. This tool also summarizes mileage, segment counts, and link type counts, and optionally checks network connectivity (both across modes and within a mode).

2.2.3 Default FTOT GIS network (Contiguous United States)

The FTOT data and documentation package currently includes two versions of the contiguous U.S. multimodal network: FTOT_Public_US_Contiguous_Network_v2025.gdb and FTOT_Public_US_Contiguous_Network_v2025_FAF4_Capacity.gdb. The main difference between the two networks is that the default 2025 version of the network (FTOT_Public_US_Contiguous_Network_v2025.gdb) contains more up-to-date road source data as listed below. However, users planning to run capacity-constrained scenarios utilizing the road network should use the alternate FAF4 Capacity version of the network, as the updated road network in the default

2025 version does not include the data to support road capacity constraints. Further information on the FAF4 Capacity network is available in Appendix A: FAF4 Capacity Network Information. See Section 2.2 of the FTOT User Guide for instructions on how to download these default networks.

A third FTOT network that includes some additional modal data for Canada, Mexico, Alaska, and Hawaii is in draft form and is available from the FTOT team by request. More information on that network is available in Appendix B: North American Network Information. In addition, a comparison of all three available networks developed by the FTOT team is provided in Appendix C: Network Comparison.

Elements of the default FTOT multimodal network include feature classes for road, rail, water, and pipeline networks covering the contiguous United States. Intermodal facilities where transfers between modes occur, waterway locks, and available tariff and capacity information for crude and petroleum 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 petroleum product pipeline networks, each pipeline category has three feature classes associated with them: 1) pipeline_trf_rts feature classes containing tariff-level information such as the cost of the movement and information on the pipeline and origin-destination pipeline stations; 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 feature classes (e.g., pipeline_crude, pipeline_prod) representing all segments associated with the pipeline network (whether or not they are associated with a confirmed tariff). FTOT primarily uses the pipeline_trf_rts feature classes to ensure that pipeline movements are restricted to the routes that tariffs allow. The latter two feature classes are only required for scenarios that need to consider capacity and/or background flows, as these simply work to ensure that overlapping tariff routes are aggregated over their corresponding segments appropriately. 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.

2.2.4 Default FTOT GIS network data sources

The default 2025 network is limited to the contiguous 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. 5—2022 (accessed via the Federal Highway Administration Office of Operations) [6]. 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)—2025 (sent directly to Volpe by FRA and also available at the National Transportation Atlas Database) [7]. 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—2024 (accessed via the USACE Geospatial Data Portal) [8]. Modified based on additional research to exclude some non-navigable waterways. The waterway network is considered comprehensive.
- Locks: USACE Locks Shapefile—2024 (accessed via the USACE Geospatial Data Portal) [8].
- Crude and Petroleum Product Pipeline Network Data: Energy Information Administration (EIA) crude oil and petroleum product pipeline data—January 2023 (accessed via EIA U.S. Energy Atlas) [9]. Modified by Volpe based on station locations in a tariff dataset compiled by Arbo (January 2024) [10] and other publicly available pipeline network data.
- Crude and Petroleum Product Pipeline Tariff Rate Data: Pipeline transportation rate and origin-destination information supplied by Arbo (January 2024). Movements on the pipeline network are constrained to the origin-destination pairs designated in the dataset.
- Intermodal facilities: Volpe-developed list of intermodal facilities. Based on public data from:
    - BTS Intermodal Freight Facilities Rail TOFC/COFC (December 2024) [11]
    - BTS Intermodal Freight Facilities Marine Roll-on/Roll-off (December 2024) [12]
    - EIA Crude Oil by Rail Dataset (October 2020) [9]
    - USACE Master Docks Database (February 2016) [13]
    - Corroboration from Surface Transportation Board Waybill Sample dataset (2013) [14]
    - Crude oil and petroleum product pipeline station data from Arbo (January 2024) [10]
    - 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 default FTOT network, FRA's North American Rail Network is subset to the main subnetwork and major industrial leads within the contiguous 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 crude oil and petroleum product pipeline network is limited to all known public pipeline tariffs compiled by Arbo. Private pipelines and public pipeline tariffs that could not be assigned to a pipeline with known geospatial data are not included.

2.2.5 Default FTOT GIS network 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 Arbo [10] 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 using publicly available and authoritative datasets when possible, with an additional focus on identifying 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 intermodal facilities dataset was created by starting from the Bureau of Transportation Statistics' public intermodal facility layers [15] and supplementing with specific agriculture-related and energy-related sources. Data attributes for intermodal facilities were standardized across the different data sources before specific criteria and visual review of satellite imagery were used 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 from other sources, 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). For agriculture-related intermodal facilities, rail intermodal points were added to the initial list based on movements of specific, raw-material-relevant commodity types in the Railway Waybill Sample [14] 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 1940 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.

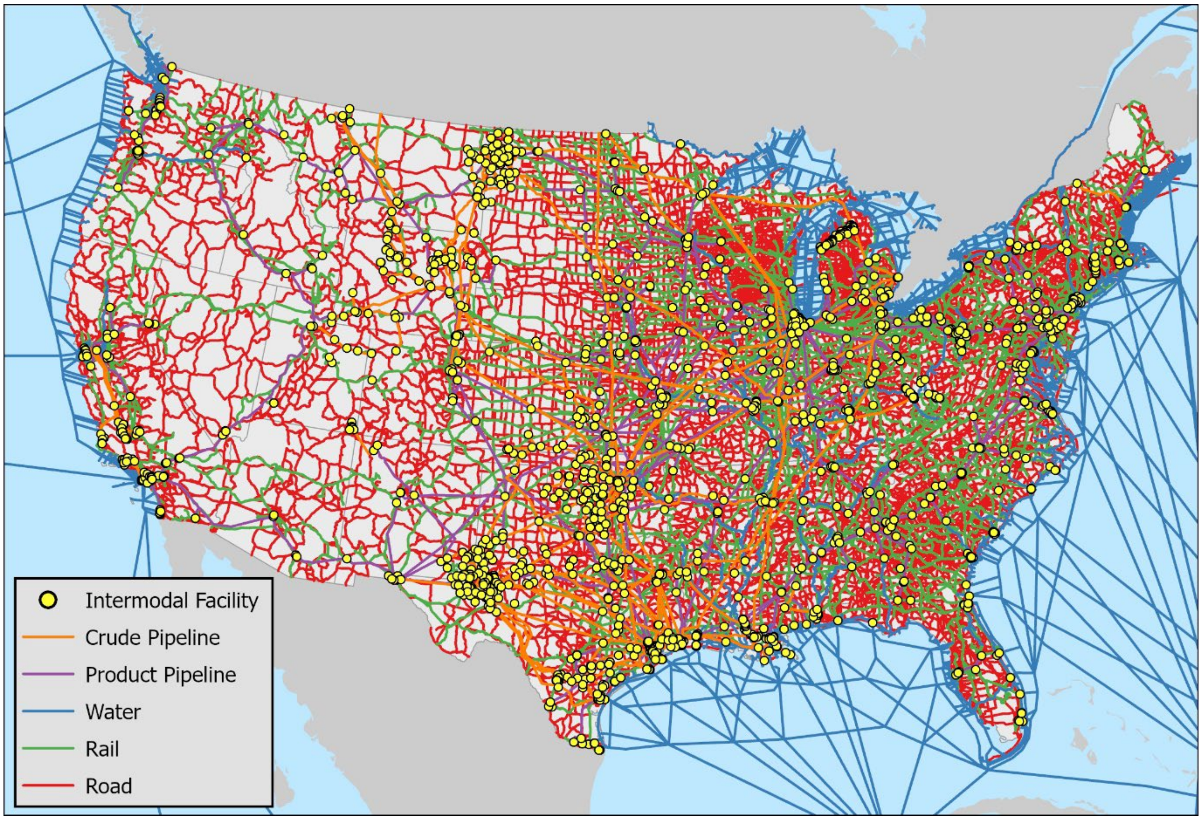

Figure 3: FTOT multimodal transportation network.

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

2.2.6 Customizing the FTOT GIS network

In some circumstances, an FTOT user may wish to customize an FTOT multimodal network. For example, a user may wish to add an additional network segment (or a handful of new network segments) to the default network. This is achievable provided the FTOT user has some existing GIS experience and the patience to customize the GIS layers contained in the default FTOT network, 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. In cases where substantial modifications to the network might be necessary or FTOT is being used to run scenarios outside of the United States, a distinct, standalone custom FTOT network may be necessary. In these cases, the user will need to process source network data (e.g., from OpenStreetMap or national, state, or local governments) into a format suitable for running with FTOT. The FTOT User Guide walks through both of these processes in more detail.

In addition, segments can be automatically removed from the FTOT network (i.e., 'disrupted') through the creation of an optional disruption data CSV. This is covered in Section 3.9 and in the User Guide.

2.2.7 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. If the facility is located directly on the modal network, a small offset is applied to the facility before an artificial link is created. After comparing to the user-specified maximum artificial link distance parameter (default is five miles), FTOT constructs new straight-line artificial links for each valid combination. If the nearest point in the modal network where the artificial link connects is not already a point where segments intersect, a new intersection point is created by splitting the network segment in two. Segment attributes are passed on 'as is' to the newly constructed segments. Since FTOT connects facilities to only the nearest point in the network for each mode, user facility placement is important; users should take care to examine routing behavior near facilities and adjust their placement as necessary. Note that facilities located beyond the artificial link distance specified in the configuration file will remain inaccessible to the corresponding mode.

When utilizing a network that distinguishes between limited access and non-limited access roadways (e.g., the default contiguous United States FTOT network), a special consideration is made for connecting facilities to the nearest road. Artificial links will prioritize a connection to the nearest non-limited access roadway located within the artificial link distance, ignoring any limited access roadways that might be closer. A facility will only connect to a limited access roadway if no non-limited access roadways are located within the artificial link distance. This reduces the risk of unusual and unrealistic routing resulting from a facility hooking into only one direction of a limited access roadway away from an interchange.

The fully connected network is stored in a feature dataset inside a scenario-specific main geodatabase, 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.

2.2.8 Network cost attributes

Network costs are calculated based on link attributes and user-defined inputs. Three types of network costs are recorded—the transport cost, carbon dioxide (CO_2) cost, and routing cost.

For road, rail, and water networks, the transport cost is a simple calculation based on the segment's length and a per-mile cost defined in the scenario configuration XML. For U.S. crude oil and liquid product pipelines, a rounded tariff cost is used based on the input data. Pipeline transportation rate and origin-destination information are supplied by Arbo. [10] 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). Private pipelines and some public pipeline tariffs that could not be assigned to pipelines with known geospatial data are not included. More information on pipeline tariffs is available from the Federal Energy Regulatory Commission (FERC) eTariff website [16] or Arbo. Pipeline tariffs are provided by Arbo in units of dollars per barrel. To convert these costs to dollars per unit of mass, an average commodity density of 874 kilograms per cubic meter [17] and 800 kilograms per cubic meter [18] is used for the crude oil pipelines and liquid product pipelines, respectively.

CO_2 emissions costs are determined similarly to transport costs. For each segment, the segment length, corresponding emissions factor, and the cost of CO_2 parameter specified in the scenario XML are multiplied to get a carbon cost per unit of commodity flowed on the segment. The emissions factors used in this calculation are the same as used to calculate emissions from transportation in the reporting section. For road segments, the emissions factor is specified in grams per vehicle distance traveled; to get a per-commodity-unit emissions factor, FTOT divides by the truck load size parameter in the scenario XML.

In addition to a transport cost and CO_2 emissions cost, a routing cost is stored for road, rail, and water segments. The routing cost is an overall generalized cost for determining optimal routing. The transport component of this cost incorporates weights based on the network mode and link attributes. These weights are configurable and are optionally set in the impedance weights CSV. 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) beyond a simpler per-distance transport cost approach. To consider a combination of transport and CO_2 costs in the routing cost, the user specifies positive scalar parameters for the weighted transport costs and the CO_2 costs. The scalars are used to calculate the routing cost as a weighted sum of transport costs and CO_2 costs. For pipeline, which by default has no CO_2 cost, the routing cost only includes the scaled transport cost.

If the scalar for the CO_2 element is greater than 0, indicating optimization should be partly based on CO_2 emissions, then the user must also specify a CO_2 unit cost in the scenario XML. If a CO_2 unit cost is not provided but the scalar is set to 0, FTOT will not consider CO_2 costs in the optimization and reporting; however, a CO_2 cost for each link will still be listed in the SQLite database assuming a unit cost of 191 USD per ton of CO_2 (and will not account for currency conversions).

Artificial links that connect user-specified facilities to each modal network have several different cost elements. First are the transport cost and routing cost, which by default are calculated using the local roads cost parameters from the XML and impedance weights CSV to model first and last-mile transport costs to/from facilities by local roads. If desired, the user can instead specify separate network costs for artificial links in the XML. These first mile/last mile costs are applied as per-distance transport, CO_2, and routing costs on artificial links. Second, additional access costs at the facility, commodity phase (e.g., solid, liquid), and input/output level can be specified on artificial links through the facility data input files (Section 3.3). These are not impeded and are incorporated into routing cost, but not transport cost. Third, the routing costs on rail and water artificial links can be additionally inflated through short-haul penalties that are configurable by the user and are designed to discourage short movements on rail and water networks. Network costs on artificial links connecting intermodal facilities, specifically, are calculated using mode-specific per-distance costs for road, rail, and water connections, in addition to transloading costs applied to all intermodal movements.

2.2.9 Exporting the data

After the facilities are connected to the multimodal network with artificial links, the network is converted into a graph of nodes and edges. The NetworkX module, described in greater detail below, is used to create a graph of the connected network.

2.3 NetworkX

NetworkX is used to create a network graph from the geospatial network and facility data. The NetworkX module reads each feature class in the network feature dataset (roads, railroads, destinations, etc.) directly from the geodatabase 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.
- Honoring directionality of network link segments (e.g., portions of the default road network are represented with dual carriageways which each operate as distinct one-way roadways, pipeline links that only flow 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. Three types of network costs are recorded. First, the transport cost is based on the distance traveled and mode. Second, the CO_2 cost is based on the distance traveled, the emissions factor, and the cost of CO_2. Finally, the routing cost is stored as a combination of transport cost, (optionally) artificial link access cost, and (optionally) CO_2 cost 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. See Section 2.2.8 for more details.

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 NetworkX shortest path algorithms. This network density reduction (NDR) pre-solve step reduces run time for computationally intensive scenarios but is currently not compatible with capacity-constrained scenarios. Reference Scenarios 2 and 5 in the Reference Scenarios documentation provide examples 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 by the shortest path algorithm in order to solve for the minimum cost path between each pair of facilities. For scenarios using candidate generation and/or maximum transport distance, NDR first identifies a subset of the graph that is reachable within the maximum transport distance from each source facility using the shortest path algorithm weighted by distance. Then, the shortest path method is used a second time with routing cost as the weight to determine minimum cost paths for each commodity and facility pair on that graph subset.

The links identified in the shortest path solution are stored in the SQLite database in two different tables.

- The ‘shortest_edges’ table stores links across all minimum cost paths; it contains a unique entry for each link. 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 (and similarly in the OC step for candidate generation). For candidate generation, the results of the simplified optimization problem are also used to identify potential locations for candidate processors.
- The ‘route_edges’ table contains an entry for each link for every minimum cost path it is in. In addition, the links are identified in each shortest path solution in order from source to target facility.

Each shortest path route, including those used in the optimal solution, is also output in a timestamped all_routes_TIMESTAMP.csv file located in the Reports folder of the scenario, with descriptive statistics of the origin and destination facility types, commodity, mode, length, routing cost per unit of commodity flowed, transport cost per unit of commodity flowed, access cost per unit of commodity flowed, CO_2 emissions per unit of commodity flowed, estimated travel time, and phase of matter of the commodity of each route, along with whether or not it was included in the optimal solution.

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, cost of CO_2 emissions, 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 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), (optionally) artificial link access costs, and (optionally) cost of CO_2 emissions.
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 mixed integer linear programming problem (Bertsimas and Tsisiklis 1997) [19]. 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 [5] and COIN-OR Linear Programming (CLP) [20] software used for the FTOT project. PuLP is a linear programming modeler, used to set up the linear program and then call the user-specified solver. FTOT is configured to use one of two open-source solvers to solve the linear programming problem: (1) the COIN-OR branch-and-cut (CBC) algorithm [21] or (2) HiGHS [22]. Any users wishing to use an alternate solver can refer to the COIN-OR documentation for configuring a solver and update the solve_pulp_problem() method in ftot_pulp.py for the solver they wish to use.

The PuLP optimizer identifies a maximum or minimum value (in this case, a minimum value for 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 all commodities. For liquid commodities, a commodity density is used to convert units of volume to units of mass. Users can specify units other than metric tons and distance units other than miles.
   - Pipeline costs are based on tariff data described in Section 2.2.
   - Facility, commodity phase, and/or input/output specific first-mile last-mile costs can be added as a cost per metric ton for all commodities by default on artificial links.
   - Transloading costs are applied as a cost per metric ton for all commodities by default at intermodal facilities.
   - Capital costs for candidate processor construction are 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. Commodity-specific modal permissions can also be set in an optional commodity mode input 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 default penalty (per unit of unmet demand in default solid units) is specified by the user in the scenario configuration file. Optional facility-specific and commodity-specific penalties can be specified by the user in the destination facility-commodity input file (per unit provided in the destination facility-commodity input 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 transport and build 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, to maximize demand fulfillment regardless of transportation cost, 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. At a lower unmet demand penalty, the user may see reduced utilization of raw material due to the optimizer electing the less expensive option of reducing flow overall instead of accumulating high transport and capital costs for a given route. Likewise, a very high unmet demand penalty may force the flow of materials in unanticipated or unrealistic ways. It is advised to run scenarios with multiple unmet demand penalties to explore the sensitivity of a given analysis. When the unmet demand penalty is set to values in the range of the actual average transportation cost, the user can evaluate the trade-off between demand fulfillment and transportation cost. This allows the user to fine tune the actual transport cost of their optimal routing solution within a desired range.

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 transport 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 movements, reflecting a reality where rail and water are typically utilized for longer commodity movements.

5) **Minimum and maximum flow requirements** applied to processing facilities
   - A processing facility may have a defined capacity. The user can provide this as a maximum input or output capacity for the processor facility and/or as a maximum capacity on a commodity entering or leaving the processor. This is defined in the max_capacity field of the proc.csv input file. If the user provides multiple max_capacity values for different commodities, the most restrictive will be used. If max_capacity is not specified, no upper restriction is placed on the processing facility.
   - The user can also provide a minimum capacity for the processor, which only applies if the processor is in use. That is, even when a minimum capacity is defined, the flow through a

processor is allowed to be zero. This is defined in the min_capacity field of the proc.csv input file.
- If no maximum or minimum processor 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. By default, the base costs represent the per metric ton-mile transport cost of traversing each mode. The impedances, or weights, act as multipliers on the transport 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 incorporated into the routing costs and used for the optimization. They are distinct from the transport cost that FTOT reports and summarizes.
   - Using an impedance weights CSV, the user can adjust the impedance weights on each type of segment in the network. If an impedance weights CSV is not provided, FTOT will assume equal weight for all link types.
   - 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 mode. Users should assign impedances to match the link types used in their network. In FTOT's default U.S. network, road network impedances are based on roadway functional class and in the water network impedances are based on actual annual freight flows as included in the National Waterway Network. For the rail network, the U.S. default network relies on a set of categories based on whether the rail segment is part of the Strategic Rail Corridor Network (STRACNET) and whether it is owned or flowed on by Class 1 freight railroads (these data are available in the publicly available version of the NARN which serves as the basis for the FTOT rail network). Users may also specify the impedances of artificial links for user-defined facilities. This impedance is applied to all facility artificial links, regardless of the network mode. More details on the exact definitions used for these impedance categories are available in Table 7.

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.

8) **Costs associated with emissions** to incorporate CO_2-based routing decisions (if applicable)

- CO_2 emissions costs for each transportation mode are created by multiplying the mode's emissions factor by the user-specified social cost of carbon (cost of CO_2 emissions) to arrive at an additional cost per unit of distance and mass. For liquids on the road network, the user-defined default density and the truck liquid capacity are used to calculate a per-ton CO_2 cost.
- The default values for CO_2 costs are significantly lower than actual transport costs on road, rail, and water per unit of distance and mass (or volume). Therefore, CO_2 costs contribute less than transport costs to the optimization if all costs are treated equally.
- To account for this and allow for greater emphasis on CO_2 emissions or transport costs per user priorities, the user may adjust the scaling factors for CO_2 cost and transport cost used in the optimization. Designating use of 100% of the total CO_2 cost and 0% of the transport cost will lead to an optimization based solely on social cost of carbon emissions. Designating use of 50% of transport costs and 100% of carbon costs will reduce the contribution of transport to the optimization by half.
- Transport by pipeline has no emissions by default, and accordingly has no associated CO_2 cost.

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

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

<u>**Variable**</u>	<u>**Explanation**</u>	<u>**Indices**</u>
x_{eat}	Flow, in units / time period, along a transport link. *Non-negative.*	▪ Transport link e in set T ▪ Commodity a ▪ During time period t
w_{nat}	Storage, in units, between time periods at an RMP or destination facility. *Non-negative.*	▪ Node n in set of RMP facilities R or destination facilities D ▪ Commodity a ▪ From time period t to time period t+1
u_{nat}	Unmet demand of a commodity at a destination facility. *Non-negative.*	▪ Node n in set of destination facilities D ▪ Commodity a ▪ During time period t
v_{nat}	Production of a commodity at a raw material producer facility. *Non-negative.*	▪ Node n in set of RMP facilities R ▪ Commodity a ▪ During time period t
z_{nbt}	Excess output commodity remaining at processing facility. *Non-negative. Counts toward upper and lower bounds on production but is not used to fulfill demand.*	▪ Node n in set of processor nodes P ▪ Processor output commodity b ▪ During time period t
y_n	Processor build variable. *Binary variable; must take value 0 or 1.*	▪ Node n in set of processor nodes P

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

Coefficient	Explanation	Indices and Components
RC_{ea}	Routing cost, in cost / unit, to flow along a transport link. Routing cost is the sum of transport cost and CO_2 cost components. $RC_{ea} = TC_{ep} + CC_{ep}$	▪ Transport link e in set T ▪ Commodity a, which has phase of matter p
TC_{ep}	Transport component of routing cost, in cost / unit, to flow along a transport link. $TC_{ep} = A_{TC} * MC_{mp} * L_e * I_{mk} + A_{TC} * AC_e + A_{TC} * PT_e$	▪ Transport link e in set T, which has mode m and link type k ▪ Phase of matter p ▪ Transport cost scalar A_{TC} ▪ Modal transport cost MC for mode m and phase p ▪ Link length L for edge e ▪ Impedance I for mode m and link type k ▪ Access cost AC on (artificial) edge e ▪ Penalty for short haul or cost of transloading PT on edge e
CC_{ep}	CO_2 component of routing cost, in cost / unit, to flow along a transport link. $CC_{ep} = A_{CC} * EF_{mk} * L_e * EC$	▪ Transport link e in set T, which has mode m and link type k ▪ Phase of matter p ▪ CO_2 cost scalar A_{CC} ▪ Modal emissions factor EF for mode m and link type k ▪ Link length L for edge e ▪ CO_2 unit cost EC
A_{TC}	Transport cost scalar, as specified by user.	▪ N/A
MC_{mp}	Modal transport cost, as specified by user.	▪ Mode, or artificial link category, m ▪ Phase of matter p
L_e	Length of transport link.	▪ Link e
I_{mk}	Impedance weight of transport link, as specified by user.	▪ Mode, or artificial link category, m ▪ Link type k
AC_e	Access cost associated with an artificial link to a user-defined facility.	▪ Link e
PT_e	Penalty associated with transportation along a link, either cost of transloading movement or short haul penalty, as specified by user.	▪ Link e
A_{CC}	CO_2 cost scalar, as specified by user.	▪ N/A
EF_{mk}	CO_2 emissions factor for transport mode, as specified by user.	▪ Mode m ▪ Link type k
EC	Unit cost of CO_2 emissions, as specified by user.	▪ N/A

BC_n	Fixed amortized cost, in scenario currency units, to build a processing facility. If the processing facility is used, this cost is paid once.	▪ Node n in set of processor nodes P
UB_{nat}	Upper bound on flow of a commodity out of origin facility, in units / time period.	▪ Node n in set of RMP nodes R ▪ Commodity a ▪ Time period t
UB_{nbt}	Upper bound on flow of a commodity out of processing facility, in units / time period.	▪ Node n in set of processor nodes P ▪ Processor output commodity b ▪ Time period t
LB_{nbt}	Lower bound on flow of a commodity out of processing facility, in units / time period, if processor is used.	▪ Node k in processor nodes ▪ Output commodity b ▪ Time period t
G_{nab}	Conversion factor at the processing facility.	▪ Node n in set of processor nodes P ▪ Input commodity a ▪ Output commodity b
PU_{na}	Penalty, in cost / unit, for not meeting demand at destination, as specified by user.	▪ Node n in set of destination nodes D ▪ Commodity a
Q_{nat}	Demand, in units / time period, for a commodity at a destination facility.	▪ Node n in set of destination nodes D ▪ Commodity a ▪ Time period t
M_e	Maximum flow (capacity), in units / time period, on an edge. Using capacity in FTOT assumes time periods are daily.	▪ Transport link e in set T

Table 3: Sets used in the FTOT optimization problem.

Set	Members of set
T	Transport edges representing flow of material along edges in the network. Each edge represents a physical link or artificial link in the network.
R	Raw material producer nodes.
P	Processor nodes.
D	Destination nodes.
N	Transportation network nodes.

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

Minimize **total cost** = $\sum_{a,t}(\sum_{e\in T} RC_{ea}x_{eat} + \sum_{n\in P} BC_n y_n + \sum_{n\in D} PU_{na}\, u_{nat})$

Subject to the constraints in Table 4. In the table, the notation $\delta^+(u)$ represents the set of edges leaving node u, while $\delta^-(u)$ represents the set of edges entering node u.

Table 4: Constraints used in the FTOT optimization problem.

Constraint	**Explanation**
Transport node conservation of flow $$\sum_{e \in \partial_n^+} x_{eat} - \sum_{e \in \partial_n^-} x_{eat} = 0$$ $\forall$ n $\in$ N \ P, $\forall$ a, $\forall$ t	Flow must be conserved at network nodes where conversion does not occur relative to time and commodity. For all transportation network nodes n, commodities a, time periods t
Processor conservation of flow $$G_{jab} \sum_{e \in \partial_j^+} x_{eat} = \sum_{e \in \partial_j^-} x_{ebt} + z_{jbt}$$ $\forall$ j $\in P, \forall a, \forall b$	Flow must be conserved at processor nodes for each pair of input and output commodities (with the appropriate conversion factors) relative to time and commodity; any excess is tracked. For all processor nodes j in the set of processor nodes P, pair of commodities a and b
Processor maximum flow $$\sum_{e \in \partial_j^-} x_{eat} \leq y_j UB_{jat}$$ $\forall$ j $\in P, \forall a, \forall t$	If a processing facility is used, flow cannot exceed the upper bound during any time period, across all commodities. Note that if the processing facility is not used ($y_j = 0$), this constraint requires the flow into the processing facility to be 0. For all processor nodes j in the set of processor nodes P, commodities a, time periods t
Processor minimum flow $$\sum_{e \in \partial_j^-} x_{eat} \geq y_j LB_{jat}$$ $\forall$ j $\in P, \forall a, \forall t$	If a processing facility is used, the flow into it over any time period must exceed the lower bound, across all commodities. Note that if the processing facility is not used ($y_j = 0$), this constraint is redundant with the non-negativity constraints on the flow variables. For all processor nodes j in the set of processor nodes P, commodities a, time periods t
RMP maximum production $$w_{jat} \leq UB_{jat}$$ $\forall$ j $\in R, \forall a, \forall t$	Storage at each origin facility does not exceed the origin facility upper bound relative to time and commodity. For all raw material producer nodes j in the set of RMP nodes R, commodities a, time periods t
RMP conservation of flow $$\sum_{e \in \partial_j^+} x_{eat} + w_{jat} = v_{jat} + w_{ja(t-1)}$$ $\forall j \in R, \forall a, \forall t$	Flow and storage out of each origin facility equals prior storage and excess supply relative to time and commodity. For all raw material producer nodes j in the set of RMP nodes R, commodities a, time periods t

Constraint	Explanation
Destination conservation of flow $\sum_{e \in \partial_j^-} x_{ebt} + w_{jb(t-1)} - w_{jbt} + u_{jbt} = Q_{jbt}$ $\forall$ j $\in$ D, $\forall$b, $\forall$t	Unmet demand plus flow into a destination plus net storage flow is equal to that destination's demand, for each commodity. For all destination nodes j in the set of destination nodes D, commodities b, time periods t
The y variables are binary (0 or 1) $y_j \in 0,1 \; \forall j \in P$	A processing facility is used, or it isn't. For all processor nodes j in the set of processor nodes P
The u, v, w, x, and z variables are non-negative $x_{eat} \geq 0 \; \forall \; e \in T, \forall a, \forall t$ $w_{nat} \geq 0 \; \forall \; n \in \{R, D\}, \forall a, \forall t$ $u_{jbt} \geq 0 \; \forall \; j \in D, \forall b, \forall t$ $v_{jat} \geq 0 \; \forall \; j \in R, \forall a, \forall t$ $z_{jbt} \geq 0 \; \forall \; j \in P, \forall b, \forall t$	No negative flows are permitted.
Transportation network edge capacity $\sum_a x_{eat} \leq M_e \; \forall$ e $\in$ T, $\forall$t	For each edge and time period, total flow (sum of all x on edge e) does not exceed capacity for that edge. For all transport links, time periods t

The optimizer then uses standard linear or mixed integer optimization techniques such as a revised simplex algorithm, an interior point algorithm, an active set quadratic programming algorithm, or a branch-and-cut 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 solver specified by the user, 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. Note: In the back end of FTOT, densities are used to convert liquid commodity quantities and units to their solid equivalents before being passed into the optimization step. After the optimization step, the optimal results are converted back to liquid units. This approach enables the optimization to differentiate liquids based on commodity-specific costs. Users can specify the density for each liquid commodity in the commodity density CSV file. If a density is not provided for a commodity, or if the commodity density file is not used, FTOT automatically applies the default density defined in the XML.

2.4.2 Candidate generation option in optimizer

FTOT can generate screening-level processor candidate locations to convert raw material to commodities demanded by facilities further downstream in the supply chain. 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. As of FTOT version 2022.4, candidate generation scenarios are compatible with the NDR pre-solve step, though the precise behavior and optimal scenario results may differ when using NDR versus when not using it, as the methodologies for candidate generation differ.

- A candidate processor input file, proc_cand.csv, must be included for FTOT to generate candidates.
- In order to get a solution, the user must set a maximum transport distance for any input material to a candidate process (whether coming from an RMP or a processor). 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. When NDR is enabled, these paths are then filtered down to the minimum cost paths to reach the next facility type in the scenario.
- Candidate generation requires two rounds of optimization. In the first optimization, commodities flow from the source of a candidate process's input commodity directly to a destination of the candidate process's output commodity. FTOT builds out paths from the source facilities and converts the input commodity into processed output commodity upon reaching the maximum transport distance, at a facility that accepts the processor's output commodity as input, or at an intermodal facility in scenarios with different allowed modes for input and output commodities of the candidate process. This conversion uses the most efficient process specified in the candidate processor input file. From this point, FTOT flows the processed output commodities through the rest of the supply chain 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 with enough material flow to support a candidate processor facility. The user can specify this amount using the minimum aggregation property for a candidate processor. Raw material producers and points where raw material flows aggregate are flagged as candidate nodes.
- 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. When NDR is enabled, the optimization problem is again run only on the subset of shortest cost paths between facilities.

Candidate generation scenarios are not currently compatible with candidate processes involving multiple inputs or for multiple processes with overlapping input or output commodities.

As of FTOT version 2024.4, FTOT can generate candidate processors at any stage in the supply chain when NDR is enabled, as long as the candidate processor's input commodity has a maximum transport distance specified.

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.
- If the user is providing a list of candidate facilities, it is optional to include a maximum transport distance for a candidate processor's input commodity. If included, the maximum transport distance will be applied in choosing optimal processors for routing.
- The user may specify both candidate processors that have been externally generated and those that are created during the FTOT candidate generation steps. A maximum transport distance for the candidate processor's input commodity is required to generate candidates and will be used for both types of candidate processors.

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.

2.4.4 End-to-end source tracking

When detailed information is required about exactly where material flowed, data from the optimal_route_segments and od_pairs tables in the SQLite database can be joined on scenario_rt_id to provide the end-to-end path of every amount of demand met. The optimal_route_segments table contains route information like mode and amount of commodity flowed, while the od_pairs table contains facility information like the origin and destination facility of each route. This tracking functionality is only accessible if NDR is enabled.

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. Commodity flow amounts for liquid commodities are converted back from units of mass to units of volume using a commodity density. 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 link type and urban/rural 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**
- **Length** of network used
- **Liquid or solid unit-distance** traveled
- **Transport cost**, based on flow, distance traveled, and base modal transport cost
- **Access cost,** based on flow
- **CO_2 cost,** based on flow, distance traveled, modal emissions factor, and cost of CO_2 emissions (for scenarios where the CO_2 cost scalar is set to greater than 0)
- **Routing cost**, based on flow, distance traveled, base modal transport cost and associated weights (impedances), and CO_2 cost
- **Routing cost from transport,** a fraction indicating the weighted transport cost component of routing cost (for scenarios where the CO_2 cost scalar is set to greater than 0)
- **Vehicle loads**, based on flow and payload, with distinct vehicles used for every loading event
- **Vehicle-distance traveled,** based on flow, distance traveled, and payload
- **Emissions**, based on vehicle-distance traveled and emissions factors (for road mode), as well as commodity flow (for non-road modes)
- **Fuel burn**, based on commodity flow, distance traveled, and fuel efficiency
- **Facility input / output amounts** and utilization (where processor utilization is based on total capacity of all processor candidates, not just those used in the optimal solution)

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.

Note that routing cost metrics always include costs on artificial links, as artificial links are considered in the optimization. Optionally, users can also include artificial links in reporting metrics for transport cost, CO_2 cost, length of network used, vehicle-distance traveled, emissions, and fuel burn by setting the \`Report_With_Artificial_Links\` toggle in the scenario XML to True. Of these metrics, routing cost, transport cost, CO_2 cost, and emissions are calculated assuming first-mile, last-mile transport on local road by default; the remaining metrics assume the vehicle load attributes of the modal network to which the artificial link connects.

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. FTOT scenario inputs 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 shows the required and optional inputs and outputs of FTOT.

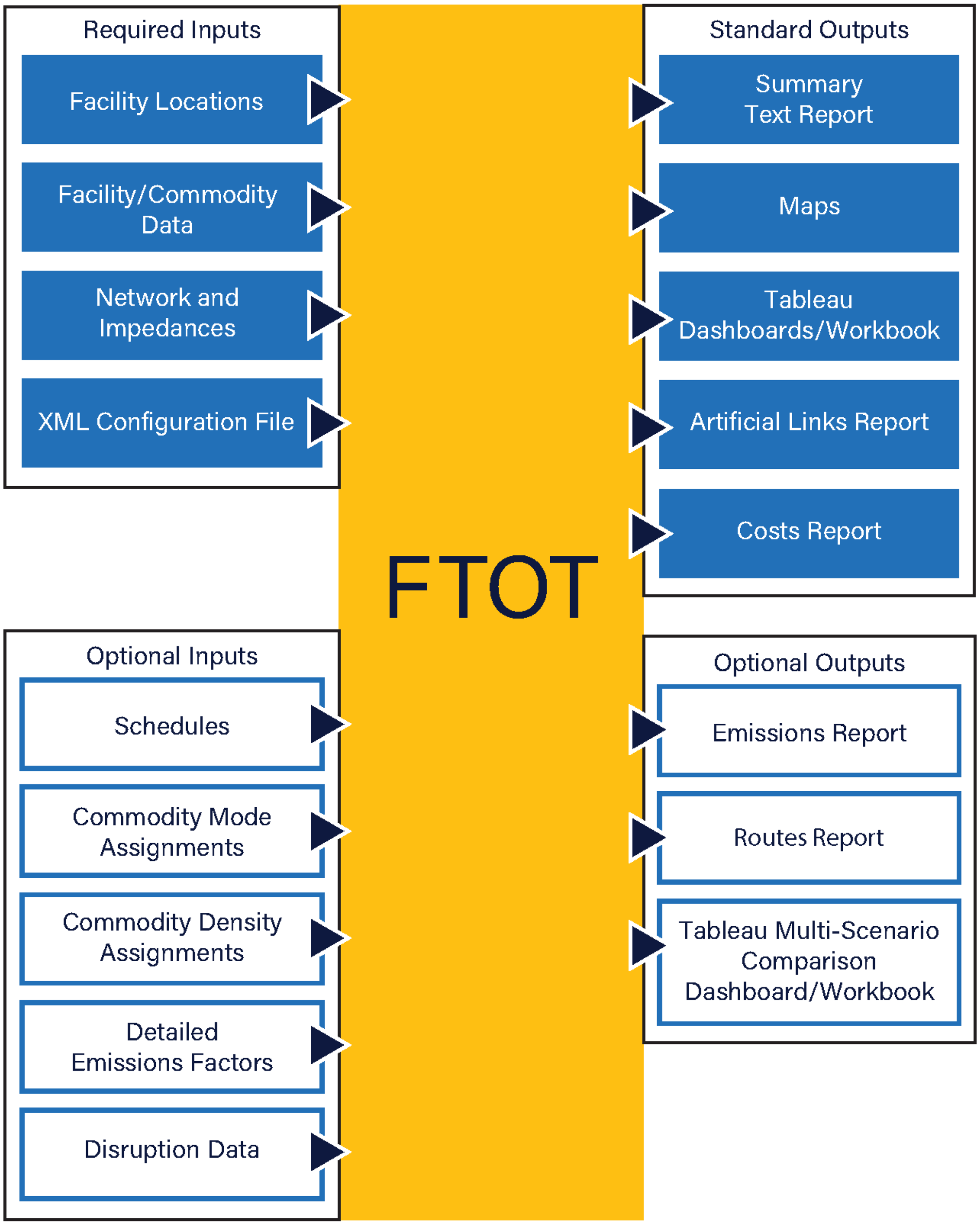

Figure 4: Inputs and outputs of FTOT.

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 file uses tags to indicate the presence 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, described in the User Guide Section 3.1.

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 defined 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 [2]
    - \`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_*\` – user-specified units for reporting. Mixed units are acceptable in input CSV files. FTOT uses the Pint module [23] in Python to track and convert quantities in the XML and CSV input files to the default units specified in the scenario configuration file. Note that default distance units must match the units

[2] The CSV file inputs can be formatted as absolute paths or as paths relative to the scenario XML file location. For example, for the Quick Start 1 scenario, the 'RMP_Commodity_Data' file path can be entered as 'C:\FTOT\scenarios\quick_start\qs1_rmp_dest\input_data\rmp.csv' or as 'input_data\rmp.csv'.

used in the “Length” attributes of the GIS network, while currency units must be consistent throughout the XML and CSV input files.

- Assumptions
  - Vehicle load capacity, fuel efficiency, and emissions factors are specified here. Users are free to modify values as necessary.
  - `Detailed_Emissions_Data` – (optional) location of the detailed emissions factors CSV file. If provided, a detailed emissions report will be generated.
  - `Density_Conversion_Factor` – (optional) default density value used to calculate emissions for liquid commodities on rail, water, and pipeline modes and fuel burn for liquid commodities on road, rail, and water modes. When a density is not provided for a liquid commodity in the commodity density CSV file, the default density value is used instead to convert amounts of that commodity from units of volume to units of mass.
  - `Speed_Time_Data` – (optional) speeds by mode and link type and traversal time for nodes (e.g., intermodal facilities and locks) for calculating travel time when NDR_On is true.
- Script Parameters
  - Network_Costs
    - Modal costs: road, railroad, and barge base costs and impedances
      - Note: pipeline costs are specified by tariffs values found in the GIS network.
      - Impedances are specified through a separate impedance weights CSV file.
    - Intermodal transloading costs
    - Artificial link costs (optional)
  - Artificial links
    - Distances to connect facility to the network
    - `Report_With_Artificial_Links` – (optional) toggle to include artificial links in summary metrics.
  - 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.
  - CO_2 optimization options – (optional) parameters to set CO_2 emissions unit cost and scaling factors for combining transport routing cost and CO_2 emissions cost in the optimization.

- Solver options – (optional) parameters to set the solver used by the optimizer and a maximum time limit for the solver to run before terminating with the best possible solution found at that time.
- Penalty for not fulfilling destination demand – default penalty value 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 four comma-separated-values input files available for facility definitions in FTOT. These files contain the facility name, facility type, commodity, quantity, units, phase of matter, 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), candidate processors (to be used in a candidate generation scenario), and destinations (locations of demand for the final commodities).

Each facility input file includes the following fields (see FTOT User Guide Section 3.3.1 for more details on how to populate these fields):

- **facility_name**; used to match the locations in the scenario geodatabase.
- **facility_type**; specifies the type of facility. Facility types are raw_material_producer, processor, or ultimate_destination.
- **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. 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 and any required ratios among inputs and/or outputs.
- **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.
- **access_cost (optional);** a cost associated with movements into or out of a facility, in addition to base transportation costs associated with the network. Access cost can be specified per facility, phase (liquid or solid), and in or out movement. The access_cost column is not compatible with candidate processors.

For raw material producers and (when NDR is enabled) processors, there is one additional optional field allowed in the input CSV file:

- **max_transport_distance (optional);** a user-specified maximum transport distance that this commodity can travel. When FTOT is generating candidate processors, this value is required for any commodity that is an input to a candidate process. This field should be excluded from scenario runs where it is not needed for performance reasons.

For destinations, there is one additional optional field allowed in the input CSV file:

- **udp (optional, facility-specific);** a user-specified unmet demand penalty applied to destination facilities where demand is not fully satisfied, for each destination facility and commodity. If not specified, FTOT will uniformly apply the default unmet demand penalty from the scenario XML.

For processors, there are three additional optional fields allowed in the input CSV file:

- **max_capacity (optional)**; here the user can specify the maximum amount of material a processor can handle. If a max_capacity or max_processor_input column does not exist, processor upper capacity is unbounded.
- **min_capacity (optional);** here the user can specify the minimum amount of material a processor can handle if it is utilized.
- **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.

The candidate processor file has a slightly different structure than the other three facility input files. For each candidate processor type specified, in addition to input and output commodities, the user must also specify the following fields, each as a row defined in the “commodity” column:

- **minsize**; the minimum facility size permitted, as total input quantity.
- **maxsize**; the maximum facility size permitted, as total input quantity.
- **cost_formula**; the amortized cost to build the processor, per unit of facility size, e.g., USD/ton.
- **min_aggregation (optional)**; the minimum threshold for flow aggregation to propose a candidate processor facility. If left blank, defaults to one-fourth of minsize.

3.4 Facility Location GIS 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 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, using 2024 Census population data. 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 5.

Table 5: 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 specification 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. Note that the vehicle types functionality is not currently compatible with CO_2-based optimization.

3.7 Commodity Density Input File (Optional)

This optional input file allows the user to specify each commodity’s density. Densities are used to temporarily convert liquid commodity quantities to mass unit equivalents during the optimization step only. This enables commodity-specific costs to be applied to liquid commodities. If a density is not specified for a liquid commodity, the default density defined in the scenario XML will be used for the conversion. Density values are also used for calculating emissions from the transport of liquid commodities on rail, water, and pipeline modes and for comparing commodity-level and facility-level capacity constraints for processors with mixed solid and liquid commodity product slates.

3.8 Detailed Emissions Factors Input File (Optional)

FTOT automatically calculates CO_2 emissions for the optimal solution using default emissions factors for each mode in the scenario XML. However, users can optionally calculate and report additional emissions information using detailed CO_2 and non-CO_2 emissions factors (for road) as well as non-CO_2 emissions factors for rail, water, and pipeline.

The common_data folder (networks subfolder) comes with a default detailed_emissions_factors.csv file 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 emissions 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 3.6). If users create a new custom vehicle, they will need to (1) assign that vehicle to a commodity and (2) add custom emissions factors to detailed_emissions_factors.csv to include that vehicle in the detailed emissions report.

The detailed_emissions_factors.csv file contains the following fields:

- vehicle_label – set to "Default" for the default vehicles assigned in the scenario XML. The user 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.
- urban – set to 1 if the emissions factor applies to urban road links, 0 if it applies to rural roads, or "NA" to not distinguish road types. Also set to "NA" if emissions factor is for water or rail.
- limited_access – set to 1 if the emissions factor applies to limited access roads, 0 if it applies to non-limited access roads, or "NA" to not distinguish road types. Also set to "NA" if emissions factor is for water or rail.
- pollutant – set to one of the following values: CO2, CO, CO2e, CH4, N2O, NOx, PM10, PM2.5, VOC. Note: pollutant is NOT case sensitive.
- value – emissions factor for this entry. Emissions factors must be in grams per unit distance for road entries (e.g., g/mile) and in grams per unit of commodity mass per unit distance (e.g., g/ton/mile) for rail and water. The vehicle emissions portion of Section 3.10 provides additional information on the prepopulated pollutant emissions factors.

Note that emissions factors for road CO_2 with "NA" in the urban and limited_access columns will override the default CO_2 factor in the XML. In contrast, providing additional information (0 or 1) in the urban and limited_access columns will enable more targeted calculations when this information is also available in the road feature class of the GIS network. If a network includes information on urban/rural and limited/nonlimited links, but the detailed emissions file doesn't have corresponding emissions factors, FTOT will use the XML default CO_2 emissions factor in place of the missing values.

In contrast, for non-CO_2 emissions factors, which do not have general default values in the scenario XML, FTOT will calculate the average of any provided emissions factors for a given mode and pollutant, and use that average in place of missing values.

For scenarios that include artificial links in reporting (`Report_With_Artificial_Links` set to True in the scenario XML), artificial link emissions are based on road emissions factors to represent first-mile last-mile transport over road, regardless of the modal network type. For CO_2 on artificial links, FTOT uses the XML default CO_2 emissions factor. For non-CO_2, FTOT uses the average of any provided emissions factors for a given pollutant on road.

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; most default parameter values are sourced from U.S. data sources. A few key default parameters and their sources are described below.

Modal Costs

Table 6: Modal cost units and default values in FTOT.

Transport mode	Cost
Roadway/Truck	0.21 USD/metric ton-mile
Railway/Railcar	0.048 USD/metric ton-mile
Waterway/Barge	0.032 USD/metric ton-mile
Pipeline	Actual tariff cost (varies by route)
Transloading cost	12.35 USD/metric ton

Source for costs—Bureau of Transportation Statistics Average Freight Revenue per Ton-Mile—2020 data. [24]

Transloading costs based on communications with Department of Energy. [25]

Modal Impedance Factors (Weights)

The default impedances are listed in Table 7. These default values are stored in the impedance weights CSV in the networks subfolder of the common_data folder. Impedance categories and weights should be modified for custom networks. Users should add the file path to the impedance weights CSV (whether the default or a custom file) to the XML. As with the input facility-commodity CSV files, the file path to the impedance weights CSV can be an absolute path or relative to the scenario XML location. If an impedance weights CSV file is not provided, FTOT will apply equal weighting across link types.

Impedances listed in the table are based on attributes contained within FTOT's source network data. Road network impedances are based on FHWA roadway functional class and water network impedances are based on actual annual freight flows as included in the National Waterway Network. The rail network relies on a set of categories based on whether the rail segment is part of the Strategic Rail Corridor Network (STRACNET) and whether it is owned or flowed on by Class 1 freight railroads (these data are available in the publicly available version of the NARN which serves as the basis for the FTOT rail network). Artificial links for user-defined facilities by default are impeded like local roads as indicated in a separate “allmodes” rows with link type “Artificial: 1”. See Section 2.2.4 for more details on how to access these source datasets.

Table 7: Modal impedance factors applied to the base modal cost based on modal link category (road) or existing flow bin (rail, waterway).

mode	link_type	weight	notes
road	1	1	Interstates get no penalty
road	2	1.1	Other freeways and expressways get 10% penalty
road	3	1.1	Principal arterials get 10% penalty
road	4	1.2	Minor arterials get 20% penalty
road	5	1.3	Major collectors get 30% penalty
road	6	1.3	Minor collectors get 30% penalty
road	7	1.3	Local roads get 30% penalty
road	--	1.3	If doesn't fall into any of the above categories give it maximum weight
rail	1: Class 1 Owned STRACNET	1	Class 1 Owned Strategical Rail Corridor Network (STRACNET) rail segments get no penalty. STRACNET rail segments are the civil rail lines most important to national defense. Class 1 railroads include the six largest freight railroads operating in the US: BNSF, CN, CPKC, CSXT, NS, and UP
rail	2: Other STRACNET	1.1	Other STRACNET rail segments (e.g., those not owned by a Class 1 carrier) get 10% penalty
rail	3: Class 1 Owned non-STRACNET	1.2	Class 1 Owned non-STRACNET rail segments get 20% penalty
rail	4: Class 1 Rights non-STRACNET	1.3	Class 1 Rights non-STRACNET get 30% penalty. These are rail segments that are not owned by a Class 1 carrier but the NARN indicates that a Class 1 carrier has rights to operate on it.
rail	5: Other Rail	1.4	All other rail (not STRACNET, not owned or flowed on by a Class 1 railroad) gets 40% penalty
rail	--	1.4	If doesn't fall into any of the above categories give it maximum weight
water	1: High Volume	1	High volume links get no penalty
water	2: Medium Volume	1.3	Medium volume links get 30% penalty
water	3: Low Volume	1.6	Low volume links get 60% penalty
water	4: No Volume	10	No freight volume links get 10x penalty
water	--	10	If doesn't fall into any of the above categories give it maximum weight
allmodes	Artificial: 1	1.3	Artificial links get same 30% penalty as local roads

Note: The default value of 10.0 for waterways with no volume is meant to discourage flows on portions of the network with little to no documented freight use. Low volume is defined as less than 1,000,000 tons. Medium volume is greater than or equal to 1,000,000 tons and less than 10,000,000 tons. High volume is defined as greater or equal to 10,000,000 tons.

Vehicle Specifications

FTOT vehicle specifications include vehicle capacities, fuel efficiencies, and CO_2 and other emissions factors. The table below lists default vehicle values. More details on emissions specifically are in the following section.

Table 8. Default vehicle specifications

Mode	Attribute	Value	Source / Method
Road	Capacity (solids)	26 tons	The Geography of Transport Systems (2024) [26]
Road	Capacity (liquids)	8 gallons	The Geography of Transport Systems (2024)
Road	Fuel efficiency	195.2 ton-mile/gallon	Using the default solid truck capacity and weighted average CO_2 emission factor below, and applying 10.21 kg CO_2/gallon diesel as described in the previous section.
Road	CO_2 emissions	1,384.06 g/mile	Calculated as the average emissions factor from the four road types in Table 8, weighted by vehicle miles traveled per MOVES3 [27].
Rail	Capacity (solids)	100 tons	The Geography of Transport Systems (2024)
Rail	Capacity (liquids)	30 thousand gallons	The Geography of Transport Systems (2024)
Rail	Fuel efficiency	471.0 ton-mile/gallon	Provided by ANL 2024 GREET [28] team
Rail	CO_2 emissions	21.5 g/ton-mile	Provided by ANL 2024 GREET [28] team
Water	Capacity (solids)	1500 tons	The Geography of Transport Systems (2024)
Water	Capacity (liquids)	454 thousand gallons	The Geography of Transport Systems (2024)
Water	Fuel efficiency	618.3 ton-mile/gallon	Derived from a 223 Btu/ton-mile value provided by ANL 2024 GREET [28] team using the calorific value and density of marine diesel oil from IMO [29].
Water	CO_2 emissions	34.7 g/ton-mile	Provided by ANL 2024 GREET [28] team
Pipeline	Capacity (liquids)	3,150 thousand gallons	Derived from largest minimum batch required across main lines of the Colonial Pipeline [30].
Pipeline	Capacity (liquids)	3,150 thousand gallons	Derived from largest minimum batch required across main lines of the Colonial Pipeline.

Vehicle Emissions

FTOT estimates total emissions for various pollutants using unique emissions factors for each mode. By default, FTOT generates CO_2 emissions as part of its main report. When cost of CO_2 emissions are included in optimization, FTOT uses a default cost of CO_2 emissions from an EPA report on the social costs of greenhouse gases. [31] Optionally, users can point to a detailed emissions factors CSV file in the scenario XML to generate a separate emissions report based on link-specific link types as well as several non-CO_2 pollutants. The following paragraphs describe FTOT's default detailed emissions factors, which are provided in the networks subfolder of the common_data folder. These emissions factors are based on the FTOT's default US network.

Road emissions. Truck emissions are calculated from emissions factors *in units of g/mile* for four district road types: urban nonlimited access, urban limited access, rural nonlimited access, and rural limited access. Emissions factor data sourced from EPA's Motor Vehicles Emission Simulator (MOVES) modeling system. Beginning with FTOT 2021.3, default truck emissions factors are from the MOVES3 model version [27], released Fall 2020. FTOT's default fuel efficiency for road in U.S. ton-miles per gallon is then calculated as the weighted average of the CO_2 emissions factors, based on each road type's vehicle miles traveled from the MOVES model, times 10.21 kg CO_2/gallon of diesel and the truck capacity (see Table 9). The CO_2 intensity of diesel fuel is sourced from EPA's Emission Factors for Greenhouse Gas Inventories. [32]

Based on research regarding the relationship between vehicle load and different emissions factors, FTOT uses a different methodology to calculate CO_2 road emissions and non-CO_2 road emissions. CO_2 emissions are calculated based on partial loads, while all non-CO_2 emissions are calculated based on nearest full truckload. Lightweighting is a common approach to reduce a vehicle's fuel consumption and subsequent CO_2 emissions. There is a body of evidence that has found a high correlation between mass and CO_2 emitted across all vehicle classes but particularly for commercial long-haul trucks like those being modeled in FTOT. [33] [34] [35] As for other regulated pollutants, this relationship with vehicle mass is usually nonlinear. Beyond differences in combustion chemistry between pollutants, the vehicle's age and aftertreatment technology are also critical to determine emissions. For example, due to recent heavy-duty standards, trucks manufactured after model year 2010 are likely to have a diesel particulate filter (DPF) and selective catalytic reduction (SCR) that would cause precipitous reductions in PM and NOx respectively. [36] [37] Criteria pollutant emissions from vehicles with these control technologies would have little to no relationship with vehicle mass. For these reasons, FTOT only uses partial truckloads to allow for a more accurate prediction of CO_2 but continues to estimate emissions from all other pollutants based on the nearest full truckload.

Non-road emissions. For rail, water, and pipeline, most emissions factors *in units of g/ton-mile* were sourced from Argonne National Laboratory's GREET 2024 Model [28]. The emissions factors for CO_2-equivalents are from the 2022 model. Since non-road emissions 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/thousand gallon (the density of kerosene). To change this conversion factor, you will need to include a commodity density CSV input file as specified in Section 3.7 or update the `Density_Conversion_Factor` element in the scenario XML.

Table 9. Detailed emissions factors by pollutant and mode.

Pollutant	Description	*Rural Limited Access* **Road (g/mi)**	*Rural Nonlimited Access* **Road (g/mi)**	*Urban Limited Access* **Road (g/mi)**	*Urban Nonlimited Access* **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	21.54	34.72	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	Nitrous oxide	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 Flows

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 FAF4 Capacity version of the FTOT Public U.S. Contiguous Network that is available for download along with other FTOT supplementary data from https://volpeusdot.github.io/FTOT-Public/data_download.html. Capacity is defined by the number of vehicles or vessels that can pass over a link in a 24-hour period as follows:

- Road (FTOT_Public_US_Contiguous_Network_v2025_FAF4_Capacity.gdb version of the network—FTOT_Public_US_Contiguous_Network_v2025.gdb version does not have capacity information)
    - 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

  - The FTOT U.S. default 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 the U.S. default network, 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). [38]
- Waterway
  - FTOT allows waterway capacity and background flows to be assigned at the link level (water feature class) and/or at the node level (locks feature class). However, only lock level capacity data are available for the default FTOT network covering the contiguous United States.
  - Background flow and capacity data is based on US Army Corps Lock Characteristics and Lock Report data. [39]
  - The default 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. [40] These data are integrated into the default FTOT network. Crude pipelines not in TRIM are left without capacity constraint.
  - A source for petroleum product pipeline background flow and capacity data has not yet been identified, so these pipelines remain without capacity constraint in FTOT.

Short Haul Penalties

FTOT uses short haul penalties to discourage short movements on the rail and waterway networks, as roads are more commonly used for short freight movements. These penalties are applied as an additional routing cost on the artificial links connecting facilities to the rail and waterway networks. A fixed penalty is calculated as the difference between a highly impeded (i.e., local) road cost and the base cost for rail or water transport cost multiplied by a user-specified number of units of distance (e.g., 100 miles), which can be adjusted in the scenario XML. If the user specifies 100 miles, this penalty will cost as much to access rail from the facility as it would to travel 100 miles on the highest cost road links; in other words, there is a threshold distance below which it will not be worth using rail or waterway if road is available. Users can adjust the short haul distance parameter in the scenario XML to increase or decrease the penalty according to their particular scenario. The larger the distance, the more impactful the penalty. The default short haul penalty parameter values in the scenario XML are 100 miles for both rail and waterway networks.

Speeds and Travel Times

FTOT includes an XML parameter to specify average speeds at the modal network- and link type-level, along with average traversal times for node feature classes such as locks and intermodal facilities. If the NDR_On XML parameter is set to True, FTOT will tabulate an estimated travel time between facility pairs and include the results in the routes CSV report. Default values shown in Table 10 are specific to the default U.S. FTOT network and sourced from U.S.-specific references.

Table 10. Average link speeds by mode and link type and node traversal times for locks and intermodal facilities for the default FTOT network.

type	mode	link_type	speed (mph)	time (hr)	notes
link	road	1	63.6	--	Calculated as weighted average by mileage of Free_Speed attribute in default FTOT network
link	road	2	59.3	--	Calculated as weighted average by mileage of Free_Speed attribute in default FTOT network
link	road	3	51.3	--	Calculated as weighted average by mileage of Free_Speed attribute in default FTOT network
link	road	4	49.1	--	Calculated as weighted average by mileage of Free_Speed attribute in default FTOT network
link	road	5	49.6	--	Calculated as weighted average by mileage of Free_Speed attribute in default FTOT network
link	road	6	48.2	--	Calculated as weighted average by mileage of Free_Speed attribute in default FTOT network
link	road	7	37.9	--	Calculated as weighted average by mileage of Free_Speed attribute in default FTOT network
link	road	--	29.1	--	Calculated as weighted average by mileage of Free_Speed attribute in default FTOT network
link	rail	--	40.8	--	Calculated as weighted average by mileage of maximum speed for freight trains by Track Class [41] in default FTOT network
link	water	--	5	--	Provided by ANL GREET team
link	pipeline_crude	--	5.5	--	From report by Institute for Agriculture and Trade Policy (2001) [42]
link	pipeline_prod	--	5.5	--	From report by Institute for Agriculture and Trade Policy (2001)
node	locks	--	--	4.02	Averaged delay and processing time across locks in USACE's Lock Usage Report
node	intermodal	--	--	50	Provided by GREENT team [43]. Applies when switching pipeline tariffs as well.

3.11 Scenario Setup Templates

The FTOT Scenario Setup Template is an Excel-based user interface to assist users in setting up new scenarios. The setup template consists of several worksheets where users enter information related to their scenario that populate the XML configuration file and facility data input files. The setup template serves as a comprehensive input file for simple supply chains. By filling out a copy of the Scenario Setup Template and running the Scenario Setup Conversion Tool (Section 5), users can generate all required input files to run an FTOT scenario.

The worksheets contained in the Scenario Setup Template are:

1. **Instructions:** Instructs users step-by-step on how to fill out the FTOT Scenario Setup Template.
2. **Configuration:** Collects scenario information for the XML configuration file.
3. **Commodities and Processes:** Defines the commodities and conversion processes, if any, used in the scenario.
4. **Facilities and Amounts:** Lists all raw material producers, processors, candidate processors, and destinations used in the scenario along with facility characteristics.
5. **Reference Tables:** Provides reference data useful to users in setting up a scenario.

Note that the Scenario Setup Template does not help create facility location GIS data or any of the optional FTOT input files. If the user needs to provide these inputs, they should be created outside of the setup template and correctly referenced in the XML configuration file.

Starting with the 2025.1 release of FTOT, a sustainable aviation fuel (SAF) specific setup template was added. The SAF Scenario Setup template focuses on SAF scenarios with the following supply chain steps: (1) feedstock sourced at the county level, (2) FTOT-generated and/or existing SAF refineries, (3) blending with jet fuel at petroleum terminals, (4) delivery to airports. The SAF Scenario Setup Template follows a similar process as the default Scenario Setup Template, with the user filling out a series of worksheets and running the Scenario Setup Conversion Tool to generate the required FTOT input files for a regional SAF supply chain. Unlike the default setup template, the SAF-specific template relies on a set of curated datasets for information instead of asking the user to provide their data. The user selects the data relevant to their scenario using a series of data slicers and drop-down menus within Excel.

3.11.1 Sustainable Aviation Fuel (SAF) Scenario Setup Template Data Sources

The SAF version of the setup template compiles several SAF-related datasets so users can easily set up a basic SAF supply chain scenario:

- **Feedstock quantities and locations:** Feedstock quantities from the U.S. DOE 2023 Billion-Ton Report [44], aggregated to counties by the Joint BioEnergy Institute (JBEI) Biositing WebTool [45] Team and accessed via API. In addition, ethanol biorefinery locations for ethanol as a feedstock are sourced from the Renewable Fuels Association [46] and accessed via the JBEI Biositing WebTool.
- **Conversion processes:** Inputs, costs, yields, and capacities for distinct SAF processing technologies sourced from the ICAO SAF Rules of Thumb [47]. Feedstock inputs were mapped to those in the U.S. DOE 2023 Billion-Ton Report. To convert the available data for use in FTOT, a

density of 0.86 kg/L was used for distillate [48] and a density of 0.8 kg/L was used for SAF [49]. Capital investment costs were amortized assuming a 30-year facility lifespan with 6% interest.

- **Existing refineries:** Sourced from Appendix B of the National Renewable Energy Laboratory (NREL) 2024 Sustainable Aviation Fuel State-of-Industry Report [50] and downloaded via the JBEI Biositing WebTool [45].
- **Blending facilities:** Sourced from EIA Petroleum Product Terminals [51].
- **Airport jet fuel demand:** 2024 demand for jet fuel by all users (e.g., airlines, general aviation, military) in millions of gallons per year, rounded to the nearest million, provided by Airlines for America [52].

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 maps and statistics for the scenario, such as relative contributions of commodities and mode to cost, emissions, and vehicle-distance traveled; the amount of supply consumed and demand met; processor capacity; 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, a permitted mode in the scenario, a commodity, and a measure. The measures reported are artificial link length, transport cost, routing cost, access cost, CO_2 emissions, fuel burn, vehicle-distance traveled, and network used. Length is the only measure reported for all facility-mode combinations; the remaining metrics are only reported for artificial links contained in the scenario's optimal solution. The network-used metric is identical to the length metric. To model first and last mile considerations, artificial link measures are calculated using local roads cost methodology for transport cost, routing cost, and CO_2 emissions, regardless of what mode the artificial link is connecting to (though transport cost and routing cost can be customized using optional XML parameters, the impedance weights CSV, and incorporation of user-specified facility, commodity phase, and input/output specific access costs). The other measures assume the vehicle load attributes of the modal network to which the artificial link connects. For the length measure, each entry in the table is the artificial link length if the facility is successfully connected to that mode, or "NA" if the facility is unable to connect to that mode within 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.
- Costs – the costs_TIMESTAMP.csv output file breaks out the cost components that feed into the objective function of the optimization problem. Cost components are grouped by cost family, namely movement (transport, transloading, first-mile last-mile costs, impedances, access costs, and mode short haul penalties), emissions (CO_2 cost of transport on network links and first-mile last-mile CO_2 cost of transport on artificial links), build cost, and unmet demand penalty. The costs report includes both the unscaled and scaled values, where scaled values account for the user-provided transport and CO_2 scaling factors that enable both transport cost and CO_2-based

optimization, respectively. The objective value found in the main text report equals the sum of all scaled cost components.

- Routes – when network density reduction is enabled to calculate shortest paths (see Section 2.3), an additional routes CSV file is generated containing information on all potential routes between facilities considered in the optimization.
- Detailed emissions – when the user provides a detailed emissions factors CSV, a separate emissions CSV report will be generated. Detailed emissions reporting can include additional CO_2 emissions factors based on road type as well as non-CO_2 emissions for transport by road, rail, and waterway. 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, vehicle-distance traveled, fuel burn, and costs. In addition, a supplementary "Routes" dashboard populates when scenarios are run with NDR enabled.

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(s), 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
- Optimal Routes by Commodity and Mode
- Facility Results by Supply and Demand

- Cost Breakdown of the Optimal Routing Solution
- Routes Summary for All Shortest Paths (if run with NDR enabled)
- 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 Scenarios 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
- Base layers (countries, states, counties) provided by the US Census Bureau and Esri's Living Atlas

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 period and commodity specific mapping is output into the scenario's Maps_Time_Commodity folder. Time step mapping is applicable for scenarios which include schedules. These m2 outputs 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 1) the FTOT template XML with default values or 2) an existing, older user-generated XML file.
- **bat_tool:** Generates a bat file for running a new FTOT scenario based on user-provided inputs.
- **generate_template_csv_files:** Generates a set of input data CSV files for RMPs, processors, candidate processors, and destinations as user requested. The user also has the option to include optional fields (e.g., max_transport_distance).
- **replace_xml_text:** Batch replaces 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.
- **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 providing the scenario directories individually.
- **aggregate_raster_data**: Aggregates grid cell production data (e.g., USDA) by county.
- **network_disruption_tool:** Generates 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, the user 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.
- **network_validation_tool:** Allows users to conduct preliminary validation on custom networks. Given a GDB file path, this tool confirms the contained feature classes use a meters-based system, have the correct mode-based naming convention, and contain all required data fields with the correct data type. This tool also summarizes mileage, segment counts, and link type counts, and optionally checks network connectivity (both across modes and within a mode).
- **scenario_setup_conversion_tool:** Converts a copy of the Scenario Setup Template file or the SAF Scenario Setup Template file (see Section 3.11), as filled out by the user to match their scenario specifications, into the bat file, XML scenario file, and input data CSV files required to run an FTOT scenario.
- **udp_sensitivity_tool:** Runs a sensitivity analysis for a user-provided FTOT scenario around the unmet demand penalty (UDP) parameter. Given a desired range for the transportation cost to deliver a unit of the finished product to its destination, the tool runs a series of scenarios varying the UDP parameter to find the level of supply chain utilization that meets that desired delivery cost.

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 693KA9-24-T-00002 under the supervision of Prem Lobo. 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

[1] Volpe National Transportation Systems Center, "Freight and Fuel Transportation Optimization Tool (FTOT)," 30 03 2023. [Online]. Available: https://volpeusdot.github.io/FTOT-Public/.

[2] V. N. T. S. Center, "General Modeling Network Specification (GMNS)," 2022. [Online]. Available: https://github.com/zephyr-data-specs/GMNS.

[3] Volpe National Transportation Systems Center, "FTOT Public Releases," [Online]. Available: https://github.com/VolpeUSDOT/FTOT-Public/releases. [Accessed 30 03 2023].

[4] NetworkX, "Network Analysis in Python v. 2.7," 2022. [Online]. Available: https://networkx.org/.

[5] S. Mitchell, A. Kean, A. Mason, M. O'Sullivan, A. Phillips and F. Peschiera, "PuLP: Optimization with Pulp," 2009. [Online]. Available: https://coin-or.github.io/pulp/.

[6] FHWA, "Freight Analysis Framework Version 5," USDOT, 2022. [Online]. Available: https://ops.fhwa.dot.gov/freight/freight_analysis/faf/.

[7] Federal Railroad Administration, "Transportation.gov: North American All Rail Network (NARN)," USDOT, 2025. [Online]. Available: https://geodata.bts.gov/datasets/usdot::north-american-rail-network-lines/explore.

[8] U.S. Army Corps of Engineers, "U.S. Waterway Data: National Waterway Network," USACE Geospatial Data Portal, 2024. [Online]. Available: https://geospatial-usace.opendata.arcgis.com/datasets/ace7645d305647448a84492a3b909d48_1/about.

[9] EIA, "U.S. Energy Atlas," [Online]. Available: https://atlas.eia.gov/search.

[10] LawIQ, "US Petroleum Product Pipeline Movements OD Pair and Base Rate Report," 2021.

[11] Bureau of Transportation Statistics, "Intermodal Freight Facilities Rail TOFC/COFC," 2024. [Online]. Available: https://data-usdot.opendata.arcgis.com/datasets/usdot::intermodal-freight-facilities-rail-tofc-cofc/about.

[12] Bureau of Transportation Statistics, "Intermodal Freight Facilities Marine Roll-on/Roll-off," 2024. [Online]. Available: https://data-usdot.opendata.arcgis.com/datasets/usdot::intermodal-freight-facilities-marine-roll-on-roll-off/about.

[13] U.S. Army Corps of Engineers, "Master Docks Database," USACE, 2016.

[14] Surface Transportation Board, "Carload Waybill Sample," STB, 2013.

[15] Bureau of Transportation Statistics, "National Transportation Atlans Database: Intermodal Terminal Facilities," USDOT, Washington, D.C., 2011.

[16] Federal Energy Regulatory Commission, "eTariff," 21 August 2023. [Online]. Available: https://www.ferc.gov/ferc-online/etariff. [Accessed 30 March 2024].

[17] U.S. Environmental Protection Agency, "Appendix A: Miscellaneous Data and Conversion Factors," in *AP-42: Compilation of Air Emissions Factors from Stationary Sources*, 1985.

[18] The Engineering ToolBox, "Fuels - Densities and Specific Volumes," 2003. [Online]. Available: https://www.engineeringtoolbox.com/fuels-densities-specific-volumes-d_166.html.

[19] D. Bertsimas and J. Tsisiklis, Introduction to Linear Optimization, Belmont, MA: Athema Scientific, 1997.

[20] COIN-OR, "COIN-OR Linear Program Solver," COIN-OR, 2007.

[21] M. Carreira-Perpinan, "Simplex Method.," Wolfram Web Resource, created by Eric W. Weisstein, 2014.

[22] Q. Huangfu and J. A. J. Hall, "Parallelizing the dual revised simplex method," *Mathematical Programming Computation,* vol. 10, no. 1, pp. 119-142, 2018.

[23] "Pint," [Online]. Available: https://pint.readthedocs.io/en/stable/.

[24] Bureau of Transportation Statistics, "Average Freight Revenue per Ton Mile," USDOT, 2018. [Online]. Available: https://www.bts.gov/content/average-freight-revenue-ton-mile.

[25] K. Moriarty, Interviewee, *Personal Communication Regarding Fuel Transloading Costs.* [Interview]. 17 November 2022.

[26] J.-P. Rodrigue, "Performance Comparison for Selected Freight Modes," in *The Geography of Transport Systems*, New York, Routledge, 2024.

[27] EPA, "MOVES3: Latest Version of Motor Vehicle Emission Simulator," 2020.

[28] M. Wang and X. Liu, Interviewees, *Emissions factor data from GREET for FTOT.* [Interview]. February 2025.

[29] International Maritime Organization, "Amendments to the 2014 Guidelines on the Method of Calculation of the Attained Energy Efficiency Design Index (EEDI) for New Ships," in *Resolution MEPC 281(70)*, 2016.

[30] M. Little and A. Kahn, "Colonial Pipeline Company: Rules and Regulations," FERC, Alpharetta, GA, 2025.

[31] USEPA, "Report on the Social Cost of Greenhouse Gasses: Estimates Incorporating Recent Scientific Advances," USEPA, 2022.

[32] EPA, "Emissions Factors for Greenhouse Gas Inventories," 2020.

[33] P. Gajendran and N. N. Clark, "Effect of Truck Operating Weight on Heavy-Duty Diesel Emissions," *Environmental Science & Technology,* vol. 37, no. 18, pp. 4309-4317, 2003.

[34] ICCT, "Footprint Versus Mass: How to Best Account for Weight Reduction in the European Vehicle CO2 Regulation," October 2017. [Online]. Available: https://theicct.org/sites/default/files/CO2-reduction-technologies_fact-sheet_10102017_vF.pdf. [Accessed 4 October 2023].

[35] S. Tsiakmakis, G. Fontaras, B. Ciuffo and Z. Samaras, "A simulation-based methodology for quantifying European passenger car fleet CO2 emissions," *Applied Energy,* vol. 199, pp. 447-465, 2017.

[36] EICB, "Selective Catalytic Reduction (SCR) + DPF (Diesel Particular Filter)," 4 January 2016. [Online]. Available: https://eibip.eu/publication/selective-catalytic-reduction-scr/.

[37] T. Ballinger, J. Cox, M. Konduru, D. De, W. Manning and P. Andersen, "Evaluation of SCR Catalyst Technology on Diesel Particulate Filters," *SAE International Journal of Fuels and Lubricants,* vol. 2, no. 1, p. 369–74, 2009.

[38] Cambridge Systematics, "National Freight Infrastructure Capacity & Investment Study," Cambridge Systematics, Cambridge, MA, 2007.

[39] U.S. Army Corps of Engineers, "U.S. Waterway Data: National Waterway Network - lockchar2018," USACE, 2018. [Online]. Available: https://usace.contentdm.oclc.org/digital/collection/p16021coll2/id/3005. [Accessed 2022].

[40] G. Oladosu and P. Leiby, "Documentation of the Transportation, Refinery and Infrastructure Model (unpubl.)," 2016.

[41] Cornell Law School, "49 CFR 213.9 Classes of track: operating speed limits," 2020. [Online]. Available: https://www.law.cornell.edu/cfr/text/49/213.9.

[42] Allegro Energy Group, "How Pipelines Make the Oil Market Work - Their Networks, Operation and Regulation," New York, 2001.

[43] EERA, "Global Routing Energy and Emissions Network for Transportation (GREENT)," 2025. [Online]. Available: https://www.greent-freight.io/.

[44] U.S. Department of Energy, "2023 Billion-Ton Report: An Assessment of U.S. Renewable Carbon Resources," 2024. [Online]. Available: https://bioenergykdf.ornl.gov/document/2023-billion-ton-report.

[45] T. Huntington, N. Baral, M. Moore, S. Nordahl, T. Hendrickson, H. Breunig, O. Kavvada, M. Yang, X. Cui, C. Scown and USDOE, "Biositing Webtool V2," JBEI, 2024. [Online]. Available: https://biositing.jbei.org/national.

[46] Renewable Fuels Association, "Ethanol Biorefinery Locations".

[47] ICAO, "SAF Rules of Thumb," [Online]. Available: https://www.icao.int/environmental-protection/Pages/SAF_RULESOFTHUMB.aspx.

[48] Oiltanking, "Middle Distillates / Gasoil," December 2015. [Online]. Available: https://www.oiltanking.com/en/news-info/glossary/middle-distillates-gasoil.html. [Accessed 7 January 2025].

[49] Air Transport Action Group, "Sustainable Aviation Fuel Metrics," February 2021. [Online]. Available: https://aviationbenefits.org/media/167233/fact-sheet_13_saf-metrics-and-conversions_4.pdf. [Accessed 7 January 2025].

[50] NREL, "Sustainable Aviation Fuel State-of-Industry Report," [Online]. Available: https://www.nrel.gov/docs/fy24osti/87803.pdf.

[51] U.S. EIA, "Petroleum Product Terminals," October 2024. [Online]. Available: https://atlas.eia.gov/datasets/eia::petroleum-product-terminals-1/about.

[52] Airlines for America, "Airport Usage Data," 2024.

[53] Bureau of Transportation Statistics, "North American Roads," 16 November 2023. [Online]. Available: https://data-usdot.opendata.arcgis.com/datasets/usdot::north-american-roads/about. [Accessed 30 March 2024].

Appendix A: FAF4 Capacity Network Information

The FAF4 Capacity version of the default FTOT network is limited to the contiguous United States. A full comparison of all FTOT networks is available in Appendix C: Network Comparison. This network is composed of features and selected attributes from the following source datasets:

- 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)—2025 (sent directly to Volpe by FRA and also available at the National Transportation Atlas Database). 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—2024 (accessed via the USACE Geospatial Data Portal). Modified based on additional research to exclude some non-navigable waterways. The waterway network is considered comprehensive.
- Locks: USACE Locks Shapefile—2024 (accessed via the USACE Geospatial Data Portal).
- Crude and Petroleum Product Pipeline Network Data: Energy Information Administration (EIA) crude oil and petroleum product pipeline data—January 2023 (accessed via EIA U.S. Energy Atlas). [9] Modified by Volpe based on station locations in a tariff dataset compiled by Arbo (January 2024) [10] and other publicly available pipeline network data.
- Crude and Petroleum Product Pipeline Tariff Rate Data: Pipeline transportation rate and origin-destination information supplied by Arbo (January 2024). Movements on the pipeline network are constrained to the origin-destination pairs designated in the dataset.
- Intermodal facilities: Volpe-developed list of intermodal facilities. Based on public data from:
    - BTS Intermodal Freight Facilities Rail TOFC/COFC (December 2024) [11]
    - BTS Intermodal Freight Facilities Marine Roll-on/Roll-off (December 2024) [12]
    - EIA Crude Oil by Rail Dataset (October 2020)
    - USACE Master Docks Database (February 2016)
    - Corroboration from Surface Transportation Board Waybill Sample dataset (2013) [14]
    - Pipeline station data from Arbo (January 2024) [10]
    - Volpe research/review of facility locations using satellite imagery and other research of public data sources

Appendix B: North American Network Information

The North American version of the default FTOT network is a variant of the default network which expands geographic coverage to include some modes for Canada, Mexico, Alaska, and Hawaii. Capacity-constrained analysis on this network is not supported. This network is in draft form and is available from the FTOT team upon request. A full comparison of all FTOT networks is available in Appendix C: Network Comparison. This network is composed of features and selected attributes from the following source datasets:

- Road: North American Roads—2020 (accessed via U.S. Department of Transportation Bureau of Transportation Statistics) [53]. The layer is a compilation of data from Transport Canada, Natural Resources Canada, USDOT's Federal Highway Administration, and the Mexican Transportation Institute. The US portion of the road network is based on FAF 4, not the newer and more comprehensive FAF 5. Note that the road network is not comprehensive and is limited to major freight corridors.
- Rail: FRA North American Rail Network (NARN)—2025 (sent directly to Volpe by FRA and also available at the National Transportation Atlas Database). 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—2024 (accessed via the USACE Geospatial Data Portal). Modified based on additional research to exclude some non-navigable waterways. The waterway network is considered comprehensive.
- Locks: USACE Locks Shapefile—2024 (accessed via the USACE Geospatial Data Portal).
- Crude and Petroleum Product Pipeline Network Data: Energy Information Administration (EIA) crude oil and petroleum product pipeline data—January 2023 (accessed via EIA U.S. Energy Atlas). [9] Modified by Volpe based on station locations in a tariff dataset compiled by Arbo (January 2024) [10] and other publicly available pipeline network data.
- Crude and Petroleum Product Pipeline Tariff Rate Data: Pipeline transportation rate and origin-destination information supplied by Arbo (January 2024). Movements on the pipeline network are constrained to the origin-destination pairs designated in the dataset.
- Intermodal facilities: Volpe-developed list of intermodal facilities. Based on public data from:
  - BTS Intermodal Freight Facilities Rail TOFC/COFC (December 2024) [11]
  - BTS Intermodal Freight Facilities Marine Roll-on/Roll-off (December 2024) [12]
  - EIA Crude Oil by Rail Dataset (October 2020)
  - USACE Master Docks Database (February 2016)
  - Corroboration from Surface Transportation Board Waybill Sample dataset (2013) [14]
  - Pipeline station data from Arbo (January 2024) [10]
  - Canadian rail intermodal and marine port facilities from Transport Canada
  - Volpe research/review of facility locations using satellite imagery and other research of public data sources.

Appendix C: Network Comparison

Table 11: Comparison of FTOT default networks.

	Default Network			Capacity Network			North American Network (Includes Canada and Mexico)		
	US	Canada	Mexico	US	Canada	Mexico	US	Canada	Mexico
Road	Yes (FAF 5)	No	No	Yes (FAF 4)	No	No	Yes (North American Roads Network)	Yes	Yes
Rail	Yes	No	No	Yes	No	No	Yes	Yes	Yes
Water	Yes	No	No	Yes	No	No	Yes	Yes	No
Pipeline	Yes	No	No	Yes	No	No	Yes	Partial	No
Intermodal Facilities	Yes	No	No	Yes	No	No	Yes	Yes	No
Compatible with Capacity	No	No	No	Yes	No	No	No	No	No