MULTI-DECADAL DECARBONIZATION PATHWAYS FOR U.S. FREIGHT RAIL

Achieving Sustainable Train Energy Pathways (A-STEP)

Award DE-FOA-00001953, LOCOMOTIVES

North Carolina State University, Raleigh, North Carolina

Probabilistic Cost Model Documentation Report

July 28, 2022

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. OVERVIEW

Here we describe the freight rail decarbonization cost model to evaluate economic feasibility of alternative powertrain decarbonization technologies. This cost model documentation report includes model structure, data sources, model integration, and model verification. The appendices include the Python codes for the cost model and a list of abbreviations, acronyms, letter variables, and definitions.

2. MODEL STRUCTURE

The freight rail probabilistic cost model is used to estimate freight rail levelized costs in U.S. dollars per freight ton-mile travel (i.e., $/ton-mile) and greenhouse gas (GHG) carbon intensity in carbon dioxide equivalent per freight ton-mile travel (i.e., CO_2eq/ton-mile) by powertrain decarbonization technology, U.S. region, assumed scenario, and technology investment year. Powertrain decarbonization technologies include biodiesel, hydrogen (H_2) fuel cell electrification, battery electrification, and overhead catenary electrification. Existing freight rail powertrain technology (e.g., diesel) is used as a baseline to benchmark and compare the levelized costs and carbon intensity among alternative powertrain decarbonization technologies. With data from the U.S. Energy Information Administration (2022), energy consumption and prices in the U.S. are reported in nine regions, including North East, Middle Atlantic, South East, North Central, Central, Texas, North West, South West, and California.

Examined scenarios include policy-driven targets or actions that can affect energy system and freight demand. Energy system scenarios and freight demand scenarios are developed in Task 4 and Task 6, respectively. Energy system scenarios include business-as-usual (i.e., no new policies), 50% U.S. GHG emissions reduction by 2050, and net-zero energy system by 2050. Freight demand scenarios include business-as-usual, agricultural production shift, decrease in agricultural production, population shift, and increase in maritime. Developments of energy system and freight demand scenarios are detailed in the model design reports for Task 4 and Task 6, respectively. In total, 15 combinations of energy system and freight demand scenarios are examined as research scenarios (Table 1). Users for the Achieving Sustainable Train Energy Pathways (A-STEP) can also define additional scenarios with user-provided data.

Table 1. Fifteen research scenarios explicitly evaluated.

Freight Demand Scenarios	**Energy System Scenarios**		
	Business as usual	**50% greenhouse gas emissions reduction by 2050**	**Net zero energy system by 2050**
Business as usual	✓	✓	✓
Agricultural production shift	✓	✓	✓
Decrease in agricultural production	✓	✓	✓
Population shift	✓	✓	✓
Increase in maritime	✓	✓	✓

2.1 Levelized Costs

To estimate freight rail levelized costs (i.e., $/ton-mile) for each powertrain decarbonization technology, region, scenario, and year, the cost model requires input data such as network-level freight train consists, energy consumption, refueling stations, energy price, freight ton-mile travel, and costs with respect to trains and infrastructure (Figure 1). Cost data with respect to trains and infrastructure are described in Section 3.1. Other required data will be an output of other modules in A-STEP. For example, network-level train consists will be provided from the multi-train simulator from Task 2. Network-level energy consumption for diesel and biodiesel freight trains will be provided by the multi-train simulator from Task 2. Network-level energy consumption for H_2 fuel cell, battery-electric, and overhead catenary electric freight trains will be provided by the infrastructure model from Task 3. Network-level infrastructure requirements for refueling or recharging stations will be provided by the infrastructure model from Task 3. Energy prices will be provided by the Tools for Energy Model Optimization and Analysis (Temoa) model from Task 4. Network-level freight ton-mile travel will be provided by the freight demand simulator from Task 6.

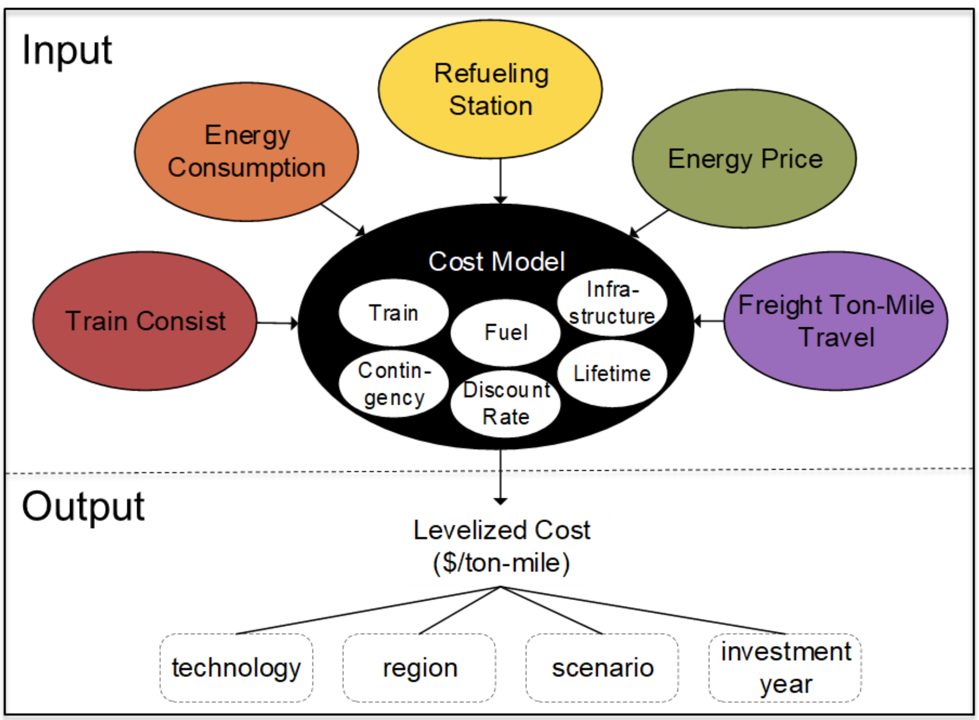

Figure 1. A conceptual diagram of the cost model inputs and outputs to estimate freight rail levelized costs in dollars per ton-mile traveled.

The freight rail levelized cost is the ratio of sum of costs to sum of freight ton-mile traveled over a technology lifetime (Bureau of Transportation Statistics, 2021):

$$\mathrm{LC}_{t,r,e,f,y} = \frac{\sum_{n=1}^{N} \frac{\mathrm{TC}_{t,r,e,f,y,n} + \mathrm{FC}_{t,r,e,f,y,n} + \mathrm{IC}_{t,r,e,f,y,n} + \mathrm{CC}_{t,r,e,f,y,n}}{(1+d)^n}}{\sum_{n=1}^{N} \frac{\mathrm{TMT}_{r,e,f,y,n}}{(1+d)^n}} \tag{1}$$

Where,

$\mathrm{LC}_{t,r,e,f,y}$	=	levelized cost for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, and technology investment year *y* ($/ton-mile);
$\mathrm{TC}_{t,r,e,f,y,n}$	=	train cost for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested ($);
$\mathrm{FC}_{t,r,e,f,y,n}$	=	fuel cost for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested ($);
$\mathrm{IC}_{t,r,e,f,y,n}$	=	infrastructure cost for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested ($);
$\mathrm{CC}_{t,r,e,f,y,n}$	=	contingency cost for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested ($);
$\mathrm{TMT}_{r,e,f,y,n}$	=	ton-mile travel for region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested (ton-mile);
d	=	annual discount rate (percentage point);
e	=	energy system scenario, such as business-as-usual, 50% U.S. GHG gas emissions reduction by 2050, and net zero energy system by 2050;
f	=	freight demand scenario, such as business-as-usual, agricultural production shift, decrease in agricultural production, population shift, and increase in maritime;
N	=	lifetime of a technology (year);
n	=	the n^{th} year since a technology is invested;
r	=	regions in the U.S.;
t	=	powertrain technology, including diesel (baseline), biodiesel, H_2 fuel cell electric, battery electric, and overhead catenary electric;
y	=	technology investment calendar year.

The freight rail levelized cost incorporates train costs, fuel costs, infrastructure costs, contingency costs, and freight ton-mile traveled in a systematic framework. The cost estimation approaches for train costs, fuel costs, infrastructure costs, and contingency costs are detailed below.

2.1.1 Train costs

The train cost includes capital costs for each freight train component, such as locomotives, powertrain systems, tender cars, and freight cars, and annual costs related to train operation and maintenance (O&M). For each freight train component, the capital cost is estimated as the product of component unit cost and number of units. For example, the capital cost for locomotives is estimated as the product of locomotive unit cost and number of locomotives. The train cost is estimated as follows:

$$\mathrm{TC}_{\mathrm{t,r,e,f,y,n}} = \mathrm{N}^{\mathrm{L}}_{\mathrm{t,r,e,f}} \times \mathrm{LOC} + \mathrm{N}^{\mathrm{P}}_{\mathrm{t,r,e,f}} \times \mathrm{PC}_{\mathrm{t,e,f,y,n}} + \mathrm{N}^{\mathrm{T}}_{\mathrm{t,r,e,f}} \times \mathrm{TCC}_{\mathrm{y,n}} + \mathrm{N}^{\mathrm{F}}_{\mathrm{t,r,e,f}} \times \mathrm{FRC} + \mathrm{OTC}_{\mathrm{t,r,e,f,y,n}} + \mathrm{OM}^{\mathrm{train}}_{\mathrm{t,r,e,f,y,n}} \quad (2)$$

Where,

$\mathrm{N}^{\mathrm{L}}_{\mathrm{t,r,e,f}}$	=	number of locomotives (*L*) for technology *t*, region *r*, energy system scenario *e*, and freight demand scenario *f* (#);
LOC	=	locomotive cost ($/unit);
$\mathrm{N}^{\mathrm{P}}_{\mathrm{t,r,e,f}}$	=	number of powertrain systems (*P*) for technology *t*, region *r*, energy system scenario *e*, and freight demand scenario *f* (#);
$\mathrm{PC}_{\mathrm{t,e,f,y,n}}$	=	powertrain system cost for technology *t*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested ($/unit);
$\mathrm{N}^{\mathrm{T}}_{\mathrm{t,r,e,f}}$	=	number of tender cars (*T*) for technology *t*, region *r*, energy system scenario *e*, and freight demand scenario *f* (#);
$\mathrm{TCC}_{\mathrm{y,n}}$	=	tender car cost for technology investment year *y* and the n^{th} year after a technology is invested ($/unit);
$\mathrm{N}^{\mathrm{F}}_{\mathrm{t,r,e,f}}$	=	number of freight cars (*F*) for technology *t*, region *r*, energy system scenario *e*, and freight demand scenario *f* (#);
FRC	=	freight car cost ($/unit);
$\mathrm{OTC}_{\mathrm{t,r,e,f,y,n}}$	=	other train cost for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested ($), such as exhaust aftertreatment system costs for diesel and biodiesel locomotives;
$\mathrm{OM}^{\mathrm{train}}_{\mathrm{t,r,e,f,y,n}}$	=	train O&M costs for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested ($);
L	=	locomotive;
P	=	powertrain system;
T	=	tender car;
F	=	freight car.

Locomotive costs and freight cars costs are assumed to be insensitive to variability in powertrain technologies because existing locomotive frames and freight cars can be adapted to alternative powertrain systems. The powertrain system refers to internal combustion engines for diesel and biodiesel trains (e.g., prime mover engines and head end power engines), H_2 fuel cell systems for H_2 fuel cell electric trains, and battery systems for battery electric trains. The capital cost for powertrain system varies with technology investment year and research scenarios. For example, the capital cost for four-hour lithium ion battery systems was projected to decrease by 57% from 2020 to 2050 because of technology improvements and deployments (Cole et al., 2021).

For diesel and biodiesel trains, exhaust aftertreatment systems are needed to control pollutant emissions from locomotives to meet increasingly stringent emission standards, such as selective catalyst reduction for nitrogen oxides control and diesel particulate filter for particulate matter control (U.S. Environmental Protection Agency, 2008; Naseri et al., 2011). For H_2 fuel cell and battery electric trains, tender cars are needed to carry H_2 storage system and battery units.

2.1.2 Fuel costs

In this work, "fuel" refers to the energy source that powers a freight train. For H_2 fuel cell electric trains, fuel refers to H_2. For battery electric and overhead catenary electric trains, fuel refers to electricity. The freight rail fuel cost is estimated as the fuel price multiplied by fuel usage:

$$\mathrm{FC}_{t,r,e,f,y,n} = \mathrm{FP}_{t,r,e,f,y,n} \times \mathrm{FU}_{t,r,e,f,y,n} \tag{3}$$

Where,

$\mathrm{FP}_{t,r,e,f,y,n}$	=	fuel price for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested (\$/gallon, \$/kg-H_2, or \$/kWh-e);
$\mathrm{FU}_{t,r,e,f,y,n}$	=	fuel usage for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested (gallon, kg-H_2, or kWh-e).

2.1.3 Infrastructure costs

The infrastructure cost includes capital costs for refueling stations and railroad tracks, and annual costs related to infrastructure O&M:

$$\mathrm{IC}_{t,r,e,f,y,n} = \mathrm{N}^{R}_{t,r,e,f} \times \mathrm{RSC}_{t,y,n} + \mathrm{RTC} + \mathrm{OIC}_{t,r,e,f,y,n} + \mathrm{OM}^{infra}_{t,r,e,f,y,n} \tag{4}$$

Where,

$\mathrm{N}^{R}_{t,r,e,f,y}$	=	number of refueling stations (*R*) for technology *t*, region *r*, energy system scenario *e*, and freight demand scenario *f* (#);
$\mathrm{RSC}_{t,y,n}$	=	refueling station cost for technology *t*, technology investment year *y*, and the n^{th} year after a technology is invested (\$/unit);
RTC	=	railroad tracks cost (\$);
$\mathrm{OIC}_{t,r,e,f,y,n}$	=	other infrastructure cost for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested (\$), such as infrastructure upgrade for overhead catenary electrification;
$\mathrm{OM}^{infra}_{t,r,e,f,y,n}$	=	infrastructure O&M costs for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested (\$);
R	=	refueling station.

Existing railroad tracks in the U.S. are assumed to be applied for alternative powertrain technologies. Thus, the railroad tracks cost is not sensitive to powertrain technologies. For overhead catenary electric trains, refueling stations may not be needed; however, infrastructure upgrades for constructing overhead catenary systems should be taken into account, such as constructing overhead wires, building electrical substations, raising bridges, raising and lowering tunnel roofs (California Air Research Board, 2016).

2.1.4 Contingency costs

Contingency costs are the allowances for expected costs that are not accounted for in the cost estimate. Contingency costs are due to uncertainties associated with the state of technology development and the lack of details of the cost estimate. Contingency costs typically account for 5% to 50% of expected costs depending on the type of cost estimates (e.g., simplified, preliminary, detailed, final) (Rothwell, 2005).

2.2 Carbon Intensity

To estimate freight rail carbon intensity (i.e., kg/ton-mile) for each powertrain decarbonization technology, region, scenario, and year, the cost model requires input data such as network-level energy consumption, upstream pollutant emission factors, and downstream pollutant emission factors. GHG pollutants are of greatest concern related to climate change (Intergovernmental Panel on Climate Change, 2022). CO_2 and methane (CH_4) are important GHG pollutant species and, thus, are the focus for this project. The upstream pollutant emission factor refers to mass of pollutant emitted per unit of energy produced. The downstream pollutant emission factor refers to mass of pollutant emitted per unit of energy consumed for combustion processes during train operations. Data sources for upstream and downstream emission factors are given later in Section 3.2.

The freight rail carbon intensity is the ratio of sum of pollutant emissions to sum of freight ton-mile traveled over a technology lifetime:

$$\mathrm{CI}_{p,t,r,e,f,y} = \frac{\sum_{n=1}^{N} \frac{\left(\mathrm{UEF}_{p,t,r,e,f,y,n} + \mathrm{DEF}_{p,t,r,e,f,y,n}\right) \times \mathrm{FU}_{t,r,e,f,y,n}}{(1+d)^n}}{\sum_{n=1}^{N} \frac{\mathrm{TMT}_{r,e,f,y,n}}{(1+d)^n}} \quad (5)$$

Where,

$\mathrm{CI}_{p,t,r,e,f,y}$	=	carbon intensity for pollutant *p*, technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, and technology investment year *y* (kg/ton-mile);
$\mathrm{UEF}_{p,t,r,e,f,y,n}$	=	upstream emission factor for pollutant *p*, technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested (kg/gallon, kg/kg-H_2, or kg/kWh-e);
$\mathrm{DEF}_{p,t,r,e,f,y,n}$	=	downstream emission factor for pollutant *p*, technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested (kg/gallon, kg/kg-H_2, or kg/kWh-e);
p	=	pollutant species, including CO_2 and CH_4.

The GHG carbon intensity in terms of CO_2 equivalent per ton-mile travel (i.e., CO_2eq/ton-mile) is estimated by weighting the carbon intensities for CO_2 and CH_4 by the 100-year global warming potential (GWP-100) for CO_2 and CH_4, respectively:

$$\mathrm{CI}_{\mathrm{GHG},t,r,e,f,y} = \mathrm{GWP}_{\mathrm{CO_2}} \times \mathrm{CI}_{\mathrm{CO_2},t,r,e,f,y} + \mathrm{GWP}_{\mathrm{CH_4}} \times \mathrm{CI}_{\mathrm{CH_4},t,r,e,f,y} \tag{6}$$

Where,

$\mathrm{CI}_{\mathrm{GHG},t,r,e,f,y}$	=	GHG carbon intensity for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, and technology investment year *y* (kg CO_2eq/ton-mile);
$\mathrm{GWP}_{\mathrm{CO_2}}$	=	Global warming potential for CO_2 (kg CO_2eq/kg CO_2). GWP-100 for CO_2 is 1 (Intergovernmental Panel on Climate Change, 2022);
$\mathrm{CI}_{\mathrm{CO_2},t,r,e,f,y}$	=	CO_2 carbon intensity for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, and technology investment year *y* (kg CO_2/ton-mile);
$\mathrm{GWP}_{\mathrm{CH_4}}$	=	Global warming potential for CH_4 (kg CO_2eq/kg CH_4). GWP-100 for CH_4 is 27.0 and 29.9 for non-fossil and fossil energy sources, respectively (Intergovernmental Panel on Climate Change, 2022);
$\mathrm{CI}_{\mathrm{CH_4},t,r,e,f,y}$	=	CH_4 carbon intensity for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, and technology investment year *y* (kg CH_4/ton-mile).

3. DATA SOURCES

This section includes data sources for costs and emission factors used in the cost model.

3.1 Cost Data

Cost data were collected for each cost component in the freight rail cost model, as summarized in Table 2. Data were collected from multiple sources, including published peer-reviewed journal papers, governmental technical reports, and publicly available online data. For a given cost component, if data are from multiple sources, minimum and maximum costs are reported in Table 2. Cost data will be updated as more available and newer data are identified.

3.2 Emission Factors

Upstream and downstream emission factors are used to quantify the GHG carbon intensity. Upstream CO_2 emission factors will be provided by Temoa from Task 4. Upstream CH_4 emission factors will be quantified based on upstream CO_2 emission factors as well as CH_4 leakage rates inferred from the Greenhouse gases, Regulated Emissions, and Energy use in Technologies (GREET) model (Argonne National Laboratory, 2021).

Table 2. Cost data for each cost component used in the freight rail cost model.

Cost Components		Costs for Powertrain Technologies: Diesel	Biodiesel	Hydrogen Fuel Cell Electric	Battery Electric	Overhead Catenary Electric	References for Data Sources [l]
Train	Locomotive [a]	$ 2,000,000-5,000,000	$ 2,000,000-5,000,000	$ 2,000,000-5,000,000	$ 2,000,000-5,000,000	$ 0	CARB, 2016; Surface Transportation Board, 2021
	Powertrain [b]	$ 320,000	$ 320,000	$ 1,390,000-6,350,000	$3,000,000-5,000,000	$ 0	CARB, 2016; ANL, 2019; Phadke and Tasar, 2019
	Tender car [c]	$ 0	$ 0	$190,000	$190,000	$ 0	ANL, 2019
	Freight car [d]	$ 100,000-150,000	$ 100,000-150,000	$ 100,000-150,000	$ 100,000-150,000	$ 100,000-150,000	Railinc Corporation, 2021
	Others [e]	$ 25,000	$ 25,000	$ 0	$ 0	$ 0	CARB, 2016
	Operation & maintenance [f]	$ 6,500,000,000 /year	$ 6,500,000,000 /year	$17,600,000,000 /year	$8,600,000,000 /year	$10,500,000,000 /year	Rocky Mountain Rail Authority, 2010; Eudy, 2018; ANL, 2019; Surface Transportation Board, 2021
Fuel [g]		$ 3.5/gallon	$ 3.8/gallon	$ 16.5/kg-H_2	$ 0.11/kWh-e	$ 0.11/kWh-e	California Energy Commission, 2019; USDOE, 2021b; EIA, 2022
Infra-structure	Refueling Station [h]	$8,000 /locomotive	$8,000 /locomotive	$8,000 /locomotive	$8,000 /locomotive	$ 0	USDOT, 2013
	Railroad tracks [i]	$ 1,000,000-2,000,000 /track-mile	$ 1,000,000-2,000,000 /track-mile	$ 1,000,000-2,000,000 /track-mile	$ 1,000,000-2,000,000 /track-mile	$ 1,000,000-2,000,000 /track-mile	International Energy Agency, 2011
	Others [j]	$ 0	$ 0	$ 0	$ 0	$ 30,000,000 /track-mile	CARB, 2016
	Operation & maintenance [k]	$ 455,000,000 /year	$ 455,000,000 /year	$ 455,000,000 /year	$ 455,000,000 /year	$ 455,000,000 /year	FRA, 2015

Continued on next page

Table 2. Continued.

Notes: a. Locomotive costs are assumed to be not sensitive to variability in powertrain technologies because existing locomotive frames can be adapted to alternative powertrain systems. Overhead catenary electric freight trains have no locomotives.

b. Existing diesel engines can be adapted to biodiesel by changing fuel filters (personal communications with rail representative at North Carolina Department of Transportation Rail Division). Overhead catenary electric freight trains are externally powered by catenary lines and, thus, have no internal powertrain systems.

c. Tender cars are not needed for diesel, biodiesel, and overhead catenary electric trains.

d. Freight car costs are the same among powertrain technologies assuming existing freight cars can be adapted to alternative powertrain systems.

e. Other capital costs for diesel and biodiesel freight trains refer to exhaust aftertreatment systems.

f. Train operation and maintenance costs are assumed to be the same for diesel and biodiesel freight trains.

g. Diesel and biodiesel fuel prices were based on U.S. national average prices between October 1 and October 15, 2021. H_2 fuel price was based on California in 2019. Electricity price was based on U.S. national average price in transportation sector in January, 2022. Details of the fuel and electricity price data will be provided from Task 4 by U.S. regions, research scenarios, and calendar years.

h. Existing diesel refueling stations are assumed to be applied for H_2 refueling and battery recharging stations. These data will be updated, as needed, if Task 3 has new findings on H_2 refueling and battery recharging infrastructure. Refueling stations are not needed for overhead catenary electric freight trains.

i. Existing railroad tracks based on diesel freight trains are assumed to be applied for alternative powertrain systems.

j. Other infrastructure costs for catenary electric freight trains refer to infrastructure upgrades for constructing overhead catenary systems.

k. Infrastructure operation and maintenance costs are assumed to be the same among powertrain technologies. These data will be updated, as needed, if more available data are identified.

l. ANL – Argonne National Laboratory; CARB – California Air Research Board; EIA – Energy Information Administration; FRA – Federal Railroad Administration; USDOE – U.S. Department of Energy; USDOT – U.S. Department of Transportation.

Downstream emission factors depend on the powertrain technology. For H_2 fuel cell, battery-electric, and overhead catenary electric freight trains, downstream CO_2 and CH_4 emission factors are zero because they do have combustion processes. For diesel and biodiesel freight trains, downstream emission factors will be obtained from the GREET model for the pump-to-wheel emission process for vehicle operation (Argonne National Laboratory, 2021).

4. MODEL INTEGRATION

The cost model was integrated into A-STEP as a stand-alone module. Figure 2 shows the main user interface for the cost model. The model was written in Python program languages and developed in a Streamlit web-based platform. Streamlit is a free and open-source web framework based on Python program languages. The Python codes for the cost model are documented in Appendix A.

There are four ribbon tabs at the top of the user interface:

- About: this ribbon tab is used when users would like to show the main user interface and download the description of the model;
- Input Files: this ribbon tab is used when users would like to upload input data files;
- Input Data Display: this ribbon tab is used when users would like to visualize the input data;
- Run Model: this ribbon tab is used when users would like to run the model.

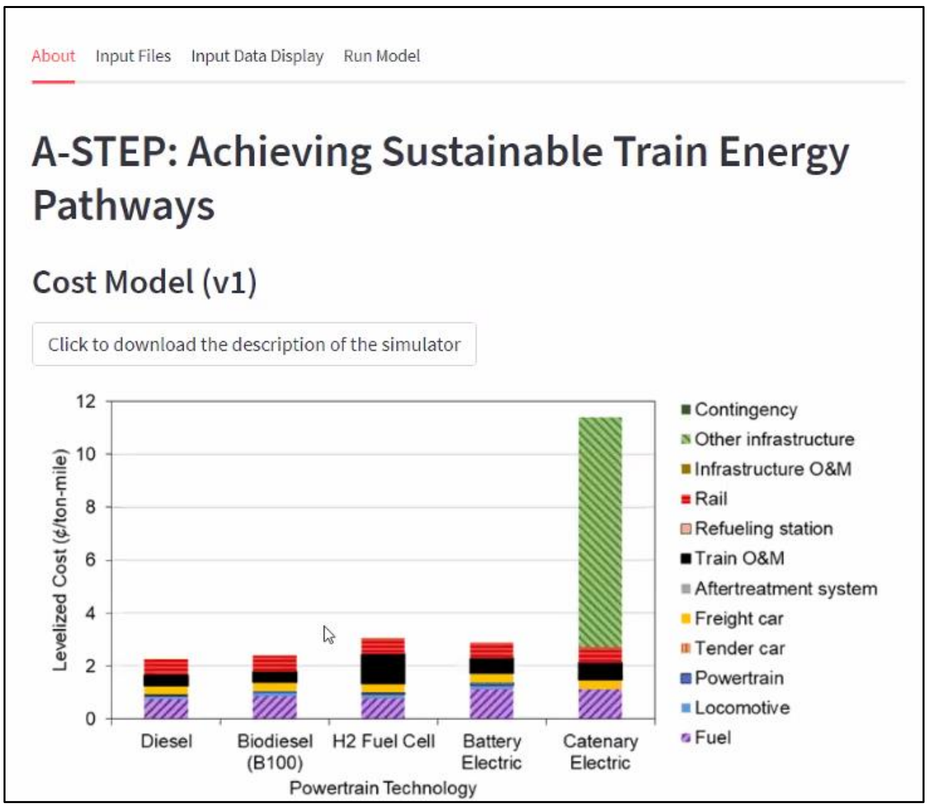

Figure 2. Main user interface for the freight rail decarbonization cost model, one module in the Achieving Sustainable Train Energy Pathways (A-STEP), shown on the Streamlit web-based platform.

4.1 Model Inputs

Figure 3 shows the user interface for Input Files for the freight rail decarbonization cost model. The model requires six input comma-separated values (csv) files, including: (1) train consist; (2) energy consumption; (3) infrastructure; (4) energy price; (5) freight demand; and (6) parameter value. For each input file, there are two options:

- Upload formatted data: this option is used when users have their own data and format the data into each data file;
- Use pre-stored data (click to download): this option is used when users do not have a full set of data files. Default data are pre-stored for each input file. Users can download the pre-stored data files and use them as inputs to run the model.

For each input file, data should be formatted following the instructions below.

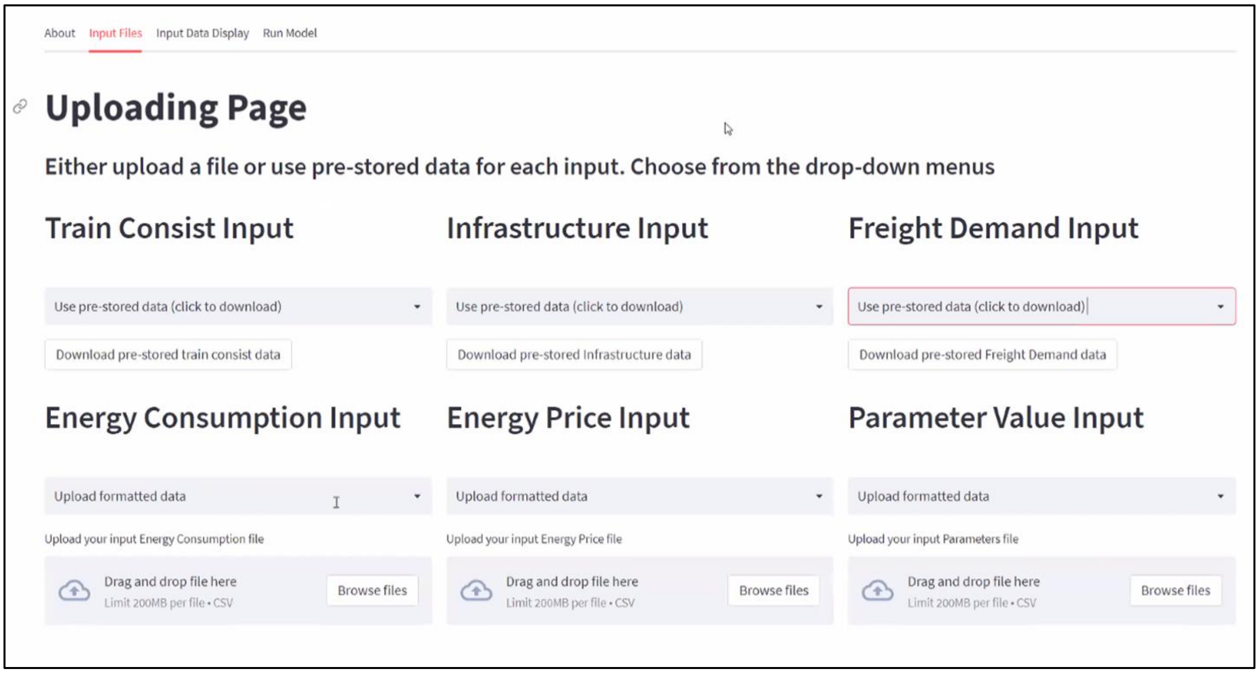

Figure 3. User interface for Input Files for the freight rail decarbonization cost model, one module in the Achieving Sustainable Train Energy Pathways (A-STEP), shown on the Streamlit web-based platform.

4.1.1 Train consist input file

The train consist input csv file has 14 columns, including technology, region, energy system scenario, freight demand scenario, year, number of locomotives, number of powertrains, number of tender cars, number of freight cars, cost per locomotive ($), cost per powertrain ($), cost per tender car ($), cost per freight car ($), and train O&M cost ($/year) (Figure 4). Users can fill in their own data or information below column headers, or download and use pre-stored data. For the train consist input file, users need to make sure:

technology	region	energy_system_scenario	freight_demand_scenario	year	number_of_locomotives	number_of_powertrains
diesel	north_east	business_as_usual	business_as_usual	2025	3241	3241
diesel	middle_atlantic	business_as_usual	business_as_usual	2025	3188	3188
diesel	south_east	business_as_usual	business_as_usual	2025	2643	2643
diesel	north_central	business_as_usual	business_as_usual	2025	2194	2194
diesel	central	business_as_usual	business_as_usual	2025	2937	2937
diesel	texas	business_as_usual	business_as_usual	2025	2979	2979
diesel	north_west	business_as_usual	business_as_usual	2025	3073	3073
diesel	south_west	business_as_usual	business_as_usual	2025	3179	3179
diesel	california	business_as_usual	business_as_usual	2025	3020	3020

(a) Columns A to G

number_of_tender_cars	number_of_freight_cars	cost_per_locomotive ($)	cost_per_powertrain ($)	cost_per_tender_car ($)	cost_per_freight_car ($)	train_O&M_cost ($/year)
0	198964	3500000	320000	0	125000	6500000000
0	201044	3500000	320000	0	125000	6500000000
0	202385	3500000	320000	0	125000	6500000000
0	165731	3500000	320000	0	125000	6500000000
0	211128	3500000	320000	0	125000	6500000000
0	165696	3500000	320000	0	125000	6500000000
0	184820	3500000	320000	0	125000	6500000000
0	179686	3500000	320000	0	125000	6500000000
0	151280	3500000	320000	0	125000	6500000000

(b) Columns H to N

Figure 4. Example of the data format for the input file for train consist: (a) columns A to G; and (b) columns H to N. Note: dummy data are shown here below column headers.

- Each row represents a unique technology, region, energy system scenario, freight demand scenario, and year;
- The data or information in the first five columns (“technology”, “region”, “energy_system_scenario”, “freight_demand_scenario”, and “year”) should be consist with those in the other four input data files, including “energy consumption”, “infrastructure”, “energy price”, and “freight demand”;
- The total number of rows should be consist with those in the other four input data files, including “energy consumption”, “infrastructure”, “energy price”, and “freight demand”.

4.1.2 Energy consumption input file

The energy consumption input csv file has 13 columns, including technology, region, energy system scenario, freight demand scenario, year, energy consumption, energy unit, upstream CO_2 emission factor, downstream CO_2 emission factor, CO_2 emission factor unit, upstream CH_4 emission factor, downstream CH_4 emission factor, and CH_4 emission factor unit (Figure 5). Users can fill in their own data or information below column headers, or download and use pre-stored data. For the energy consumption input file, users need to make sure:

- Each row represents a unique technology, region, energy system scenario, freight demand scenario, and year;
- The data or information in the first five columns (“technology”, “region”, “energy_system_scenario”, “freight_demand_scenario”, and “year”) should be consist with those in the other four input files, including “train consist”, “infrastructure”, “energy price”, and “freight demand”;
- The total number of rows should be consist with those in the other four input files, including “train consist”, “infrastructure”, “energy price”, and “freight demand”.

4.1.3 Infrastructure input file

The infrastructure input csv file has 10 columns, including technology, region, energy system scenario, freight demand scenario, year, number of charging stations, cost per station ($), track mile (mi), rail track cost ($/mi), and infrastructure O&M cost ($/year) (Figure 6). Users can fill in their own data or information below column headers, or download and use pre-stored data. For the infrastructure input file, users need to make sure:

- Each row represents a unique technology, region, energy system scenario, freight demand scenario, and year;
- The data or information in the first five columns (“technology”, “region”, “energy_system_scenario”, “freight_demand_scenario”, and “year”) should be consist with those in the other four input files, including “train consist”, “energy consumption”, “energy price”, and “freight demand”;
- The total number of rows should be consist with those in the other four input files, including “train consist”, “energy consumption”, “energy price”, and “freight demand”.

technology	region	energy_system_scenario	freight_demand_scenario	year	energy_consumption	energy_unit
diesel	north_east	business_as_usual	business_as_usual	2025	423297794	gallon
diesel	middle_atlantic	business_as_usual	business_as_usual	2025	418551370	gallon
diesel	south_east	business_as_usual	business_as_usual	2025	443117856	gallon
diesel	north_central	business_as_usual	business_as_usual	2025	393368333	gallon
diesel	central	business_as_usual	business_as_usual	2025	359308812	gallon
diesel	texas	business_as_usual	business_as_usual	2025	372358702	gallon
diesel	north_west	business_as_usual	business_as_usual	2025	342369942	gallon
diesel	south_west	business_as_usual	business_as_usual	2025	351496905	gallon
diesel	california	business_as_usual	business_as_usual	2025	323293933	gallon

(a) Columns A to G

upstream_CO2_emission_factor	downstream_CO2_emission_factor	CO2_emission_factor_unit	upstream_CH4_emission_factor	downstream_CH4_emission_factor	CH4_emission_factor_unit
1.342	10.9994	kgCO2/gallon	0.01328522	0.000992586	kgCH4/gallon
1.361	8.7506	kgCO2/gallon	0.0130031	0.000742172	kgCH4/gallon
1.331	8.4669	kgCO2/gallon	0.01450882	0.000959472	kgCH4/gallon
1.705	8.5194	kgCO2/gallon	0.01356289	0.000793753	kgCH4/gallon
1.391	11.3477	kgCO2/gallon	0.01358435	0.000827824	kgCH4/gallon
1.333	9.4956	kgCO2/gallon	0.01145932	0.000909551	kgCH4/gallon
1.558	10.4765	kgCO2/gallon	0.01333145	0.000919213	kgCH4/gallon
1.955	8.9973	kgCO2/gallon	0.01169799	0.000723642	kgCH4/gallon
1.931	8.7565	kgCO2/gallon	0.01483722	0.000773719	kgCH4/gallon

(b) Columns H to M

Figure 5. Example of the data format for the input file for energy consumption: (a) columns A to G; and (b) columns H to M. Note: dummy data are shown here below column headers.

technology	region	energy_system_scenario	freight_demand_scenario	year	number_of_charging_stations
diesel	north_east	business_as_usual	business_as_usual	2025	54
diesel	middle_atlantic	business_as_usual	business_as_usual	2025	55
diesel	south_east	business_as_usual	business_as_usual	2025	56
diesel	north_central	business_as_usual	business_as_usual	2025	50
diesel	central	business_as_usual	business_as_usual	2025	51
diesel	texas	business_as_usual	business_as_usual	2025	56
diesel	north_west	business_as_usual	business_as_usual	2025	54
diesel	south_west	business_as_usual	business_as_usual	2025	43
diesel	california	business_as_usual	business_as_usual	2025	52

(a) Columns A to F

cost_per_station ($)	track_mile (mi)	rail_track_cost ($/mi)	infrastructure_O&M_cost ($/year)
460833.7237	12934	1500000	455000000
460833.7237	16458	1500000	455000000
460833.7237	17089	1500000	455000000
460833.7237	13999	1500000	455000000
460833.7237	15298	1500000	455000000
460833.7237	15440	1500000	455000000
460833.7237	18175	1500000	455000000
460833.7237	18333	1500000	455000000
460833.7237	15016	1500000	455000000

(b) Columns G to J

Figure 6. Example of the data format for the input file for infrastructure: (a) columns A to F; and (b) columns G to J. Note: dummy data are shown here below column headers.

4.1.4 Energy price input file

The energy price input csv file has 7 columns, including technology, region, energy system scenario, freight demand scenario, year, energy price, and energy price unit (Figure 7). Users can fill in their own data or information below column headers, or download and use pre-stored data. For the energy price input file, users need to make sure:

- Each row represents a unique technology, region, energy system scenario, freight demand scenario, and year;
- The data or information in the first five columns ("technology", "region", "energy_system_scenario", "freight_demand_scenario", and "year") should be consist with those in the other four input files, including "train consist", "energy consumption", "infrastructure", and "freight demand";
- The total number of rows should be consist with those in the other four input files, including "train consist", "energy consumption", "infrastructure", and "freight demand".

technology	region	energy_system_scenario	freight_demand_scenario	year	energy_price	energy_price_unit
diesel	north_east	business_as_usual	business_as_usual	2025	3.35	$/gallon
diesel	middle_atlantic	business_as_usual	business_as_usual	2025	2.91	$/gallon
diesel	south_east	business_as_usual	business_as_usual	2025	3.32	$/gallon
diesel	north_central	business_as_usual	business_as_usual	2025	4.17	$/gallon
diesel	central	business_as_usual	business_as_usual	2025	3.06	$/gallon
diesel	texas	business_as_usual	business_as_usual	2025	3.14	$/gallon
diesel	north_west	business_as_usual	business_as_usual	2025	2.92	$/gallon
diesel	south_west	business_as_usual	business_as_usual	2025	2.98	$/gallon
diesel	california	business_as_usual	business_as_usual	2025	2.93	$/gallon

Figure 7. Example of the data format for the input file for energy price. Note: dummy data are shown here below column headers.

4.1.5 Freight demand input file

The energy price input csv file has 6 columns, including technology, region, energy system scenario, freight demand scenario, year, and freight ton-mile travel (Figure 8). Users can fill in their own data or information below column headers, or download and use pre-stored data. For the freight demand input file, users need to make sure:

- Each row represents a unique technology, region, energy system scenario, freight demand scenario, and year;
- The data or information in the first five columns ("technology", "region", "energy_system_scenario", "freight_demand_scenario", and "year") should be consist with those in the other four input files, including "train consist", "energy consumption", "infrastructure", and "energy price";
- The total number of rows should be consist with those in the other four input files, including "train consist", "energy consumption", "infrastructure", and "energy price".

technology	region	energy_system_scenario	freight_demand_scenario	year	freight_ton_mile_travel
diesel	north_east	business_as_usual	business_as_usual	2025	1.96E+11
diesel	middle_atlantic	business_as_usual	business_as_usual	2025	2.14E+11
diesel	south_east	business_as_usual	business_as_usual	2025	1.58E+11
diesel	north_central	business_as_usual	business_as_usual	2025	1.80E+11
diesel	central	business_as_usual	business_as_usual	2025	1.71E+11
diesel	texas	business_as_usual	business_as_usual	2025	1.87E+11
diesel	north_west	business_as_usual	business_as_usual	2025	2.09E+11
diesel	south_west	business_as_usual	business_as_usual	2025	1.75E+11
diesel	california	business_as_usual	business_as_usual	2025	1.98E+11

Figure 8. Example of the data format for the input file for freight demand. Note: dummy data are shown here below column headers.

4.1.6 Parameter value input file

Three parameter values are required to fill in the parameter value input csv file, including annual discount rate (percentage point), technology lifetime (year), and contingency factor (%) (Figure 9). Users can fill in parameter values based on their own assumptions, or download and use pre-stored data.

Parameters	
Annual discount rate (percentage point) =	0.1
Technology lifetime (year) =	30
Contingency factor (%) =	20

Figure 9. Example of the data format for the input file for parameter value. Note: dummy data are shown here for parameter values.

4.2 Model Outputs

Figure 10 shows the user interface for Run Model for the freight rail decarbonization cost model. After uploading all six required input files, users can go to the Run Model ribbon tab and click the button for "Press to run the simulation" to run the cost model. The model processing time is about 4 to 6 seconds and displayed under "Press to run the simulation" when the model run is completed. The cost model generates two output files, including levelized cost data and carbon intensity data. Users can download both files and save them in csv formats.

4.2.1 Levelized cost output file

The levelized cost output csv file has 10 columns, including technology, region, energy system scenario, freight demand scenario, year, levelized train cost (cents/ton-mile), levelized fuel cost (cents/ton-mile), levelized infrastructure cost (cents/ton-mile), levelized contingency cost (cents/ton-mile), and total levelized cost (cents/ton-mile) (Figure 11). The total levelized cost is the sum of levelized costs for each cost component, including train, fuel, infrastructure, and contingency. Each row in the output file represents levelized costs for a given technology, region, energy system scenario, freight demand scenario, and year.

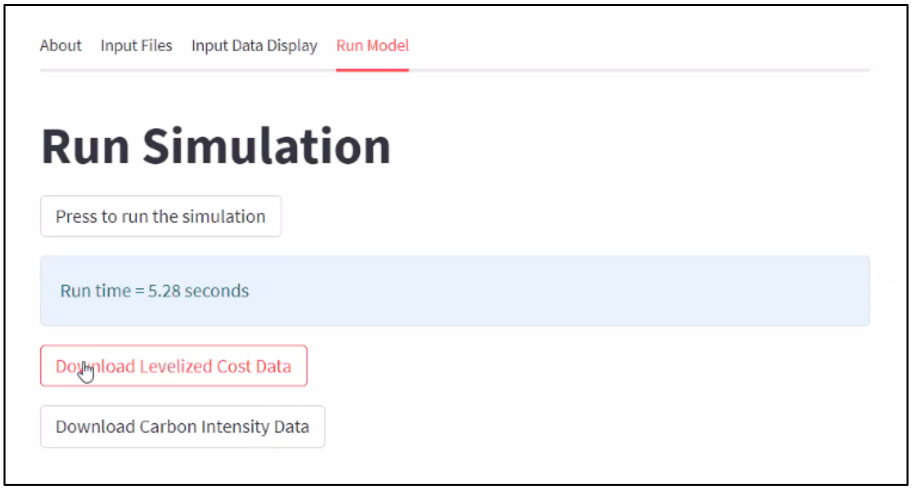

Figure 10. User interface for Run Model for the freight rail decarbonization cost model, one module in the Achieving Sustainable Train Energy Pathways (A-STEP), shown on the Streamlit web-based platform.

4.2.2 Carbon intensity output file

The carbon intensity output csv file has 8 columns, including technology, region, energy system scenario, freight demand scenario, year, CO_2 carbon intensity (kg CO_2/ton-mile), CH_4 carbon intensity (kg CH_4/ton-mile), and CO_2eq carbon intensity (kg CO_2eq/ton-mile) (Figure 12). Each row in the output file represents carbon intensities for a given technology, region, energy system scenario, freight demand scenario, and year for CO_2, CH_4, and GHG, respectively.

5. MODEL VERIFICATION

A demo test was performed to verify the functionality of the cost model. The demo test was run based on dummy input data files. Although the input data were dummy, the order of magnitude of the input data was based on fair assumptions. The output levelized cost and carbon intensity data downloaded from the cost model are 100% match to those calculated independently from a Microsoft Excel Workbook. Thus, the Python codes used in the cost model are verified functionally correct.

To evaluate the validity of the methodological framework for the cost model, the output data based on dummy input data were benchmarked to independent studies. The order of magnitude of the output data is valid in comparison to independent studies. For example, for diesel freight trains, the model-estimated total levelized costs range from 4.2 to 8.7 ¢/ton-mile depending on the region, scenario, and year. This range encloses the U.S. average levelized cost of 4.4 ¢/ton-mile for freight rail in 2020 (Bureau of Transportation Statistics, 2021). For diesel freight trains, the model-estimated CO_2 carbon intensities range from 0.015 to 0.041 kg CO_2/ton-mile depending on the region, scenario, and year. This range encloses the U.S. average CO_2 carbon intensity of 0.022 kg CO_2/ton-mile for freight rail in 2021 (U.S. Environmental Protection Agency, 2022). Thus, the methodological framework for the cost model is deemed valid.

technology	region	energy_system_scenario	freight_demand_scenario	year	levelized_train_cost (cents/ton-mile)
diesel	north_east	business_as_usual	business_as_usual	2025	3.955439064
diesel	middle_atlantic	business_as_usual	business_as_usual	2025	3.625766486
diesel	south_east	business_as_usual	business_as_usual	2025	4.846072146
diesel	north_central	business_as_usual	business_as_usual	2025	4.147441364
diesel	central	business_as_usual	business_as_usual	2025	4.523521656
diesel	texas	business_as_usual	business_as_usual	2025	4.050797881
diesel	north_west	business_as_usual	business_as_usual	2025	3.659061074
diesel	south_west	business_as_usual	business_as_usual	2025	4.366283864
diesel	california	business_as_usual	business_as_usual	2025	3.787092938

(a) Columns A to F

levelized_fuel_cost (cents/ton-mile)	levelized_infrastructure_cost (cents/ton-mile)	levelized_contingency_cost (cents/ton-mile)	total_levelized_cost (cents/ton-mile)
0.724517614	0.563311159	1.048653568	6.291921406
0.570292214	0.598741812	0.958960102	5.753760614
0.928317578	0.826824561	1.320242857	7.921457142
0.910751794	0.641677016	1.139974035	6.839844209
0.641443203	0.712150785	1.175423129	7.052538772
0.625678988	0.657066529	1.06670868	6.400252077
0.477464637	0.651720787	0.957649299	5.745895796
0.597572411	0.782897966	1.149350848	6.896105089
0.477363367	0.608060503	0.974503362	5.847020169

(a) Columns G to J

Figure 11. Example of the data format for the output file for levelized cost: (a) columns A to F; and (b) columns G to J. Note: the output data here are based on dummy input data used in Figures 4 to 9.

technology	region	energy_system_scenario	freight_demand_scenario	year
diesel	north_east	business_as_usual	business_as_usual	2025
diesel	middle_atlantic	business_as_usual	business_as_usual	2025
diesel	south_east	business_as_usual	business_as_usual	2025
diesel	north_central	business_as_usual	business_as_usual	2025
diesel	central	business_as_usual	business_as_usual	2025
diesel	texas	business_as_usual	business_as_usual	2025
diesel	north_west	business_as_usual	business_as_usual	2025
diesel	south_west	business_as_usual	business_as_usual	2025
diesel	california	business_as_usual	business_as_usual	2025

(a) Columns A to E

CO2_carbon_intensity (kgCO2/ton-mile)	CH4_carbon_intensity (kgCH4/ton-mile)	CO2eq_carbon_intensity (kgCO2eq/ton-mile)
0.026691229	3.09E-05	0.027611428
0.019816381	2.69E-05	0.020619119
0.027396273	4.33E-05	0.028685167
0.022330673	3.14E-05	0.023265074
0.026703113	3.02E-05	0.027603403
0.021577158	2.46E-05	0.022311618
0.019678247	2.33E-05	0.020372647
0.02196239	2.49E-05	0.022704673
0.017412358	2.54E-05	0.018170284

(a) Columns F to H

Figure 12. Example of the data format for the output file for carbon intensity: (a) columns A to E; and (b) columns F to H. Note: the output data here are based on dummy input data used in Figures 4 to 9.

6. REFERENCES

Argonne National Laboratory, 2021. Summary of Expansions and Updates in GREET 2021 (Technical Report No. ANL/ESD-21/16). Energy Systems Division, Argonne National Laboratory, Argonne, IL.

Argonne National Laboratory, 2019. Total Cost of Ownership for Line Haul, Yard Switchers, and Regional Passenger Locomotives.

Bureau of Transportation Statistics, 2021. Average Freight Revenue per Ton-Mile [WWW Document]. Bureau of Transportation Statistics, U.S. Department of Transportation. URL https://rosap.ntl.bts.gov/gsearch?ref=docDetails&related_series=National%20Transportation%20Statistics%20%28NTS%29&hdnSortBy=pub_date_key%20desc&maxResults=50&start=0 (accessed 4.5.22).

California Air Research Board, 2016. Technology Assessment: Freight Locomotives (Technical Report). California Air Resources Board, Sacramento, CA.

California Energy Commission, 2019. Joint Agency Staff Report on Assembly Bill 8: 2019 Annual Assessment of Time and Cost Needed to Attain 100 Hydrogen Refueling Stations in California (Staff Report No. CEC-600-2019-039). California Energy Commission, Sacramento, CA.

Cole, W., Frazier, A.W., Augustine, C., 2021. Cost Projections for Utility-Scale Battery Storage: 2021 Update (Technical Report No. NREL/TP-6A20-79236). National Renewable Energy Laboratory, Golden, CO.

Energy Information Administration, 2022. Average Price of Electricity to Ultimate Customers by End-Use Sector [WWW Document]. Electric Power Monthly. URL https://www.eia.gov/electricity/monthly/epm_table_grapher.php?t=epmt_5_6_a (accessed 4.23.22).

Eudy, L., 2018. Technology Validation: Fuel Cell Bus Evaluations.

Federal Railroad Administration, 2015. Operations and Maintenance (O&M) Costs Technical Memorandum (Internal Draft (final version)). Federal Railroad Administration, Washington, DC.

Intergovernmental Panel on Climate Change, 2022. AR6 Synthesis Report: Climate Change 2022 (Final Report No. IPCC AR6). Intergovernmental Panel on Climate Change, Geneva, Switzerland.

International Energy Agency, 2011. Rail Infrastructure (Technology Brief No. T15). International Energy Agency, Paris, France.

Naseri, M., Chatterjee, S., Castagnola, M., Chen, H.-Y., Fedeyko, J., Hess, H., Li, J., 2011. Development of SCR on Diesel Particulate Filter System for Heavy Duty Applications. SAE Int. J. Engines 4, 1798–1809. https://doi.org/10.4271/2011-01-1312

Phadke, A., Tasar, E., 2019. Big Batteries on Wheels - The economic, environmental, and resilience case for rapidly converting diesel locomotives to battery-electric (Working Paper No. 004). Lawrence Berkeley National Laboratory, Berkeley, CA.

Railinc Corporation, 2021. The North American Freight Railcar Review 2021 (Technical Report). Railinc Corporation, Cary, NC.

Rocky Mountain Rail Authority, 2010. High-Speed Rail Feasibility Study Business Plan (Technical Report). Rocky Mountain Rail Authority, Georgetown, CO.

Rothwell, G., 2005. Cost Contingency as the Standard Deviation of the Cost Estimate. Cost Engineering 47, 22–25.

Surface Transportation Board, 2021. Surface Trainsportation Reports Annual R-1 Report [WWW Document]. Surface Transportation Board. URL https://www.stb.gov/reports-data/economic-data/ (accessed 3.31.22).
U.S. Department of Energy, 2021. Clean Cities Alternative Fuel Price Report (Technical Report). Alternative Fuels Data Center, U.S. Department of Energy, Washington, DC.
U.S. Department of Transportation, 2013. Cost-Benefit Analysis Alternative Fuels and Motive Designs (Final Report No. DOT/FRA/ORD-13/21). Federal Railroad Administration, U.S. Department of Transportation, Washington, DC.
U.S. Energy Information Administration, 2022. Annual Energy Outlook 2022 (Technical Report). U.S. Energy Information Administration, Washington, DC.
U.S. Environmental Protection Agency, 2022. Emission Factors for Greenhouse Gas Inventories (Data Report). Center for Corporate Climate Leadership, U.S. Environmental Protection Agency, Washington, DC.
U.S. Environmental Protection Agency, 2008. EPA Finalizes More Stringent Emissions Standards for Locomotives and Marine Compression-Ignition Engines (Technical Report No. EPA420- F- 08– 004). Office of Transportation and Air Quality, U.S. Environmental Protection Agency, Ann Arbor, MI.

APPENDIX A. PYTHON CODES FOR THE COST MODEL

```
# -*- coding: utf-8 -*-
"""
Created on Fri Jul 22 15:07:15 2022

@authors: Ishtiak Ahmed; Tongchuan Wei
"""

# description:         The freight rail decarbonization cost model is used to
estimate freight rail levelized costs and carbon intensity for alternative
powertrain decarbonization technologies
# version:             1.0.0
# python_version:      3.8
# status:              in progress
# requirements
import time
from functools import reduce
import numpy as np
import pandas as pd
import streamlit as st
import plotly.express as px
from PIL import Image

version = 1
#set page config
st.set_page_config(page_title=None, page_icon=None, layout="wide",
initial_sidebar_state="auto", menu_items=None)

#define tab names and numbers
tab1, tab2, tab3, tab4 = st.tabs(["About", "Input Files", "Input Data
Display", "Run Model"])

# %% dataframe conversion function
#this function converts the output data to utf-8 encoded data. It is
important to download the outputs
@st.cache
def convert_df(df):

   return df.to_csv().encode('utf-8')

# %% visualization function
# Only needed if you want to show a graph of the input file
def interactive_plot(df):

    x_axis_val = st.selectbox('Select the X-axis', options=df.columns)
    y_axis_val = st.selectbox('Select the Y-axis', options=df.columns)

    plot = px.line(df, x=x_axis_val, y=y_axis_val)
    st.plotly_chart(plot, use_container_width=True)

# %% Model function
def model(df1,df2,df3,df4,df5,df_parameter):
    dfs = [df1, df2, df3, df4, df5]
    df_final = reduce(lambda left, right: pd.merge(left, right,
on=['technology',
'region','energy_system_scenario','freight_demand_scenario','year']), dfs)
```

```
    i = np_parameter[0, 1] # annual discount rate (percentage point)
    J = np_parameter[1, 1] # technology lifetime (year)
    c = np_parameter[2, 1] # contingency factor (%)

    for n in df_final.index:
        LN = df_final.loc[n, 'number_of_locomotives'] # network-level number
of locomotive
        PN = df_final.loc[n, 'number_of_powertrains'] # network-level number
of powertrains
        TN = df_final.loc[n, 'number_of_tender_cars'] # network-level number
of tender cars
        FN = df_final.loc[n, 'number_of_freight_cars'] # network-level number
of freight cars
        LC = df_final.loc[n, 'cost_per_locomotive ($)'] # locomotive unit
cost ($/unit)
        PC = df_final.loc[n, 'cost_per_powertrain ($)'] # powertrain unit
cost ($/unit)
        TC = df_final.loc[n, 'cost_per_tender_car ($)']  # tender car unit
cost ($/unit)
        FC = df_final.loc[n, 'cost_per_freight_car ($)']  # freight car unit
cost ($/unit)
        OMT = df_final.loc[n, 'train_O&M_cost ($/year)']  # train operation &
maintenance cost ($/year)
        FP = df_final.loc[n, 'energy_price'] # fuel or energy price
($/gallon, $/kg-H2, or $/kWh-e)
        FU = df_final.loc[n, 'energy_consumption'] # fuel or energy usage
(gallon, kg-H2, or kWh-e)
        RN = df_final.loc[n, 'number_of_charging_stations'] # network-level
number of refueling or recharging stations
        RSC = df_final.loc[n, 'cost_per_station ($)'] # refueling or
recharging station unit cost ($/unit)
        TRM = df_final.loc[n, 'track_mile (mi)'] # freight rail track miles
        RTC = df_final.loc[n, 'rail_track_cost ($/mi)'] # rail track cost
($/mile)
        OMI = df_final.loc[n, 'infrastructure_O&M_cost ($/year)'] #
infrastructure operation & maintenance cost ($/year)
        TMT = df_final.loc[n, 'freight_ton_mile_travel'] # freignt ton-mile
travel (ton-mile/year)
        LTC = (LN * LC + PN * PC + TN * TC + FN * FC + OMT * J)/(1 +
i)**J/((TMT * J)/(1 + i)**J) * 100 # estimate levelized train cost (¢/ton-
mile)
        df_final.loc[n, 'levelized_train_cost (cents/ton-mile)'] = LTC
        LFC = FP * FU * J/(1 + i)**J/((TMT * J)/(1 + i)**J) * 100 # estimate
levelized fuel cost (¢/ton-mile)
        df_final.loc[n, 'levelized_fuel_cost (cents/ton-mile)'] = LFC
        LIC = (RN * RSC + TRM * RTC + OMI * J)/(1 + i)**J/((TMT * J)/(1 +
i)**J) * 100 # estimate levelized infrastructure cost (¢/ton-mile)
        df_final.loc[n, 'levelized_infrastructure_cost (cents/ton-mile)'] =
LIC
        LCC = (LTC + LFC + LIC) * c/100 # estimate levelized contingency cost
(¢/ton-mile)
        df_final.loc[n, 'levelized_contingency_cost (cents/ton-mile)'] = LCC
        TLC = LTC + LFC + LIC + LCC # estimate total levelized cost (¢/ton-
mile)
        df_final.loc[n, 'total_levelized_cost (cents/ton-mile)'] = TLC
        CO2U = df_final.loc[n, 'upstream_CO2_emission_factor'] # upstream CO2
```

```
emission factor (kgCO2/gallon, kgCO2/kg-H2, or kgCO2/kWh-e)
        CO2D = df_final.loc[n, 'downstream_CO2_emission_factor'] # downstream
CO2 emission factor (kgCO2/gallon, kgCO2/kg-H2, or kgCO2/kWh-e)
        CH4U = df_final.loc[n, 'upstream_CH4_emission_factor'] # upstream CH4
emission factor (kgCH4/gallon, kgCH4/kg-H2, or kgCH4/kWh-e)
        CH4D = df_final.loc[n, 'downstream_CH4_emission_factor']  #
downstream CH4 emission factor (kgCH4/gallon, kgCH4/kg-H2, or kgCH4/kWh-e)
        LCO2 = (CO2U + CO2D) * FU * J / (1 + i) ** J / ((TMT * J) / (1 + i)
** J)  # estimate levelized CO2 carbon intensity (kgCO2/ton-mile)
        df_final.loc[n, 'CO2_carbon_intensity (kgCO2/ton-mile)'] = LCO2
        LCH4 = (CH4U + CH4D) * FU * J / (1 + i) ** J / ((TMT * J) / (1 + i)
** J)  # estimate levelized CH4 carbon intensity (kgCH4/ton-mile)
        df_final.loc[n, 'CH4_carbon_intensity (kgCH4/ton-mile)'] = LCH4
        if df_final.loc[n, 'technology'] == 'biodiesel':
            LCO2eq = LCO2 + 27.0 * LCH4 # estimate levelized CO2-equivalent
carbon intensity based on GWP-100 for CH4-non fossil from IPCC AR6
        else:
            LCO2eq = LCO2 + 29.8 * LCH4 # estimate levelized CO2-equivalent
carbon intensity based on GWP-100 for CH4-fossil from IPCC AR6
        df_final.loc[n, 'CO2eq_carbon_intensity (kgCO2eq/ton-mile)'] = LCO2eq

    levelized_cost =
df_final.drop(df_final.columns[[5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21
,22,23,24,25,26,27,28,29,30,31,37,38,39]],axis=1)
    carbon_intensity = df_final.drop(df_final.columns[5:37],axis=1) # drop
unnecessary columns for the carbon intensity output file
    return levelized_cost, carbon_intensity

# %% This section is to avoid name error until either all necessary data are
uploaded or pre-stored data are chosen
try: df1
except NameError: df1 = None
try: df2
except NameError: df2 = None
try: df3
except NameError: df3 = None
try: df4
except NameError: df4 = None
try: df5
except NameError: df5 = None
try: df_parameter
except NameError: df_parameter = None

# %% Read hard-coded files
# working directory. Change if necessary
direct = "G:\My Drive\\ITRE_Postdoc\\Rail Decarbonization\\Task 5 Cost
Analysis\\Python code\\cost_model_v1\\Input\\"

train_pre = pd.read_csv(f"{direct}a_data_file_train_consist.csv")
energ_pre = pd.read_csv(f"{direct}b_data_file_energy_consumption.csv")
inf_pre = pd.read_csv(f"{direct}c_data_file_infrastructure.csv")
price_pre = pd.read_csv(f"{direct}d_data_file_energy_price.csv")
dem_pre = pd.read_csv(f"{direct}e_data_file_freight_demand.csv")
param_pre = pd.read_csv(f"{direct}f_data_file_parameter.csv")

# Read hard-coded pdf files
with open(f"{direct}A dummy pdf for cost model.pdf", "rb") as pdf_file:
```

```
    PDFbyte = pdf_file.read()

# add sample output image. This image is hardcoded. Please change the
directory based on where this image files is stored in your pc.
image = Image.open(f"{direct}CM Sample Output.png")

# %% About page
page_title = "A-STEP: Achieving Sustainable Train Energy Pathways"
body = "A-STEP: Achieving Sustainable Train Energy Pathways"
subhead = f"Cost Model (v{version})"

with tab1:
    #st.set_page_config(page_title=page_title)
    st.header(body)
    st.subheader(subhead)

    #add Description pdf
    st.download_button(
       label = "Click to download the description of the simulator",
       data = PDFbyte,
       file_name = "Description.pdf",
       mime='application/octet-stream',
       key='Description'
    )

    # add sample output image
    st.image(image, caption='Sample output: Technology vs Cost')

# %%input page
with tab2:
    st.title("Uploading Page")
    st.subheader("Either upload a file or use pre-stored data for each input.
Choose from the drop-down menus")

    col1, col2, col3 = st.columns(3)

    # create a dropdown meny to choose upload file or use pre-stored data.
One for each of the six inputs, i.e., option1 through option6

    # Dropdown options for train consist data
    with col1:
        st.header("Train Consist Input")
        option1 = st.selectbox(
             label = '', options = ('Upload formatted data', 'Use pre-stored
data (click to download)'),key = 1)
        if option1 == 'Upload formatted data':
            traindf = st.file_uploader("Upload your input Train Consist
file", type=["csv"])
            if traindf is not None:
                df1 = pd.read_csv(traindf)
        elif option1 == 'Use pre-stored data (click to download)':
            st.download_button(
                "Download pre-stored train consist data",
                convert_df(train_pre),
                "train consist.csv",
                "text/csv",
                key='ps1'
```

```
            )
            df1 = train_pre

    # Dropdown options for energy consumtion data
        st.header("Energy Consumption Input")
        option2 = st.selectbox(
             '',
             ('Upload formatted data', 'Use pre-stored data (click to 
download)'),key = 2)
        if option2 == 'Upload formatted data':
            energ = st.file_uploader("Upload your input Energy Consumption 
file", type=["csv"])
            if energ is not None:
                df2 = pd.read_csv(energ)
        elif option2 == 'Use pre-stored data (click to download)':
            st.download_button(
                "Download pre-stored Energy Consumption data",
                convert_df(energ_pre),
                "evenrgy consume.csv",
                "text/csv",
                key='ps2'
            )
            df2 = energ_pre
    with col2:
        # Dropdown options for infrastructure data
        st.header("Infrastructure Input")
        option3 = st.selectbox(
             '',
             ('Upload formatted data', 'Use pre-stored data (click to 
download)'),key = 3)
        if option3 == 'Upload formatted data':
            inf = st.file_uploader("Upload your input Infrastructure file", 
type=["csv"])
            if inf is not None:
                df3 = pd.read_csv(energ)
        elif option3 == 'Use pre-stored data (click to download)':
            st.download_button(
                "Download pre-stored Infrastructure data",
                convert_df(inf_pre),
                "Infrastructure.csv",
                "text/csv",
                key='ps3'
            )
            df3 = inf_pre

        # Dropdown options for energy price data
        st.header("Energy Price Input")
        option4 = st.selectbox(
             '',
             ('Upload formatted data', 'Use pre-stored data (click to 
download)'),key = 4)
        if option4 == 'Upload formatted data':
            price = st.file_uploader("Upload your input Energy Price file", 
type=["csv"])
            if price is not None:
                df4 = pd.read_csv(price)
        elif option4 == 'Use pre-stored data (click to download)':
```

```
            st.download_button(
                "Download pre-stored Energy Price data",
                convert_df(price_pre),
                "Energy price.csv",
                "text/csv",
                key='ps4'
            )
            df4 = price_pre
    with col3:
        # Dropdown options for Freight Demand data
        st.header("Freight Demand Input")
        option5 = st.selectbox(
             '',
             ('Upload formatted data', 'Use pre-stored data (click to
download)'),key = 5)
        if option5 == 'Upload formatted data':
            dem = st.file_uploader("Upload your input Freight Demand file",
type=["csv"])
            if dem is not None:
                df5 = pd.read_csv(dem)
        elif option5 == 'Use pre-stored data (click to download)':
            st.download_button(
                "Download pre-stored Freight Demand data",
                convert_df(dem_pre),
                "Freight demand.csv",
                "text/csv",
                key='ps5'
            )
            df5 = dem_pre

        # Dropdown options for Parameters data
        st.header("Parameter Value Input")
        option6 = st.selectbox(
             '',
             ('Upload formatted data', 'Use pre-stored data (click to
download)'),key = 6)
        if option6 == 'Upload formatted data':
            param = st.file_uploader("Upload your input Parameters file",
type=["csv"])
            if param is not None:
                df_parameter = pd.read_csv(param)
        elif option6 == 'Use pre-stored data (click to download)':
            st.download_button(
                "Download pre-stored Parameters data",
                convert_df(param_pre),
                "Parameters.csv",
                "text/csv",
                key='ps6'
            )
            df_parameter = param_pre

# %% design  display page

with tab3:
    st.title("Show Uploaded Data")

    if df1  is None or df2 is None or df3 is None or df4 is None or df5 is
```

```
None or df_parameter is None:
        st.warning("Upload all necessary data first. Go back to the Input
Page")
    else:
        st.header('Show input train consist data')
        interactive_plot(df1)

        st.header('Show Energy Consumption Data')
        st.write(df2)

        st.header('Show Infrastructure Data')
        st.write(df3)

        st.header('Show Evergy Price Data')
        st.write(df4)

        st.header('Uploaded Frieght Demand Data')
        st.write(df5)

        np_parameter = np.array(df_parameter)
        st.header('Uploaded Parameter Data')
        st.write(df_parameter)

# %% design model page
with tab4:
    st.title("Run Simulation")

    @st.experimental_memo(suppress_st_warning = True)
    def computation():
        start = time.time()
        levelized_cost, carbon_intensity =
model(df1,df2,df3,df4,df5,df_parameter)
        end = time.time()
        levelized_cost_df = convert_df(levelized_cost)
        carbon_intensity_df = convert_df(carbon_intensity)
        return levelized_cost_df,
carbon_intensity_df,start,end,levelized_cost,carbon_intensity

    if df1  is None or df2 is None or df3 is None or df4 is None or df5 is
None or df_parameter is None:
        st.warning("Upload all necessary data first. Go back to the Input
Page")
    else:
        if st.button('Press to run the simulation'):
            levelized_cost_df,
carbon_intensity_df,start,end,levelized_cost,carbon_intensity = computation()
            st.info(
               f"Run time = {round(end-start,ndigits = 2)} seconds"
            )
            st.download_button(
                "Download Levelized Cost Data",
                levelized_cost_df,
                f"levelized cost_{version}.csv",
                "text/csv",
                key='download-csv'
            )
            st.download_button(
```

```
    "Download Carbon Intensity Data",
    carbon_intensity_df,
    f"carbon intensity_{version}.csv",
    "text/csv",
    key='download-profile'
)
st.header("Visualize Levelized Cost output")
interactive_plot(levelized_cost)
st.header("Visualize Carbon Intensity output")
interactive_plot(carbon_intensity)
```

APPENDIX B. ABBREVIATIONS, ACRONYMS, LETTER VARIABLES, AND DEFINITIONS

ANL	Argonne National Laboratory
A-STEP	Achieving Sustainable Train Energy Pathways
CARB	California Air Research Board
$\mathrm{CC}_{t,r,e,f,y,n}$	Contingency cost for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested
CH_4	Methane
$\mathrm{CI}_{CO_2,t,r,e,f,y}$	CO_2 carbon intensity for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, and technology investment year *y*
$\mathrm{CI}_{CH_4,t,r,e,f,y}$	CH_4 carbon intensity for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, and technology investment year *y*
$\mathrm{CI}_{GHG,t,r,e,f,y}$	Greenhouse gas carbon intensity for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, and technology investment year *y*
$\mathrm{CI}_{p,t,r,e,f,y}$	Carbon intensity for pollutant *p*, technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, and technology investment year *y*
CO_2	Carbon dioxide
CO_2eq	Carbon dioxide equivalent
csv	Comma-separated values
d	Annual discount rate
$\mathrm{DEF}_{p,t,r,e,f,y,n}$	Downstream emission factor for pollutant *p*, technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested
e	Energy system scenario
EIA	Energy Information Administration
F	Freight car
f	Freight demand scenario
$\mathrm{FC}_{t,r,e,f,y,n}$	Fuel cost for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested
$\mathrm{FP}_{t,r,e,f,y,n}$	Fuel price for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested
FRA	Federal Railroad Administration

(continued on next page)

Continued from the previous page

FRC	Freight car cost
$FU_{t,r,e,f,y,n}$	Fuel usage for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested
GHG	Greenhouse gas
GREET	Greenhouse gases, Regulated Emissions, and Energy use in Technologies
GWP-100	100-year global warming potential
GWP_{CH_4}	Global warming potential for CH_4
GWP_{CO_2}	Global warming potential for CO_2
H_2	Hydrogen
$IC_{t,r,e,f,y,n}$	Infrastructure cost for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested
L	Locomotive
$LC_{t,r,e,f,y}$	Levelized cost for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, and technology investment year *y*
LOC	Locomotive cost
N	Lifetime of a technology
n	The n^{th} year since a technology is invested
$N^F_{t,r,e,f}$	Number of freight cars (*F*) for technology *t*, region *r*, energy system scenario *e*, and freight demand scenario *f*
$N^L_{t,r,e,f}$	Number of locomotives (*L*) for technology *t*, region *r*, energy system scenario *e*, and freight demand scenario *f*
$N^P_{t,r,e,f}$	Number of powertrain systems (*P*) for technology *t*, region *r*, energy system scenario *e*, and freight demand scenario *f*
$N^R_{t,r,e,f}$	Number of refueling stations (*R*) for technology *t*, region *r*, energy system scenario *e*, and freight demand scenario *f*
$N^T_{t,r,e,f}$	Number of tender cars (*T*) for technology *t*, region *r*, energy system scenario *e*, and freight demand scenario *f*
$OIC_{t,r,e,f,y,n}$	Other infrastructure cost for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested
O&M	Operation and maintenance

(continued on next page)

Continued from the previous page

$\mathrm{OM}^{\mathrm{infra}}_{\mathrm{t,r,e,f,y,n}}$	Infrastructure O&M costs for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested
$\mathrm{OM}^{\mathrm{train}}_{\mathrm{t,r,e,f,y,n}}$	Train O&M costs for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested
$\mathrm{OTC}_{\mathrm{t,r,e,f,y,n}}$	Other train cost for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested
P	Powertrain system
p	Pollutant species
$\mathrm{PC}_{\mathrm{t,e,f,y,n}}$	Powertrain system cost for technology *t*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested
R	Refueling station
r	Regions in the U.S.
$\mathrm{RSC}_{\mathrm{t,y,n}}$	Refueling station cost for technology *t*, technology investment year *y*, and the n^{th} year after a technology is invested
T	Tender car
t	Powertrain technology
Temoa	Tools for Energy Model Optimization and Analysis
$\mathrm{TC}_{\mathrm{t,r,e,f,y,n}}$	Train cost for technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested
$\mathrm{TCC}_{\mathrm{y,n}}$	Tender car cost for technology investment year *y* and the n^{th} year after a technology is invested
$\mathrm{TMT}_{\mathrm{r,e,f,y,n}}$	Ton-mile travel for region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested
$\mathrm{UEF}_{\mathrm{p,t,r,e,f,y,n}}$	Upstream emission factor for pollutant *p*, technology *t*, region *r*, energy system scenario *e*, freight demand scenario *f*, technology investment year *y*, and the n^{th} year after a technology is invested
USDOE	U.S. Department of Energy
USDOT	U.S. Department of Transportation
y	Technology investment calendar year