U.S. Department of Transportation
Volpe Center

Maritime Analytics for Research and Innovation on Noise and Energy – Tool (MARINE-T): Methods and Validation

soleg / 123RF Stock Photo

Prepared by:
Dan F.B. Flynn, Kevin Zhang, Brandon Loesch, Erika A. Sudderth, Arielle Herman, Kirby Ledvina, Kristin C. Lewis, Lyle Tripp, Jaewoong Yun, Juwon Drake, Meghan Ahearn, Eric Boeker

DOT Report Number: DOT-VNTSC-USDOE-26-01

Prepared for:
U.S. Department of Energy: Bioenergy Technologies Office

Contents

List of Figures iv
List of Tables iv
List of Abbreviations v
1. Introduction 1
2. Scope 3
3. Reference Data Sources 4
4. MARINE-T Methods 5
4.1 Vessel Types 6
4.2 Vessel Characteristics 6
4.2.1 Vessel Characteristic Database Preparation 6
4.2.2 Ship-level Data 8
4.2.3 Substitution Mapping 9
4.3 Vessel Movements Data 9
4.3.1 Interpolation 10
4.3.2 Operating Mode Assignment 14
4.4 Fuel Burn Computation Methods 16
4.4.1 Vessel Characteristics for Fuel Burn 16
4.4.2 Speed to Power Calculation 16
4.4.3 Auxiliary Engine and Boiler Power Lookup 17
4.4.4 Power to Fuel Burn Calculation 18
4.5 Aggregation and Reporting 18
4.5.1 Fuel and Energy Metrics 18
4.5.2 Trip Inference and Port Movements 20
5. Model Validation 23
6. Conclusion and Next Steps 26
7. References 27

List of Figures

Figure 4-1. Schematic of MARINE-T, the Volpe Maritime Analytics for Research and Innovation on Noise and Energy – Tool. This report focuses on the fuel burn components of the model; the underwater noise module methods will be presented in detail in a separate report. 5

Figure 4-2 Left: a standard example of potential misattribution, where 100% of the metrics are attributed to the top cell. Right: an example where a grid cell is not attributed any metrics. 12

Figure 4-3 Left: One point is interpolated between the original points. Right: Two points are interpolated between the original points, creating a segment that now properly aggregates a portion of the metrics to a cell that was previously not included. 13

Figure 4-4 Operating mode decision tree. The left half of the decision tree is for vessels operating in open ocean, coastwise, in Great Lakes, or at port........ 15

Figure 4-5 Operating mode decision tree. The right half of the decision tree is for vessels operating in inland waterways. 15

Figure 4-6 View of MARINE-T U.S. EEZ total fuel consumption output for the year 2022. Each pixel here is approximately 1 square kilometer. 19

Figure 4-7. View of Gulf Coast regional scenario outputs, focused on technologies and fuels for offshore activities for harbor craft and offshore supply vessels. 20

Figure 4-8. MARINE-T Trip Inference and Port Movement dashboard image. Chords between regions represent the metric tons of marine diesel oil equivalent fuel consumed on trips in the U.S. EEZ in 2022. Arcs show total fuel for port-to-port movements within a region. 22

Figure 5-1. Estimated and reported fuel consumption for 267 records in 2019, 2021, and 2022 combined, representing 156 unique vessels. The dashed line indicates the ideal 1:1 relationship. 24

List of Tables

Table 4-1. Number of trips that occurred between U.S. ports versus non-U.S. ports. 22

Table 5-1. Average model performance for fuel consumption validation by vessel type for 267 records in 2019, 2021, and 2022 combined, representing 156 unique vessels. 23

List of Abbreviations

Abbreviation	Term
ABS	American Bureau of Shipping
AIS	Automatic Identification System
BETO	Bioenergy Technologies Office
DOE	Department of Energy
DOT	Department of Transportation
ECA	Emissions Control Area
EEZ	Exclusive Economic Zone
EPA	Environmental Protection Agency
ESRI	Environmental Systems Research Institute, Inc.
GISIS	Global Integrated Shipping Information System
GSHHG	Global, Self-consistent, Hierarchical, High-resolution Geography Database
GT	Gross Tonnage
HFO	Heavy Fuel Oil
IHS	Information Handling Services
IMO	International Maritime Organization
IMO DCS	IMO Data Collection System
kWh	Kilowatt-hour
MARAD	Maritime Administration
MARINE-T	Maritime Analytics for Research and Innovation on Noise and Energy – Tool
MARPOL	IMO International Convention on the Prevention of Pollution from Ships
MDO	Marine Diesel Oil
MMSI	Maritime Mobile Service Identity
MSSIS	Maritime Safety and Security Information System
MT	Metric Tons
NGA	National Geospatial-Intelligence Agency
NOAA	National Oceanic and Atmospheric Administration
NOx	Nitrogen Oxides
O-D	Origin-Destination
OGV	Ocean-going Vessel
PM	Particulate Matter
PoC	Ports of Call
RMSE	Root Mean Square Error
RPM	Revolutions per Minute
S&P	Standard and Poor's
SFOC	Specific Fuel Oil Consumption
SOG	Speed Over Ground
SOx	Sulfur Oxides
STEAM	Ship Traffic Emissions Assessment Model
TEU	Twenty-foot Equivalent Unit
UN	United Nations
URN	Underwater Radiated Noise
USCG	United States Coast Guard

Abbreviation	Term
UTC	Coordinated Universal Time
VLSFO	Very Low Sulfur Fuel Oil
WPI	World Port Index

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE (26-11-2025)	**2. REPORT TYPE** White Paper	**3. DATES COVERED (From - To)**

4. TITLE AND SUBTITLE

Maritime Analytics for Research and Innovation on Noise and Energy – Tool (MARINE-T): Methods and Validation

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Dan F.B. Flynn, Kevin Zhang, Brandon Loesch, Erika A. Sudderth, Arielle Herman, Kirby Ledvina, Kristin C. Lewis, Lyle Tripp, Jaewoong Yun, Juwon Drake, Meghan Ahearn, Eric Boeker

5d. PROJECT NUMBER
VR44

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

U.S. Department of Transportation
John A. Volpe National Transportation Systems Center
220 Binney Street
Cambridge, MA 02142-1093

8. PERFORMING ORGANIZATION REPORT NUMBER

DOT-VNTSC-USDOE-26-01

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Department of Energy Bioenergy Technologies Office
15013 Denver West Blvd
Golden, CO 80401

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The Volpe National Transportation Systems Center has developed the *Maritime Analytics for Research and Innovation on Noise and Energy – Tool* (MARINE-T) to quantify and spatially map fuel use, energy consumption, and underwater radiated noise from U.S. maritime vessels. MARINE-T employs a bottom-up modeling approach integrating vessel movement data from Automatic Identification System (AIS) records with vessel characteristics to estimate outputs at consistent spatial and temporal scales. The tool supports analytical priorities of the U.S. Department of Energy and U.S. DOT Maritime Administration by enabling assessment of relationships between transportation efficiency and underwater noise, and identifying potential synergies for economic competitiveness. Primary outputs include gridded maps of fuel burn and noise source levels across the U.S. Exclusive Economic Zone, along with a dashboard providing visual and tabular summaries by vessel type, geographic region, and time period. Current capabilities use an inventory approach and source level noise metrics. Scenario analysis capabilities allow evaluation of operational, technological, and fueling interventions. This document outlines methods for the fuel burn components of the model and validation outputs; noise components of the model will be described in a separate report. Validation against IMO Data Collection System fuel reports shows estimation accuracy within 6% for large vessel movements. MARINE-T provides a robust decision-support resource for promoting maritime transportation efficiency and economic competitiveness.

15. SUBJECT TERMS

Marine Fuel, Fuel Burn, Vessel Energy, Vessel Noise, MARINE-T, AIS

16. SECURITY CLASSIFICATION OF: Unclassified			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT Unclassified	**b. ABSTRACT** Unclassified	**c. THIS PAGE** Unclassified	Unlimited	36	**19b. TELEPHONE NUMBER (include area code)**

Standard Form 298 (Rev. 8-98)

1. Introduction

The U.S. Department of Transportation (DOT) Volpe National Transportation Systems Center (Volpe) has developed the Maritime Analytics for Research and Innovation on Noise and Energy – Tool (MARINE-T) to quantify and map fuel use, energy use, and noise levels from maritime transportation in the United States. MARINE-T is a bottom-up model based on vessel movements and vessel characteristics. The tool produces reports and visuals of energy and fuel use and underwater radiated noise source levels resulting from maritime vessel traffic in the U.S. Exclusive Economic Zone (EEZ), with capacity to produce these outputs at a global scale. The MARINE-T model enhances the analytical capabilities of the Department of Energy (DOE) and the Maritime Administration (MARAD) in support of the economic development and efficiency of the maritime transportation sector priorities. The model also supports MARAD efforts on synergy between approaches to improve transportation efficiency and those to manage underwater noise to foster economic competitiveness of the U.S. maritime sector.

Maritime vessels conventionally utilize heavy fuel oil (HFO) in ocean-going transit, and marine diesel oil (MDO) near coastal waters and in ports. MARINE-T estimates fuel consumption from vessel movements to map and analyze total fuel consumption in the waters of the U.S. From these fuel consumption estimates, for DOE purposes these values can be used to estimate particulate matter (PM) and sulfur oxide (SO_x) amounts. For MARAD purposes, quantifying and mapping fuel consumption of different fuels is important for promoting the economic competitiveness of U.S. flag vessels, and supports administration priorities on maritime competitiveness (1).

The International Maritime Organization (IMO) International Convention on the Prevention of Pollution from Ships (known as MARPOL) established emissions control areas (ECAs), limiting maximum allowable sulfur concentrations in marine fuels and maximum nitrogen oxide (NO_x) emissions rates in engine exhaust. The requirements apply to vessels operating in U.S. waters as well as ships operating within 200 nautical miles of the coast of North America, the region known as the North American ECA (2). Sulfur oxide from fuel oil used onboard ships in specific designated ECAs have been limited to 0.10% m/m (mass by mass) since 2006-2014, depending on the region (3). A limit of 0.5% m/m is in force for ships operating outside designated ECAs, a significant reduction from the previous limit of 3.5%. An amendment to Annex VI of MARPOL made the new limit compulsory, and most ships now use very low sulfur fuel oil (VLSFO), even outside of the specific designated ECAs (4). Scrubbers can also be used with traditional heavy fuel oil to remove sulfur oxide from emissions and may be cost effective as VLSFO has a price premium over traditional heavy fuel (5). Understanding how much and where fuels of different properties are being consumed is important for supporting the U.S. maritime sector.

Fluctuating and increasing prices of traditional maritime fuels have generated strong interest in alternative maritime fuels that are low in sulfur and may help the shipping industry meet target goals. To support estimation of the potential demand for maritime fuels, including the potential use of domestically produced, alternative fuels at U.S. ports, a research team at the Volpe Center developed MARINE-T to estimate maritime vessel fuel burn and associated metrics. This project was sponsored by the U.S. DOE Bioenergy Technologies Office (BETO), in collaboration with U.S. DOT MARAD, and in consultation with the National Oceanic and Atmospheric Administration (NOAA). From the initial prototype, MARINE-T has added features such as energy use mapping, a noise module, and scenario analysis features. The models and tools will support analyses of the potential for domestically produced alternative fuels to meet VLSFO fuel demand for U.S.-based marine vessel traffic; provide a unified platform to assess fuel consumption by vessel type and operating mode; and conducting scenario analyses for vessel operations, technology improvements, and fueling options.

Maritime fuel use is commonly tracked through fuel sales and required vessel reporting to the IMO Data Collection System (IMO DCS) for vessels over 5,000 gross tons (GT) (6). In contrast, MARINE-T uses a bottom-up approach based on automatic identification system (AIS) data and vessel characteristics. This type of model has proven successful in the past, pioneered by the Ship Traffic Emissions Assessment Model (STEAM) (7). STEAM demonstrated that sufficiently complete records of ship location, heading, and speeds from AIS data, coupled with sufficient details on the installed power and design speed of vessels, can be used to estimate both the fuel burn and resulting emissions outputs with enough accuracy for regional scale estimates (e.g., modeled versus observed fuel consumption was within 6% accuracy in 2009). This type of model demonstrated as well that these estimates could be generated at a more granular spatiotemporal resolution (e.g., 100 x 100 meter grid (10)) than previous methods which relied on simplified information and were based on aggregate data. STEAM has been extended to include multi-engine setups, incorporate wave action, and consider additional output metrics (9) as well as provide estimates at a global scale (8). STEAM has been extended as well to estimate underwater radiated noise source levels (10). Similar efforts have used real-time AIS data to estimate emissions at a port level (11) and to estimate carbon intensity for international shipping (12). MARINE-T provides U.S. federal partners, namely DOT, DOE, and NOAA, with a fuel burn, total energy use, and underwater radiated noise model based on data available within U.S. DOT. The tool uses open-source data science tools and best available science, enhancing U.S. government access to important analytical tools and data. In addition, MARINE-T provides the opportunity for scenario analyses to evaluate the impacts of different fueling scenarios, technology innovations, and operational strategies to enhance efficient and economically competitive maritime operations in the waters of the U.S.

This report describes MARINE-T, including available AIS and vessel characteristics data, data preparation, the model process, assumptions, and applications. This report also includes a summary of validation results for the fuel burn model using 2019, 2021, and 2022 annual fuel consumption data reported to the IMO DCS. The validation results demonstrate that the model produces outputs within 6% of reported fuel consumption for global movements of large vessels.

2. Scope

The goals of MARINE-T are to:

1. Provide metrics on fuel burn, total energy use, and underwater noise source levels for maritime vessels in the waters of the U.S., using consistent input data and processing steps, and outputting at consistent spatial and temporal scale.
2. Provide a dashboard with visual and tabular summaries of fuel use, energy use, and noise across the U.S. EEZ by vessel type, geographic area, and time period.
3. Develop scenario analysis options for decision-makers to evaluate the potential impact of energy efficiency and noise technologies and operational interventions.

The current phase of the project produces outputs using an inventory approach to fuel burn, total energy use, and underwater noise source levels. For the noise module in particular, the model in this current phase produces source level metrics, with propagation to be considered for addition at a later phase. At the highest level, MARINE-T provides an estimation of total fuel burn and noise source levels for the reference year for the U.S. EEZ, and has capacity to provide such outputs at a global level. Additional outputs over other time periods will be available in future versions of the model.

The primary visual representations of model results are gridded outputs showing locations with the greatest quantity of fuel burn and underwater noise source levels. These heat maps can enable decision-makers to evaluate patterns in operations based on the locations, vessel types, and time periods with different amounts of fuel burn, energy use, and noise source levels. Future features will show individual vessel traces to help decision-makers view the output metrics for a specific trip. In the first phase of this work, the Volpe team has used MARINE-T to conduct scenario analyses for selected regions of the U.S. to assess consequences of operational changes, technology innovation, and fueling scenarios. These scenario analyses can be used to facilitate innovation and efficient movement in the maritime transportation sector, showing potential cost savings from different operational, technological, and fueling strategies. The Volpe team also developed a dashboard for federal partners that includes base year data and scenario analyses.

3. Reference Data Sources

MARINE-T references several existing data sources, collated and stored in databases, to process marine vessel movements into fuel burn and total energy use estimates. Vessel movements (AIS data) are provided by the Volpe Maritime Safety and Security Information System (MSSIS) SeaVision team, which compiles land-based and satellite AIS data across the globe (13). Vessel characteristics for vessels with IMO numbers are drawn from the Standard and Poor's (S&P) Global Information Handling Services (IHS) database.

MARINE-T analyses encompass a broader scope of data coverage and model outputs than prior federal analyses. Important features of MARINE-T that distinguish it from other related federal work include:

- MARINE-T covers all vessel classes, including recreational vessels in addition to all other Class I, II, and III vessels.
- The geographic scope of MARINE-T includes the Great Lakes and U.S. inland waterways, as well as the complete U.S. EEZ (including territorial waters).
- The input AIS data include satellite as well as shore-based AIS data, compiled by the Volpe MSSIS team.
- MARINE-T incorporates ports of call data for AIS-based records of vessels entering and exiting global ports based on the World Port Index (14), compiled by the Volpe MSSIS team.

4. MARINE-T Methods

MARINE-T uses a bottom-up, AIS-based approach to estimating fuel burn, energy use, and underwater radiated noise from vessel movements and vessel characteristics. This approach was first pioneered for fuel burn and air pollutants in the STEAM model (7). MARINE-T builds on an initial fuel burn model developed by Volpe in the Port of Seattle area. The Volpe team validated the model with global, full-year data reported to the IMO DCS from selected U.S. flag vessels. MARINE-T uses AIS data compiled and processed by Volpe's MSSIS team, combined with vessel characteristic data from registrations and other U.S. vessel databases, and federal data resources. The MSSIS data coverage is global and near-real time. MARINE-T currently uses archival data for full-year analysis of user-defined study areas, from individual ports to the full waters of the U.S. EEZ.

The core pieces of the model are:

1. Vessel Characteristics
2. Vessel Movements
3. Ship-specific Power Propulsion Model
4. Fuel Burn and Emissions Calculation
5. Underwater Radiated Noise (URN) Source Level Calculation
6. Spatiotemporal Aggregation and Reporting
7. Trip Inference and Port Movements

Figure 4-1 shows a high-level schematic of MARINE-T, including both fuel burn and emissions modules, as well as the noise module. The detailed methods for the noise module will be documented in a separate report.

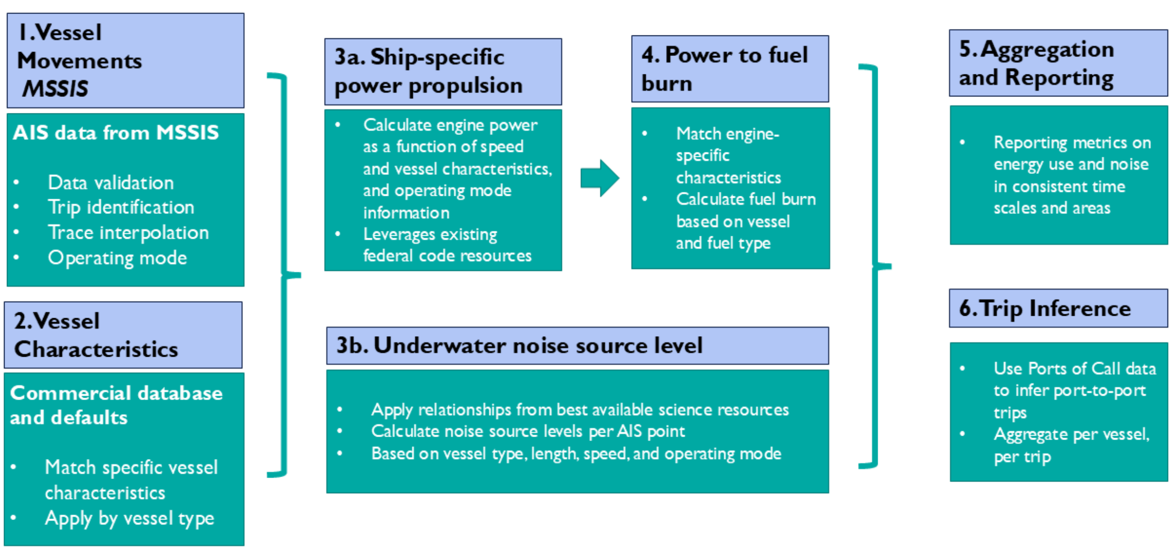

Figure 4-1. Schematic of MARINE-T, the Volpe Maritime Analytics for Research and Innovation on Noise and Energy – Tool. This report focuses on the fuel burn components of the model; the underwater noise module methods will be presented in detail in a separate report.

4.1 Vessel Types

Vessel type identification is critical for ensuring vessel characteristic data are assigned properly for emissions and noise calculations, as well as accuracy in reporting. The vessel types used by MARINE-T are compiled across all vessels operating in the U.S. EEZ, including workboats, ocean-going and coastwise vessels, and those vessels operating in the Great Lakes and inland waterways.

Due to the range of source databases used, several vessel type classifications are maintained and cross-walked within MARINE-T, including the United States Coast Guard (USCG) NavCen vessel types (used by MSSIS), the S&P Global StatCode 5 ship-type coding system (15), ship subtypes based on class and capacity in the Third IMO Greenhouse Gas Study (16) for auxiliary engine and boiler loads, the JOMOPANS-ECHO model vessel types for noise (17), and the Environmental Protection Agency (EPA) vessel types used for reporting (18). The AIS data source used by MARINE-T contains USCG NavCen vessel types for the vast majority of vessels; this attribute along with vessel characteristic data on length and beam are used to infer a StatCode 5 ship-type for every vessel, which is then used to associate all vessel characteristic data needed for the fuel and noise computation methods.

4.2 Vessel Characteristics

Vessel characteristic data are a crucial input to make accurate determinations of fuel usage, as these measurements are highly dependent on information like engine specifications, fuel type, and load factor. These data may come from a variety of sources, both publicly available and proprietary. A key factor in the performance of a model lies in the accuracy and specificity of the data it is built on; the Volpe team conducted a wide review of available sources for vessel characteristic data, resulting in the databases and lookup tables detailed below.

4.2.1 Vessel Characteristic Database Preparation

4.2.1.1 Vessel Identification

IMO types 1 and type 2 vessels are directly matched to the S&P Global IHS database via an IMO number when available. In the absence of an IMO number, vessels are matched by the Maritime Mobile Service Identity (MMSI) of the AIS transponder. The primary key in MARINE-T is the vessel MMSI.

The commercial vessel characteristics database was used to generate default vessel characteristic data for each StatCode 5 ship-type. Using vessels found in the 2021 commercial vessel characteristics database, median values for service speed, twenty-foot equivalent units (TEU), deadweight, gross tonnage, depth, total kilowatts of main engines, and maximum speed were calculated for each NavCen vessel type and StatCode 5 ship-type combination and stored as a static lookup table.

Whenever possible, vessel-specific information – including the vessel type category – is determined by matching AIS data to vessel characteristic data by MMSI or IMO number. When this matching process is unsuccessful, a basic crosswalk of NavCen vessel type to StatCode 5 ship type is used to infer vessel type and assign default vessel characteristic data for the vessel. Based on exploratory data analysis on the intersection of vessels where both NavCen vessel type and StatCode 5 ship type were known, the following mappings were chosen:

- Case 1: A single StatCode 5 ship type is assigned for a NavCen vessel type. In this case, the vast majority (e.g., >80%) of the exploratory dataset falls within the single ship type. This is applied to:
    - 25-OSV
    - 30-Fishing
    - 31-Towing
    - 32-BigTow
    - 33-Dredging
    - 34-Diving
    - 35-MilOps
    - 36-Sailing
    - 37-Pleasure
    - 50-Pilot
    - 51-SAR
    - 52-Tug
    - 54-AntiPol
    - 55-Law
- Case 2: For certain NavCen vessel types with greater divergence of vessel characteristics within the category, classification trees were built based on vessel length and beam to map to a small set of three to four StatCode 5 ship types. This is applied to:
    - 4-HSC – mapped to "Passenger Ship", "Crew/Supply Vessel", or "Passenger/Ro-Ro Ship (Vehicles)"
    - 6-Passenger – mapped to "Passenger/Ro-Ro Ship (Vehicles)", "Passenger Ship", "Passenger/Cruise", or "Crew/Supply Vessel"
    - 7-Cargo – mapped to "Bulk Carrier", "General Cargo Ship", or "Container Ship (Fully Cellular)"
    - 8-Tanker – mapped to "Chemical/Products Tanker", "Products Tanker", "Crude Oil Tanker", or "LPG Tanker"
- Case 3: For the remainder of NavCen vessel types, where limited data were available, a StatCode 5 ship type of "Default" is assigned.
- Case 4: For vessels with no NavCen vessel type identified, a StatCode 5 ship type of "Unknown" is assigned.

Once a StatCode 5 ship type is assigned, all required vessel characteristics (whether vessel-specific, vessel type specific, or default) are able to be assigned. For reporting, these StatCode 5 ship types are categorized as one of the EPA reporting categories based on a one-to-one mapping.

4.2.1.2 Database Sources

When possible, individual vessels are matched by MMSI and IMO number to their specific vessel characteristics. When not possible (because a vessel is not required to have an IMO number), a series of lookup tables are used to fill in the most appropriate vessel characteristics. MARINE-T assigns vessel characteristics using the below data sources.

Information Handling Services (S&P Global IHS, 15). IHS is a broad, commercial source of data on global ships, with access provided through a purchased government seat in MSSIS. The fuel burn model uses IHS Fairplay data to match to specific vessels when possible, and also to aggregate a set of vessel characteristics defaults per vessel type when individual vessels are not possible to match directly. The resulting lookup table identifies the average total installed power, design speed, deadweight, depth, and gross tonnage per vessel type contained in the database. When possible, engine characteristics are applied to vessels with IMO numbers using exact matches to the commercial vessel characteristic database, down to the main engine make and model. When not possible, e.g., because the vessel is not of a class required to be assigned an IMO number, lookup values based on vessel class are applied.

EPA Port Emissions Inventory Guidance (20)**.** The EPA Port Emissions Inventory Guidance, updated in 2022, includes data tables with compiled vessel characteristic defaults, organized by ocean-going vessels (OGV), harbor craft, and other ship types. The fuel burn model incorporates these defaults for a wider coverage of vessel types than is available in IHS Fairplay. This secondary lookup table identifies the average total installed power and design speed for each ship type, which have been mapped to the MSSIS vessel types. Separately, the EPA tables are used to find the most common engine category (e.g., C3, C2) and engine tier (e.g., Tier 0, Tier 1) for each ship type.

The above data sources contain the following granular, vessel-level data preferred by MARINE-T:

- Engine builder, maker, and model
- Engine displacement per cylinder, stroke type, stroke length – measures of engine size, which are used to categorize engine type and tier
- Engine model year – taken to be the mean of all years in which a model was built; used to look up categorical data (era of manufacture)

MARINE-T will use the following categorical data from the above data sources when vessel-level data are not available:

- EPA engine type
- Engine model era of manufacture
- Specific fuel consumption in grams per kilowatt-hour (kWh).

4.2.2 Ship-level Data

The following vessel characteristic data fields are the most critical for MARINE-T:

- Vessel type – Different data sets have varying vessel types (e.g., fishing, cargo, tanker) that are mapped to each other based on definitions and a series of lookup tables and decision trees. MARINE-T primarily uses StateCode 5 ship type.
- Design speed (knots) – Represents the speed at which the vessel was designed to be operated. At a given data point, the relationship between calculated speed and design speed is crucial to the power calculation.

- Total installed main engine(s) power (kilowatts) – The main engine(s) of a vessel are used for ship propulsion, so the power usage is taken as a fraction of the total installed power of the main engine(s).
- Main engine manufacturer and model – a specific model of main engine is assigned to IMO types 1 and 2 vessels (provided that the engine's stroke type is known[1]). When available, these are used to determine engine category. Engine category (C1, C2, or C3) is defined by the displacement of the engine in terms of liters per cylinder (21).
- Total installed auxiliary engine power (kilowatts) – Maximum power output from an auxiliary engine on a ship; when available, this value will be used to cap the estimated power output for a given operating mode.
- Total installed boiler power (kilowatts) – Maximum power output from a boiler.

Several other vessel characteristic data fields are useful to have (i) as proxies or independent variables in an inference model for more critical characteristics like design speed or installed engine power, (ii) as characteristics to assign the appropriate vessel type, or (iii) for ship-level validation purposes:

- Length, beam, depth/draft – Vessel dimensions, particularly within a vessel type, can be used to infer other vessel characteristics.
- Deadweight, gross tonnage – Measures of vessel mass and load size, respectively.

4.2.3 Substitution Mapping

MARINE-T uses approaches similar to other published work for filling in missing characteristics. As noted in the EPA Port Emissions Inventory Guidance: "Missing data is a common occurrence in vessel characteristic data sets. The best practice for filling gaps in vessel characteristic data is to use information from vessels with known values operating in the geographical domain" (20). MARINE-T applies default / lookup table values when individual vessel matching is not possible, i.e., when only MMSI and USCG vessel type are available from the source AIS data, but no IMO number is available. This process is detailed in Section 4.1.

4.3 Vessel Movements Data

The core components in an AIS data set are position and speed data per vessel (with a unique identifier) and per timestamp. The AIS data used as input for MARINE-T are provided by SeaVision (13), which includes land-based and satellite AIS data across the globe. The SeaVision AIS data include the following fields:

1. MMSI code (vessel identifier)
2. Time of AIS data point (in Coordinated Universal Time (UTC))

[1] Stroke type is required by the EPA Toolkit for looking up engine category. IMO 4th GHG Study method appears to use revolutions per minute (RPM) only.

3. Latitude
4. Longitude
5. Speed over ground (SOG) in knots
6. Heading

In addition, some metadata are provided: source of AIS data point, port index, and EEZ, which has approximately the same coverage as the ECA along coasts, although is more extensive across all the waters of the U.S.

The input data set is validated at a vessel trajectory level; the AIS data points for a specific vessel are pulled from the database, sorted in chronological order, and separated into trips using the logic below for which fuel burn calculations are done. Distance as determined by the great circle distance and time between consecutive data points are calculated. Before trip identification, data validation is run to filter out any AIS data points that are flagged as potential outliers or incorrect data using the following decision: if the average speed traveled since the previous data point is larger than the maximum speed threshold (set at 1.5x the design speed of the vessel from the vessel characteristics table), the data point is marked invalid.

Once the AIS data have been validated for a vessel, trip starts are identified by their starting time, using the following decisions:

1. The first chronological data point in the database is assumed to be a trip start.
2. The first data point after a time interval larger than the gap threshold (set at 72 hours, same as used in the STEAM model (8)) is considered a trip start.
3. All other points are considered non-trip-start points. Further inference on port-to-port trips is applied as a post-hoc analysis after the full model run is complete, described in section 4.6.

At this point, interpolated points are added to a vessel's trajectory as needed to ensure adequate granularity of vessel movement data for the fuel burn computation. Distance as determined by the great circle distance and time between consecutive data points are re-calculated as needed.

4.3.1 Interpolation

For cases where the travel time between two consecutive AIS data points is greater than one hour (for non-trip-start points), intermediary points are linearly interpolated in MARINE-T until the maximum elapsed time is one hour. This process improves the spatial accuracy of fuel use and noise reporting.

First, the pair of consecutive AIS points are used to determine the great circle distance, or solve the inverse geodesic problem. Given two coordinates, the Haversine distance and azimuth (angle measured clockwise from north) are calculated:

$$a = \sin^2\left(\frac{\Delta\varphi}{2}\right) + \cos\varphi_1 \cdot \cos\varphi_2 \cdot \sin^2(\frac{\Delta\lambda}{2})$$

$$c = 2 \cdot atan2(\sqrt{a}, \sqrt{(1-a)})$$

$$Haversine\ distance = d = R \cdot c$$

$$Azimuth = \theta = atan2(sin\Delta\lambda \cdot cos\varphi_2, cos\varphi_1 \cdot sin\varphi_2 - sin\varphi_1 \cdot cos\varphi_2 \cdot cos\Delta\lambda)$$

Where a is the square of half the chord length between points, c is the angular distance (in radians), φ_x is the latitude at point x, $\Delta\lambda$ is the difference in longitude between points, and R is the Earth's radius, where *atan2* is the 2-argument arctangent.

Second, using the calculated Haversine distance and elapsed time, a speed is calculated. Using this speed, the distance between the first point and each consecutive interpolated point is calculated. This is performed until the max elapsed time between points is one hour:

$$Calculated\ speed = \frac{Haversine\ distance}{Elapsed\ time}$$

$$\#\ of\ points = \lfloor Elapsed\ time/3600s \rfloor$$

$$Interpolated\ distance = Calculated\ speed * 3600s$$

Third, the direct geodesic problem is solved. Given the coordinates of a point, the angle of travel, and distance, coordinates of interpolated points are calculated:

$$\varphi_{interpolated} = \text{asin}\,(sin\varphi_1 \cdot cos\delta + cos\varphi_1 \cdot sin\delta \cdot cos\theta)$$

$$\lambda_{interpolated} = \lambda_1 + atan2(sin\theta \cdot sin\delta \cdot cos\varphi_1, cos\delta - sin\varphi_1 \cdot sin\varphi_{interpolated})$$

Where δ is the angular distance, defined as the Haversine distance divided by the Earth's radius.

The geodesic problems are solved using functions within the *geosphere* R package (22); the above equations are approximations of the actual methods used. Notably, all AIS points are kept throughout this step; the above process only adds points to ensure the max travel time for non-trip-start points is one hour.

4.3.1.1 Grid Interpolation

The MARINE-T model spatially aggregates fuel burn and noise metrics to a grid cell level. For example, when applied to the U.S. EEZ, the analysis uses a one kilometer by one kilometer grid. This gridded aggregation is necessary for analysis of fuel burn and noise metrics in a consistent fashion, and is necessary for the noise metrics as the spatial component of noise source levels is critical.

Grid interpolation is used for cases where two consecutive AIS points are within different grid cells. The input data represent a vessel's movement along a path with consecutive AIS points. If spatial aggregation is only based on the metrics at the point level, values can be incorrectly attributed to the wrong grid cells, or even miss grid cells that a vessel traveled through. Figure 4-2 shows examples of potential misattributions when grid interpolation is not implemented.

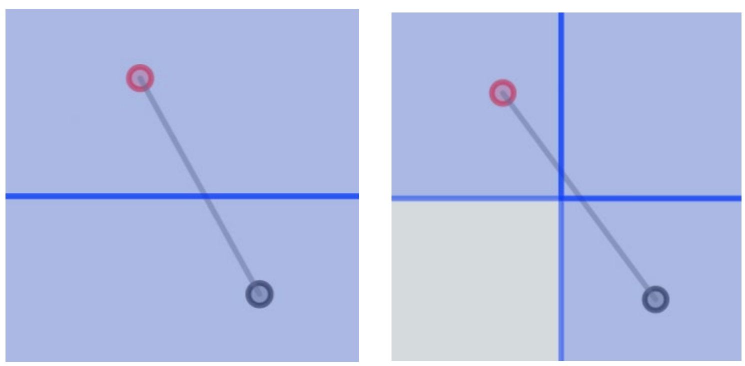

Figure 4-2 Left: a standard example of potential misattribution, where 100% of the metrics are attributed to the top cell. Right: an example where a grid cell is not attributed any metrics.

The grid interpolation method is used to account for this potential misattribution. All AIS points are spatially joined with the reporting grid, such that each AIS point is associated with a grid cell ID. In all instances in the vessel's trip trajectory where the grid cell ID of the current AIS point (black in the above example) is not equal to the previous AIS point (red in the above example), straight lines are created to connect the current points to the previous points. Along this assumed path of the vessel, one or more points are interpolated in each intersection with a grid cell boundary, as shown in Figure 4-3.

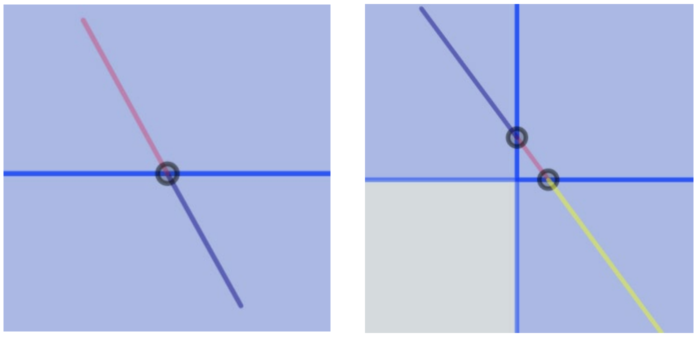

Figure 4-3 Left: One point is interpolated between the original points. Right: Two points are interpolated between the original points, creating a segment that now properly aggregates a portion of the metrics to a cell that was previously not included.

Traveled time between the original point and the interpolated points is then calculated:

$$t_{interpolated} = \Delta t * \frac{d_{interpolated}}{d_{total}}$$

Where Δt is the original traveled time between current and previous point, $d_{\text{interpolated}}$ is the distance traveled between original and interpolated point, and d_{total} is the original distance traveled between current and previous point.

Timestamps are then generated for each interpolated point:

$$T_{interpolated} = T_{previous} + t_{interpolated}$$

The interpolated points are flagged as interpolated and combined with the original points table. The final step is readjusting all distances and traveled times in the points table to account for the newly created interpolated points. This table is validated once again, using the method described above. This method is also used for adding points when a vessel enters or exits the U.S. EEZ.

4.3.1.2 Inland Waterway Alignment

The final step of the interpolation process involves adjusting interpolated points within inland waterways. The linear grid interpolation method works without issue for vessels operating on the open ocean, but fails to account for the complexity of river navigation. When interpolation is required between points along a bend or curve in a river, the linear method may place points on land. To handle this oversimplification within rivers, a snapping method is used to adjust misplaced points. Interpolated points are spatially joined with land and river polygons to determine which points need adjustment. Correctly placed points fall outside of land boundaries (ocean-going) or within the boundaries of a river (inland). Misplaced points are those within land boundaries, outside the boundaries of an inland waterway, and within a certain distance of a river. The centerline of the rivers are granulated into points, and the interpolated points inherit the coordinates of the closest river centerline point, identified using a nearest neighbor search algorithm.

4.3.2 Operating Mode Assignment

Operating modes capture the state of a vessel during a trip, and are critical for both the fuel burn and noise components of MARINE-T. These operational modes include: at berth, at anchorage near ports, at anchorage off-shore, maneuvering, and transit. Additional special operational modes include: dredgers dredging, tug boats towing near ports, and tow boats towing near inland waterways. Operating modes depend on a vessel's geographic location, proximity to port, and speed.

The current MARINE-T operational mode assignment logic relies on three spatial inputs: the World Port Index (WPI) from the National Geospatial-Intelligence Agency (NGA) (14), USA Detailed Water Bodies database from the Environmental Systems Research Institute, Inc. (ESRI) (23), and the global coast and lake geography database from the NOAA National Centers for Environmental Information (24). WPI is a database of worldwide maritime port information which serves as a general reference and planning tool for mariners. The detailed water bodies database is a shapefile with polygons that represent major water features (rivers, lakes, reservoirs etc.) in the United States. The Global Self-consistent, Hierarchical, High-resolution Geography Database (GSHHG) (24) from NOAA provides shapefiles containing polygons representing many categories of shorelines in varying spatial resolutions. MARINE-T uses the boundaries between land and ocean (L1) in high resolution (4 of 5 levels) and the boundaries between lake and land (L2) in full resolution (5 of 5).

There are two distinct decision trees that are used for assigning operating modes: inland and ocean-going/coastal/Great Lakes travel. AIS points are spatially joined with the detailed water bodies layer, creating a flag for points that are within a river. Points within rivers do not require additional spatial analysis, so the operating mode can be assigned based on SOG (and vessel class), using the decision tree from Figure 4-5. The inland waterways decision tree matches methods proposed by the International Council on Clean Transportation (25).

If a point is not within a river, the alternate decision tree is used. These points are matched with the nearest port from WPI, utilizing a nearest neighbor function. The Haversine distance between points and the nearest port is then calculated, representing the great circle distance between two points. Points are then determined to be within the boundaries of the L1 (ocean) and L2 (lake) shapefiles from GSHHG. If an AIS point is within both L1 and L2, the nearest neighbor is found from L2 (lake). If an AIS point is within neither L1 nor L2, the nearest neighbor is found from L1 (ocean). The nearest neighbor function requires both input datasets to be points, so the L1 and L2 shapefiles were cast from polygons to points – generating points at vertices. Again, the Haversine distance is calculated from AIS points to the nearest coast vertex. Using these spatial checks and SOG, the operating mode can be determined using the decision tree from Figure 4-4. The ocean-going decision tree is based closely on methods from the IMO (12).

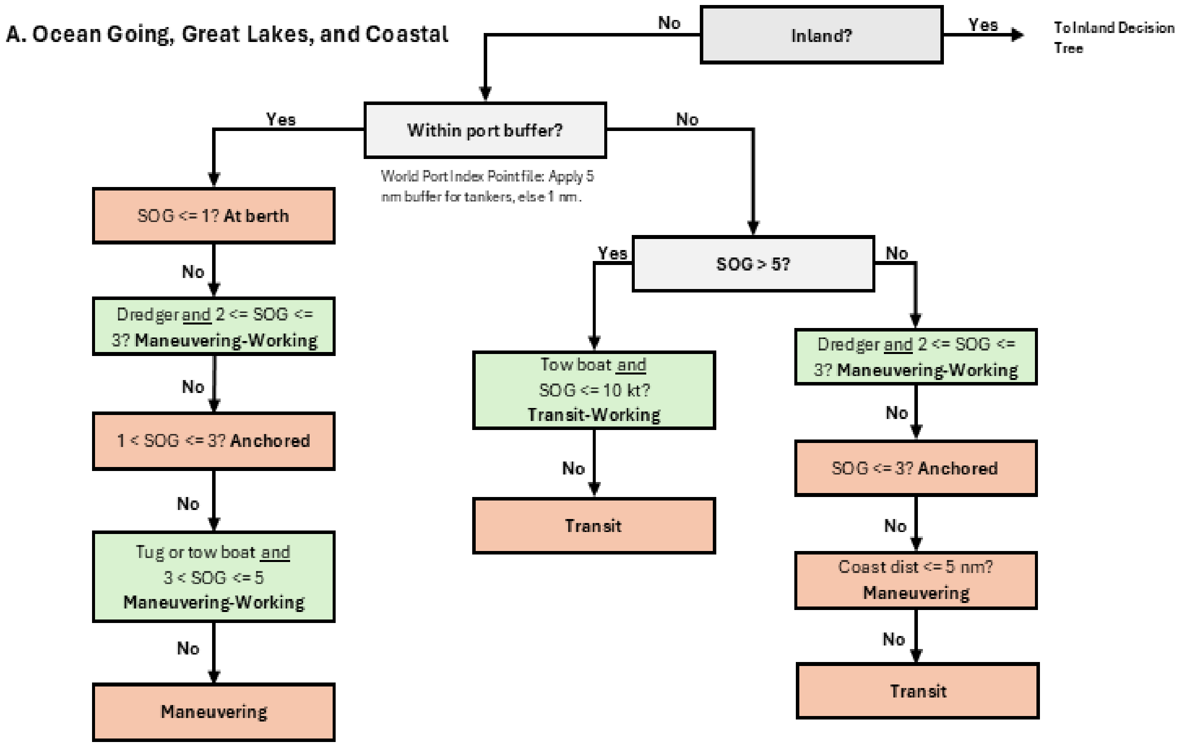

Figure 4-4 Operating mode decision tree. The left half of the decision tree is for vessels operating in open ocean, coastwise, in Great Lakes, or at port.

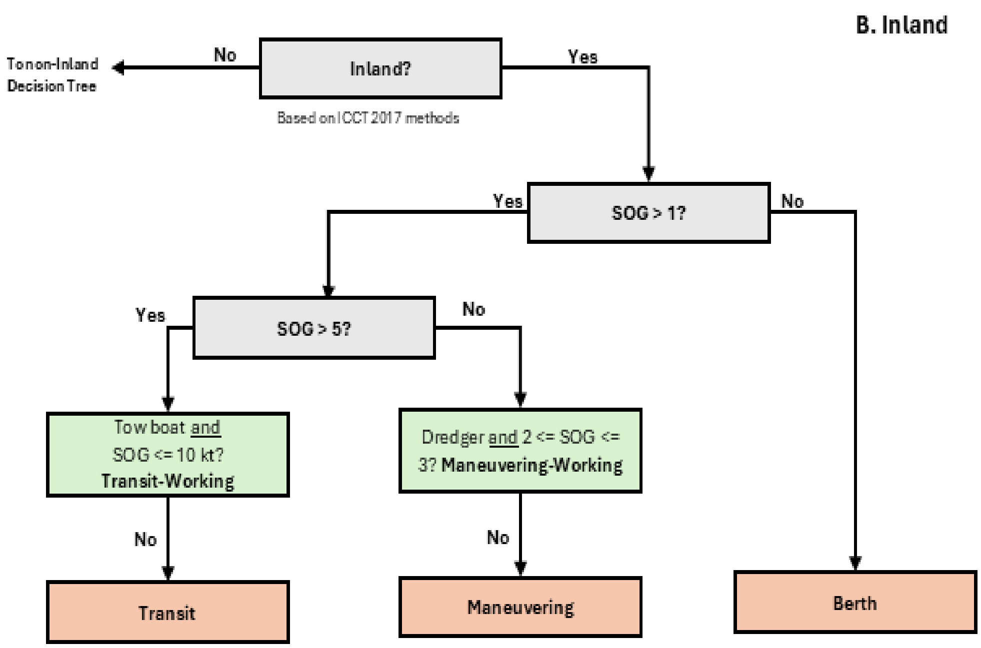

Figure 4-5 Operating mode decision tree. The right half of the decision tree is for vessels operating in inland waterways.

4.4 Fuel Burn Computation Methods

Once vessel characteristics and vessel movement data have been prepared and validated for all vessels in the analysis, these data are passed to the fuel burn computation module, represented in Steps 3a and 4 in Figure 4-1. Based on specific vessel and engine characteristics, power and subsequently fuel burn are calculated for each vessel at every AIS point in the vessel's trajectory.

4.4.1 Vessel Characteristics for Fuel Burn

MARINE-T uses the following vessel characteristics for fuel burn and emissions calculations:

- Vessel type
- Vessel engine class (C1, C2, C3)
- Design speed
- Number of main engines
- Installed power of each main engine
- Number of auxiliary engines
- Installed power of each auxiliary engine
- Number of boilers
- Main fuel type

These values are pulled from the vessel characteristics database.

4.4.2 Speed to Power Calculation

Given speed values from the AIS data, there are established relationships within the literature (7,20) to translate speed to the required engine power needed to move a vessel at that speed. MARINE-T currently uses the Propeller Law formulation, with an adjustment to account for the sea margin. Use of the Admiralty Formula, which can account for the realized draft (depth of the vessel in the water) compared to the vessel design draft, has been used in studies such as the Fourth IMO Greenhouse Gas Study (12). However, this formulation assumes draft data are updated frequently and accurately; the Volpe team has assessed that these assumptions are not likely to be met for the MARINE-T data source, since the vessel draft fields in AIS data rely on manual entry. Further assessment of the Admiralty Formula and draft data quality is ongoing.

Based on the SOG data provided by AIS and the vessel characteristics from the commercial database and lookup tables, transient power for vessel movements is calculated from transient speed (SOG) using the following formula:

$$Power_{transient} = Power_{installed} \times \left(\frac{Speed_{transient}}{Speed_{design}}\right)^3$$

This formulation is based on the cubed ratio of the transient speed of the vessel (SOG) compared to the speed at which the vessel is designed to operate optimally, times the total installed power of the main engine or engines. A sea margin is added to account for resistance of 10% for coastal movements and 15% for at-sea conditions, following established research.

Once power (in kilowatts) is calculated at each AIS data point, interpolation between consecutive data points is used to calculate a total power usage (in kilowatt-hours) across a vessel's journey. Finally, power is assumed to be zero between identified trips (e.g., power used since the previous AIS data point is set to zero for each identified trip start).

In further work, this model may be expanded by considering refinements to both transient speed (e.g., by incorporating a wave effect model or effect of frictional resistance on speed) and design speed.

4.4.3 Auxiliary Engine and Boiler Power Lookup

Auxiliary engine and boiler loads are also considered, accounting for distinct operating modes. Auxiliary engine and boiler power outputs are especially significant near port areas as auxiliary engines assist in maneuvering and generate electricity while at anchor or at berth, in the absence of shore power connections. In addition, boilers support steam-powered pumps which have significant energy needs when offloading tanker ships.

Auxiliary engine and boiler power lookups use EPA's Ship Power Model and built-in data tables, which assign a kilowatt load value based on subtypes from the Third IMO Greenhouse Gas Study. Subtypes are defined by IMO ship class and capacity; see Tables 8 and 9 in the Third IMO study for associated loads (16). MARINE-T maps the NavCen vessel types to an IMO subtype considering vessel deadweight tonnage, gross tonnage, or TEU capacity, as applicable, and attributes the associated kilowatt load based on operating mode. In cases where NavCen vessels do not uniquely map to IMO subtypes, MARINE-T uses a modified subtype and load that averages over the corresponding IMO subtype loads. Harbor craft and vessels with C1 and C2 engines are assumed not to have boilers.

In further work, this model may be improved by considering the following:

- Ensuring missing NavCen vessels are assigned default power outputs. Currently, missing vessel type is not handled in the auxiliary and boiler lookups, resulting in zero power output.
- Capping vessel-specific auxiliary engine or boiler loads at the total installed auxiliary engine power from the S&P Global IHS database.

4.4.4 Power to Fuel Burn Calculation

The primary data requirements to translate main engine power to fuel burn are engine specifications (e.g., C1/C2/C3 engine, engine tier), fuel type used, and the specific fuel oil consumption (SFOC) rate for that particular engine and fuel type. Different models of fuel burn can also be constructed based on ship activity. In MARINE-T, several assumptions are made in this step:

1. Fuel switching is calculated to occur exactly at boundary of the ECA.
2. Smaller ships with single engines operate entirely within the ECA using low-sulfur fuels.

SFOC is based on the make and model of the installed main engine for each ship, if vessel-specific data are available. Otherwise, lookup values using the EPA methodology for SFOC are used. As a final backstop, a value of 185 grams per kWh is used. For boilers and auxiliary engines, fuel burn is calculated with the default SFOC of 185 grams per kWh.

4.5 Aggregation and Reporting

Once outputs from fuel burn calculations are provided at both the AIS data point level and between consecutive data points, the fuel burn model can aggregate these values based on (1) vessel type, (2) timeframe of interest, (3) geographical region of interest. For instance, the model can provide values of total fuel burn by vessel type for each week in 2019 for the Pacific Northwest ECA region, total fuel consumption by ocean-going vessels for a given day in the Northeast, or fuel consumption within the Great Lakes for all of 2022.

MARINE-T outputs include cumulative fuel burn, energy use, and noise source levels for the reference year across the U.S. EEZ. For visualization, outputs of fuel burn, energy use, and noise source levels are aggregated to a spatial grid. The default spatial grid has a one kilometer by one kilometer resolution, and is based on the U.S. EEZ, including Great Lakes and inland waterways.

4.5.1 Fuel and Energy Metrics

A primary goal of MARINE-T is to be able to analyze fuel burn, energy use, and noise source levels for a region, for a time period, and by vessel types. An example from Port of Seattle area fuel burn is shown in Figure 4-6, for cargo vessels in 2022. The full MARINE-T model is used to produce aggregations of national, full year estimates of metric tons (MT) of fuel consumption of maritime vessel activity in the U.S. EEZ, as well as summaries by region, vessel type, season or month of year, and flag carrier. With the noise module, MARINE-T also produces reports of underwater radiated noise (URN) source levels at the vessel type level, by region, season or month of year, and flag.

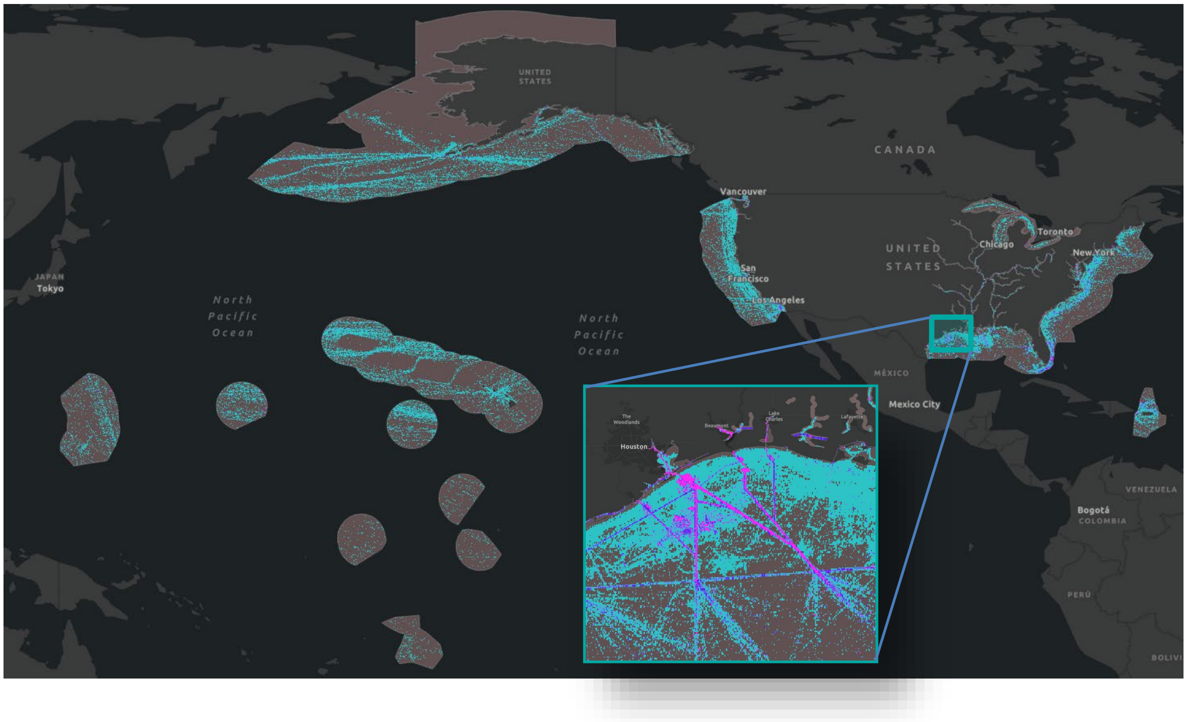

Figure 4-6 View of MARINE-T U.S. EEZ total fuel consumption output for the year 2022. Each pixel here is approximately 1 square kilometer.

Aggregation to these levels allows visualization of differences from base year when conducting scenario analyses. In the first phase of work on MARINE-T, the Volpe team completed scenarios in three regions, assessing different fuels including methanol and biofuels, operational changes such as speed modifications, and technology changes such as noise management technologies. Additional metrics include vessel counts in an area, vessel types active in an area, and total hours and distance traveled by vessel, per vessel type, in an area and time selection. Example outputs from one regional scenario shown in Figure 4-7.

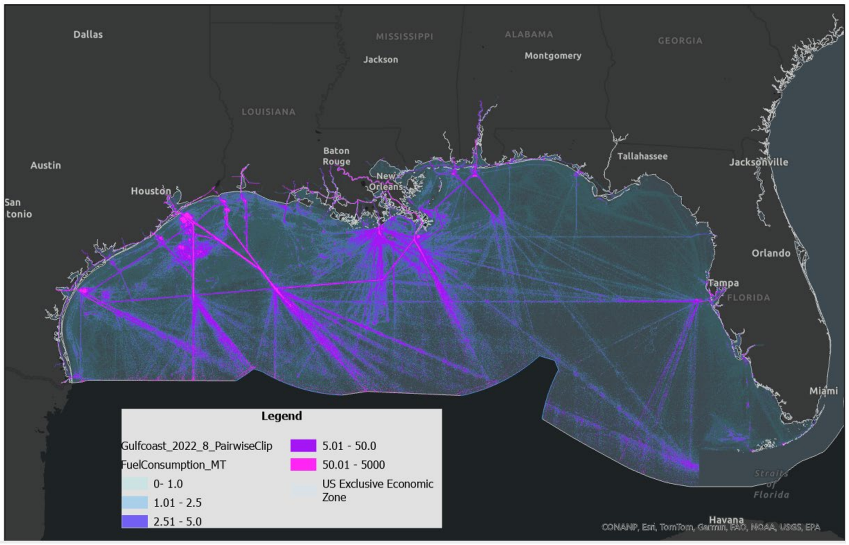

Figure 4-7. View of Gulf Coast regional scenario outputs, focused on technologies and fuels for offshore activities for harbor craft and offshore supply vessels.

4.5.2 Trip Inference and Port Movements

After metrics from vessel movements are calculated, MARINE-T applies a trip inference and port-to-port movement analysis. This step allows detailed assessment of trips to specific ports and between regions. This section describes the data sources and methods used to undertake the trip inference, how these are aggregated to origin and destination port, and the output dashboard and port profile reports.

4.5.2.1 Trip Inference Data: SeaVision Ports of Call Table

In SeaVision, the Ports of Call (PoC) table covers the period 2016–2022 at a global scale and contains about 150 million rows, with four fields: MMSI, EntryTime, ExitTime, and PortIndex. The table is derived from global AIS data and captures when a vessel is observed entering and exiting a port in the World Port Index. Vessels are considered to have entered a port when it is within 10 km of a port and has SOG under 0.3 knots. Vessels are considered to have exited a port when any of those conditions are no longer true. In cases where multiple ports are within 10 km, the assignment favors the largest port. If a record shows the same entry and exit time, only one observation satisfied the call conditions. If the PortIndex value is 0, vessels are not calling at a port.

Vessel movements are from the SeaVision AIS data (combining satellite and shore-based AIS data) and commercial vessel characteristics data, as described above.

4.5.2.2 Trip Inference Method

In MARINE-T, the PoC table is joined to model outputs using the following steps:

1. Extract port calls for a given month (including a day before and after as a buffer).

2. Filter out any observations where PortIndex = 0.
3. Generate a unique CallNumber after grouping by MMSI and arranging by EntryTime. This is used to distinguish between consecutive calls made to the same port.
4. Perform a full join between MARINE-T outputs and PoC data using the AIS TimeOfFix field. The TimeOfFix value of the model outputs must be between the PoC EntryTime and ExitTime. Port calls made outside of the U.S. EEZ are included due to the full join (rather than a left join). These cases do not have associated AIS data in the model outputs, but TimeOfFix is set to the port call EntryTime to keep calls in chronological order.
5. For every modeled AIS ping, assign PortIndex and CallNumber for observations that are not contained by a port call by filling from the previous non-null observation.
6. For each MMSI, trip IDs are assigned by incrementing when the CallNumber changes. Trip IDs are tracked across months so a ship that had 10 trips in January would start at trip ID 11 in the next month where data is available.

Trip assignments are stored as part of MARINE-T outputs.

4.5.2.3 Summarizing Trip Statistics

The result from the trip inference step is a table which contains all AIS records plus empty AIS records representing port calls made outside the U.S. EEZ. MARINE-T uses this table to generate a trip summary table by grouping on the MMSI and Trip ID fields. A list of the calculated fields is below:

- Destination Port – the port where the trip ends (can be the same as the origin)
- Sum Fuel – total fuel burned (in MT) during the trip
- Travel Time – total duration of trip (in seconds)
- Distance Traveled – total distance traveled (in m) during trip

These aggregated values only summarize data that happened within the U.S. EEZ, because the AIS source data used by MARINE-T are only available for vessel movements within the EEZ. All portions of a trip that occur outside the U.S. EEZ cannot be calculated.

Flags are generated to categorize the trips:

- SamePort – are the origins and destinations the same?
- US_to_US – are the origins and destinations both in the U.S.?
- US_to_nonUS – did the trip start in the U.S. and end outside the U.S.?
- nonUS_to_US – did the trip start outside the U.S. and end in the U.S.?
- nonUS_to_nonUS – did the trip start and end outside the U.S.?

The trip inference module generates outputs for MARINE-T, with a focus on fuel consumption, distance travelled, and time underway. These are aggregated to origin and destination ports. MARINE-T outputs from the full year U.S. EEZ model run were used to develop the MARINE-T Trip Inference and Port Movements dashboard (Figure 4-8).

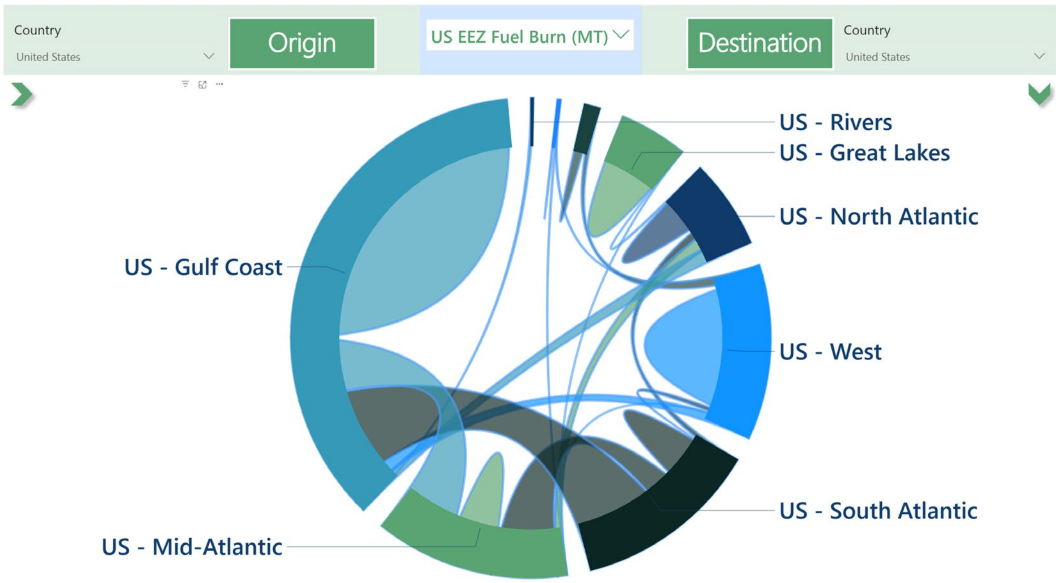

Figure 4-8. MARINE-T Trip Inference and Port Movement dashboard image. Chords between regions represent the metric tons of marine diesel oil equivalent fuel consumed on trips in the U.S. EEZ in 2022. Arcs show total fuel for port-to-port movements within a region.

This dashboard has multiple filters, including by month of year, vessel type, region, and specific port of origin or destination. The metric being visualized can toggle from metric tons of marine diesel oil fuel equivalent to the number of trips.

In addition to the dashboard, the aggregated data on trips and port movements can be visualized by additional metrics, for instance showing the dwell time per vessel per port, and dwell times in port by vessel type.

For the full year U.S. EEZ model run, an origin-destination (O-D) matrix was derived from the outputs of the trip inference module. Using the Origin Port and Destination Port fields, the total number of trips between each pair of ports was computed. Among 2,514 unique ports that have at least one record in the trips tables, there are 6,320,186 possible O-D pairs. However, 98.9% of these pairs have no trips between them, indicating that most port pairs are inactive in the observed data. Of those 2,514 ports, 629 are located in the United States. The O-D matrix can be used to understand activity between specific port pairs. Table 4-1 below summarizes the number of trips that occurred between U.S. ports vs non-U.S. ports.

Table 4-1. Number of trips that occurred between U.S. ports versus non-U.S. ports.

Trip Type	Total Trips
US to US	14,274,926
US to non-US	119,299
non-US to US	91,253
non-US to non-US	2,281,125

5. Model Validation

MARINE-T is validated on reported fuel consumption data; reported metrics on other outputs are not available. The Volpe team used the IMO Global Integrated Shipping Information System (GISIS) Ship Fuel Oil Consumption database, as part of the Data Collection System (6). This database includes records for ships of 5,000 GT and above, with fuel oil consumption data for each type of fuel oil they use, as well as annual, global hours underway and miles traveled. Thus, for each vessel, there is a single annual total quantity of fuel consumed for each fuel type, representing the total fuel consumption for all activities of that vessel in that year. Member states can view the non-anonymized data for their flag vessels; this allows U.S. government agencies to view U.S. flag vessel data. Additional validation work is ongoing for non-US flag vessels, in collaboration with the American Bureau of Shipping (ABS).

Validation was performed for three years (2019, 2021, and 2022) on U.S. flag vessels. The number of vessels in each year ranged from 112 to 128. In each of the three validation years, for each vessel, the full year of global AIS data were run through MARINE-T. Only vessels which had the same date range available in both the IMO GISIS validation data and in the AIS data were kept; for some vessels, the date ranges in the AIS data were not compatible with the date ranges in the IMO GISIS database. After filtering by date and checking for other issues (see below), the count of vessels used for validation in the three years was 78 (2019), 102 (2021), and 87 (2022). The validation data represented over 19 million AIS records over the three years, 52 million nautical miles of travel, and over 2.8 million hours of operation.

Overall, MARINE-T generates estimates of fuel consumption within 5.81% of the reported values as seen in Table 5-1. Across all vessels, the root mean square error (RMSE) of fuel consumption was 726 MT of fuel, out of an average reported value of 12,513 MT. By vessel type, the closest match between reported and estimated fuel consumption is for containerships, tankers, and general cargo ships.

Table 5-1. Average model performance for fuel consumption validation by vessel type for 267 records in 2019, 2021, and 2022 combined, representing 156 unique vessels.

Vessel Type	Vessel Count	Average IMO Reported Fuel (MT)	Average MARINE-T Estimated Fuel (MT)	RMSE	Percent Error Fuel Estimation
Containership	101	20,225	19,129	1095.2	-5.41
Passenger ship	3	14,913	18,770	3857.4	25.87
Ro-ro cargo ship	60	9,905	8,441	1464.7	-14.79
Tanker	57	7,180	7,591	410.3	5.71
General cargo ship	16	6,100	5,781	318.7	-5.22
Bulk carrier	8	5,839	6,433	594.2	10.18
Others	22	4,807	3,437	1369.5	-28.49
Overall	**267**	**12,513**	**11,787**	**726.6**	**-5.81**

Looking at the reported and estimated fuel consumption data in Figure 5-1, the trend for underestimation of fuel consumption is driven in part by some outliers. In addition, larger discrepancies arise for containerships with high values for reported fuel consumption.

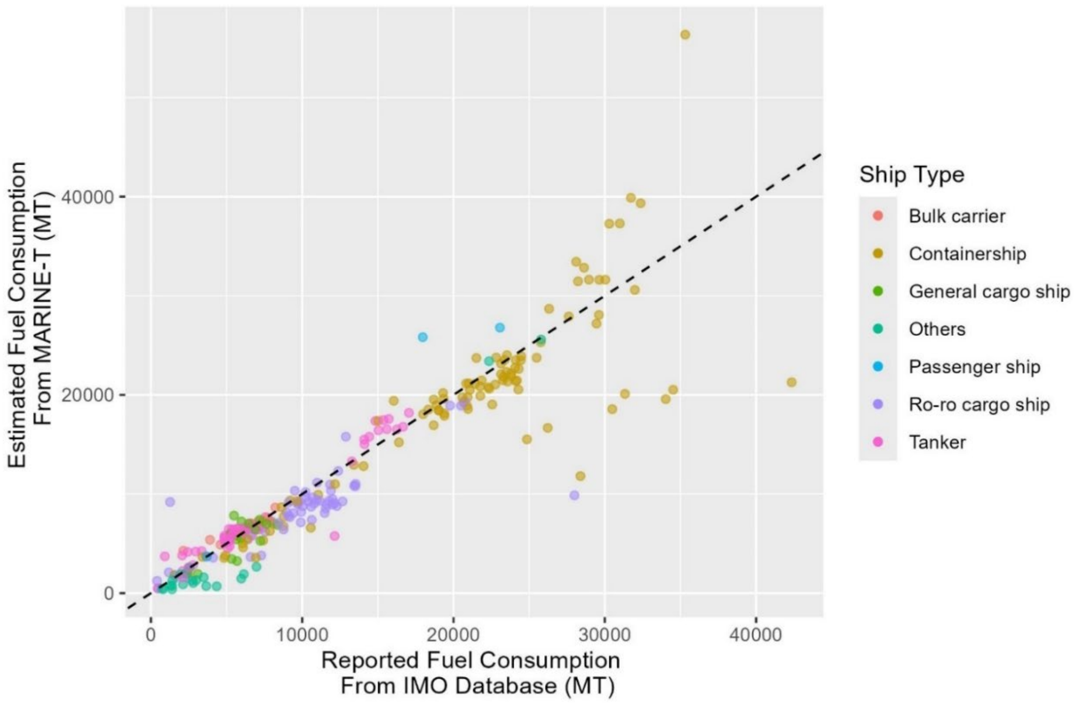

Figure 5-1. Estimated and reported fuel consumption for 267 records in 2019, 2021, and 2022 combined, representing 156 unique vessels. The dashed line indicates the ideal 1:1 relationship.

Two outliers were excluded, where the reported values in the IMO GISIS database were substantially higher than comparable vessels of similar type, installed power, and miles traveled. For example, the greatest reported fuel consumption in 2019 was for a tanker vessel with a reported consumption of 61,299 MT of MDO. However, this vessel only travelled a reported 64,005 nautical miles. The estimated distance from AIS data was quite close at 63,013 nautical miles; however, the estimated fuel consumption was far less, at 4,643 MT of MDO-equivalent fuel. This record in 2019 was excluded from the validation data, as was another record where fuel consumption was over five times greater than other vessels with similar installed power and miles traveled in a year. This demonstrates the need for closer scrutiny of the reported fuel consumption data.

While distance travelled is an important component of fuel consumption, fuel consumption is not expected to be linearly related to distance in the model. In the current version of the model, ship dimensions, installed main engine power, auxiliary engine and boiler power, and SFOC of engines all have a substantial influence on fuel consumption outputs. Factors such as wave action and in-port maneuvering assistance by tugs—which are not currently in the model—are expected to have substantial influence on fuel use.

Further, the received AIS data that MARINE-T is based on vary in completeness for individual ships. Many of the larger ships have long gaps between reported location, where several days of travel are not reported. Future work on interpolation improvements and operating mode assignment will continue to improve MARINE-T.

As an additional validation effort, the American Bureau of Shipping (ABS) conducted a comprehensive verification and validation of the MARINE-T model, confirming its high accuracy in estimating annual vessel fuel consumption against "ground truth" data from the IMO GISIS database. For the U.S. fleet validation cohort, the model's fuel consumption estimates were within -5.81% of reported values, with particularly robust performance for containerships, tankers, and general cargo vessels. Analysis of a broader non-U.S. fleet dataset of 2,947 vessels showed a nearly ideal 1:1 correlation for fuel consumption, though secondary metrics like distance and hours revealed systemic overestimations and underestimations, respectively. These discrepancies likely stem from unmodeled environmental factors such as wave action and limitations in current AIS interpolation techniques. Overall, the report validates the MARINE-T methodology as a sound platform for national-scale energy transition strategies while identifying specialized operational modules for niche vessels, like passenger ships, as a priority for future refinement.

6. Conclusion and Next Steps

Volpe has completed the first phase of development of MARINE-T to calculate spatiotemporally explicit fuel burn, energy use, and underwater radiated noise for individual vessels based on position data and vessel characteristics. MARINE-T has been run on complete global annual data for select sets of vessels for validation against complete-year, global fuel consumption data. The validation using global data for select vessels showed that the fuel burn estimates are reasonably accurate for many vessel types. The validation work identified opportunities for model improvement. Continued improvement will come from:

- Further enhancing the trip identification logic to enable more detailed modeling of ship activity in and near ports;
- Better mapping of the installed power and design speed of vessels, as well as vessel-specific information for the SFOC of installed main engines, especially for vessels of unknown types;
- Enhanced specificity and accuracy of vessel characteristics such as dimensions, load factors, engine size, cylinder displacement range, and model year to enhance power-to-fuel burn modeling.

In parallel, Volpe is continuing to investigate additional areas for improvement of model performance, including increasing the scope of validation with vessel-owner provided fuel consumption data. More broadly, Volpe is using MARINE-T to support analysis of future vessel fuel demand, changing operational strategies such as speed modification, and assessment of noise management technologies on vessels to assess opportunities for synergy between energy efficiency, noise management, and cost savings.
The power of MARINE-T lies not just in the ability to produce summary statistics on an annual or more frequent basis, but rather, to allow the user to implement 'what if' scenarios that change operational parameters, technology applications, and fuel use of specific vessels or entire fleets of vessels to understand the potential impact through MARINE-T results. Future reports will detail the scenarios run on the 2022 base year outputs of MARINE-T, where specific 'what-if' scenarios were run to investigate the energy efficiency and cost savings potential to maritime fleets in the U.S., using over 600 million AIS positions for 2022. These scenario analyses demonstrate how MARINE-T can be used to quantify and map fuel burn, energy use, and noise from maritime transportation in the U.S. to promote efficient and cost-effective movement of vessels in the U.S. maritime sector.

7. References

1. The White House [Internet]. 2025 [cited 2025 Dec 2]. Restoring America's Maritime Dominance. Available from: https://www.whitehouse.gov/presidential-actions/2025/04/restoring-americas-maritime-dominance/

2. US EPA. EPA Collaboration on International Air Pollution Standards for Ships [Internet]. 2016 [cited 2024 Nov 5]. Available from: https://www.epa.gov/regulations-emissions-vehicles-and-engines/epa-collaboration-international-air-pollution-0

3. Special Areas under MARPOL [Internet]. [cited 2024 Nov 5]. Available from: https://www.imo.org/en/OurWork/Environment/Pages/Special-Areas-Marpol.aspx

4. IMO 2020 – cutting sulphur oxide emissions [Internet]. [cited 2024 Nov 5]. Available from: https://www.imo.org/en/MediaCentre/HotTopics/Pages/Sulphur-2020.aspx

5. Lasek M. Ship & Bunker. [cited 2024 Nov 4]. S&B ANALYSIS: Latest on Brent / VLSFO / HSFO Relationship, Bunker Price Outlook as Low Sulfur Premium Nears 2-Year High. Available from: https://shipandbunker.com/news/world/460972-sb-analysis-latest-on-brent-vlsfo-hsfo-relationship-bunker-price-outlook-as-low-sulfur-premium-nears-2-year-high

6. IMO Data Collection System (DCS) [Internet]. [cited 2025 Nov 20]. Available from: https://www.imo.org/en/ourwork/environment/pages/data-collection-system.aspx

7. Jalkanen JP, Brink A, Kalli J, Pettersson H, Kukkonen J, Stipa T. A modelling system for the exhaust emissions of marine traffic and its application in the Baltic Sea area. Atmospheric Chem Phys. 2009 Dec 7;9(23):9209–23.

8. Johansson L, Jalkanen JP, Kukkonen J. Global assessment of shipping emissions in 2015 on a high spatial and temporal resolution. Atmos Environ. 2017 Oct 1;167:403–15.

9. Jalkanen JP, Johansson L, Kukkonen J, Brink A, Kalli J, Stipa T. Extension of an assessment model of ship traffic exhaust emissions for particulate matter and carbon monoxide. Atmospheric Chem Phys. 2012 Mar 12;12(5):2641–59.

10. Jalkanen JP, Johansson L, Andersson MH, Majamäki E, Sigray P. Underwater noise emissions from ships during 2014–2020. Environ Pollut. 2022 Oct 15;311:119766.

11. Huang L, Wen Y, Zhang Y, Zhou C, Zhang F, Yang T. Dynamic calculation of ship exhaust emissions based on real-time AIS data. Transp Res Part Transp Environ. 2020 Mar 1;80:102277.

12. International Maritime Organization. Fourth IMO GHG Study 2020 [Internet]. 2021. Available from: https://wwwcdn.imo.org/localresources/en/OurWork/Environment/Documents/Fourth%20IMO%20GHG%20Study%202020%20-%20Full%20report%20and%20annexes.pdf

13. SeaVision - Maps [Internet]. [cited 2024 Nov 4]. Available from: https://seavision.volpe.dot.gov/

14. Maritime Safety Information [Internet]. [cited 2025 Nov 19]. Available from: https://msi.nga.mil/Publications/WPI

15. S&P Global StatCode 5 Shiptype Coding System [Internet]. [cited 2025 Nov 19]. Available from: https://cdn.ihs.com/www/pdf/Statcode-Shiptype-Coding-System.pdf

16. Third IMO Greenhouse Gas Study [Internet]. [cited 2025 Nov 19]. Available from: https://wwwcdn.imo.org/localresources/en/OurWork/Environment/Documents/Third%20Greenhouse%20Gas%20Study/GHG3%20Executive%20Summary%20and%20Report.pdf

17. MacGillivray A, de Jong C. A Reference Spectrum Model for Estimating Source Levels of Marine Shipping Based on Automated Identification System Data. J Mar Sci Eng. 2021 Apr;9(4):369.

18. USEPA/Marine_Emissions_Tools: Tools for estimating emissions from commercial marine vessels. [Internet]. [cited 2025 Nov 19]. Available from: https://github.com/USEPA/Marine_Emissions_Tools

19. S&P Global Market Intelligence [Internet]. [cited 2024 Nov 4]. Sea-Web The Ultimate Marine Online Database. Available from: https://www.spglobal.com/market-intelligence/en/solutions/sea-web-maritime-reference

20. EPA Office of Transportation and Air Quality. Ports Emissions Inventory Guidance: Methodologies for Estimating Port-Related and Goods Movement Mobile Source Emissions. 2022.

21. 40 CFR 1042.901 -- Definitions. [Internet]. [cited 2025 Nov 19]. Available from: https://www.ecfr.gov/current/title-40/part-1042/section-1042.901

22. Spherical computation — R Spatial [Internet]. [cited 2025 Nov 19]. Available from: https://rspatial.org/raster/sphere/index.html

23. USA Detailed Water Bodies - Overview [Internet]. [cited 2025 Nov 19]. Available from: https://www.arcgis.com/home/item.html?id=48c77cbde9a0470fb371f8c8a8a7421a

24. GSHHG - A Global Self-consistent, Hierarchical, High-resolution Geography Database [Internet]. [cited 2025 Nov 19]. Available from: https://www.soest.hawaii.edu/pwessel/gshhg/

25. Olmer N, Comer B, Roy B, Mao X, Rutherford D. Greenhouse gas emissions from global shipping, 2013–2015: Detailed methodology.

U.S. Department of Transportation
John A. Volpe National Transportation Systems Center
220 Binney Street
Cambridge, MA 02142-1093

617-494-2000
www.volpe.dot.gov

U.S. Department of Transportation
Volpe Center