D1.3 Data Management Plan

Programme: EMFAF-2023-PIA-FLAGSHIP
Grant agreement number: 101124925
Project acronym: GreenMED
Project title: Green Shipping Pathways Towards a Clean Energy Transition in the Mediterranean

Prepared by: MT

Date: 31/01/2024
Deliverable version: v1.0

Co-funded by the European Union

Funding acknowledgement

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

DELIVERABLE DETAILS	
Grant Agreement No.	101124925
Project acronym	GreenMED
Project full title	Green Shipping Pathways Towards a Clean Energy Transition in the Mediterranean
Dissemination level	Public
Due date of deliverable	M4
Actual submission date	M4
Deliverable name	Data Management Plan
Type	DMP – Data Management Plan
Document status	Draft
WP contributing to the deliverable	WP1
Author(s)	Konstantina Bereta (MT), Eleni Tsili (MT), Manolis Annetis (NTUA), Christoforos Diamantis (NTUA), Ioannis Kopsacheilis (NTUA), Aurelio Lázaro Chueca (VPF)
Other Contributor(s)	None
Reviewer(s)	Manolis Annetis (NTUA)
Keywords	Data Management, Data Collection, Data Requirements, FAIR Data, Data Security, Ethics

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

VERSION HISTORY		
VERSION	**COMMENTS**	**DATE**
0.1	First draft prepared by MT and sent to partners for input	24/11/2023
0.2	VPF input in Section 3, The EU Blue Economy Observatory, European Alternative Fuels Observatory, The MedPort Association database	05/12/2024
0.3	VPF input incorporated into document	12/12/2023
0.4	NTUA input in Section 3, CO2 emission data, The MarineTraffic database, The Sea-web database, AIS Tracking – Ship Tracker for Maritime Traffic, The European Marine Observation and Data network (EMODnet), THETIS-MRV, Alternative Fuels Insight (AFI) platform; Section 5; Section 6; NTUA also reviewed the entire document and proposed amendments	11/01/2024
0.5	NTUA input and amendments incorporated into document, Revised version	15/01/2024
0.6	CMMI suggestions incorporated into document, Revised version	24/01/2024
1.0	Final version	31/01/2024

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

TABLE OF CONTENTS

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

LIST OF ABBREVIATIONS

Abbreviation	Meaning
AFI	Alternative Fuels Insight
AIS	Automatic Identification System
BDN	Bunker Delivery Notes
DCS	Data Collection System
DL	Deliverable Leader
EAFO	European Alternative Fuels Observatory
EMODnet	European Marine Observation and Data Network
EMSA	European Maritime Safety Agency
ETS	Emission Trading System
EU	European Union
GHG	Greenhouse gases
GT	Gross Tonnage
IMO	International Maritime Organization
MEPC	Marine Environment Protection Committee
MSSO	Mediterranean Sustainable Shipping Observatory
PC	Project Coordinator
TEN-T	Trans-European Transport Network
WP	Work Package

LIST OF FIGURES

LIST OF TABLES

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

1. EXECUTIVE SUMMARY

This document describes the data management plan that will be implemented in the course of the GreenMED project. The document will be updated in light of important changes to data management with respect to project needs. The primary goal of the project's data management plan is to ensure the effective collection, storage, sharing, and preservation of project data throughout its lifecycle. By implementing robust data management practices, the Consortium aims to enhance the quality, accessibility, and long-term value of the research outcomes of the GreenMED project. By adopting this comprehensive data management plan approach, project partners aim to maximize the impact of the conducted research, fostering collaboration, and ensuring the integrity and accessibility of the data used for current and future stakeholders. The primary objective of this complete version of the data management plan is to describe the adopted approach, along with an overview of the datasets which will be used as input at this early stage of the project. AIS datasets are highlighted, since they are crucial for the project's studies. An updated version of this document will be delivered no later than the end of the project, in case of important changes with respect to project needs, and will probably include detailed descriptions and access rights for all datasets that will be used and produced in the project, including intermediate datasets (e.g., produced by AI models).

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

2. INTRODUCTION

This document is deliverable D1.3 “Data Management Plan” of Work Package 1 (WP1) and outlines a comprehensive plan of how data will be collected and handled during the GreenMED project and the methods for sharing and making it accessible. This information is crucial for determining the most suitable sustainability model for the project's outcomes and for distributing the project's open data in accordance with this plan.

2.1 Purpose and Scope of the document

The current Data Management Plan defines in detail the adopted approach, but also includes an overview of the requirements and specifications of data types that will be used. In addition, the deliverable investigates their availability and examines the data sources, either publicly available, or introduced by consortium partners. This task is closely linked to all data-related activities within GreenMED. T2.1 is particularly important, as it will define in more detail the data specifications and the respective reporting.

Specifically, the deliverable includes information concerning:

- How data will be managed both during and following the project's conclusion.
- Description of how data will be generated, collected and processed.
- Explanation of the methodology and standards that will be employed.

2.2 Deliverable Structure

This deliverable is divided in the following sections:

Section 1 introduces the executive summary.

Section 2 includes an introduction to the document.

Section 3 provides an overview of the datasets that will be used during the project's lifecycle; particular focus is given on AIS datasets, since they are crucial for the project’s studies.

Section 4 outlines the steps the consortium intends to implement to align with the FAIR data management principles.

Section 5 describes the allocation of roles and the resources employed.

Section 6 is dedicated to data security issues.

Section 7 describes in detail all issues related to ethics and legal compliance.

Section 8 provides the conclusions of the current document.

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

3. DATA SUMMARY

The focus of the initial phases regarding the MSSO (Mediterranean Sustainable Shipping Observatory) knowledge capacity development is region-specific and centers on the energy demand of the Mediterranean maritime transportation activity and the maritime fuels supply chain. The regionally-oriented studies employed to support this knowledge base are primarily categorized into the supply and the demand axes. In order to support both, the GreenMED project will be using region-specific data for port and bunkering infrastructure and fuel supply chain as well as for regional shipping activity. Additionally, energy efficient solutions and decarbonization measures, which will substantially impact the ship energy demand, will be recorded and employed in the project's research. These will set the basis for implementing the project's methodology and ensuring the success and reliability of its final results. The selection of the appropriate data, their correct and systematic collection, and their classification is therefore a task of major importance.

In this chapter, the document describes the data that will be generated or collected within the scope of the GreenMED project as well as the collection methods and processing steps that will be followed. An overview of most data is presented. A particular focus is given on AIS datasets, since they are a crucial part of the project's studies.

3.1 Data requirements and specifications

The development of the knowledge capacity of the GreenMED project and MSSO requires a range of data sets. During the first year of the project several datasets will be employed for the regionally-focused studies, with only descriptive information and study outcomes being subsequently transferred and housed into the MSSO tool. Datasets can be distinguished in two main categories. The first will primarily center on maritime transportation activities in the Mediterranean region and refers mostly to ship databases and AIS data. The second category that will complement the GreenMED knowledge base pertains to data related to the fuel supply chain within the Mediterranean region. This can also benefit from AIS data, but its primary focus is on business and market data regarding bunkering in ports and supply chain.

3.1.1 Ship database

The ship database will contain information related to the fleet of scope that includes all ships operating in the Mediterranean Sea, either they are exclusively operating in the region or simply passing through, provided they have spent a minimum amount of time within the area. Datasets will support the definition of the baseline fleet of all commercial ship types (e.g., Bulk Carriers, Oil Tankers, Containers, Chemical Tankers, General Cargo, Liquified Gas Tankers, Cruise, Ro-Ro), considering at least all vessels exceeding 5,000 GT, which aligns with the minimum ship limit for the IMO Data Collection System (DCS). However, the project aims to include ships below this threshold, for a more accurate approximation of the Mediterranean fleet energy demand.

Regarding ship data, the project will adopt an anonymized approach and will provide information on a set of ship-related parameters which will include: ship type, dimensions (e.g., length overall, breadth extreme, design draught, depth), capacity (deadweight, gross tonnage, TEU), design speed, year built, main engine power and speed, fuel type and AIS data (e.g., speed over ground, coordinates). The above parameters depend also on their availability and

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

their applicability to the various ship types. This inventory of ship parameters is not exhaustive and can be adjusted as necessary to align with the project's requirements, during the following stages of the project. Subsequently, in relation to the established data sets, main engine type, fuel type and auxiliary engines and boilers power demand will be assigned to the identified fleet. Ships with complete or partial information will have fuel and machinery types assigned accordingly. For vessels with missing information, fuel and machinery types will be allocated based on statistics per ship type and size information from the available dataset. To ensure comprehensive data coverage, vessels outside the primary fleet of interest may be included for statistical purposes in defining missing data. In a final step, an operational profile will be established for all identified vessels. This profile will include factors such as time spent at sea, distance traveled, average speed, bunkering log, unique voyages and ports visited. This information is essential for calculating or estimating the fleet's energy demands.

3.1.2 AIS data

AIS datasets are particularly described here as a crucial aspect for the project's studies. AIS is an autonomous identification and tracking system that is carried by the majority of vessels at sea. As from 2002 it is mandatory for all vessels over 300 gross tons according to the IMO (International Maritime Organization). It provides both dynamic and static information about the vessel that operates it. Kinematic information is mainly relevant to the vessel's mobility and navigational status. A more detailed description of this type of information can be found in Table 1. Static information is mostly structural and registrational information of the vessel as well as voyage-related information that normally remains fixed throughout a voyage. A more detailed description of static information can be found in Table 2.

Table 1: Dynamic AIS information

Attribute name	Description
MMSI	Maritime Mobile Service ID - a nine-digit number sent in digital form over a radio frequency that identifies the vessel's transmitter station.
Position coordinates	Vessel's latitude and longitude. Latitude ranges from -90 to 90 and longitude from -180 to 180. Both with up to 0.0001 minutes accuracy.
Speed over ground (SoG)	The ship's speed, measured in knots. Ranges from 0 to 102 knots with 0.1 knot resolution.
Course over ground (CoG)	Vessel's motion direction relative to the magnetic north pole. Ranges from 0 to 359 degrees with 0.1 minute resolution.
Heading	Vessel's heading direction relative to the magnetic north pole. Ranges from 0 to 359 degrees.
Navigational Status	The AIS Navigational Status of the subject vessel as input by the vessel's crew. There might be discrepancies with the vessel's detail page when vessel speed is near zero (0) knots.
Rate of turn	Right or left turn angle of vessel. Ranges from 0 to 720 degrees with minute resolution.

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

Table 2: Static AIS information

Attribute name	Description
IMO Number	International Maritime Organisation number - a seven-digit number that uniquely identifies vessels
Callsing	A uniquely designated identifier for the vessel's transmitter station
Name	The vessel's name that is manually inserted by crew members.
Type	The vessel type id. It is 0-255 code that is mapped to its type (e.g. tanker, passenger, etc.).
Dimensions	Dimensions of ship in meters. Four integers indicating dimension to bow, dimension to stern, dimension to port (i.e., left side of the vessel when facing the bow), and dimension to starboard (i.e., right side of the vessel when facing the bow).
Location of antenna	Location of the antenna in the vessel
Draught	The Draught (in metres x10) of the subject vessel according to the AIS transmissions
Destination	The Destination of the subject vessel according to the AIS transmissions
ETA	The Estimated Time of Arrival to Destination of the subject vessel according to the AIS transmissions
Type of positioning system	The positioning system used to represent the coordinates of the location of the vessel

3.1.3 CO2 emission data

Another crucial aspect regarding the project's studies are the emissions data that may be available for the fleet of scope. These data will be used mainly for validating calculated emissions and calibrating the calculation model accordingly. However, other uses may be applicable. Such data could be made available through:

- Ship owners
- Ship databases
- Emissions reporting systems

In more detail, through a variety of methods, emissions can be reported, either internally, within a shipping company, or externally, to ship databases or emissions reporting systems. These methods are usually the following:

a. Estimating emissions using recorded fuel consumption, from Bunker Delivery Notes (BDNs) and period stock takes of fuel tanks.

b. Estimating emissions using recorded fuel consumption, by monitoring fuel tanks onboard.

c. Estimating emissions using recorded fuel consumption, through flow meters, only in

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

case of applicable combustion processes.

d. Direct CO2 emissions measurements in the exhaust fumes.

Most popular methods yet are a. and b., which are adopted by the majority of the shipping companies, as they do not require any modifications or additional equipment.

For internal reporting, companies can keep record freely, use these data for producing insights and managing their fleet effectively. They may also provide such data on demand, for relevant studies and research. However, such reporting, maybe in its simplest form yet, has become mandatory in the context of IMO and EU. In more detail, in October 2016, MEPC 70 adopted, by resolution MEPC.278(70), the IMO ship fuel oil consumption system (IMO DCS), consisting of requirements for ships to record and report their fuel oil consumption with a view to inform further IMO measures to reduce GHG emissions from ships. Starting from 1 January 2019, ships of 5,000 GT and above (which produce approximately 85% of the total CO2 emissions from international shipping) are required to collect consumption data for each type of fuel oil they use as well as other specified data. On the other hand, at the EU level, the EU-MRV system is used to report CO2 emissions from ships, according to the EU Regulation 2015/757. This Regulation applies to ships above 5,000 gross tonnage in respect of CO2 emissions released during their voyages from their last port of call to a port of call under the jurisdiction of a Member State and from a port of call under the jurisdiction of a Member State to their next port of call, as well as within ports of call under the jurisdiction of a Member State. All relevant information is accessible through the search tool or can be exported in a spreadsheet for further analysis. Since 30 June 2020, all the verified information submitted by companies to the European Commission for the reporting year 2019 is accessible. The Regulation was amended recently (Regulation (EU) 2023/957) for incorporating the aspects of EU Emission Trading System (ETS).

Based on the EU-MRV, Table 3 summarises some important attributes from the variety of information and data provided, that are useful for the project studies. These are freely available online and can be accessed by all partners. For ships not subjected to the EU Regulation 2015/757, data may be scarce. However, since the fleet of scope is limited to the Mediterranean, enough data may be available for calibrating the calculation model and validating the results. Either from other sources, or on demand through shipping companies, such data usually follow the same format.

Table 3: CO2 emission data

Attribute name*	**Description**
IMO number	International Maritime Organisation number - a seven-digit number that uniquely identifies vessels. It is used as an identification attribute across datasets for vessels.
Ship name	The vessel's name that is usually manually inserted by ship personnel.
Ship type	The type of the vessel among several prefixed types following the IMO's categorisation.
Reporting period	The year that refers to the specific record for the ship under consideration.

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

Technical efficiency	The estimated index value (EIV), which is reported instead of the EEDI, where the latter is not applicable. It is measured in gCO2/t·nm and is calculated based on relevant guidance.
Monitoring Methods	From the aforementioned monitoring and reporting methods for CO2 emissions, this field is fixed with the method employed, either a., b., c. or d.
Total fuel consumption	The total fuel consumption of the vessel under consideration, within the reporting period, measured in metric tons.
Total CO2 emissions	The total CO2 emissions of the vessel under consideration, within the reporting period, based on the selected monitoring method (i.e., a, b, c or d, as mentioned above), measured in tons.
Annual time spent at sea	The annual time spent at sea, in hours, by the vessel under consideration.
Annual average fuel consumption per distance	The average fuel consumption of the vessel under consideration per unit of distance made, measured in kg/nm.

* The attributes presented in the table are indicative, with respect to the data that are reported through EU-MRV. These can be either source data, or calculated data from other attributes. Reporting systems may provide such data, but shipping companies may have other data attributes or in different formats.

3.1.4 Supply chain, ports and infrastructure data

Datasets refer to all ports within the Mediterranean region, including their bunkering facilities, and also other elements of the supply chain with respect to conventional and alternative maritime fuels. Ports will be registered along with various parameters, such as their name, location, sizing, area, historical traffic data for the last four years (e.g., daily arrivals and departures, vessels in ports), available technologies (e.g., cold ironing) and their facilities and will qualify for consideration only if they meet minimum traffic criteria, such as daily arrivals and departures and the number of vessels in port. This criterion aims to exclude inactive and small provincial ports. The categorization of supply chain elements, infrastructures and ports will be based on their location, the aforementioned parameters where applicable, and their participation in multimodal logistics networks and supply chains, such as the Trans-European Transport Network (TEN-T) and other future networks. Regarding bunkering, all three methods, namely truck-to-ship, shore-to-ship, and the most popular ship-to-ship approach, will be taken into account. Additionally, the infrastructure at each bunkering site will be mapped. Bunkering volumes will be determined using a model that takes into account the calculated fuel consumption of the target fleet, bunkering-related information (i.e., bunkering time) obtained from AIS traffic data, port data, and available research studies. Additional data related to the supply chain include details on bunkering quantities and costs, as well as information concerning logistic networks, imports, and production capacities will be defined based on available data sources, studies and information from the established stakeholder network.

3.2 Data collection

The GreenMED project adopts a tri-fold strategy in data collection in order to address the issues of missing or non-uniform data. The data gathered through this process will be analyzed and will serve as input for the development, assessment and evaluation of the developed scenarios.

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

The data sources that will be used for the project are the following:

1. The main source of data will be historical data (e.g., vessel routes and traffic density, miles traveled). The majority of the data required for the energy demand study will be provided by MarineTraffic and also from NTUA, through its access to the Sea-web database[1]. For the supply chain study, publicly available data will be explored in-depth, while all consortium partners will seek out sources, at national and regional levels, to support quantification. VPF will provide a significant portion of data, as a member also of the MedPort Association, while AASTMT will provide data related to the northern Africa supply chain mapping. It is noted that all data from the MarineTraffic database will be accessed and processed only by MarineTraffic.
2. Complementary data from extensive literature review (e.g., for infrastructure, fuel storage capacity, green shipping technologies) will be sought out by all partners during the regionally oriented studies, particularly in cases of significant missing data. Publicly available data will be explored in-depth, while insight information from commercial sources will be considered as well.
3. Expert insights and contribution through the MSSO and dedicated project workshops. In particular stakeholders, with a wide knowledge over the status quo, but also for the future trends, will support the data collection. Their engagement will be ensured through surveys, questionnaires and interviews, if applicable. CMMI will play a vital role in their approach, as the project's dissemination manager.

Most of the collected data will not be structured into a database, but they will surely provide direct results customized to meet the requirements of subsequent tasks. Each partner will maintain separate, in-house databases, and data will be anonymized for research purposes. These may be reconsidered within the course of the project, as some results and datasets may be required to be made available to stakeholders of the MSSO.

The required datasets will be collected during T2.2. During this task all baseline sets of data, for the fleet of scope and the bunkering facilities and ports in the Mediterranean will be provided, while also delivering the inventories of conventional and alternative fuels, as well as green shipping technologies.

The following paragraphs summarise the data sources identified so far by the project partners, and they will be used during the data collection process, under T2.2.

3.2.1 The MarineTraffic database

For the energy demand study of the Mediterranean fleet of scope, the main source of data will be provided by MarineTraffic, regarding both ship data and AIS data. The MarineTraffic database has the capacity for supporting the study objectives through historic data. The datasets will be accessed and processed only by MarineTraffic, in order to provide aggregated results and statistics, based on the scope definition and established assessment framework (T3.1).

3.2.2 The Sea-web database

In the same sense, NTUA, through the access to the Sea-web database[2] and its various modules, will be able to complement any missing or additional data, depending on the needs of the

[1] https://www.spglobal.com/marketintelligence/en/mi/products/sea-web-maritime-reference.html
[2] [Online] Available at: https://www.spglobal.com/marketintelligence/en/mi/products/sea-web-maritime-reference.html

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

studies. The database is mainly identified as a ship database, with a large variety of information and data per ship. The available information and data through Sea-web are presented in Table 4. It is noted that these data will be accessed and processed mainly by NTUA, which will possibly share portions of data to project partners for the needs of the studies.

Table 4: Sea-web database information and data description

Categories of Data	Description of Data
Registration data	This category of data includes attributes such as IMO number, MMSI number, callsign, ship type, port of registry, flag state, status and ownership.
Dimensions and Tonnages	This refers to attributes describing the dimensions and tonnages of the vessel, such as length over all, breadth, draught, displacement, gross tonnage, deadweight and other relevant information.
Construction data	Apart from the ship type, as constructed, this category focuses on construction information such as the year built, shipyard, design details, main engine Tier and EEDI.
Cargo and Capacities	This category of data refers to cargo gear (e.g., pumps, cranes), tank capacities, for both cargo and fuel.
Machinery data	All available information and data regarding ship machinery are included in this category, such as main engine data (e.g., power, speed, type), auxiliary engines data, boilers and thrusters information.
Movements	This category presents all the ship movements, including data referring to position, trading area, destination, origin, status, port calls, estimated time of arrival and other relevant data.
Performance data	This category refers to ship performance data, including time spent at sea, distance made, average speed, total time remaining still and other relevant data.

* The table describes categories of data and not actual attributes. A more detailed description of the data required and used for the project will be presented in later deliverables (i.e., D1.2).

3.2.3 AIS Tracking - Ship Tracker for Maritime Traffic[3]

These AIS data may complement any missing data, or offer additional information. However, they will be mainly used by NTUA for calibrating the calculation model for the fuel consumption and emissions, since the main source of AIS data for the implementation of project studies on energy demand will be the in-house database of MarineTraffic (refer to 3.2.1). The available information and data are following the AIS data format as presented in 3.1.2. It is noted that these data will be accessed and processed mainly by NTUA, which will possibly share portions of data to project partners for the needs of the studies.

3.2.4 The European Marine Observation and Data Network (EMODnet)

Data from the marine environment are crucial for addressing threats related to it. However, marine data collection, storage and access in the European area has been carried out in a

[3] [Online] Available at: https://www.spglobal.com/marketintelligence/en/mi/products/ais-live-ship-tracker.html

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

fragmented way for many years, serving individual purposes of private or public organisations, often isolated from each other. Thus, EMODnet was presented as a solution to collectively observe the sea, process the data according to international standards and make them freely available as interoperable data layers and data products. EMODnet is actually a network of organisations that materialises this effort, supported by the EU's integrated maritime policy. Through an integrated web-platform[4], it provides access to marine data across the following seven disciplines; bathymetry, biology, chemistry, geology, human activities, physics and seabed habitats. The EMODnet offers a variety of data products, but also makes most of them visible through the interactive Map Viewer, organised in layers of data with spatial characteristics. An indicative screenshot is shown in Figure 1.

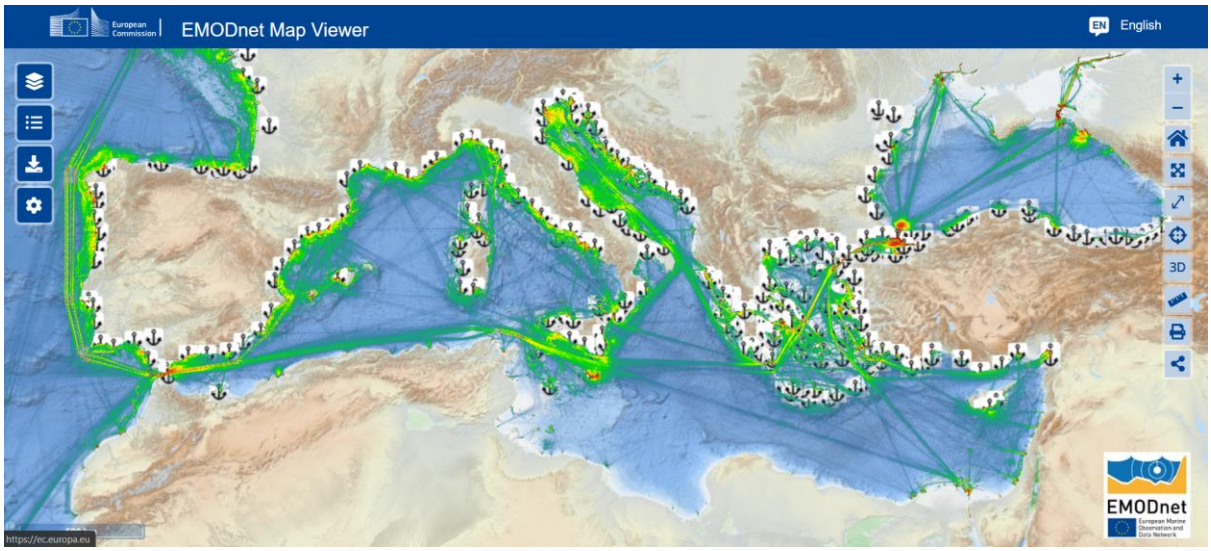

Figure 1: EMODnet Map Viewer[5]

In the case of this project's interests, the human activities discipline is useful for extracting data and for using them accordingly in the project's studies. All information and data are freely available, and thus, can be accessed by all project partners. Several useful categories of data are summarised in Table 1.

Table 5: EMODnet information and data description

Categories of Data	Description of Data
Environment	These datasets may refer to marine protected areas, emission control areas (ECA) and sulfur emission control areas (SECA), and other national or international environmental restrictions.
Energy	This category includes information regarding ocean energy facilities, either project locations or test sites, offshore wind farms and other relevant data.
Main Ports	Includes all the main ports in the EU, with data covering a period from 1997 to 2023, referring either to goods traffic, passenger traffic or vessel traffic.

[4] [Online] Available at: https://emodnet.ec.europa.eu/en
[5] [Online] Available at: https://emodnet.ec.europa.eu/geoviewer/

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

Oil and Gas	This category refers to information related to offshore installations, boreholes and active licenses.
Pipelines	This information refers strictly to pipelines covering the European area, both at sea and ashore.
Route Density	This category includes data describing the route densities at sea, covering a period from 2019 to 2023, at annual, monthly or seasonal aggregated totals.
Vessel Density	In the same sense, this category presented data of vessel densities across European waters, at annual or monthly aggregated totals, covering a period from 2017 to 2022.

* The categories of data and information presented in the table above are indicative. A more detailed description of the data required and used for the project will be presented in later deliverables (i.e., D1.2).

3.2.5 THETIS-MRV

Following the implementation of the EU Regulation 2015/757 regarding emissions reporting, the EU, in collaboration with the European Maritime Safety Agency (EMSA), created the THETIS-MRV platform[6] for this exact purpose, i.e., for reporting and validating the CO2 emissions of all vessels reaching a port of an EU member state. Respective data of the EU-MRV are indicatively described in Section 3.1.3 and Table 3, and focus on the following attributes:

- IMO number
- Reporting period
- Technical efficiency
- Monitoring Methods
- Total fuel consumption
- Total CO2 emissions
- Annual time spent at sea
- Annual average fuel consumption per distance

3.2.6 The EU Blue Economy Observatory

The EU Blue Economy Observatory[7] compiles and provides detailed information on various aspects related to the blue economy in Europe. Some types of information that can be extracted from this source include the data presented in Table 6. All information and data are freely available and can be accessed by all project partners.

Table 6: EU Blue Economy Observatory data description

Categories of Data	Description of Data
Economic and Financial Data	Information on the economic contribution of specific sectors within the blue economy, such as fishing, aquaculture, maritime transport, coastal tourism, among others.

[6] [Online] Available at: https://mrv.emsa.europa.eu/
[7] [Online] Available at: https://blue-economy-observatory.ec.europa.eu/index_en

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

Maritime Transport and Logistics	Statistics related to maritime transport, ports, and associated logistics infrastructure. This could include data on cargo volumes, shipping routes, and trends in maritime freight transport.
Marine Renewable Energies	Information on the development and production of renewable energies in marine environments, such as offshore wind farms and wave energy projects.
Marine Research and Technology	Data on scientific research, technological advancements, and innovation projects in the marine field. This could include information on new technologies for ocean monitoring, underwater research, etc.
Governance and Regulation	Information on legislation, regulations, and policies related to the blue economy in the European Union.
Sustainable Development and Environment	Data on sustainable practices, marine conservation initiatives, and efforts to address environmental challenges associated with marine activities.

3.2.7 European Alternative Fuels Observatory

The European Alternative Fuels Observatory[8] (EAFO) provides comprehensive information related to alternative fuels across Europe. The dataset includes various aspects of alternative fuels adoption and usage. Some types of information and data that can be extracted from EAFO are presented in Table 7. The data are freely available and can be accessed by all project partners.

Table 7: European Alternative Fuels Observatory data description

Information and Data categories	**Information and data description**
Alternative Fuels Usage	Detailed statistics on the consumption and adoption of alternative fuels, including but not limited to electric vehicles, hydrogen, natural gas, and biofuels.
Charging Infrastructure	Information about the availability and distribution of charging infrastructure for electric vehicles / vessels, including charging station locations, types of chargers, and charging capacities.
Emission Reduction Impact	Insights into the environmental impact of alternative fuels, including data on emission reductions, air quality improvements, and contributions to overall sustainability goals.
Policy and Regulation	Information on regulatory frameworks, incentives, and policies promoting the use of alternative fuels and sustainable mobility solutions across European countries.
Research and	Data on research initiatives, technological advancements, and innovative projects in the field of alternative fuels, with a focus on

[8] [Online] Available at: https://alternative-fuels-observatory.ec.europa.eu/

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

Information and Data categories	Information and data description
Innovation	fostering sustainable and low-emission transportation.
Market Trends and Analysis	Analysis of market trends, consumer preferences, and industry developments related to alternative fuels and vehicles, providing valuable insights for strategic decision-making.
Collaborative Initiatives	Information on collaborative efforts, partnerships, and initiatives aiming to advance the adoption of alternative fuels and create a more sustainable transportation ecosystem.

3.2.8 Alternative Fuels Insight (AFI) platform

Driven by the decarbonization goals set by IMO, there is a variety of efforts for providing the shipping industry with current developments and insights for alternative fuels and their uptake. One of the most concrete efforts is the Alternative Fuels Insight (AFI) platform[9] from DNV. Its goal is to support the industry's clean energy transition by providing a complete overview of the developments regarding alternative fuels and technologies, covering both investments on ships, bunkering infrastructure and fuel production facilities. The data are organised in interactive dashboards, while the platform offers an interactive map viewer depicting spatial data for ships and infrastructure. An indicative screenshot is shown in Figure 2.

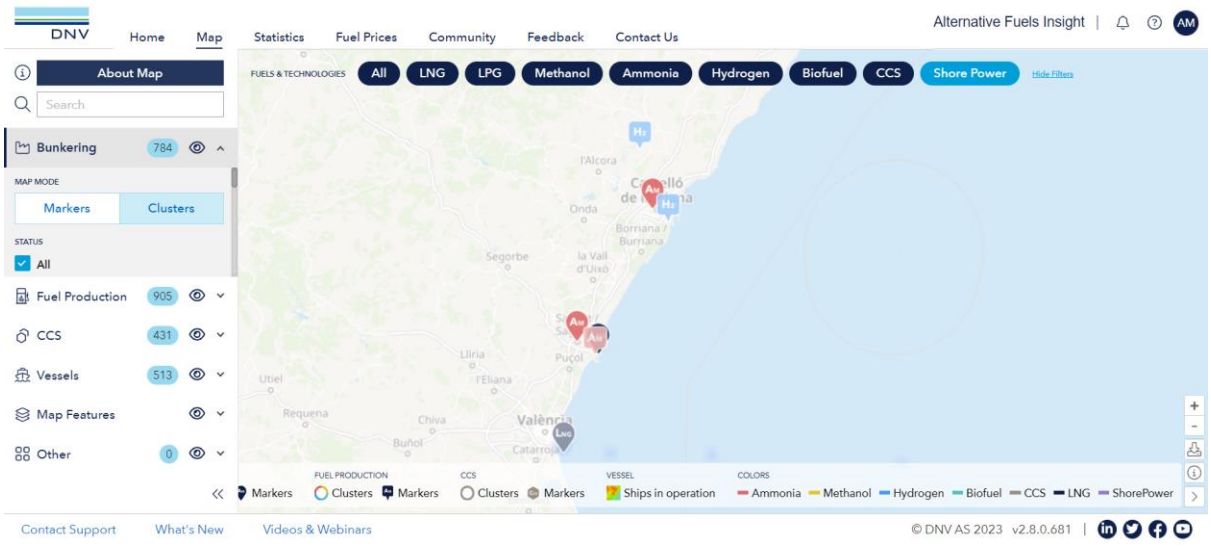

Figure 2: Interactive AFI map viewer

A wide portion of data and information are freely available, while a more detailed view, with aggregated datasets, can be accessed through a variety of subscriptions. The necessary level of detailed for AFI data will be explored and consortium partners may proceed with subscription options (i.e., NTUA) for the needs of the project's studies. In such a case, the data will be accessed and processed individually by the respective partner, and portions of data will

[9] [Online] Available at: https://afi.dnv.com/

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

possibly be shared to other project partners for the needs of the studies. The publicly available information and data from AFI can be accessed and used by all partners. The various data that may be useful for the project's studies are presented in Table 8.

Table 8: DNV's AFI data and information description

Categories of Data	Description of Data
Fuel Production	This category presents all alternative fuel production sites, either under discussion, decided, under development or in operation, for a variety of alternative fuels, i.e., Methanol, Ammonia, Hydrogen and Biofuels.
Bunkering	This category presents data and information regarding bunkering infrastructure, either under discussion, decided, under development or in operation. These include bunker vessels, trucks loading sites (i.e., ports), bunker vessels sites, local storages, tank to ship infrastructure, and other bunkering infrastructure. It also includes shore power infrastructure, i.e., cold ironing.
Vessels	It includes a variety of ship types, and whether they are equipped with batteries or scrubbers. Additional technologies may be considered in the future (e.g., wind assisted propulsion). Carbon Capture Systems (CCS) are considered here as well.
Fuel Prices	This module reports on fuel prices across regions at an international level, providing useful information about ship bunkering quantities as well.

3.2.9 The MedPort Association database

The MedPort Association[10] plays a pivotal role in providing comprehensive datasets related to its platform focused on collaboration, production and exchange between the Mediterranean ports on common issues. This dataset encompasses diverse aspects within the maritime and port sector. Several types of information and data that can be extracted from the MedPort Association's database are presented in Table 9.

Table 9: MedPort Association datasets and information description

Information and data categories	Information and data description
Maritime Traffic and Shipping Patterns	In-depth data on maritime traffic patterns, shipping routes, and traffic intensity within the regions covered by member ports.
Operational Standards and Best Practices	Information on operational standards, best practices, and guidelines endorsed by the MedPort Association to ensure efficient and secure port operations.
Trade and Economic Impact	Statistics on trade volumes, economic contributions, and the overall impact of member ports on regional and international commerce.

[10] [Online] Available at: https://medports.org/

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

Information and data categories	Information and data description
Infrastructure Development	Updates on port infrastructure development projects, expansions, and technological advancements implemented by member ports under the guidance of the MedPort Association.
Environmental Sustainability Initiatives	Data on environmental initiatives and sustainability practices undertaken by member ports, aligning with global efforts to reduce the maritime industry's ecological footprint.
Safety and Security Measures	Information on safety protocols, security measures, and incident reports to ensure the well-being of personnel and assets within member ports.
Regulatory Compliance and Advocacy	Details on regulatory compliance standards set by the MedPort Association and advocacy efforts aimed at influencing maritime policies for the benefit of member ports.

3.3 Data cleaning, evaluation and pre-processing

Following the collection of the data thoroughly described in 3.2, the cleaning and preprocessing of data will be executed in T2.3 and will mainly focus on the two main data categories: ship and AIS data, and bunkering and supply chain data, which includes information related to ports and other infrastructure.

Initially, regarding the ship dataset and related parameters, the process will involve clearing and filtering inaccurate or erroneous data, as well as filling in missing information (e.g., through multilinear regression models, statistically, or other). In the meantime, information like the main engine type, fuel type, and auxiliary engines and boilers power demand will be assigned to the identified fleet. Additionally, the processing of AIS data (e.g., time, latitude, longitude, speed over ground, draught) will include a series of data analysis techniques, like data cleaning, filtering, resampling, filling in missing information, and extrapolation into annual hourly observations. Furthermore, a distinction between regionally operating ships and those engaged in international voyages will be performed during the pre-processing stage. This will be made possible by employing a methodology that utilizes data of port calls for each ship in order to identify distinct voyages. The above processing steps related to AIS data will be carried out exclusively by MarineTraffic, delivering only the necessary aggregated results.

Regarding the supply chain, respective processes will be utilized during the data pre-processing stage. Data pertaining to ports and bunkering facilities will be cleaned and filtered, whereas for unstructured data concerning the remainder of the supply chain, a simpler processing and documenting procedure will be adopted.

The above data cleaning and preprocessing steps are assigned to T2.3 and the outcomes of it will progressively offer insights for T2.4 and the techno-economic analysis, while also closely linked to the WP3, and in particular T3.1 and T3.2, at least in the first phase of the project.

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

4. FAIR DATA

The GreenMED consortium is dedicated to enhancing the sustainability and impact of the project by promoting the reuse of research data, both throughout the project's implementation and after its conclusion. In accordance with H2020 data management guidelines[11], the GreenMED data management plan will ensure that the project's research data is Findable, Accessible, Interoperable, and Reusable (FAIR principles). Following FAIR principles, it improves the utility of the data used and produced by the project consortium for others and enhances the credibility and impact of the datasets. This section presents perspectives regarding several key issues pertinent to FAIR data management and outlines the consortium's plans to address these issues.

4.1 Data findability

Every dataset that will be shared within the consortium or which will be made publicly available, will be accompanied by metadata. All datasets will also follow the respective standards. For example, AIS datasets provided by MarineTraffic, and other similar data (e.g., data from IHS databases) where applicable, will follow the IMO's specifications and standards.

4.2 Data accessibility

For every dataset used within the consortium (or shared publicly), clear access protocols will be defined. For example, MarineTraffic AIS raw data will be exclusively processed by MarineTraffic, or similar ship data provided by NTUA through the HIS databases. More specifically on AIS raw data, since AIS data are part of MarineTraffic's commercial activities before and beyond the scope of the project, their access will be limited to MarineTraffic. However, the output of processing will be shared with the consortium. Samples of data will also be shared as open datasets so that they can be used by the research community. Detailed access protocols (e.g., licenses) will be defined for these samples.

4.3 Data interoperability

To facilitate the integration of datasets used in the project with other datasets, common vocabularies and terminology should be used. To achieve this, project partners will make sure that standard formats will be employed for each dataset and a common vocabulary and terminology will be developed, where applicable. In addition, for the interoperability of datasets across in-project coding, project partners will ensure appropriate file types.

4.4 Data reusability

To improve the reusability of the datasets used and developed in the consortium, each dataset will be accompanied by detailed documentation, describing the data attributes as well as a license defining access rights to the datasets. As mentioned, all datasets will be accompanied by metadata. Examples of metadata can be data owner, file name, version, file type, classification (unrestricted/restricted), project code, acronym and timestamp. In cases where not all metadata are defined by the file name, additional metadata will be documented as text in a separate document (ReadMe-file), such as details regarding software or procedures for access for each data set.

[11] Guidelines on FAIR data management in Horizon 2020, Version 3, July 2016. [Online] Available at: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

5. ROLES AND RESOURCES

The Role of the Data and Ethics Manager has been assigned to this deliverable's leader (DL), MarineTraffic (MT). The responsibility will be fulfilled in collaboration with the Project Coordinator (PC), who is also the Data Controller.

For the duration of the project, partners will cover curation and storage costs through their own infrastructure. For AIS datasets, which are a main aspect of the project's studies, they are preserved and curated in accordance with best practices for database management systems by MT, as recommended by the respective vendors and open-source communities. Other datasets, related to other partners, are following similar standards.

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

6. DATA SECURITY

For the duration of the project, partners will use their own infrastructure for data storage. In general, it is the policy of the project that project data should be stored safely with automatic backup. Collaboration and sharing within the project are achieved through Microsoft 365 collaboration tools like MS Teams. More specifically, the shared data that will be made available for the long-term by deposition in trusted, free-of-charge data-sharing platforms offered by partners (i.e., MS Teams by NTUA), and on the project's website (i.e., the MSSO website), when it will be developed and deployed.

Data shared or stored through the MS Teams benefit from Microsoft Defend, which provides protection against threats from links (URLs), files or collaboration tools. For the data stored and used internally (and exclusively) by partners, regarding the project, each partner is responsible for data security aspects. Similar (or of higher data security standards) infrastructure is proposed for the project.

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

7. ETHICS AND LEGAL COMPLIANCE

When feedback needs to be collected from experts, users or other stakeholders, data collected will undergo anonymization procedures and only aggregated results of responses on questionnaires and interviews are included in the respective reports and deliverables, especially when these are publicly shared. Stakeholders participate in the collection of such data on a voluntary basis after having obtained a complete description about the project, its objectives and vision. All participants need to first provide their explicit consent by signing respective consent forms in compliance with existing regulatory frameworks. In general, any relevant activity will be carried out in compliance with international and national legislation but also the European Charter of Fundamental Rights and several of its principles relevant to this project (Art. 3, 7, 8, GDPR regulation & Regulation 2016/67).

Hard or electronic copies of gathered data are collected in person, or electronically, by a responsible GreenMED consortium member (also mentioned on the signed consent form) and are safely kept in shielded envelops or password protected files. It is the responsibility of the GreenMED partner, affiliated with the corresponding individual, to ensure abidance by the relevant regulatory frameworks until data comes at the possession of the data controller partner. The Coordinator receives these copies in person, or electronically, before the end of the project. In case this is not possible for hard copies, it is the responsibility of the respective GreenMED researcher to create password protected files of electronic copies of all data that remain at their possession and communicate them via secure partner-specific institutional repositories until these data come at the possession of the data controller. As soon as the data comes into the possession of the data controller, the GreenMED partner erases all collected data.

The data then moves to the premises of the data controller, where only electronic copies are created and are kept at a server without internet connection. The data controller has secure access to these data granted by the coordinator. Each set of data items (consent form, questionnaires, recorded interview or responses to surveys) for a single participant is assigned a code to identify their data after anonymization procedures. This is necessary, for instance, so that an expert user's data can be withdrawn and erased upon their request. This code is made known to the corresponding questionnaire participant at the time of data collection. The key-file containing identity information is kept separately from the de-identified, pseudonymized parts of the data on the same server. The data containing information about the participants' identity will be stored in a secure file to which only the data controller will have access.

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

8. CONCLUSIONS AND DISCUSSION

GreenMED project aims at setting up the MSSO by utilizing valuable data and insights and focuses on the region-specific aspects of the energy demand in Mediterranean maritime transportation and the maritime fuels supply chain. Energy-efficient solutions and decarbonization measures, which impact the ship energy demand, are incorporated into the project's research, forming the basis for its methodology and ensuring reliable results. The proper selection, systematic collection, and classification of data are highlighted as crucial tasks for project success.

This deliverable outlines the datasets that will be collected and used throughout the lifecycle of the project and describes the initial steps that will be followed for pertinent data management purposes, adhering to the FAIR principles. This document constitutes the first version of deliverable D1.3 “Data Management Plan” of Work Package 1 (WP1). The deliverable might be revised and updated, based on the needs of the project and the data challenges through its duration. A possible updated version will be re-submitted no later than the end of the project (M24).

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925

REFERENCES

Not applicable

The project has received funding from the European Maritime, Fisheries and Aquaculture Fund (EMFAF) under grant agreement No. 101124925