REPRESENTATIVE DATASET
EO4SEM

Prepared by: GMV
FMI
GHGSat
LSCE
ONERA
Science Partners

Code: D2 - Representative Dataset
Version: 1.0
Date: 2025/10/29
Internal code: GMV 26097/25 V1/25

THIS PAGE IS INTENTIONALLY LEFT BLANK

APPROVAL

Title: Representative Dataset	
Issue Number: 1	**Revision Number: N/A**
Author(s): EO4SEM Team	**Date:** 2025/10/29
Approved by: GMV	**Date of Approval:** 2025/10/29

CHANGE LOG

Reason for change	Issue	Revision	Date

CHANGE RECORD

Issue Number:	**Revision Number:**		
Reason for change	**Date**	**Pages**	**Paragraph(s)**
N/A			

DISTRIBUTION

Name/Organisational Unit:

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The maritime transport sector is pivotal to global trade, facilitating 90% of goods movement worldwide and underpinning critical industries like tourism and fisheries. Within the European Union (EU), it supports 75% of international trade, making it integral to the EU's economic framework. However, the sector is a significant contributor to environmental challenges, responsible for 3-4% of global anthropogenic CO_2 emissions and 14% of the EU's transport emissions (~135.9 million tonnes in 2022).

To mitigate this impact, the EU has introduced ambitious regulatory frameworks under the Green Deal[1] and Fit-for-55[2] package. These include the extension of the Emission Trading System[3] (ETS) to maritime transport, the FuelEU Maritime regulation[4], and revisions to the Monitoring, Reporting, and Validation [5](MRV) regulation. Together, they aim to:

- Establish carbon pricing for maritime emissions starting in 2026.
- Reduce the Greenhouse Gas (GHG) intensity of ship energy use.
- Integrate stricter monitoring for GHG such as methane (CH_4) and nitrous oxide (N_2O).

Despite these initiatives, effective implementation hinges on robust, independent data collection and monitoring capabilities to ensure compliance and track progress towards emission reduction targets.

To this end, key priorities include monitoring individual vessel emissions (CO_2, CH_4, and N_2O) per voyage, aggregated emissions at regional and national levels, and fugitive emissions from ships. Additionally, there is a need to track other contributors to global warming, such as black carbon and Volatile Organic Compounds (VOCs), to support comprehensive enforcement of EU and international policies.

Given the existing regulatory context, in the context of EO4SEM, Earth Observation (EO) data is being leveraged to assess the potential of current and future EO instruments (and requirements for the future EO instruments) to support such an improvement of the knowledge on shipping emissions.

EO data is being used as a complementary source to existing engine-driven emission models, such as FMI's STEAM v4, which relies on AIS (Automatic Identification System) data streams to estimate emissions at ship level. While effective, these models are limited by AIS availability and cannot account for non-reporting ships, such as those with disabled AIS or naval vessels.

EO imagery can help bridge these gaps by providing independent data for emission monitoring and validation, and in the scope of this project, is being used with two interconnected goals in mind:

- TROPOMI (on-board Sentinel-5), PRISMA, EnMAP and GHGSat observations (together composing the reference dataset) are being processed with inversion models to refine emission estimates for pollutants and GHGs (NO_x, CO_2, and CH_4) over the sea. These emission estimates are being aggregated temporally and spatially to create accurate maps of ship-driven pollutants and GHGs across maritime corridors, local regions and ports.
- To create a reference dataset and the respective new developments and methodologies to be used as benchmark, helping to guide the development of technical specifications and consolidate the added value that the Copernicus Expansion Missions could represent in the future for this high priority area, more precisely CO2M and CHIME.

To support stakeholders' needs, the EO-driven approach will provide a benchmark for automated and scalable emission analysis, reducing manual effort while increasing reliability. It will also help assess emission patterns near ports and industrial areas, even if as a secondary objective. Finally, the entire service will be deployed on an ESA Network of Resources (NoR) platform, ensuring interoperability and ease of access for users.

[1] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en
[2] https://www.consilium.europa.eu/en/policies/fit-for-55/
[3] https://eur-lex.europa.eu/eli/reg/2023/957/oj
[4] https://eur-lex.europa.eu/eli/reg/2023/1805/oj
[5] https://climate.ec.europa.eu/eu-action/eu-emissions-trading-system-eu-ets/monitoring-reporting-and-verification-eu-ets-emissions_en

1.1. SCOPE

This document, Deliverable 2 – Representative Dataset, provides an overview of the EO4SEM representative dataset, including information on the data there contained and as well, instructions on how to access it.

1.2. REFERENCE DOCUMENTS

The following documents, although not part of this document, amplify or clarify its contents. Non-academic documents are referenced in the form [RD.x]. Academic references are listed in **Error! Reference source not found.**.

Table 1.1: Reference documents

Ref.	Document ID.	Title	Rev.
[RD1]	ESA Contract 4000145881/24/I-DT-bgh – EO4SEM_signed	ESA contract No. 4000145881/24/I-DT-bgh with GMV Skysoft S.A. Sentinel Users Preparation (SUP) initiative - [SUP-1] - Applications Preparedness with stakeholder and end-users participation. Project: EO for Ship Emission Monitoring (EO4SEM).	NA
[RD2]	GMV 10933/24 V1/24	Full Proposal for EO4SEM Project – "EO for Ship Emission Monitoring".	1.0

1.3. ACRONYMS

Table 1.2: List of acronyms

Acronym	Description
AIS	Automatic Identification System
AWS	Amazon Web Services
CNMF	Constrained Nonnegative Matrix Factorisation
CDS	Copernicus Climate Data Store
CH_4	Methane
CMEMS	Copernicus Marine Environmental Monitoring Service
CO	Carbon Monoxide
CO_2	Carbon Dioxide
EnMap	Environmental Mapping and Analysis Program
EO	Earth Observation
ETS	Emission Trading System
EU	European Union
FMI	Finnish Meteorological Institute
GHG	Greenhouse Gas
MRV	Monitoring, Reporting, and Validation
NoR	Network of Resources
N_2O	Nitrous Oxide

Acronym	Description
NO_x	Nitrogen Oxides
ppm.m	parts per million by mass
PRISMA	PRecursore IperSpettrale della Missione Applicativa
RD	Reference Document
SO_x	Sulphur Oxides
SRON	Space Research Organisation Netherlands
STEAM	Ship Traffic Emission Assessment Model
SWIR	Short-Wave InfraRed
TROPOMI	TROPOspheric Monitoring Instrument
VOC	Volatile Organic Compound
WFMD	Weighting Function Modified Differential

2. REPRESENTATIVE DATASET

This chapter details the datasets contained if the EO4SEM Representative Dataset. This is a foundational component of the project and aims to enhance the measurement of GHG emissions from maritime sources. It will be updated and expanded throughout the project.

The Representative Dataset will integrate data from TROPOMI, PRISMA and EnMap hyperspectral, Sentinel-2, and GHGSat to assess the potential added value of future satellite missions under the Copernicus Expansion programme, specifically CO2M and CHIME.

At this stage of the project, the following datasets are already available:

Dataset	Description	Unit	Parameters/Bands/ Variables
Maritime Emission Methane Monthly Maps (SORN)	Monthly maps of the CH4 a priori and a posteriori maritime emissions over Europe at the 0.5° spatial resolution for the year 2019. The a priori estimates of CH4 anthropogenic emissions in this study are based on emission estimates from the CAMSREG-v8 gridded inventory. We assimilate observations from the TROPOMI-SRON-v2.4 product.	ktCH4/month	CH4_Prior, CH4_Post
Maritime Emission Methane Monthly Maps (WFMD)	Monthly maps of the CH4 a priori and a posteriori maritime emissions over Europe at the 0.5° spatial resolution for the year 2019. The a priori estimates of CH4 anthropogenic emissions in this study are based on emission estimates from the CAMSREG-v8 gridded inventory. We assimilate observations from the TROPOMI-WFMD-v1.8 product.	ktCH4/month	CH4_Prior, CH4_Post
Ship emission estimates produced by STEAM Model	This dataset provides gridded binary data at an hourly temporal resolution and at a finer spatial resolution of 0.05° × 0.1° for the year 2023, covering the European maritime domain (15.25°W to 40°E; 25°N to 74.25°N). In principle, STEAM produces detailed ship emission estimates by integrating vessel activity data from the Orbcomm Ltd AIS archive with technical specifications from the S&P Global databases. It models ship fuel consumption and emissions by incorporating its internal database of ship-specific parameters, such as engine load and abatement technologies, and environmental influences, including wind, sea ice, and ocean currents from the Copernicus Climate Data Store (CDS) and Copernicus Marine Environmental Monitoring Service (CMEMS). Recent refinements of STEAM include updated emission factors for NOx, CO2, SOx, PM2.5, and alternative fuels like methanol and ammonia, ensuring alignment with evolving maritime regulations and technologies.	KG	CH4(Methane), CO (Carbon monoxide) NOx (NOx), CO2 (Carbon dioxide)
GHGSAT Methane plume concentration	The CH4PL GeoTIFF shows the methane concentration reported in the CH4 layer, only at the location where a plume was detected.	parts per billion (PPB)	CH4 anomaly
Shipping Lanes Binary Masks	This dataset provides binary masks for all defined shipping lanes derived from STEAM 2023 dataset through geospatial analysis of maritime traffic patterns. Each lane is represented by a variable 'shipping_lane_X_mask' (e.g., 'shipping_lane_17_mask', 'shipping_lane_18_mask'), where 1 indicates pixels within the lane and 0 elsewhere. These masks support spatial analysis and emission calculations for maritime routes.	Unitless (boolean)	Shipping_lane_X_mask (X = 1 to 59, e.g., shipping_lane_17_mask, shipping_lane_18_mask)

Dataset	Description	Unit	Parameters/Bands/ Variables
Images of NO2 plumes from individual ships	Each image file contains the TROPOMI data to visualize an individual ship plume. The filename follows the format plume_YYYY_MM_DD_XXX.nc where YYYY_MM_DD corresponds to the date of detection of the plume and XXX is a unique incremental index for that date.	molec/cm²	latitude, longitude, NO2, latitude_bounds, longitude_bounds
Marine Regions	These regions are often used in conjunction with national boundaries, such as Exclusive Economic Zones (EEZs), to delineate specific areas	Marine Regions Gazetteer's Marine Region ID (MRGID)	Geojson
EnMap SWIR Radiance	Image extracted from EnMap database	W/m2/st/µm	Radiance
EnMap CH4 Detection	Methane detection : Value of Methane spectral correlator applied to data pixel by pixel. False alarms can occured.	Value from 0-1	Detection map
EnMap CH4 quantification (ppm.m)	When detection mask are applied to methane correlation mask, an integrated column concentration (ppm.m) is then estimated. This value is assumed to be an enhancement compared to the local background.	ppm.m	Methane concentration

3. ACCESSING THE REPRESENTATIVE DATASET

3.1. AWS S3 BUCKET

The Representative Dataset is hosted on Amazon Web Services (AWS) via an S3 bucket, offering direct cloud-based access to raw and processed data. This method is ideal for users who prefer programmatic access or need to integrate the dataset into automated workflows using tools like Python, R, or cloud computing platforms.

Step 1: Use the credentials to access the AWS cloud dashboard.
S3 web link:https://991490111768.signin.aws.amazon.com/console
Username: IAMAdminUser.
Password : It will be shared securely via email. Please contact the project manager at fbrandao@gmv.com.

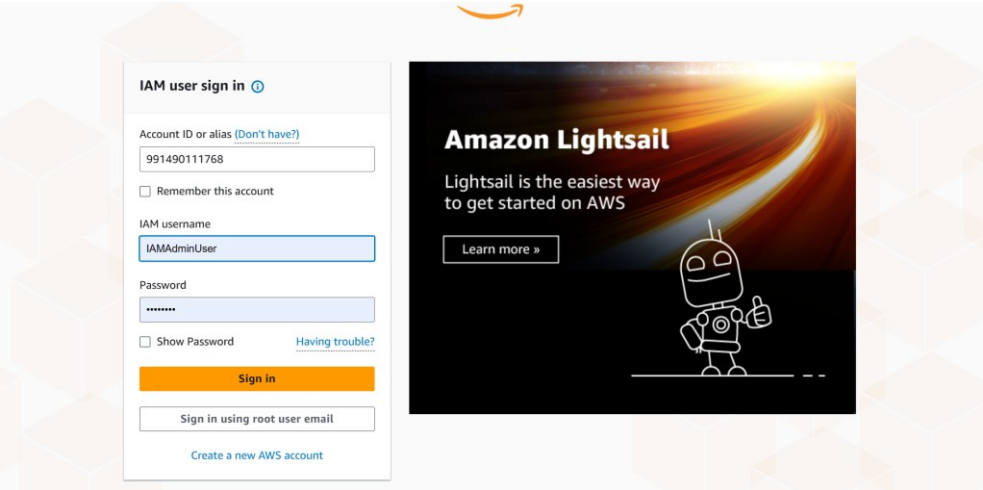

Figure 3.1: Accessing the AWS cloud dashboard

Step 2: Access the S3 bucket using the search tab.

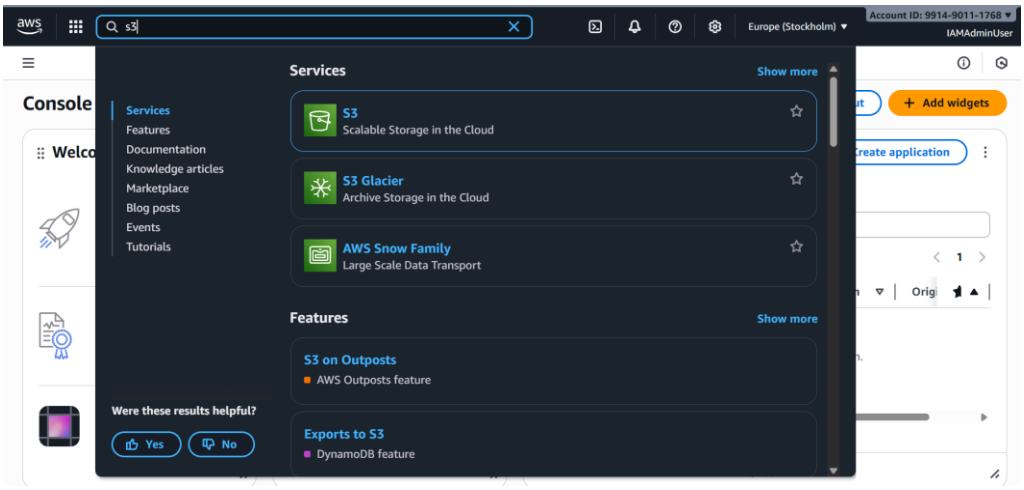

Figure 3.2: Accessing the S3 bucket using search tab

Step 3: Access the eo4sem S3 bucket and download the data by selecting from various products as displayed below.

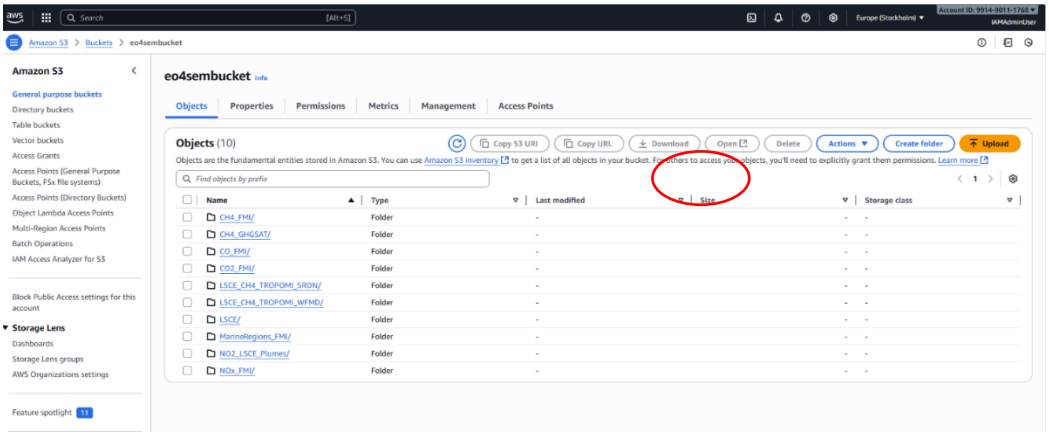

Figure 3.3: Datasets in S3 bucket

3.2. ESA APEX GEOSPATIAL EXPLORER

The ESA APEX EO4SEM Geospatial Explorer (https://explorer.eo4sem.apex.esa.int/) is a web-based visualization tool provided by ESA that allows users to interactively browse, explore, and analyse the Representative Dataset.

It's particularly useful for researchers and policymakers who want to inspect spatial patterns, zoom into specific regions, and overlay additional geospatial layers without needing to download the data. However, if user can download the Geotiff images by the clicking the Download icon under each layer as shown the screenshot below:

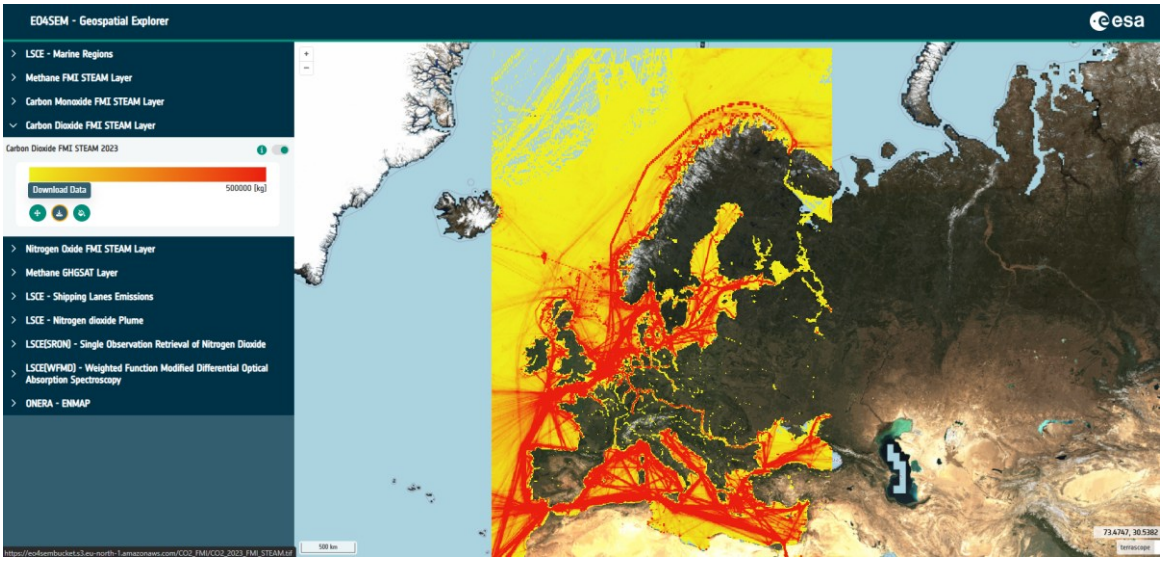

Figure 3.4: EO4SEM Geospatial Explorer

3.3. PRODUCT CATALOGUE

The Product Catalogue offers a structured and searchable interface to discover metadata and download specific components of the Representative Dataset. Users can filter by region, time period, data type, and other parameters, making it a convenient entry point for targeted data retrieval and documentation.

Step 1 Access the EO4SEM data catalogue https://browser.eo4sem.apex.esa.int/?.language=en

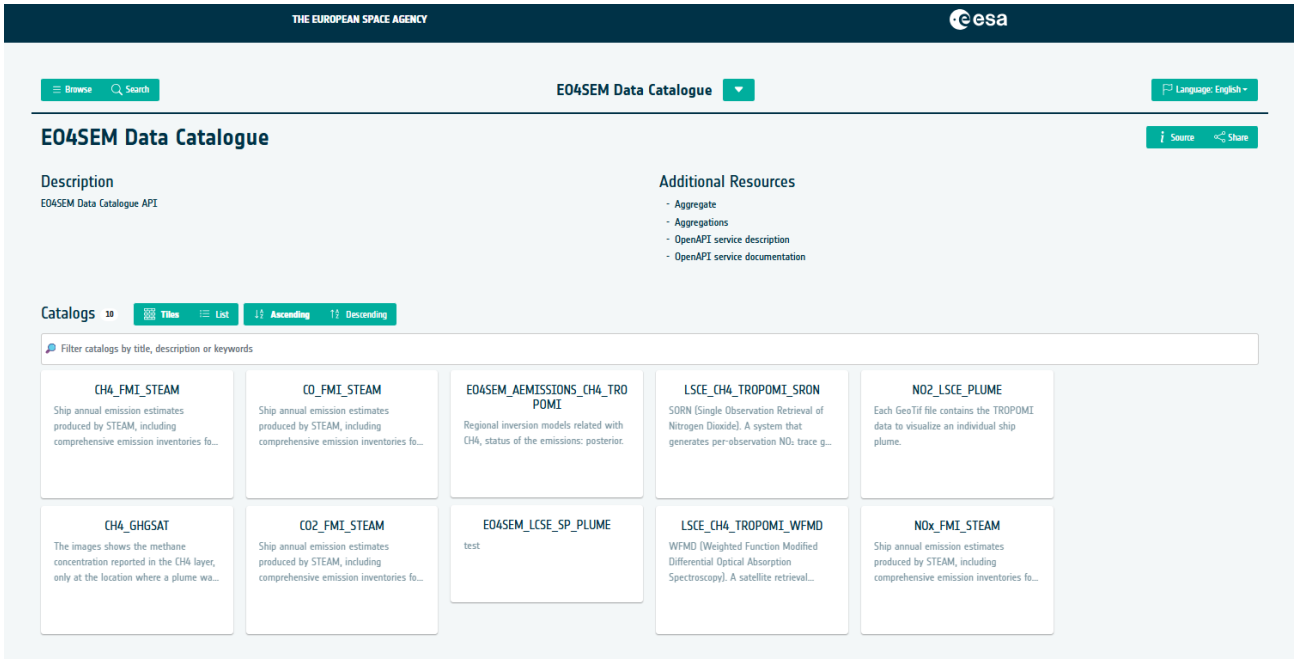

Figure 3.5: Catalogues available in the EO4SEM Product Catalogue

Step 2 Click on any data thumbnail to explore. Here we selected **EO4SEM_AEMISSIONS_CH4_TROPOM**

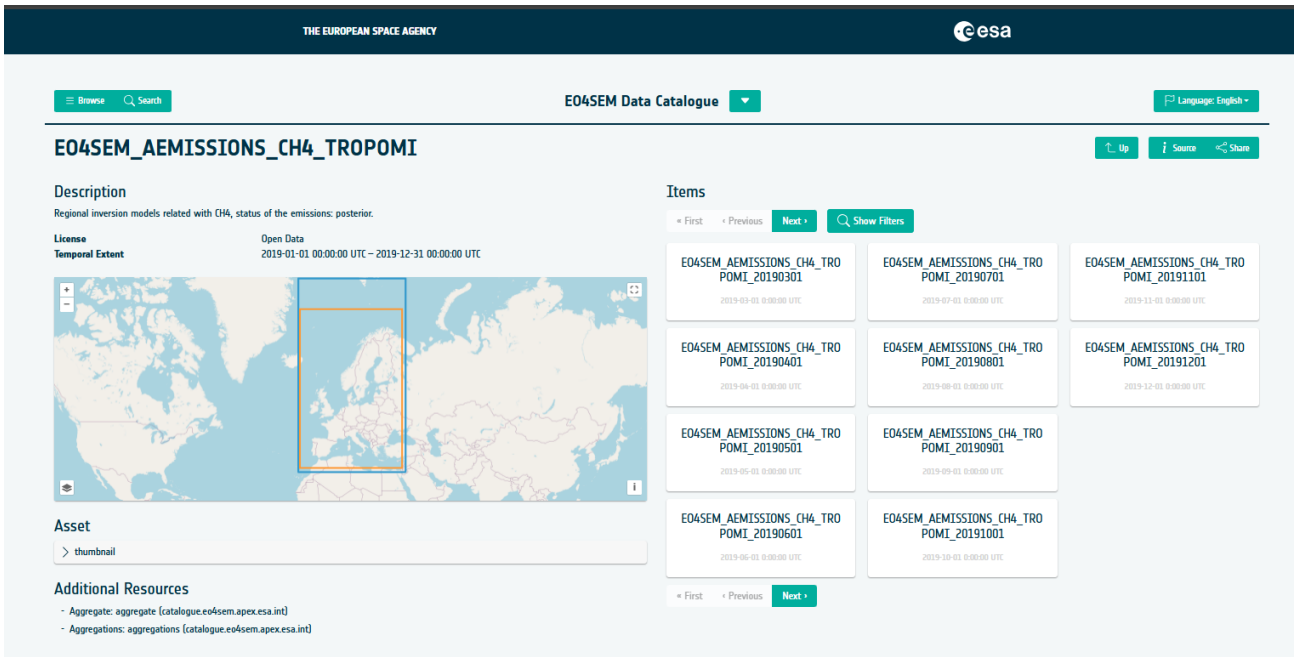

Figure 3.6: Details of a given catalogue

Step 3 Below is a screenshot of the catalogue showing monthly time-series of CH4 from TROMI data. Click one of them to explore and download.

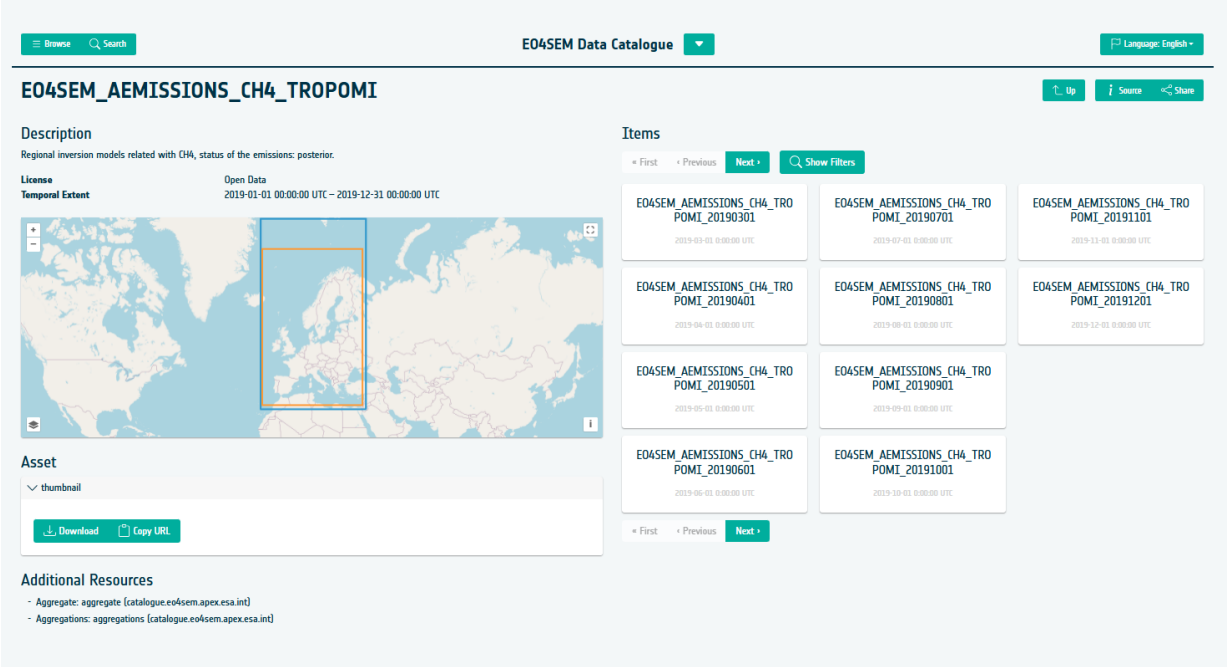

Figure 3.7: Details of a time-series catalogue

Step 4 Use the drop-down under Asset to display metadata of the raster imagery and use the Download icon to get the data in Geotiff format.

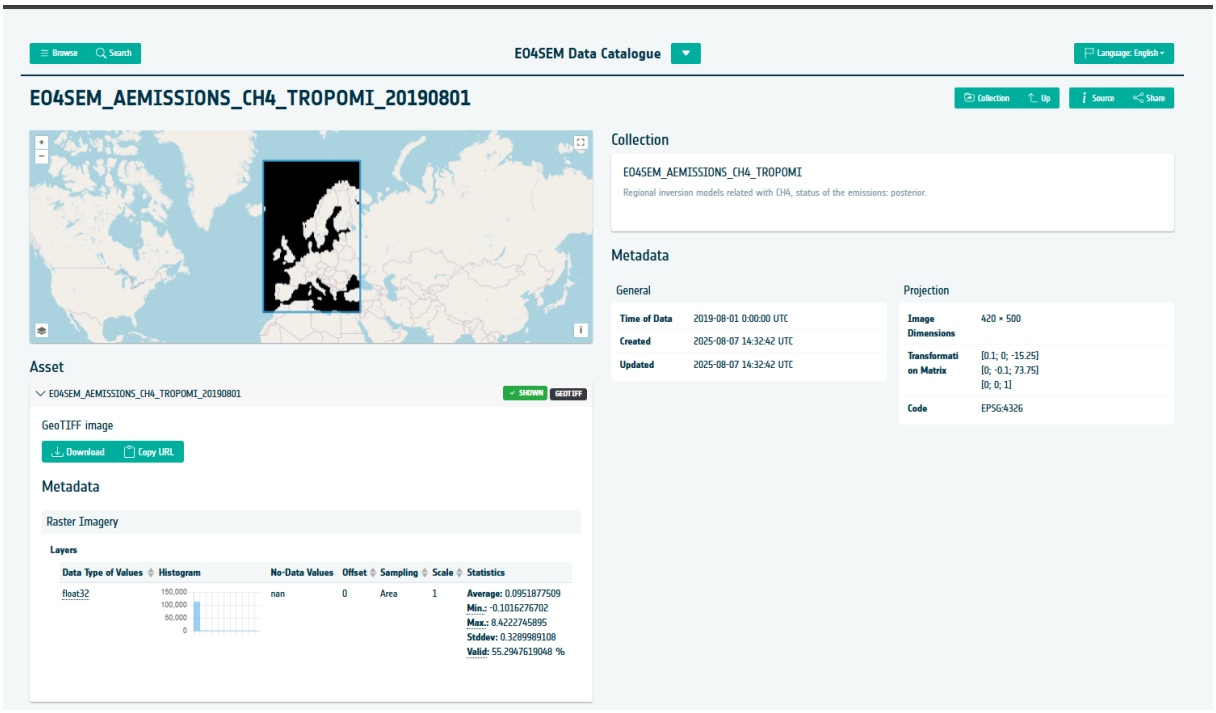

Figure 3.8: Metadata of a given collection

END OF DOCUMENT