Quantifying Greenhouse Gas (GHG) Emissions Associated with Global Seafood Production

Technical Documentation

A capstone project submitted in partial satisfaction of the requirements for the degree
of Master of Environmental Data Science for the
Bren School of Environmental Science & Management

By

CARMEN HOYT
JOSHUA MULL
NICOLE PEPPER
STEPHEN CARROLL

Committee in charge:
CARMEN GALAZ GARCÍA
GAVIN MCDONALD

May 2025

Signature Page

Seamissions

As developers of this Capstone Project documentation, we archive this documentation on the Bren School's website such that the results of our research are available for all to read. Our signatures on the document signify our joint responsibility to fulfill the archiving standards set by the Bren School of Environmental Science & Management.

Carmen Hoyt

Joshua Mull

Nicole Pepper

Stephen Carroll

The Bren School of Environmental Science & Management produces professionals with unrivaled training in environmental science and management who will devote their unique skills to the diagnosis, assessment, mitigation, prevention, and remedy of the environmental problems of today and the future. A guiding principle of the School is that the analysis of environmental problems requires quantitative training in more than one discipline and an awareness of the physical, biological, social, political, and economic consequences that arise from scientific or technological decisions.

The Capstone Project is required of all students in the Master of Environmental Data Science (MEDS) Program. The project is a six-month-long activity in which small groups of students contribute to data science practices, products or analyses that address a challenge or need related to a specific environmental issue. This MEDS Capstone Project Technical Documentation is authored by MEDS students and has been reviewed and approved by:

Carmen Galaz García

DATE

Gavin McDonald

Acknowledgements

We would like to extend our deepest gratitude to our clients and advisors whose support and expertise have made this project possible. Your willingness to guide us, share ideas, provide insightful feedback, and commit precious time has been invaluable in ensuring that our project yields meaningful results while reflecting thoughtful analysis.

We are especially grateful for our client and faculty advisor:

Gavin McDonald, Senior Project Scientist, Environmental Markets Lab (emLab), Bren School of Environmental Science & Management

We also thank the following individuals for their support:

Bren School:
Dr. Steve Gaines, Dean & Distinguished Professor;
Dr. Carmen Galaz García, Assistant Teaching Professor;
Dr. Bruce Kendall, Professor & Associate Dean;
Dr. Sandy Sum, Teaching Assistant

emLab:
ES Burns, Project Scientist;
Pol Carbo Mestre, Data Scientist

Global Fishing Watch:
Tyler Clavelle, Senior Data Scientist

Abstract

Global fisheries are heavily reliant on fossil fuels, contributing significantly to the rise in global greenhouse gas (GHG) emissions driving climate change. While satellite technology is commonly used to monitor land-based emissions (and ocean-based emissions of shipping vessels), studies primarily estimating ocean-based emissions remain limited in the fishing sector. In collaboration with the Environmental Markets Lab (emLab) and Global Fishing Watch (GFW), this project leverages novel, high-resolution, satellite-based datasets to provide precise insights into the GHG emissions associated with global fisheries. We develop a reproducible, extensible, and open-source data processing pipeline to connect emissions data with seafood production data, along with an interactive dashboard to explore the resulting dataset. Our findings will enable novel research opportunities, offer actionable data to identify major GHG contributors, and facilitate new policy and market-based interventions to reduce fisheries-related emissions at scale.

Table of Contents

1) Executive Summary

Food systems contribute roughly a quarter of GHG emissions, but the contributions from seafood remain poorly understood. Estimates of fishery emissions have relied on only a few case studies and broad generalizations (Parker et al., 2018). Increasingly, Automatic Identification System (AIS) is being used to estimate emissions from industrial vessels (IMO 2020), but this approach has not yet been used to specifically analyze fisheries-related emissions on a global scale. However, roughly 75% of industrial vessels operate without broadcasting their location with AIS, meaning their emissions would not be tracked with any method that relies solely on AIS (Paolo et al., 2024). Existing total emissions for global fisheries are therefore unreliable, making it difficult to identify high-emitting fleets, countries, or regions, or carbon-intensive species (Greer et al., 2019). Utilizing recent advances in remote sensing and machine learning, Global Fishing Watch (GFW) and emLab have generated high-resolution estimates of GHG emissions that includes over 150,000 broadcasting fishing vessels, along with a brand new dataset of emissions from non-broadcasting fishing vessels (Kroodsma et al., 2018; GFW & emLab, 2024). Because these data are not associated with catch data for individual species, it's unclear which fisheries, or species, contribute the most emissions. This project aims to close this knowledge gap by integrating the novel emissions datasets with Food and Agriculture Organization (FAO) Seafood Production data to create the first comprehensive quantification of emissions associated with global seafood production. This will enable market and policy interventions to mitigate the impact of these GHG emissions on the environment (McDonald et al., 2024).

Problem statement or knowledge gap

Previous approaches to assessing GHG emissions from fisheries have relied upon generalized assumptions or limited, fleet-specific case studies. This has led to broad uncertainties about the true emissions of the industry (Parker et al., 2018). Factors such as fishing vessel size, gear type, distance to fishing grounds, and traceability through AIS all contribute to varied emissions estimates for fleets across the globe. The difficulty in modeling these factors contributing to emissions has led to overly simplified models which have yielded coarse and potentially misleading estimates of fisheries-related emissions, masking significant variability among fleets and target species (Parker et al., 2018).

A critical knowledge gap persists due to the scarcity of high-resolution, spatially explicit, and vessel-level emissions data, making it challenging to pinpoint specific vessel types, regions, or countries responsible for disproportionately high emissions (Kroodsma et al., 2018). Without more detailed data, policymakers and fisheries managers cannot effectively identify or regulate the most impactful fisheries operations, inhibiting potential strategies to mitigate the industry's climate impacts (Paolo et al., 2024) while potentially alleviating stress on fisheries at the same time. Additionally, without detailed emissions data, potential market-driven interventions such as low-carbon seafood labeling or certification schemes are limited, as these approaches depend on transparent, verifiable emissions data to inform consumer choices and incentivize sustainable practices (McDonald et al., 2024). Addressing these critical data gaps is crucial to facilitate policy measures and market-based solutions aimed at reducing seafood-based carbon emissions.

Objectives

The primary objective of this project is to develop a reproducible pipeline to quantify GHG emissions associated with global seafood production, linking fishing vessel emissions data to species-specific FAO seafood production statistics. This dataset aims to provide emissions estimates for nine GHG and non-GHG pollutants by FAO region, year, country, and species. Additionally, the project seeks to enhance the usability and accessibility of these data through an interactive dashboard, enabling targeted regulatory, policy, and market-based interventions to reduce the carbon footprint associated with global seafood production.

Products & Deliverables

The deliverables for the project include:

- **Emissions Processing Pipeline:** Reproducible, extensible, and open-source data processing pipeline to estimate GHG emissions from global seafood production by harmonizing emLab's high-resolution vessel emissions data, including both AIS-broadcasting and non-broadcasting vessels, with FAO catch records.
- **Seamissions Dashboard:** Interactive public dashboard to visualize the relationships between GHG emissions and fleets, flag, catch species, and FAO region.
- **Results Report:** Comprehensive written assessment comparing the project's emissions estimates with existing published data.

As open-source products, our deliverables can be viewed and utilized by the public, research communities, and other working groups who may wish to build upon our findings. The dashboard can be used as an exploratory tool for policymakers and regulators interested in developing emissions reduction strategies as well as for consumers concerned with selecting more sustainable seafood products. Seafood certification or labeling bodies may be interested in using our findings to incorporate emissions standards into their programs.

2) Approach

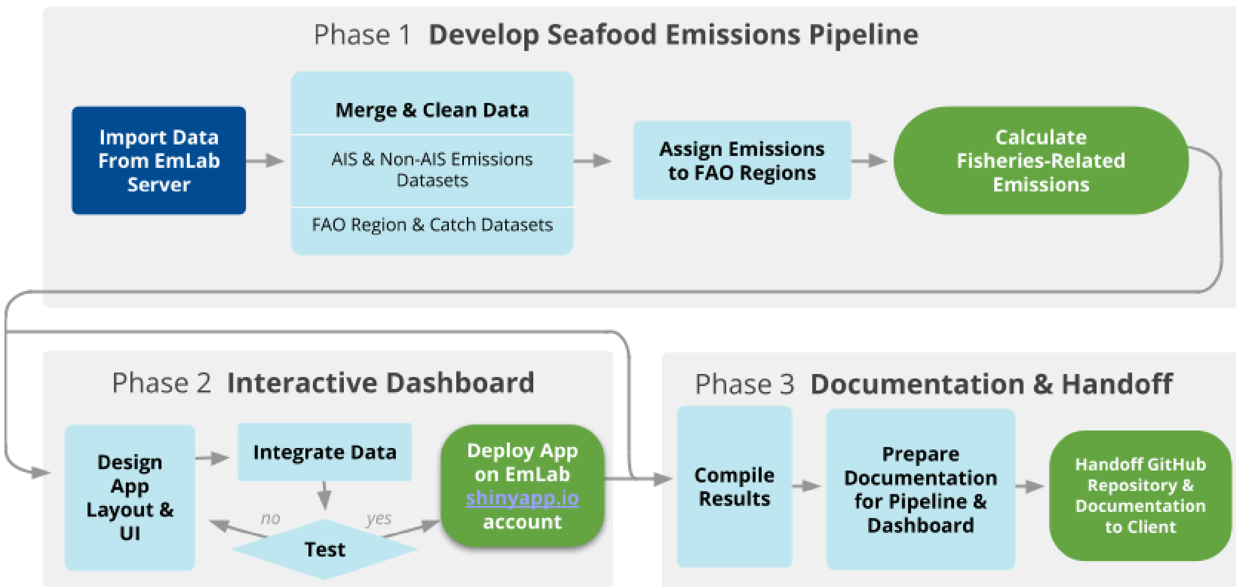

Figure 1. The three-phase approach to the analysis.

This project develops a Seafood Emissions Pipeline by importing and merging broadcasting and non-broadcasting emissions datasets from the emLab server. The merged emissions data are then proportionally distributed by FAO region, year, country, and species, enabling the estimation of fisheries-related emissions. The project integrates the resulting dataset into an interactive dashboard. After successful testing, the dashboard will be deployed using emLab's shinyapps.io account. Finally, the project results will be compiled, and comprehensive documentation for both the data pipeline and interactive dashboard will be produced. This documentation, the corresponding products, and the GitHub repository will be delivered to our clients.

3) Methods

This analysis was conducted in R. For a full description of each package, see Appendix B. For a full description of each dataset, see Appendix C.

3.1) Seafood Emissions Pipeline

Packages: {targets}, {tidyverse}, {janitor}, {here}, {lubridate}, and {sf}.

The Seafood Emissions Pipeline comprises many smaller, computationally demanding, and time-consuming steps. Therefore, the {targets} package was used to help manage the development and implementation of the pipeline by breaking down each step into a separate function, documented in the "R/functions.R" file. The "_targets.R" file was used to assign each function an input and output so that changes could be isolated to a particular step. Each time the pipeline runs, {targets} assesses any changes to inputs and outputs, running only steps that have been updated. This avoids unnecessarily re-running the entire pipeline every time a change is made, cutting down on time and helping to ensure the expected inputs and outputs are consistent, maintaining reproducibility.

The {tidyverse}package was used for data wrangling and cleaning as well as plotting.
The {janitor} package was used to convert all dataframe column names to snakecase, allowing for consistent naming conventions. The {here} package was used to assist in loading local files.
The {lubridate} package was used for working with date-time objects.
The {sf} package was used for geospatial analysis, namely the emissions data.

All analyses were performed using R and hosted on the emLab server. An "emissions-pipeline" was created on GitHub for version control and collaboration, and team members worked from individual branches.

In addition to the datasets described in Appendix C, two keys were built for the analysis and included in the emissions-pipeline repository under `data-keys/`:

1. `Full_species_key.csv`

The full_species_key.csv was adapted from the master_species_key.csv from Danielle Ferraro and Gordon Blasco by adding about 1,000 rows from the FAO species groups .csv for species that were recorded in the final FAO dataset but not in the key. The resulting .csv was saved to the repo for future reference.

2. `Flag_key.csv`

The flag key was manually created by assigning ISO3 codes to countries where spelling or special character use was not equivalent between FAO and SAU datasets. The resulting .csv was saved to the repo for future reference.

3.1.1) Processing

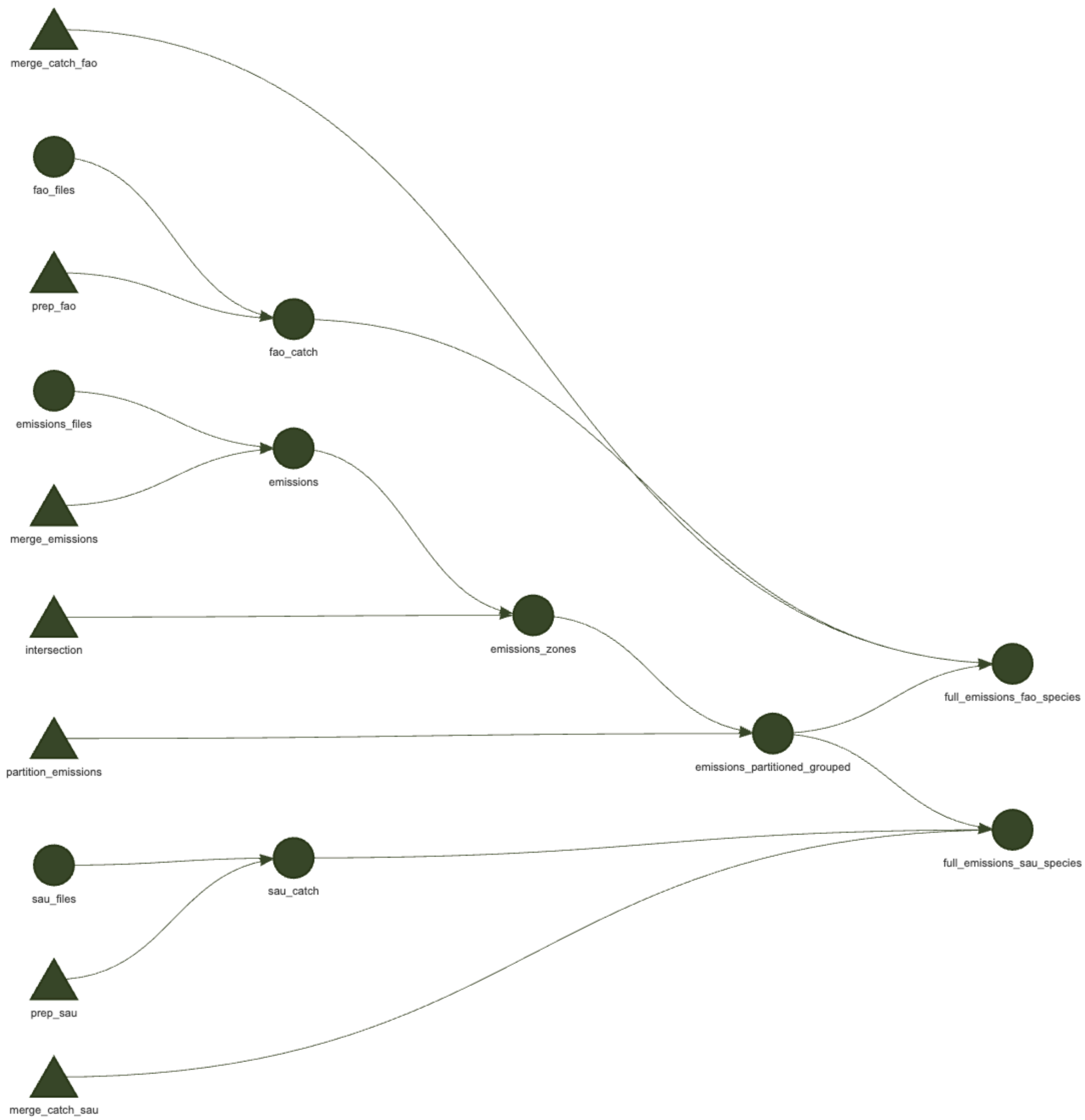

Figure 2. The target network for the emissions pipeline.

Emissions datasets were read into the pipeline, the column names were converted to snake case, and a new `year-month` column was created for both datasets. In the broadcasting dataset, NA values in the `flag` column are filled with "UNK" to represent flag unknown. Additionally, `vessel_class` was filtered for fishing gear types of high confidence of correct identification ("fishing", "squid_jigger", "drifting_longlines", "pole_and_line", "other_fishing", "trollers", "fixed_gear", "pots_and_traps", "set_longlines", "set_gillnets", "trawlers", "dredge_fishing", "seiners", "purse_seines", "tuna_purse_seines", "other_purse_seines", "other_seines", and "driftnets").

In the non-broadcasting dataset, emissions estimate columns for each of the nine pollutants (CO_2, CH_4, N_2O, NOX, SOX, CO, VOCS, PM2.5, and PM10) are renamed to match the broadcasting dataset, and a `flag` column is created and populated with "DARK" to distinguish non-broadcasting emissions from the broadcasting emissions. Then, the datasets were concatenated.

A `year` column was created, and the combined dataset was filtered to 2016-2024 to match the available data for the non-broadcasting dataset. Emissions estimates are then aggregated (summed) by year and flag for each one-by-one degree pixel (distinguished by `lat_bin` and `lon_bin`). Then, spatial attributes (points) were created for each `lat_bin` and `lon_bin` in the WGS coordinate reference system (unit: degrees). An empty grid was generated from the point geometry, the emissions data were joined back to the empty grid, and the geo-dataframe was transformed to Equal Earth projection. Every grid cell was assigned a unique ID. Using the FAO shapefile, an intersection was run on the emissions grid cells to assign each to an FAO region. Some grid cells overlapped multiple regions, resulting in multipolygons for those grid cell IDs. Multipolygons were broken down into individual sub-polygons. The area was calculated for each sub-polygon, and the individual sub-polygon areas were summed for each grid cell ID. Emissions from each grid cell ID were partitioned out based on the proportion of sub-polygon area to total grid cell area associated with each grid cell ID. The emissions partitioning was validated using a check to trigger a warning if more than 0.001% of emissions were lost in comparing the total emissions estimates before and after partitioning. Some emissions are expected to be lost due to floating point error and rounding, and 0.001% was arbitrarily selected as a threshold (though the actual number of lost emissions is likely much smaller).

3.1.2) Merging

The FAO seafood catch data was imported and assembled by joining flag and species codes to the catch data and filtering out mammals, plants, reptiles, corals, and freshwater species (for more information, see section 4.4), with species distinguished by a unique numeric code in the `identifier` column. For each FAO Major Fishing Area of interest (zones 18, 21, 27, 31, 34, 47, 48, 51, 57, 58, 61, 67, 71, 77, 81, 87), non-broadcasting emissions were partitioned out amongst each fishery (flag-species combination) that reported FAO catch by proportion of catch (in metric tonnes) for that region. To gain species-level resolution of broadcasting emissions, for countries that reported both broadcasting emissions and catch, broadcasting emissions were proportionally distributed across species by catch (in metric tonnes); otherwise, broadcasting emissions were left as one total for a species of 'no_fao_species_catch_data' value. These broadcasting emissions were fully joined with the non-broadcasting emissions because some flags overlapped both FAO catch (receiving non-broadcasting emissions) and broadcasting emissions, while others were either broadcasting only (no reported catch) or non-broadcasting only (reported catch but no associated broadcasting emissions). For countries with both broadcasting and non-broadcasting emissions, estimates for each pollutant were summed. The tables for each FAO region were concatenated into one final dataset, assigning emissions estimates for each `zone`, `year`, `flag`, and `species` (with species as 'no_fao_species_catch_data' for countries with no reported catch but broadcasting emissions) with associated catch (in metric tonnes) for each of the nine pollutants. The full_species_key.csv (key 1) was then joined to the dataset by `identifier`, providing more information about each species, such as common name, scientific name, and FAO International Standard Statistical Classification of Aquatic Animals and Plants (ISSCAAP) group.

The final dataset was used to generate summary tables and figures for the final report. This includes trends in global emissions by gear type, trends in global emissions by the top 10 flags, trends in global emissions by FAO region, trends in emissions-per-unit-catch by the top 10 flags, trends in emissions-per-unit-catch by ISSCAAP species group, trends in emissions-per-unit-catch by FAO region. The concatenated emissions dataset intersected by FAO regions (but not proportionally distributed) was used to create a map of broadcasting and non-broadcasting emissions (aggregated across all gear types and flags for any single year).

Throughout the pipeline, various checkpoints were used to download and save intermediate datasets. These serve as places to check outputs, but are not necessary for the overall function of the pipeline. There are many lengthy steps, so checkpoints allow for data to be read in at any point to resume analysis without starting over.

3.1.3) Validation

The Sea Around Us (SAU) catch time series data was used to validate our final dataset. SAU data includes underreported, IUU, artisanal, etc., catch that isn't included in FAO reporting. Therefore, it may give us an upper bound of catch to compare against FAO data, which is self-reported and may be more of a lower bound for catch estimates.

SAU data includes catch (in metric tonnes) of species by country and FAO region over time (downloaded from the SAU website by FAO region as .csv files). These data were assembled and processed similarly to the FAO seafood catch data in step 8A.4. The flag_key.csv (key 2) was used to reconcile SAU country with the ISO3 code used by FAO to ensure consistency despite minor differences in spelling. This allowed us to compare emissions per unit catch between the two datasets by flag and by ISSCAAP group. Validation tables for both metrics were created and referenced in our final report.

Additionally, emLab and GFW provided us with an additional dataset (dataset 5 above) that contained `year`, `flag`, top visited country (the country the vessel visited the most times between 2015-2024), and annual CO_2 emissions for every fishing vessel in the larger broadcasting emissions dataset. This dataset contains emissions estimates for all industrial ocean vessels tracked by AIS and is further reduced to vessels exhibiting fishing activity for this analysis. It is assumed that if the top country value is NA for a vessel (potentially due to turning off AIS before entering port), it's the same value as `flag`. This provides some insight as to how many vessels are flagged to other countries (as is the case in using flags of convenience) and explores the potential over-/under-estimation of emissions for a country as a result. To do this, 4 smaller analyses were conducted.

First, the entire dataset (2015-2024) was assessed for match or mismatch between the registered flag and the top-visited country. It was found that 22% of vessels were a mismatch. The second analysis quantified emissions associated with each vessel match or mismatch, attributing about 71% of emissions to mismatched flags, totaling 6,835,022,045 metric tonnes of CO_2 from 2015-2024. The third analysis aimed to assess the overestimation of emissions for each country by quantifying how many vessels associated with their AIS flag had a different top country. Not surprisingly, Panama ("PAN") had the largest amount of emissions CO_2 by weight attributed to mismatched vessels (flagged to Panama but visiting other countries), accounting for over 96% of their reported emissions. The fourth analysis aimed to assess the underestimation of emissions for each country by quantifying how many vessels associated with a particular top country were AIS-flagged to a different country (or countries) to contextualize how many emissions are missing from each country's total estimates. Perhaps also unsurprisingly, China ("CHN") had the largest amount of emissions CO_2 by weight missing from their overall estimates (vessels were visiting China the most but flagged to other countries), accounting for about 75% of their perceived total emissions.

4) Assumptions & Limitations

4.1) Inherent Dataset Assumptions

4.1.1) FAO Dataset

FAO Catch Data refers to the global fisheries and aquaculture statistics compiled by the Food and Agriculture Organization of the United Nations. These datasets serve as a primary source for understanding global fishery production trends, as FAO collects and publishes official fisheries catch reports submitted by national governments. The data includes reported landings from marine and inland fisheries, species-level catch statistics organized by FAO statistical areas, commercial fisheries data from both industrial and artisanal operations, and records from

Regional Fisheries Management Organizations (RFMOs), particularly for highly migratory species such as tuna and billfish

Despite its widespread use, FAO Catch Data has several limitations that can lead to an underestimation of actual fishing activity. Some of FAO assumptions include

- Catch reported by governments is accurate, even though many nations may underreport or manipulate data for political or economic reasons.

- Unreported fisheries do not contribute significantly to total catch, despite growing evidence that small-scale and IUU fishing can be substantial

Even with these assumptions, FAO data is still considered the global gold standard for tracking catch, as is even used as a baseline for organizations like the Sea Around Us (SAU).

4.1.2) SAU Dataset

The Sea Around Us (SAU) is an organization dedicated to assessing the impact of global fisheries on marine ecosystems by expanding upon existing fisheries data. While the Food and Agriculture Organization compiles official catch statistics from national governments, these reports often fail to account for illegal, unreported, and unregulated (IUU) fishing, as well as artisanal and subsistence fishing, discards, and underreported industrial fishing. To address these gaps, SAU reconstructs missing fisheries catch data using alternative sources such as local fisheries reports, historical records, trade and market data, and expert interviews. Their methodology is rooted in the rationale established by Pauly (1998) and first implemented by Zeller et al. (2007), which argues that estimating missing fisheries data is preferable to reporting zeroes. When official reports lack documentation, recording a zero incorrectly suggests that no fishing occurred, leading to misleading conclusions by researchers and policymakers. Instead, SAU provides rough estimates to acknowledge the presence of fishing activity, ensuring that even unreported efforts are accounted for in larger datasets. The reconstruction process relies on a combination of assumptions and extrapolated data sources to present a more comprehensive view of global fisheries, helping researchers understand the full scope of overfishing and ecosystem decline. However, this approach is not without challenges, assumption-based extrapolations introduce uncertainty, data gaps in remote regions make estimations difficult, and conflicts may arise between SAU's findings and government or industry reports. Some big assumptions include:

- Anchor point data is expanded to country-wide estimates, assuming similar patterns across regions and time periods. Anchor points are catch estimates tied to a specific year and sector, often covering only a fraction of a country's coastline rather than its full

Exclusive Economic Zone (EEZ) or Inshore Fishing Area (IFA). While SAU conservatively scales these estimates to national levels with factors like fisher or population density, shelf area, or relative IFA size, this cautious approach may still lead to underestimations of total catch rather than overestimations

- It is assumed that fishing activity continues between anchor points unless there is evidence of a major environmental or socio-political disruption. Commercial catches (industrial and artisanal) are interpolated linearly, while non-commercial catches (subsistence and recreational) are estimated based on population or fisher trends. This assumption simplifies reconstruction efforts but may not fully account for variations in fishing effort, changes in catch per unit effort, or other external influences.
- If a species or fishing gear known to be used in a given country is missing from official reported catch data, it is assumed that its catch was overlooked rather than nonexistent. This assumption helps identify gaps in data collection, ensuring that unreported fisheries activity is accounted for in reconstruction efforts. For example, if a coastal region is known to have small-scale fisheries using traditional weirs, but official reports do not include catches from weir fishing, it is inferred that the official data collection failed to capture this activity rather than assuming no weir fishing occurred. Similarly, if reef fishes are absent from a Pacific Island's reported data, researchers may conclude that local reef fishing was underreported or ignored, rather than believing those species were never caught. While this assumption helps improve the completeness of reconstructed fisheries data, it introduces uncertainty, as it relies on external knowledge of fishing practices instead of direct reported data. If a country has poor documentation or inconsistent fisheries reporting, estimates based on known fishing activity might not be entirely accurate.

Despite these limitations, SAU believes this is a critical step toward capturing the true scale of fishing impacts and improving fisheries management worldwide.

4.1.3) Emissions Dataset

Global Fishing Watch is a nonprofit organization dedicated to increasing transparency in global fisheries through open-access monitoring of commercial fishing activity. Using satellite data, machine learning, and vessel tracking technologies like the Automatic Identification System (AIS), Global Fishing Watch provides real-time insights into fishing operations worldwide. This dataset is the first of its kind and includes every single AIS vessel in the world.

However, a fundamental challenge in utilizing GFW data is that it focuses on effort, not catch, making it difficult to directly link vessel activity to specific species being harvested. This distinction becomes critical when comparing GFW effort data to FAO catch records, as fishing intensity does not necessarily correlate with actual landings. Additionally, the presence of flags of convenience, where vessels register under foreign jurisdictions to avoid regulations, complicates efforts to match GFW's tracking data with official catch reports. Vessels operating under these flags often evade scrutiny, making it harder to determine accurate fishing patterns and legal compliance.

Another key challenge is that many vessels do not broadcast their location with AIS and whose activity can only be detected through Sentinel-1 satellite imaging. While these detections provide insights into untracked fleet locations and sizes, they do not reveal vessel identities, flags, or fishing operations, making it difficult to partition emissions data or link effort to specific flags. Additionally, small-scale fishing vessels often do not appear in AIS tracking and may not even be detected by Sentinel-1 satellites, leading to potential underestimations of emissions from small-scale fisheries.

4.2) Pipeline Assumptions

In preparing the FAO seafood catch data in step 3.1.1, mammals, plants, and reptiles were filtered out of the major groups, and "Corals," ISSCAAP group 82, were also removed, as these were assumed not to be the target species of the gear types included in the emissions estimates. Additionally, ISSCAAP groups 41 and 51, representing freshwater crustaceans and freshwater molluscs, respectively, as well as "River eels" were removed. These were removed because any emissions associated with freshwater collection would have been eliminated during the intersection of the FAO regions shapefile and the emissions grid. Therefore, it is assumed that none of the resulting emissions can be attributed to fishing for freshwater species.

Step 3.1.1 includes filtering `vessel_class` for gear types that we are confident are identified by the machine learning algorithm as fishing. In doing this, we are removing data from the broadcasting emissions (from gear types that may be mis-identified) that would normally factor into calculating the non-broadcasting emissions. Therefore, non-broadcasting emissions may be overestimated.

Also, in step 3.1.1, NA values in the `flag` column in the broadcasting dataset are assumed to be unknown and filled with "UNK." Emissions from each grid cell ID are partitioned out based on the proportion of sub-polygon area to total grid cell ID area, assuming that emissions are uniform throughout the entire one-by-one degree pixel.

In step 3.1.2, non-broadcasting emissions are partitioned out to FAO reporting fisheries (flag-species combinations) under the assumption that they are actively emitting in the region since they are reporting catch in the region, but that they may not necessarily be using AIS on all (or any) of their fishing vessels. Additionally, this assumes that non-broadcasting emissions are directly proportional to the proportion of catch weight (in metric tonnes) by fishery (species-flag combo) and that the different gear types used to target the various fisheries have the same rate of emissions-per-unit-catch. For broadcasting emissions that are divided out proportionally amongst reported catch (by weight), we assume that all of a country's catch is reported and that emissions rates are the same for each species (when in reality, emissions estimates may vary the different gear types used to target individual fisheries).

Throughout the pipeline, the validation check of 0.001% is used as an arbitrary threshold, and the assumption is that no emissions are truly "lost" (but rather are the result of a discrepancy due to floating-point error and rounding).

5) Results Comparison Report

The carbon footprint of marine fisheries is a critical aspect of global ocean sustainability. Our project examines CO_2 emissions from fishing activities, integrating FAO data, Sea Around Us estimates, and calculations derived from AIS and Sentinel-1 satellite tracking. By leveraging satellite-based vessel identification, we aim to improve emissions estimates, especially for vessels that do not use AIS, which have been underrepresented in previous assessments.

This research builds upon previous studies, particularly Greer et al. (2018) and Parker et al. (2018), which analyzed global fisheries emissions using effort-based and fuel-based approaches, respectively. Our findings suggest that total CO_2 emissions from global fisheries in 2016 were 146 million metric tonnes, about 9% lower than the estimates for industrial fleets alone in the Greer et al. study. These discrepancies likely result from differences in methodology, including our use of direct satellite tracking rather than reconstructed effort-based estimates.

Furthermore, our analysis aligns with broader trends showing that wild fish landings have declined since the late 1990s, as observed in FAO data and Sea Around Us catch records. While total seafood production continues to rise due to aquaculture expansion, wild capture fisheries have stagnated or declined, reinforcing concerns about overfishing, habitat degradation, and climate-related impacts.

One of the most striking findings in my study is the rapid growth of emissions intensity. We observed a 52.6% increase in emissions from 2016 to 2023, rising from 148 million metric tonnes CO_2e to 226 million metric tonnes CO_2e, a much steeper rate of increase than the 28% growth reported by Parker et al. (2018) between 1990 and 2011. Additionally, emissions per tonne of landed fish rose from 1.86 metric tonnes CO_2e in 2016 to 2.47 metric tonnes CO_2e in 2022, surpassing the 2.2 metric tonnes CO_2e per tonne reported by Parker et al. (2018) for 2011.

Fuel-intensive fisheries, particularly crustacean fisheries, continue to contribute disproportionately to total emissions. If this trend persists, emissions from global fisheries could more than double by 2032, underscoring the urgent need for improved fuel efficiency and sustainable fishing practices to mitigate environmental impacts.

Overall, our study highlights the value of integrating satellite-based tracking into emissions assessments, improving accuracy, and addressing gaps in previous estimates. As global fisheries become more fuel-intensive, there is a pressing need to enhance efficiency and implement sustainability measures to curb the industry's growing carbon footprint.

The full comparison report can be found in our [Quarto Book]

6) User manuals

6.1) Emissions Pipeline

The emissions pipeline is built to incorporate each of the datasets addressed in Appendix C.

The two emissions datasets must be .csv files accessible in the "/home/emlab/projects/current-projects/ocean-ghg-fisheries/data/raw/emissions" folder on the emLab server, and they must include the following columns:

- `month` with data as "yyyy-mm-dd"
- `flag` with ISO3 codes for each broadcasting country (*does not apply to non-broadcasting data)
- `lon_bin` with longitude values (range: -180 to 180)
- `lat_bin` with latitude values (range: -90 to 90)
- `emissions_{pollutant}_mt` with emissions estimates for each pollutant (in metric tonnes)

The FAO Major Fishing Areas (Regions) shapefile must be a .shp/.shx/.sbn/.prj/.dbf file accessible in the "/home/emlab/projects/current-projects/ocean-ghg-fisheries/data/raw/fao_region_shapefile" folder on the emLab server.

The FAO Seafood Production data must be accessible in the "seamissions/data/fao_seafood_production/" folder on the emLab server. At a minimum, the folder will need to contain the following .csv files:

- Capture_Quantity.csv (catch records)
- CL_FI_COUNTRY_GROUPS.csv (country codes)
- CL_FI_SPECIES_GROUPS.csv (species codes)

As of April 2025, the FAO Seafood Production data is only available through 2022. Since emissions estimates are available from 2016 onward (in near-real time), the analysis will be limited by the availability of FAO data. As new emissions or FAO data are published, the pipeline can be re-run to update the final dataset.

The emissions-pipeline repository contains two files: "functions.R" and "_targets.R". The pipeline can be run using the targets::tar_make() function. An RDS dataset of emissions partitioned and summarized by year, flag, and pollutant for one-by-one degree pixel is saved as an *intermediate* in the following filepath:
"/home/emlab/projects/current-projects/ocean-ghg-fisheries/data/processed/dashboard.RDS"

The *final* dataset is saved to the following filepath as `full_emissions_fao_species.csv`:
"/home/emlab/projects/current-projects/ocean-ghg-fisheries/data/processed/full_emissions_fao_species.csv"

The pipeline will stop if it encounters the following errors:

- Extra files in the "/home/emlab/projects/current-projects/ocean-ghg-fisheries/data/raw/emissions/" folder on the emLab server
- A mismatch in column names between the broadcasting and non-broadcasting datasets
- Lost rows as a result of the concatenation of the broadcasting and non-broadcasting datasets
- A non-zero difference in emissions before and after the concatenation
- Different CRSs for the emissions grid and the FAO regions shapefile
- A mismatch in rows after joining emissions data to spatial grid
- Emissions losses over 0.001% after partitioning based on the proportion of polygon areas within a grid ID

- Non-broadcasting emissions losses over 0.001% during partitioning to FAO reporting countries in a particular region
- Emissions losses over 0.001% during full join for broadcasting and non-broadcasting emissions in a particular region

The stop messages are written to help identify where the error may have occurred, allowing for specific workshopping to adjust the pipeline as needed. The {targets} package will identify which function returned the error message.

The flag analysis includes a couple of assumptions as well. It is assumed that if the top country value is NA for a vessel (potentially due to turning off AIS before entering port), it's the same value as the AIS-registered flag. Additionally, it assumes that catch is reported to the AIS-registered flag country and not the top visited country.

6.2) Dashboard

6.2.1) Dashboard Description and Intended Use

The Seamissions Explorer dashboard is a public-facing, interactive and reactive dashboard built in R using the Shiny web application framework. It enables users to explore greenhouse gas emissions from industrial fishing vessels alongside catch estimates through a global map and a collection of graphs. Users can filter results by location, year, flag (country), species group, among other dimensions. The application is hosted on emLab's shinyapps.io account to enable public access.

The dashboard is intended as an educational and exploratory tool, not an official regulatory product. All results should be interpreted to care due to several key limitations:

- **Uncertainty in emissions estimates** – particularly from the reattribution of emissions from non-broadcasting vessels to broadcasting vessels, a necessary step to assign emissions to flag states and align with FAO-reported catch data.
- **Ambiguity in catch location and reporting** – stemming from underreporting and inconsistencies between where fish were actually, the vessel's flag state, and the country where landings were reported to the FAO.
- **Dependence on harmonized datasets with different formats and assumptions** – including the integration of FAO catch records and GFW's fishing vessel emissions dataset, which originate from distinct sources with differing scopes, structures, and assumptions.

This tool is intended to support decision makers in exploring pathways for reducing emissions in the fishing sector, consumers interested in understanding the carbon footprint of different seafood products, and researchers and educators looking to examine emissions trends across time and regions.

6.2.2) Overview of Dashboard Layout

Dashboard User Interface Overview

The dashboard is organized as a Shiny Navigation Bar Page (navbarPage), it contains four primary tabs:

1. **Home**: This is the landing page for the application. It provides an overview of the application and quick links and a short description of the two main tools and a link to the learn more page.
2. **Fishing Vessel Emissions Map**: The Fishing Vessel Emissions Map contains an interactive global map. It displays annually aggregated fishing vessel emissions on a one-by-one degree pixel from 2016 - 2024. Users can visualize AIS-broadcasting data and have the option to filter to a specific year, flag. And also can visualize non-broadcasting data. They can also overlay FAO major fishing regions.

   It was created using 'mapdeck', an R package which lets you plot large datasets using Mapbox GL. The map consists of two primary datasets: 'broadcasting_emissions' and 'non-broadcasting_emissions", these datasets were prepared by grouping by `lat_bin`, `lon_bin`, `year`, and `flag`to calculate the total emissions per country (flag) and year combination per pixel. It also incorporates other background layers including FAO region boundaries. All spatial layers were projected to EPSG:4326 Geodetic coordinate system, which is required by Mapbox.
3. **Compare Seafood Emissions:** The second tab of the dashboard, Compare Seafood Emissions, contains a compilation of plots displaying fishing vessel emissions associated with catch using the dataset that we generated in the Emissions Pipeline.

   The plots were created using the 'ggplot' package, a component of the popular 'tidyverse' collection of packages, in R that is used for visualizing data. The plots also integrated 'ggimage' and 'ggflags' packages to visualize images next to each of the plot elements.

4. **Learn more:** This tab provides an overview of the purpose of the dashboard and its intended use. It provides key details regarding the data, the data sources, and a high-level overview of how the datasets were produced. The tab will also include a disclaimer that outlines important assumptions and limitations to keep in mind while interacting with the tool.

More information regarding the structure of the application, including a full list of packages used and details on how to run, update, and maintain the application are located in the GitHub repo.

8) References

8.1 Government and Organizational Reports

Environmental Protection Agency (EPA), 2024. Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2000. EPA 430-R-24-004.

Intergovernmental Panel on Climate Change (IPCC), 2023. Summary for Policymakers. In: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC, Geneva, Switzerland, pp. 1-34.

International Maritime Organization. (2020). Fourth IMO GHG Study 2020: Full report and annexes. https://www.imo.org/en/OurWork/Environment/Pages/Fourth-IMO-Greenhouse-Gas-Study-2020.aspx

Food and Agriculture Organization of the United Nations (FAO). FAOSTAT. Retrieved from https://www.fao.org/faostat/en/#home.

8.2 Journal Articles

Greer, K., Zeller, D., Woroniak, J., Coulter, A., Winchester, M., Deng Palomares, M.L., and Pauly, D., 2019. Global trends in carbon dioxide (CO2) emissions from fuel combustion in marine fisheries from 1950 to 2016. *Marine Policy*, 107(103382), pp. 1-9.

Kroodsma, D.A., Mayorga, J., Hochberg, T., Miller, N.A., Boerder, K., Ferretti, F., Wilson, A., Bergman, B., White, T.D., Block, B.A. and Woods, P., 2018. Tracking the global footprint of fisheries. *Science*, 359(6378), pp.904-908.

Parker, R.W.R., Blanchard, J.L., Gardner, C., Green, B.S., Hartmann, K., Tyedmers, P.H., and Watson, R.A., 2018. Fuel use and greenhouse gas emissions of world fisheries. *Nature Climate Change*, 8, pp. 333–337.

Halpern, B.S., Frazier, M., Verstaen, J., Rayner, P., Clawson, G., Blanchard, J.L., Cottrell, R.S., Froehlich, H.E., Gephart, J.A., Jacobsen, N.S., Kuempel, C.D., McIntyre, P.B., Metain, M., Moran, D., Nash, K.L., Tobben, J., and Williams, D.R., 2022. The environmental footprint of global food production. *Nature Sustainability* 5, pp. 1027–1039.

Paolo, F.S., Kroodsma, D., Raynor, J., Hochberg, T., Davis, P., Cleary, J., Marsaglia, L., Orofino, S., Thomas, C., and Halpin, P., 2024. Satellite mapping reveals extensive industrial activity at sea. *Nature* 625, pp. 85–9.

ICES Journal of Marine Science. "Bias in Global Fishing Watch AIS Data Analyses Results in Overestimate of Northeast Atlantic Pelagic Fishing Impact." Oxford Academic, Retrieved from https://academic.oup.com/icesjms/article//82/3/fsaf033/8090016

8.3 Online Sources & Research Tools

Global Fishing Watch (GFW), 2024. What is AIS?. https://globalfishingwatch.org/faqs/what-is-ais/

GFW and Environmental Markets Lab, 2024. Quantifying Ocean-based Greenhouse Gas Emissions. https://emlab-ucsb.github.io/ocean-ghg/

Max, L., Parker, R., Tyedmers, P., nd. Seafood Carbon Emissions Tool. http://seafoodco2.dal.ca/
Sea Around Us. "Catch Reconstruction and Allocation Methods." Sea Around Us. Retrieved from https://www.seaaroundus.org/catch-reconstruction-and-allocation-methods/#_Toc421534360.

9) Appendices

Appendix A. Timeline

Week/Date	**Component of Technical Documentation**	**Deadlines**
Week 2 | April 7th	**All**: Decide Pipeline Direction **Stephen**: TD [Outline] **Josh**: TD [Outline] **Carmen**: Pipeline work **Nicole:** Dashboard	**Outline of Sections** (1pm, April 9th)
Week 3 | April 14th	**All**: Pipeline Finalization **Stephen**: Executive Summary	**First draft** (Friday 18th)

	[Draft] **Josh**: TD [Draft] **Carmen**: Pipeline work **Nicole**: Dashboard **All:** Preliminary draft of Technical Documentation to the EDS 411B instructor and faculty advisor	
Week 4 | April 21st	**All:** Repo, documentation, README	Feedback on outline given 25th
Week 5 | April 28th	**All:** Incorporate feedback from outline on respective sections. Look at reach goals	Dashboard testing
Week 6 | May 5th	**All**: Abstract and Executive Summary	
Week 7 | May 12th	**Stephen:** Technical Documentation, GitHub repo, & Dashboard tuning **Josh:** Comparison report **Carmen & Nicole:** Presentation planning & practice	**Technical Documentation Draft Due** (Monday) Practice presentation I **Due May 17** Submit final presentation program abstract and acknowledgements to The abstract should be approved by your faculty advisor
Week 8 | May 19th	**Stephen & Josh:** Incorporate feedback into Technical Documentation & GitHub repo. **Stephen:** Dashboard tuning **Josh:** Convert Technical Documentation to a Quarto document **Carmen & Nicole:** Presentation planning & practice, merge repositories	Feedback on draft from faculty advisor & EDS 411B instructor due back to group **Due May 23 (Friday)** Submit editable file of closed captioning script Submit data to data repository and send DOI to CP Coordinator Practice presentation II
Week 9 | May 26th	**All:** Final edits & polishing	Practice presentation III

	(expected to primarily be on User Documentation)	(Wednesday) **Final Presentations** (Friday) **Due May 29 (Thursday)** Share link or file of Final Presentation slides (final version to be shown at the event) **Due May 30 (Friday)** **Capstone Project Final Presentations**
Week 10 | June 2nd	**All:** Submit Technical Documentation & Repository for signing	Submit **Technical Documentation** and **Repository** for signing (Monday) **All final deliverables** (Friday)

Appendix B. Summary of Software & Tools

Data Processing Pipeline		
Details on the project environment, including package versions, can be found here.		
R Package(s)	**Citation(s)**	**Use**
tidyverse: including **dplyr**, **tidyr**, **readr**. **janitor** **here**	Wickam et al., 2019 Wickam et al., 2023 Wickam et al., 2023 Firke, S., 2023 Müller, K., 2020	Cleaning, wrangling, tidying data, and reproducible filepaths
sf	Pebesma E, Bivand R, 2023	Geospatial data processing and manipulation, reading in shapefiles
targets	Landau WM, 2021	Keep track of pipeline elements, save time on re-running

Seamissions Explorer Dashboard		
Details on the dashboard project environment, including package versions, can be found here.		
R Package(s)	**Citation(s)**	**Use**
here readr dpyr	Müller, K., 2020 Wickam et al., 2023 Wickam et al., 2023	Reproducible filepaths
rsconnect shiny shinydashboard shinyWidgets shinyjs shinyBS shinycssloaders bs4Dash later bslib	Atkins et al., 2025 Chang et al., 2024 Chang et al., 2021 Perrier et al., 2025 Attali, D., 2021 Bailey, E., 2022 Attali, D., 2024 Granjun D., 2024 Chang et al., 2025 Sievert et al., 2025	Dashboard design and layout
tidyverse ggflags ggimage	Wickam et al., 2019 Auguie et al., 2025 Yu, G., 2023	Plots
mapdeck sf	Cooley, D., 2024 Pebesma E, Bivand R, 2023	Mapping
RColorBrewer scales	Neuwirth, E., 2022 Wickham et al., 2025	Color palettes & label formatting

Appendix C. Datasets

Data Processing Pipeline - Raw Datasets		
Data	**Source**	**Use**
Broadcasting Emissions: AIS-based emissions estimates for one-by-one degree pixels (.csv)	emLab and GFW	Emissions pipeline

Non-broadcasting Emissions: Sentinel-1-based emissions estimates for one-by-one degree pixels (.csv)	emLab and GFW	Emissions pipeline
FAO Major Fishing Areas (Regions) shapefile: FAO zones (.shp)	marineregions.org	Emissions pipeline
FAO Seafood Production Data: FAO reported catch, country codes, and species codes (.csv)	FAO website	Emissions pipeline
Flag Data: AIS-tracked vessel information including registered country and top visited country (.csv)	emLab and GFW	Emissions pipeline
Master Species Key (.csv)	Danielle Ferraro and Gordon Blasco (via emLab)	Emissions pipeline

Seamissions Explorer Dashboard - Processed Datasets			
This table describes the data used in our dashboard. Shiny apps require data to be stored directly within the app to run and deploy the app locally you will need to download and upload the dashboard data folder locally.			
File Name	**Description**	**Source**	**Use**
broadcasting_emissions.rds	Polygon boundaries for 1°×1° gridded CO_2 emissions for all broadcasting vessels, aggregated annually by country.	Original Data Source: emLab and GFW With additional processing by Seamissions team	Emissions dashboard
nb_emissions.rds	Polygon boundaries for 1°×1° gridded CO_2 emissions for all non-broadcasting vessels, aggregated annually.	Original Data Source: emLab and GFW With additional processing by Seamissions	Emissions dashboard

species_data.rds	Tabular data representing CO_2 emissions, aggregated annually by species group,	Original Data Sources: emLab and GFW FAO With additional processing by Seamissions team	Emissions dashboard
top_flags.rds	Tabular data representing CO_2 emissions, aggregated annually by flag and filtered to the top 10 emitting countries for each year.	Original Data Sources: emLab and GFW FAO With additional processing by Seamissions team	Emissions dashboard
top_isscaap.rds	Tabular data representing CO_2 emissions, aggregated annually by species group and filtered to the top 10 emitting species_groups for each year.	Original Data Sources: emLab and GFW FAO With additional processing by Seamissions team	Emissions dashboard