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

Last Updated on May 20, 2025

Table of contents

1 About

This Quarto book documents the steps in the Emissions Pipeline.

2 FAO Data Assembly

This chapter describes the process used to assemble FAO Seafood Catch data.

2.1 Datasets

FAO Global Capture Production data was downloaded as .csv files in a zipped folder. The following .csv files were used in this analysis:

- `Capture_Quantity.csv` (catch quantity)
- `CL_FI_COUNTRY_GROUPS.csv` (country information)
- `CL_FI_SPECIES_GROUPS.csv` (species information)

At the time of this analysis (Spring 2025), FAO data was only avaialable through 2022.

2.2 Packages

- {tidyverse}
- {janitor}

3 Methods

3.0.1 Remove non-target species

Prior to joining the .csv files, “PISCES”, “CRUSTACEA”, “MOLLUSCA”, and “INVERTEBRATA AQUATICA” were filtered out of the major groups represented in the `CL_FI_SPECIES_GROUPS.csv`. Additionally, ISSCAAP group 82 (Corals) was removed. These species were assumed not to be the target species of the fishing gear types (“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”) represented in the broadcasting emissions dataset. This, however, does have implications for the non-broadcasting dataset (**see...**).

ISSCAAP groups 41 and 51, representing freshwater crustaceans and freshwater molluscs respectively, as well as the species “River eels”, were filtered out because any emissions associated with freshwater collection are 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.

3.0.2 Add species information

The analysis is conducted with species distinguished by a unique numeric code in the `identifier` column, and then additional species information is joined back to the final table. Species information was obtained from a modified version of the `data-keys/master_species_key.csv` created by Danielle Ferraro and Gordon Blasco and provided by emLab. Some species in the resulting FAO dataset were not represented in the master_species_key.csv, so the missing species were added from the `FAO CL_FI_SPECIES_GROUPS.csv`.

4 Merge Emissions Datasets

This chapter describes the process used to merge emissions datasets.

4.1 Datasets

Two emissions datasets, obtained from emLab, were used in this analysis:

- 1. Broadcasting emissions: `meds_capstone_ais_emissions_data_v20241121.csv`
- 2. Non-broadcasting emissions: `meds_capstone_non_broadcasting_emissions_data_v20250116.csv`

The data were pre-filtered by emLab from a larger emissions dataset to select for fishing vessels.

The following columns are required:

- `month`
- `flag`
- `vessel_class`
- `lon_bin`
- `lat_bin`
- `emissions_{pollutant}_mt`

define broadcasting vs. non-broadcasting?

4.2 Packages

- {tidyverse}
- {janitor}
- {lubridate}

5 Methods

5.1 Join Emissions Data

Emissions datasets (datasets 1 and 2 above) 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 were filled with "UNK" to represent flag unknown, and `vessel_class` was filtered for gear types identified with a high degree of confidence (i.e. "squid_jigger", "drifting_longlines", "pole_and_line", "trollers", "pots_and_traps", "set_longlines", "set_gillnets", "trawlers", "dredge_fishing", "tuna_purse_seines", "other_purse_seines", "other_seines"). This eliminated gear types such as "passenger" that were likely mis-identified as "fishing" or as "passenger" by GFW's machine learning algorithm.

```
 [1] "trawlers"           "set_longlines"      "drifting_longlines"
 [4] "trollers"           "squid_jigger"       "pots_and_traps"
 [7] "other_seines"       "pole_and_line"      "other_purse_seines"
[10] "tuna_purse_seines"  "set_gillnets"       "dredge_fishing"
[13] NA
```

In the non-broadcasting dataset, emissions estimate columns for each of the 9 pollutants (CO2, CH4, N2O, NOX, SOX, CO, VOCS, PM2.5, 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`).

5.2 Assumptions

By filtering out certain gear types,... implications for non-broadcasting

6 Intersection by FAO Region

This chapter describes the process of spatially diving emissions amongst FAO Major Fishing Area (Region).

6.1 Datasets

A zip file containing the following shapefiles was downloaded from Marineregions.org:

- `World_Fao_Zones.dbf`
- `World_Fao_Zones.prj`
- `World_Fao_Zones.sbn`
- `World_Fao_Zones.shp`
- `World_Fao_Zones.shx`

6.2 Packages

- {tidyverse}
- {sf}

7 Methods

7.1 Intersection

Spatial attributes (points) were created for each `lat_bin` and `lon_bin` in the native 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 geodataframe was transformed to Equal Earth projection. Every grid cell was assigned a unique ID. Using the FAO shapefile (dataset 3), 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.

7.2 Partition

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).

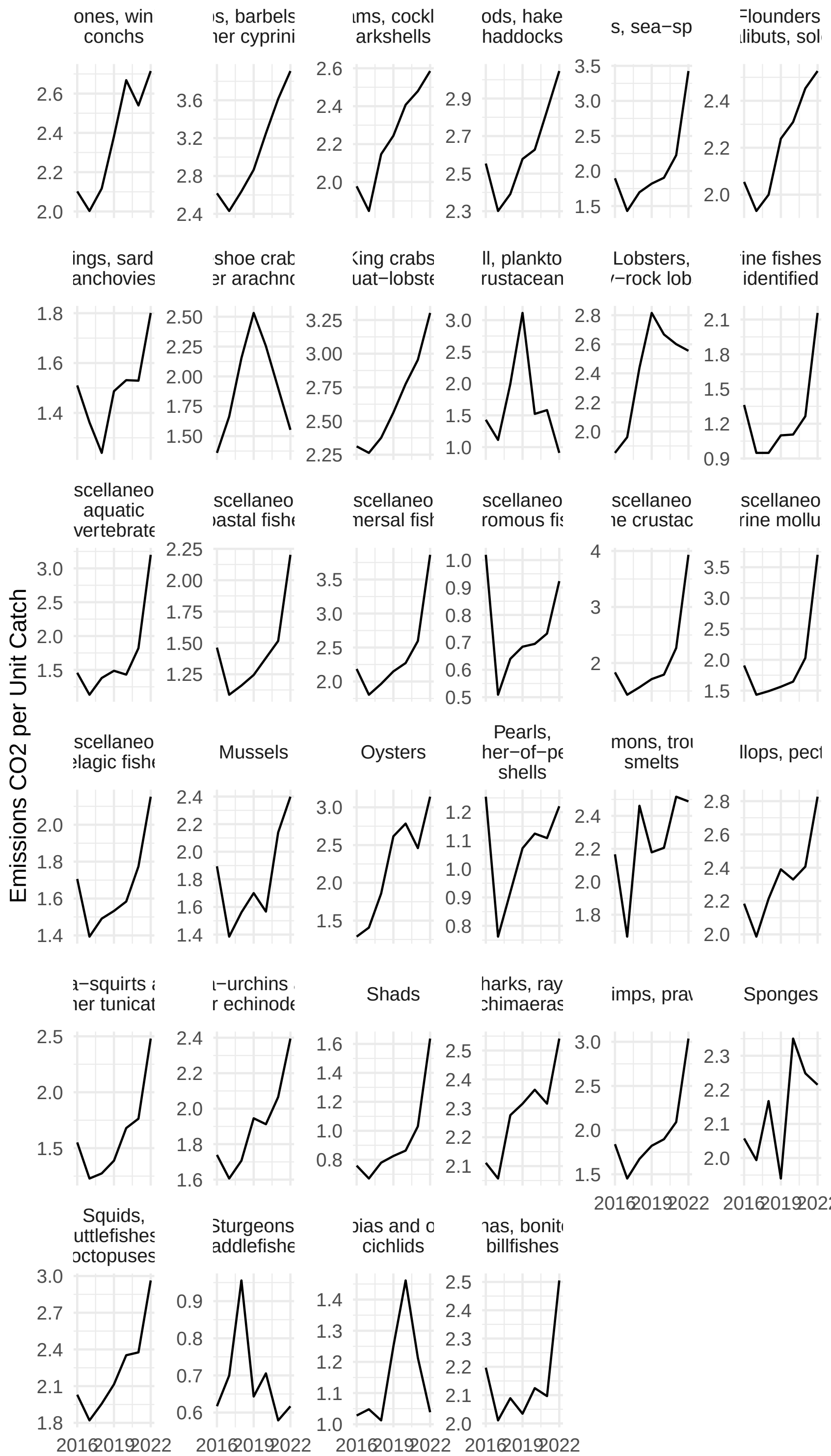
7.3 Assumptions

This assumes a uniform distribution of emissions within the 1 x 1 degree pixel.

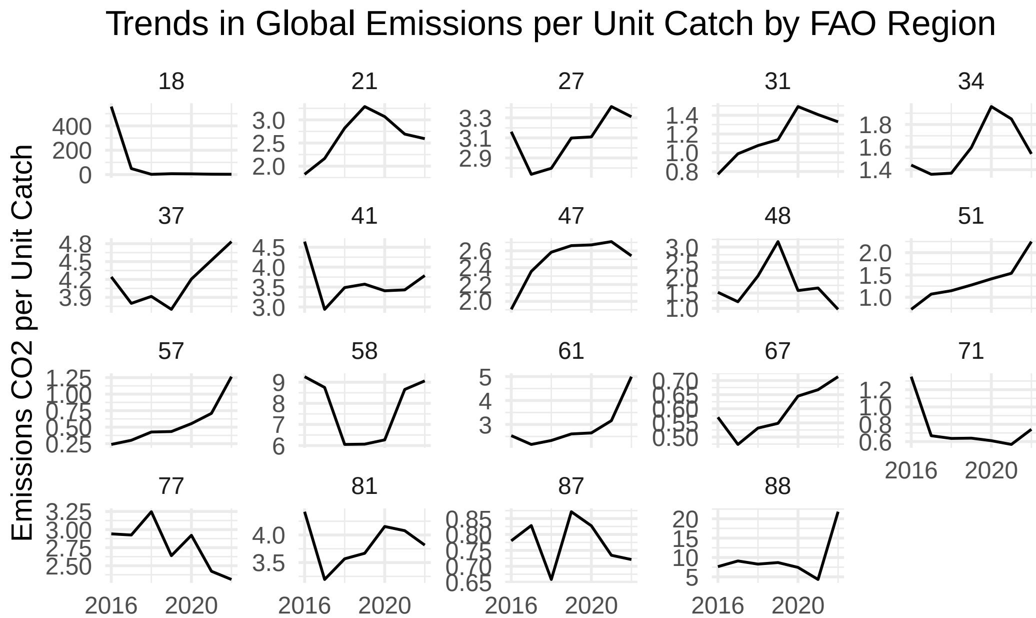
8 Results

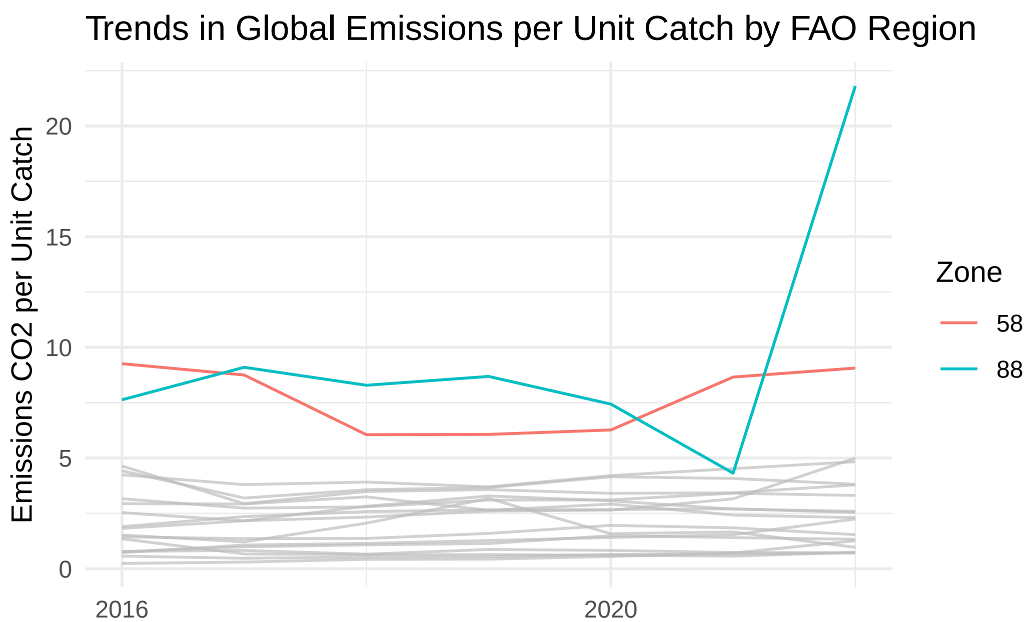
This chapter details the results from the primary analysis.

8.1 Trends in emissions-per-unit-catch by ISSCAAP species group



8.2 Trends in emissions-per-unit-catch by FAO region





8.3 Species Spotlight: Region 77 Salmons, Trouts, Smelts

9 SAU Validation

This chapter describes how we validated our findings by comparison with Sea Around Us catch data.

9.1 Datasets

Sea Around Us data was downloaded for each FAO Region (18, 21, 27, 31, 34, 37, 41, 47, 48, 51, 57, 58, 61, 67, 71, 77, 81, 87, and 88).

9.2 Packages

- {tidyverse}
- {janitor}

10 Methods

Sea Around US (SAU) data was prepped using the same methods as FAO catch data (see "FAO Data Assembly").

check that SAU doesn't contain mammals, plants, etc.

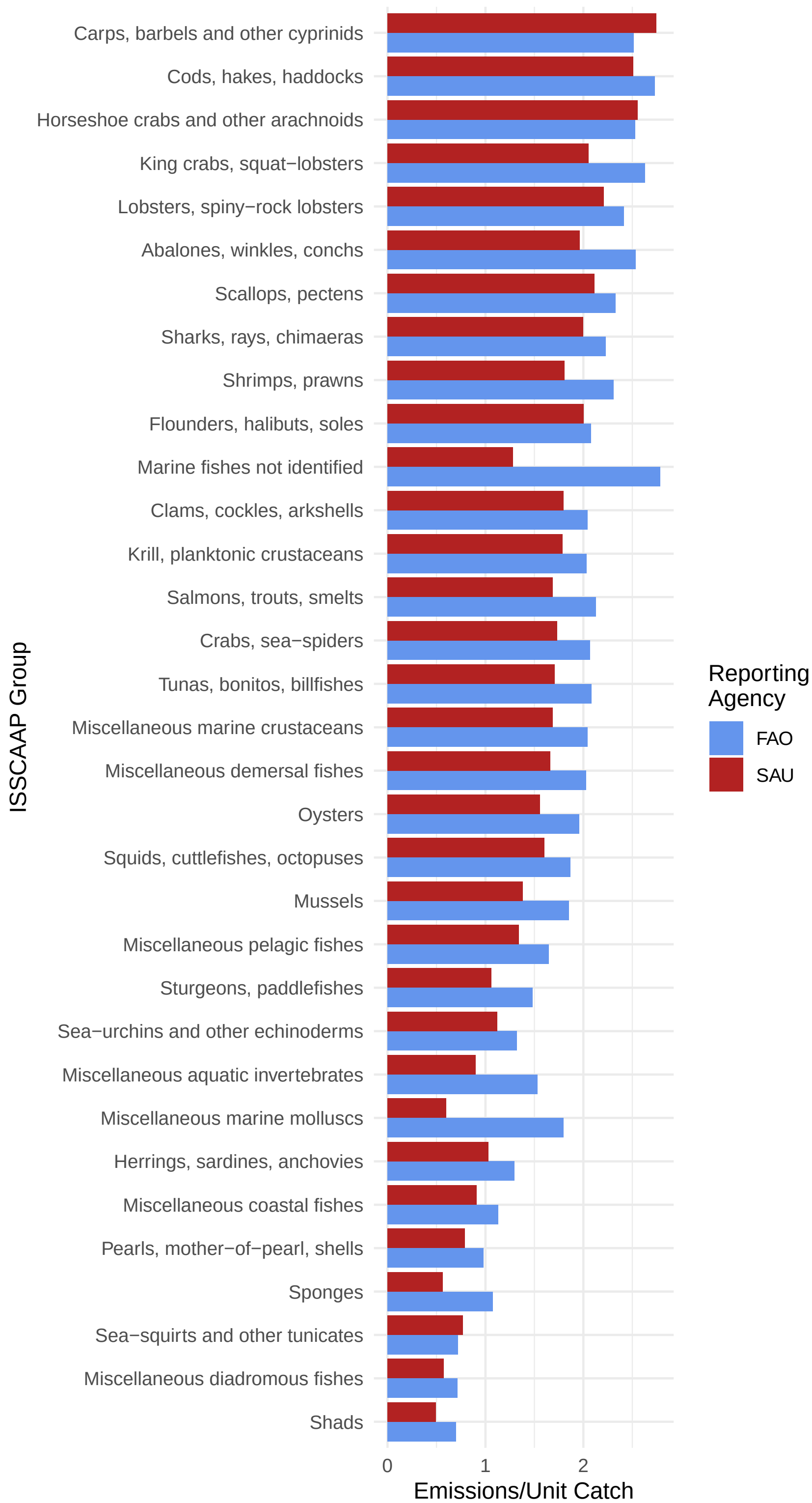
11 Results

11.1 Overall Comparisons

```
[1] "FAO emissions-per-unit-catch: 4.56."
```

```
[1] "SAU emissions-per-unit-catch: 3.76."
```

11.2 Comparisons by ISSCAAP Group



12 Flag ID

This chapter details the results of our flag analysis.

13 About

The following analysis looks into the relationship between a vessel's AIS-registered flag and its top-visited (home port) country using data from emLab and Global Fishing Watch.

13.1 Datasets

`meds_meds_capstone_annual_vessel_flag_country_emissions_v20241121.csv`

- `ssvid`: unique vessel ID
- `flag`: AIS-registered country
- `top_visited_country_iso3`: country the vessel visited the most number of times; home port
- `year`: year
- `emissions_co2_mt`: annual CO_2 emissions (MT)

13.2 Packages

- {tidyverse}
- {kableExtra}

14 Methods

14.1 1. Full Dataset Mismatch

To give a high level sense of just how much of an issue flags of convenience could be, we assessed AIS-registered flag and top-visited country mismatch for the dataset as a whole.

To do this, we created a `match` column, populated with the following values:

- `TRUE`: match
- `FALSE`: mismatch
- `NA`: no value in AIS-flag, cannot determine.

14.2 2. Full Dataset Emissions

Next, we looked at how much emissions (MT), on aggregate, could be affected by this flagging issue. We summarized emissions by match (TRUE), mismatch (FALSE), and NA for each year and overall.

14.3 3. Overestimating Emissions

To quantify overestimation, or emissions attributed to a AIS-flag that visit different top-country, we assessed the fraction of emissions that end up in a different country (mismatch/FALSE) for each flag by year.

Table 14.1: Flag Mismatch

match	percent
FALSE	21
TRUE	75
NA	4

Table 14.2: Flag Mismatch Emissions

match	sum_emissions
FALSE	6835022045
TRUE	2784977001
NA	29475094

Table 14.3: Overestimation of Emissions

flag	total_emissions	percent_emissions_leaving_by_flag
PAN	1206570122	96.9
LBR	864314868	99.9
MHL	804641557	99.9
HKG	624671969	99.0
SGP	563593064	94.9
MLT	424942851	98.8
DEU	398070016	85.6
BHS	307567222	98.9
GRC	193952596	83.2
CYP	136205494	98.7

Ex. If for the flag of Panama (PAN) 75% of emissions are from vessels that have a different home port country, this could mean that we are *overestimating* emissions for Panama by upwards of 75% due to flagging issues.

14.4 4. Underestimating Emissions

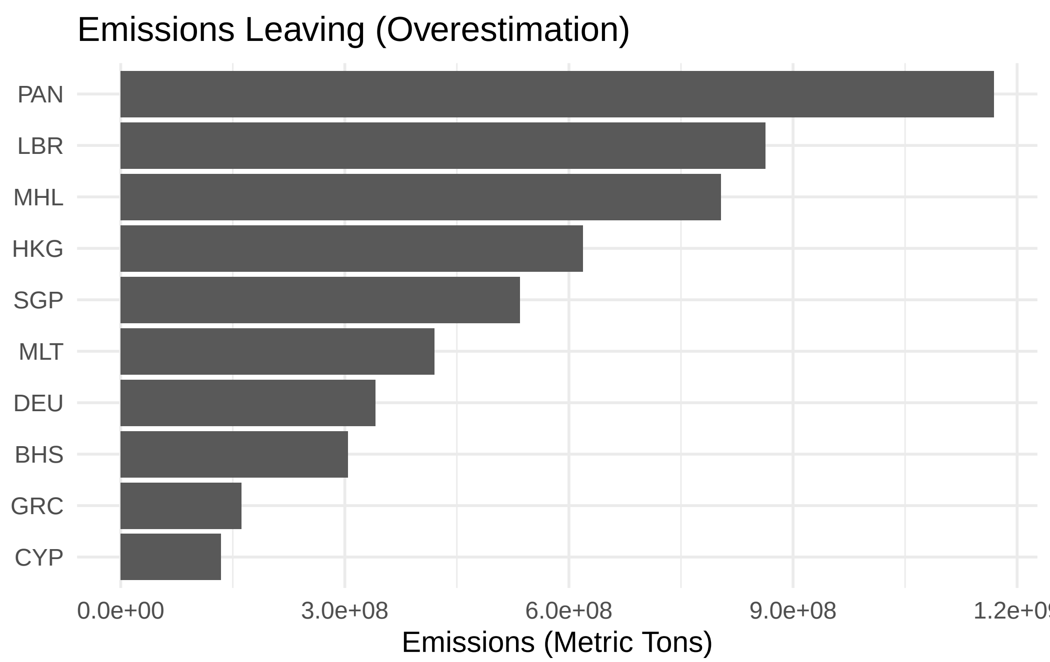
To quantify underestimation, we assessed emissions for each top-visit country that come from a different AIS-flag.

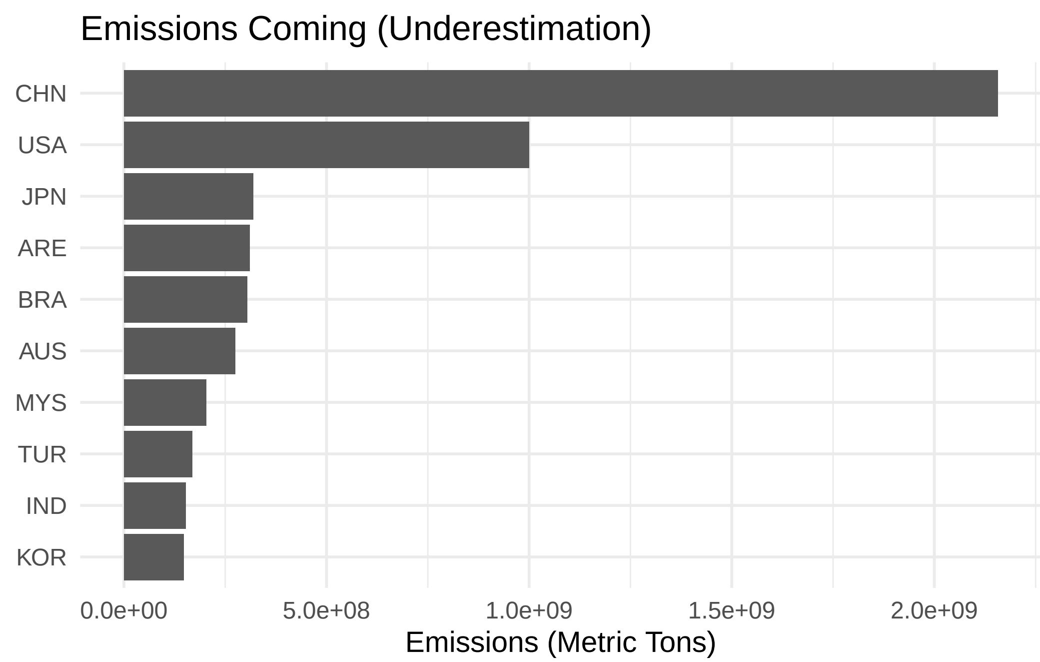
Ex. If for the country of China 25% of emissions are from vessels that have a different AIS-flag, this could mean that we are *underestimating* emissions for China by upwards of 25% due to flagging issues.

Table 14.4: Underestimation of Emissions

flag	total_emissions	percent_emissions_coming_by_flag
CHN	2873213170	75.0
USA	1263414706	79.1
JPN	576401931	55.3
ARE	319120513	97.2
BRA	331487795	91.7
AUS	329806075	83.2
MYS	232364455	87.3
TUR	246145472	68.5
IND	194510391	78.1
KOR	248005266	59.6

14.5 Visualize





14.6 Assumptions

Where top-visited country is NA, it was filled with AIS-flag value (assumes proper registration).

Vessels are assumed to land catch in their top-visited country, and that catch gets reported to the FAO by that country.

15 Comparison Report

This chapter details the results of our comparison report.