

Analysis for: Satellites reveal global extent of forced labor in the world's fishing fleet

Gavin McDonald - Environmental Market Solutions Lab (emLab)

10/09/2020

Contents

Analysis	2
Load data and necessary functions	2
Define pre-processing steps	2
Specify models	2
Cross-validation	2
Cross-validation performance	2
Choose optimized model	2
Final model building using full dataset and optimized model	2
Robustness checks	3
Use BigQuery	4
High risk fishing effort	4
High risk port visits	5
Time at sea statistic	5
Known registry vessel characteristics	5
Figures	6
Figure 1: Training data model feature summary	6
Figure 2 - Forced labor risk by fishing fleet	7
Figure 3 - Spatial forced labor risk	9
Figure 4: Port visits by high-risk vessels	10
Figure S1 - Positive vessel cases	11
Figure S3 - Cross-validation performance	12
Figure S4: Classification by training label	13
Figure S5: Variable importance	14
Figure S6 - Summary of known vessel registry information	15
Figure S7 - Cross-validation results from robustness checks	16
Figure S8 - Final model results from robustness checks	17
Statistics	18
Using base model assumptions	18
Using range of results from robustness checks	20

Analysis

Load data and necessary functions

Define pre-processing steps

Specify models

Cross-validation

Cross-validation performance

Choose optimized model

Final model building using full dataset and optimized model

Robustness checks

Use BigQuery

High risk fishing effort

High risk port visits

Time at sea statistic

This query generates the fraction of total time at sea by included vessels in the analysis. By running this query, we find that “These vessels represent 33% of the total time at sea spent by all fishing vessels operating in this time period tracked by Global Fishing Watch.”

Known registry vessel characteristics

Figures

Figure 1: Training data model feature summary

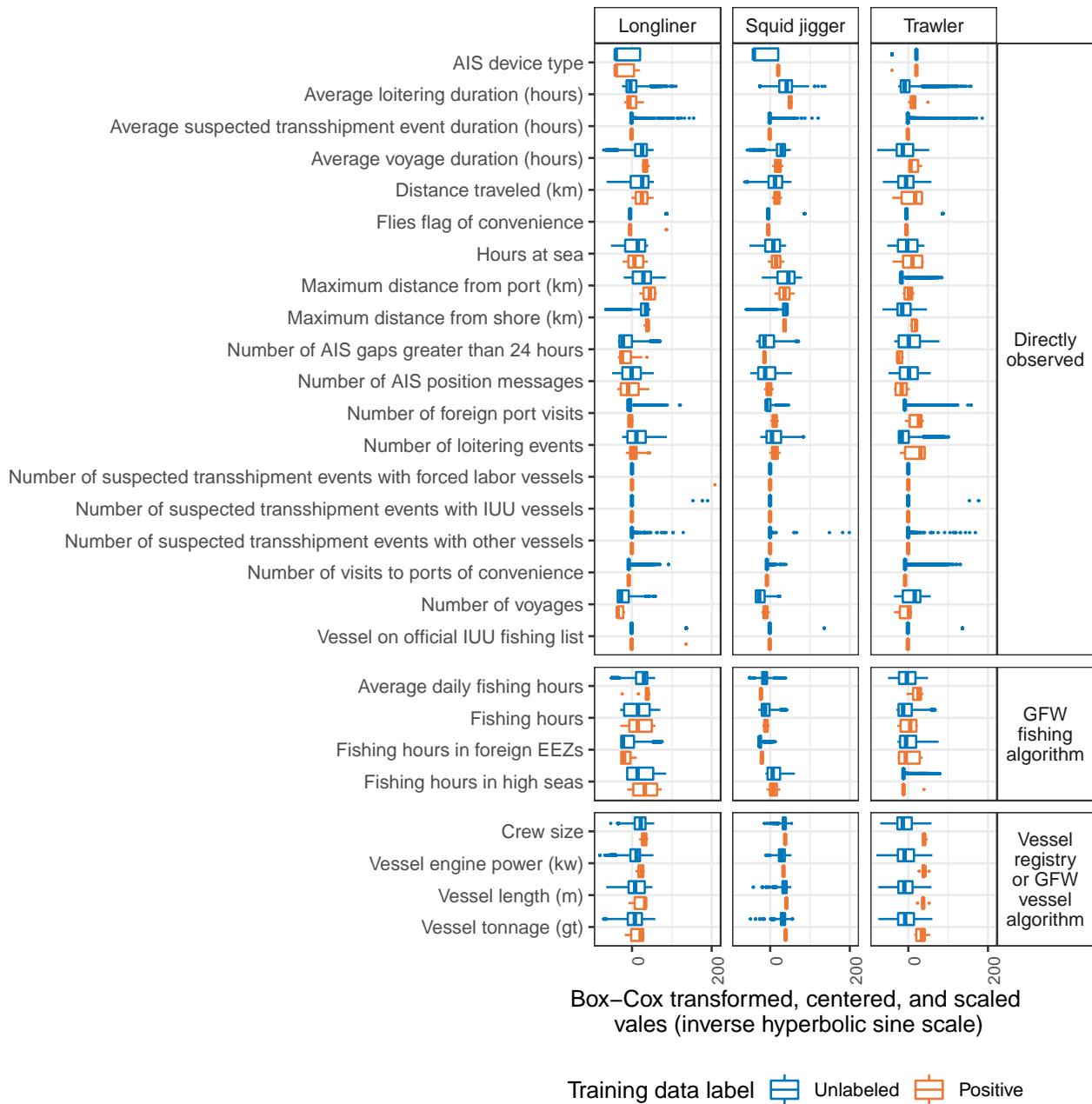
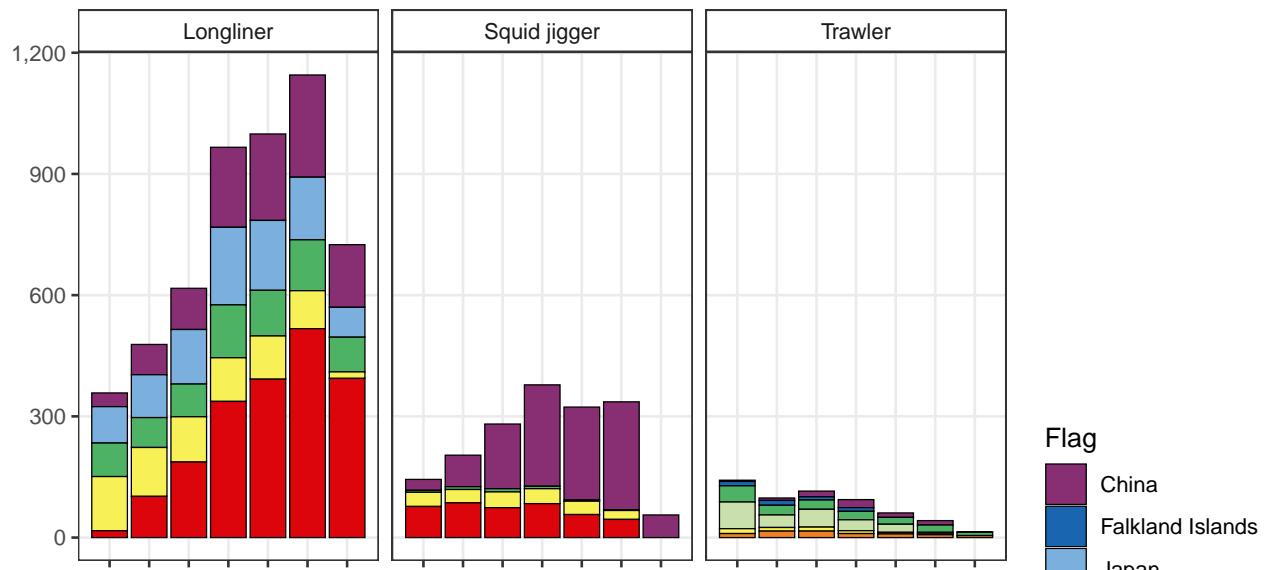


Figure 2 - Forced labor risk by fishing fleet

Figure 2 using point estimates

(A) Number of high-risk vessels



(B) Percentage of total vessels that are high-risk

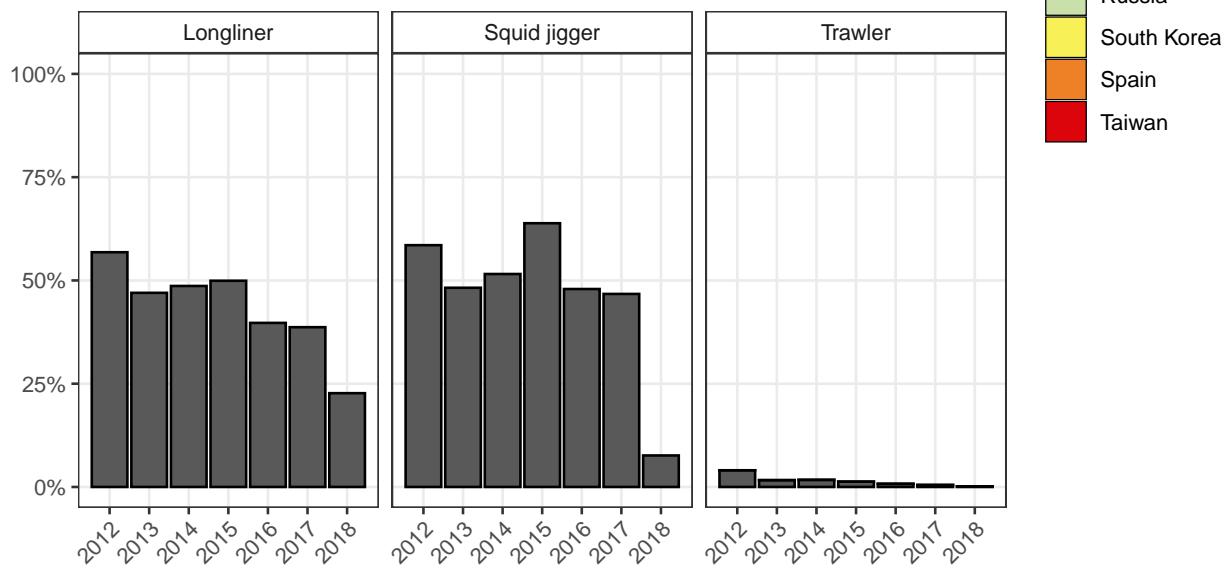
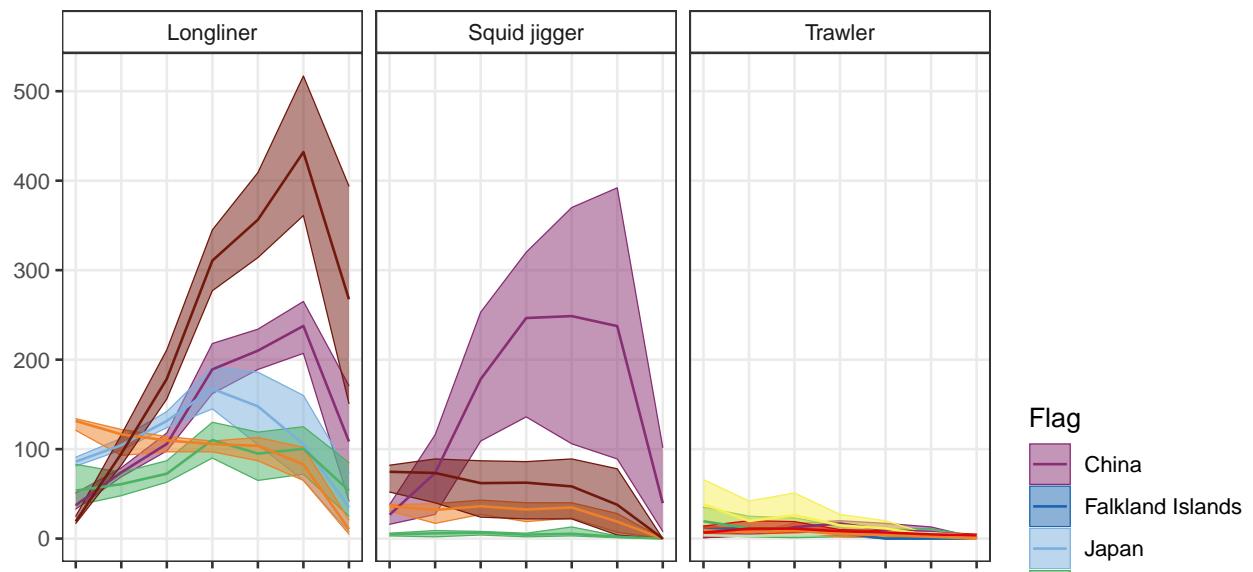


Figure 2 using range estimates from robustness checks

(A) Number of high-risk vessels



(B) Percentage of total vessels that are high-risk

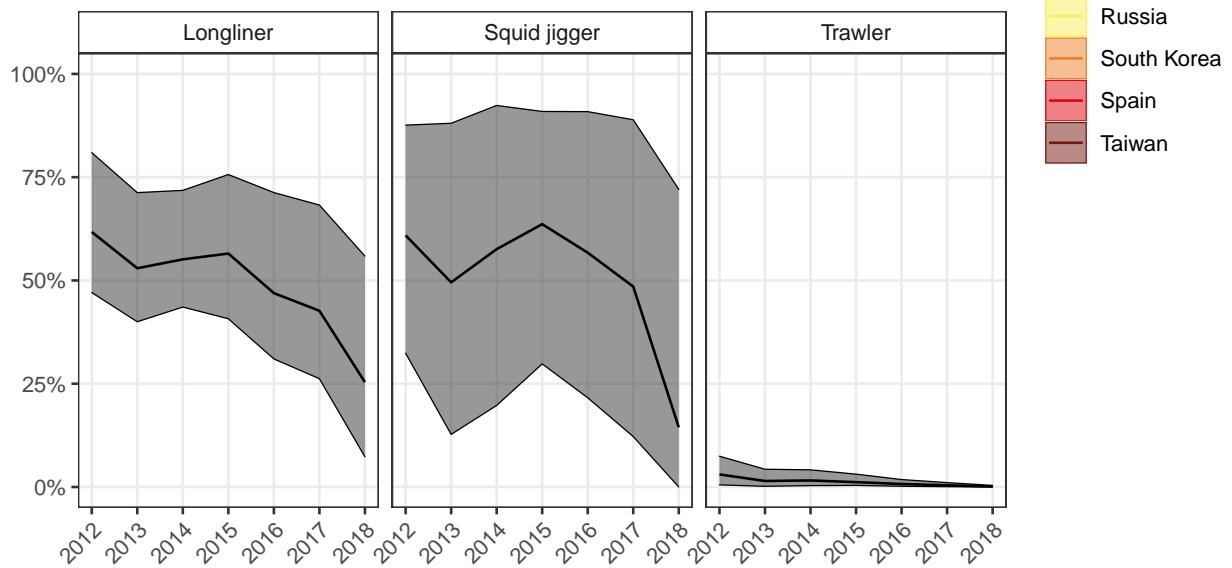
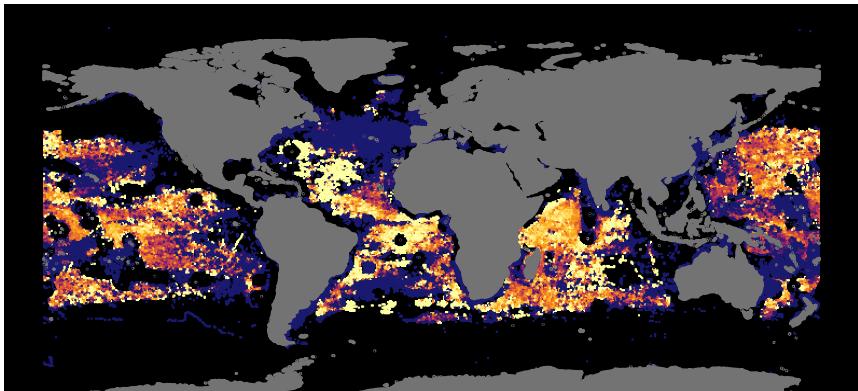
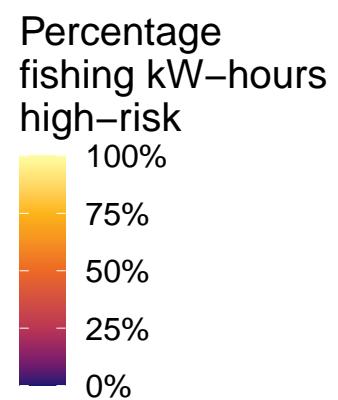
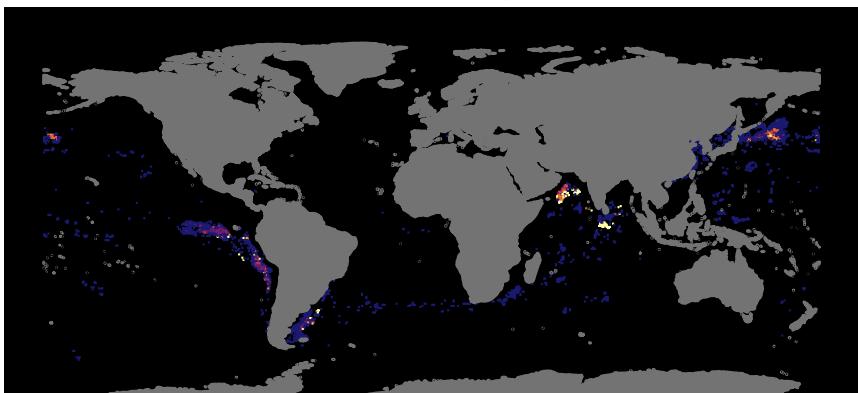


Figure 3 - Spatial forced labor risk

(A) Longliner



(B) Squid jigger



(C) Trawler

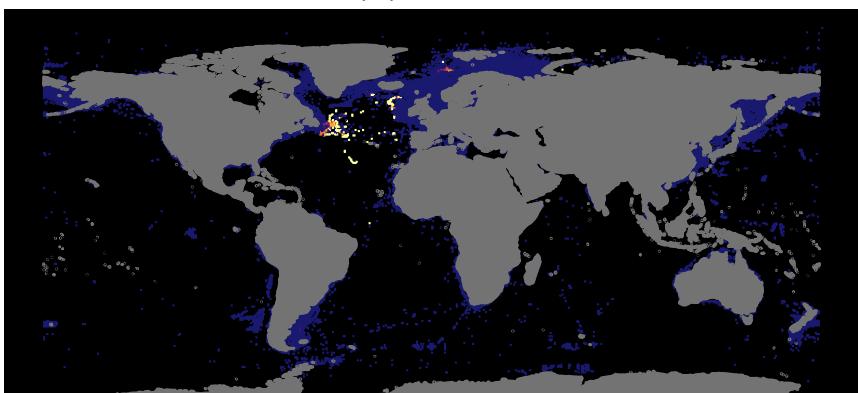
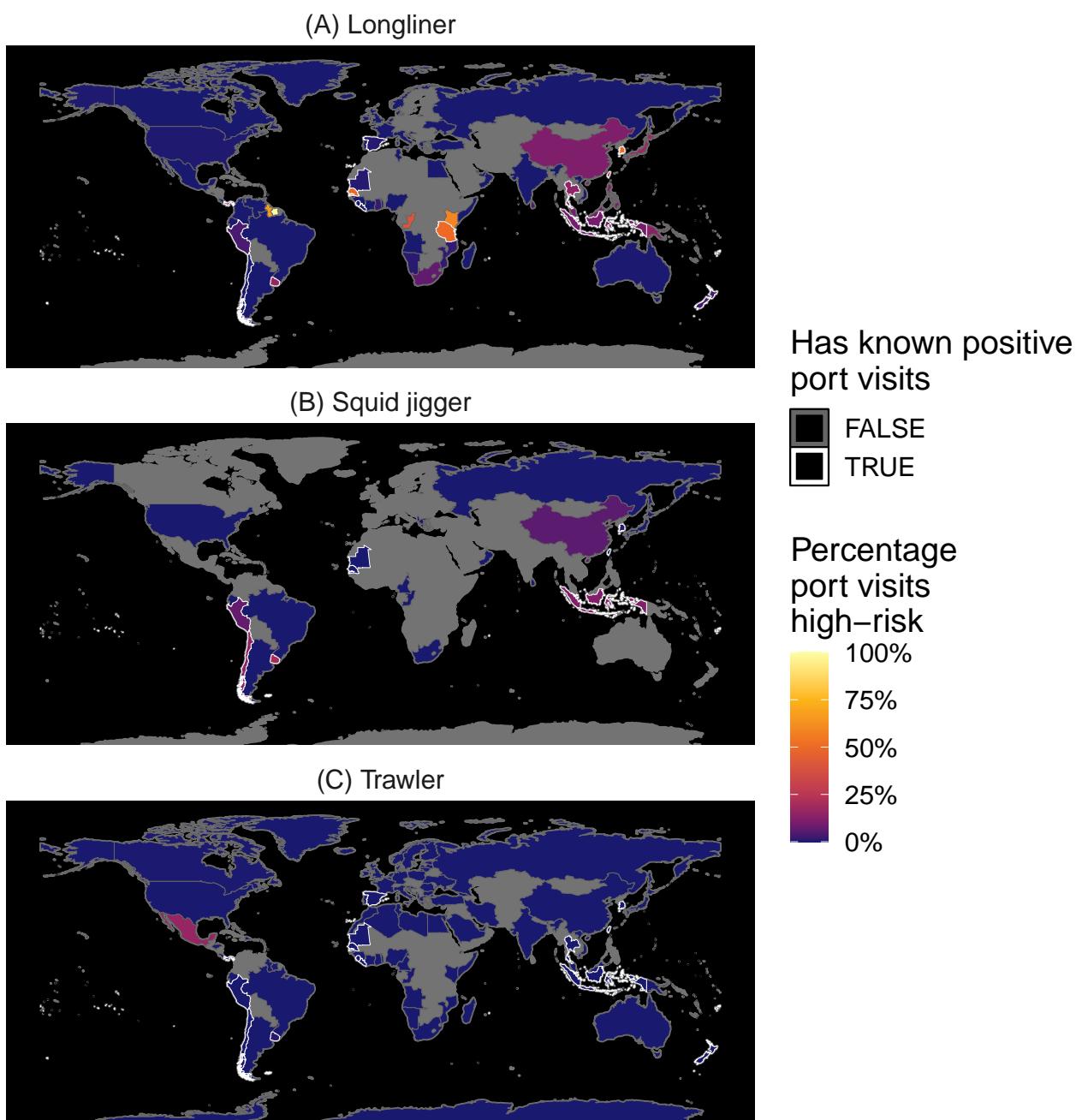


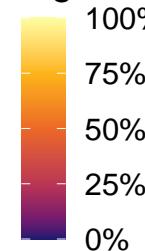
Figure 4: Port visits by high-risk vessels



Has known positive
port visits

■	FALSE
■	TRUE

Percentage
port visits
high-risk



(C) Trawler

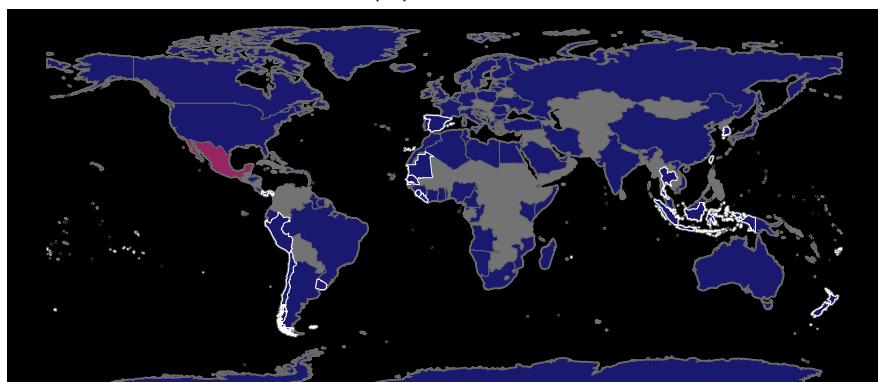


Figure S1 - Positive vessel cases

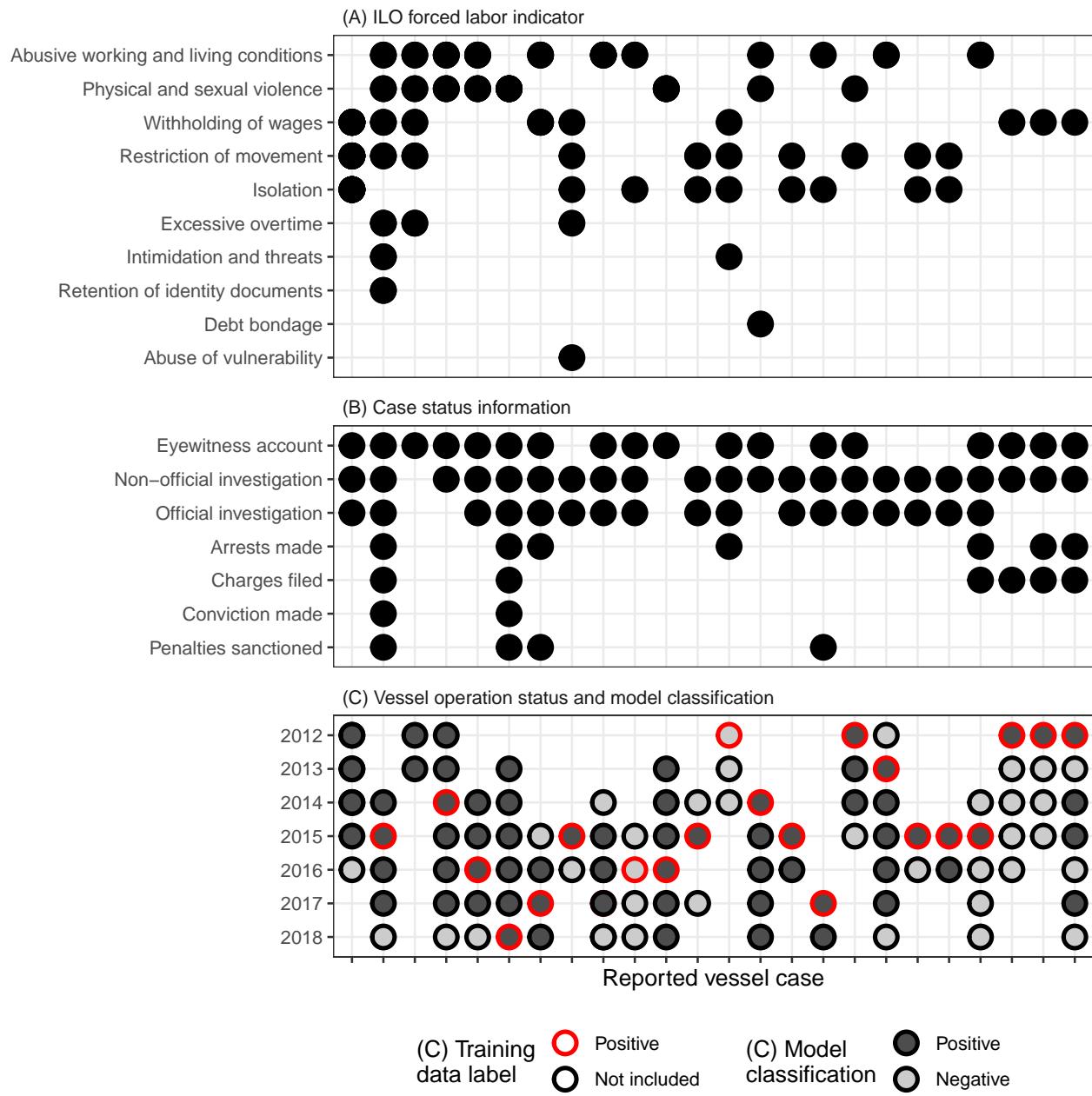


Figure S3 - Cross-validation performance

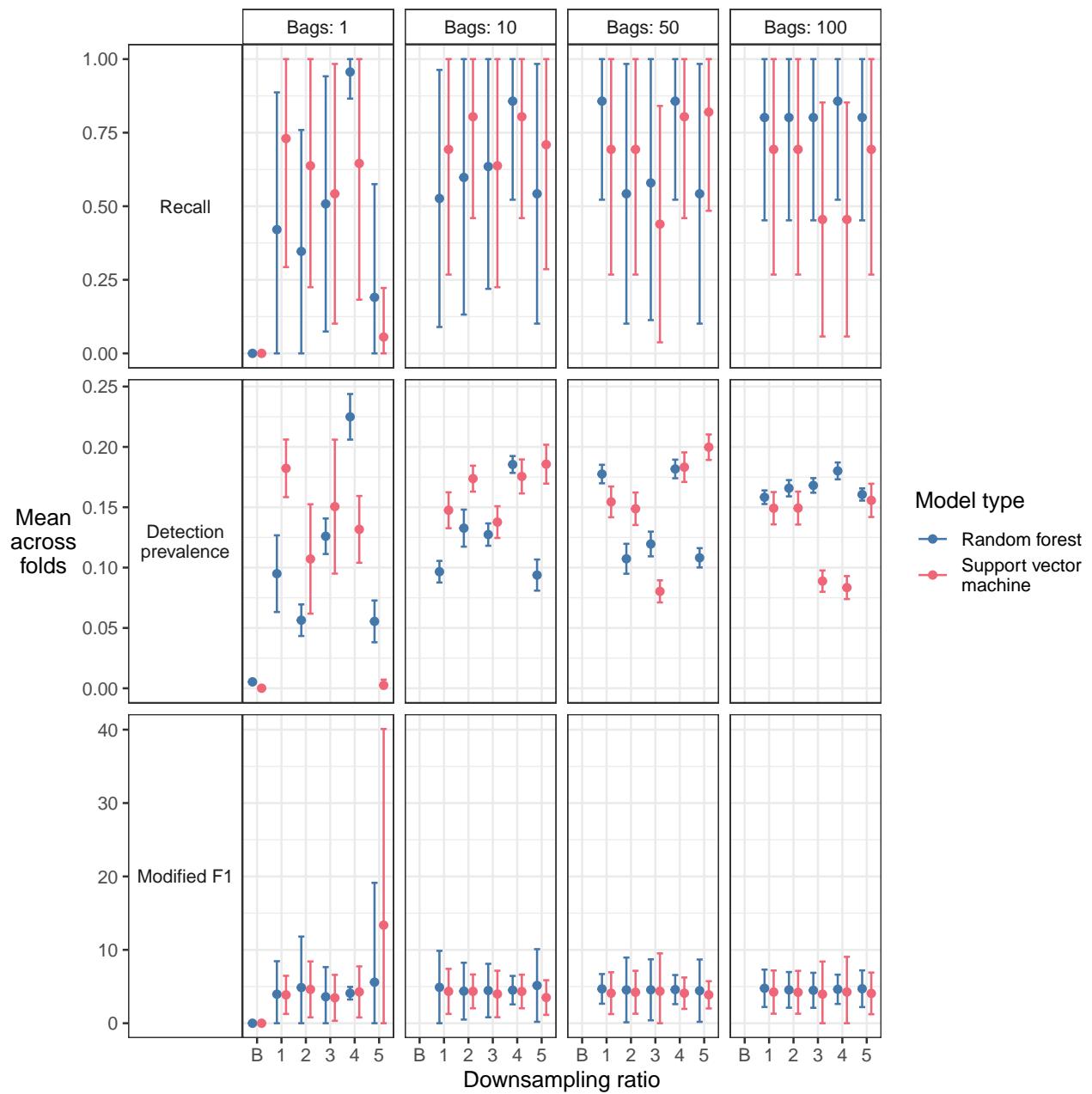


Figure S4: Classification by training label

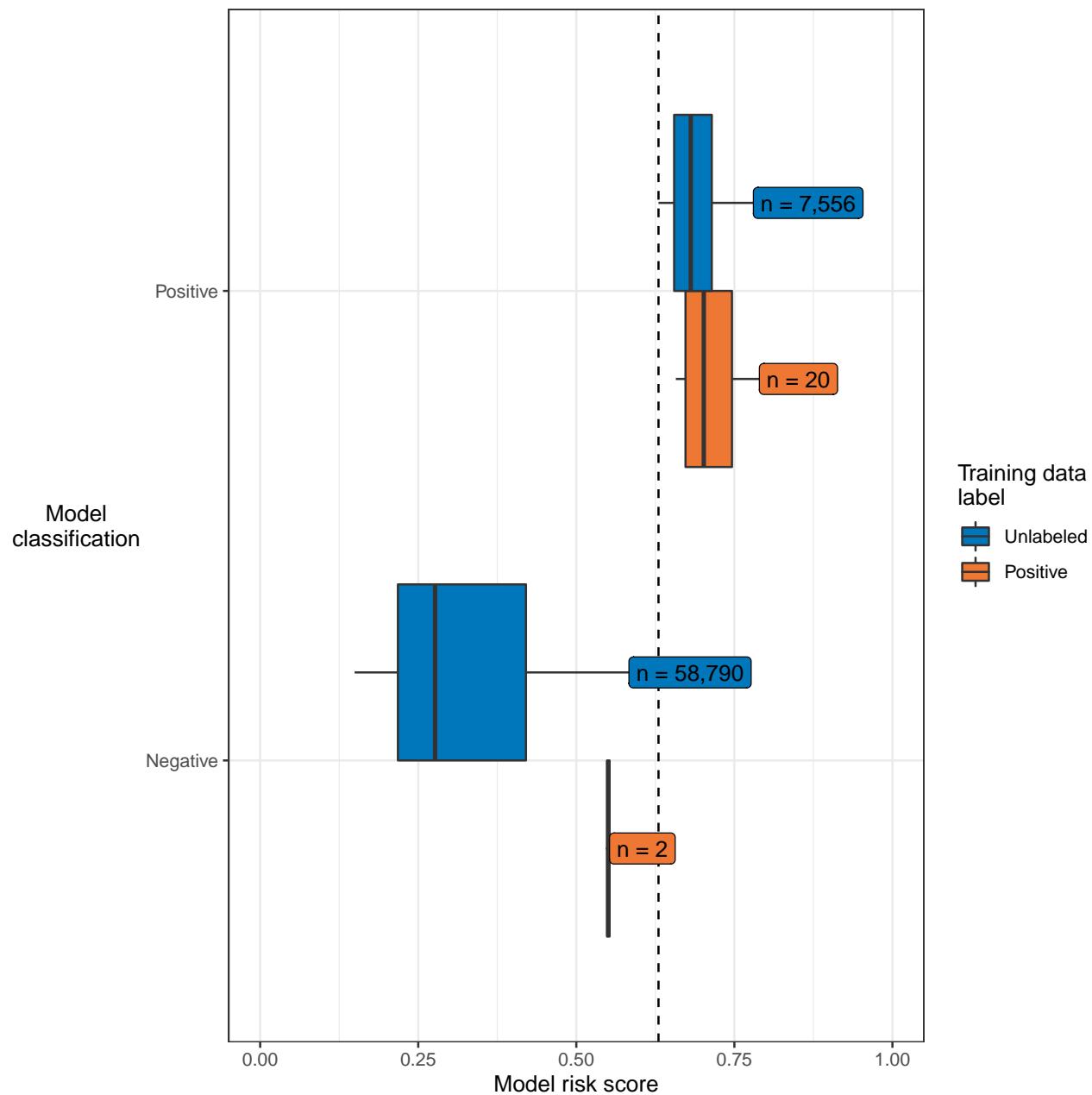


Figure S5: Variable importance

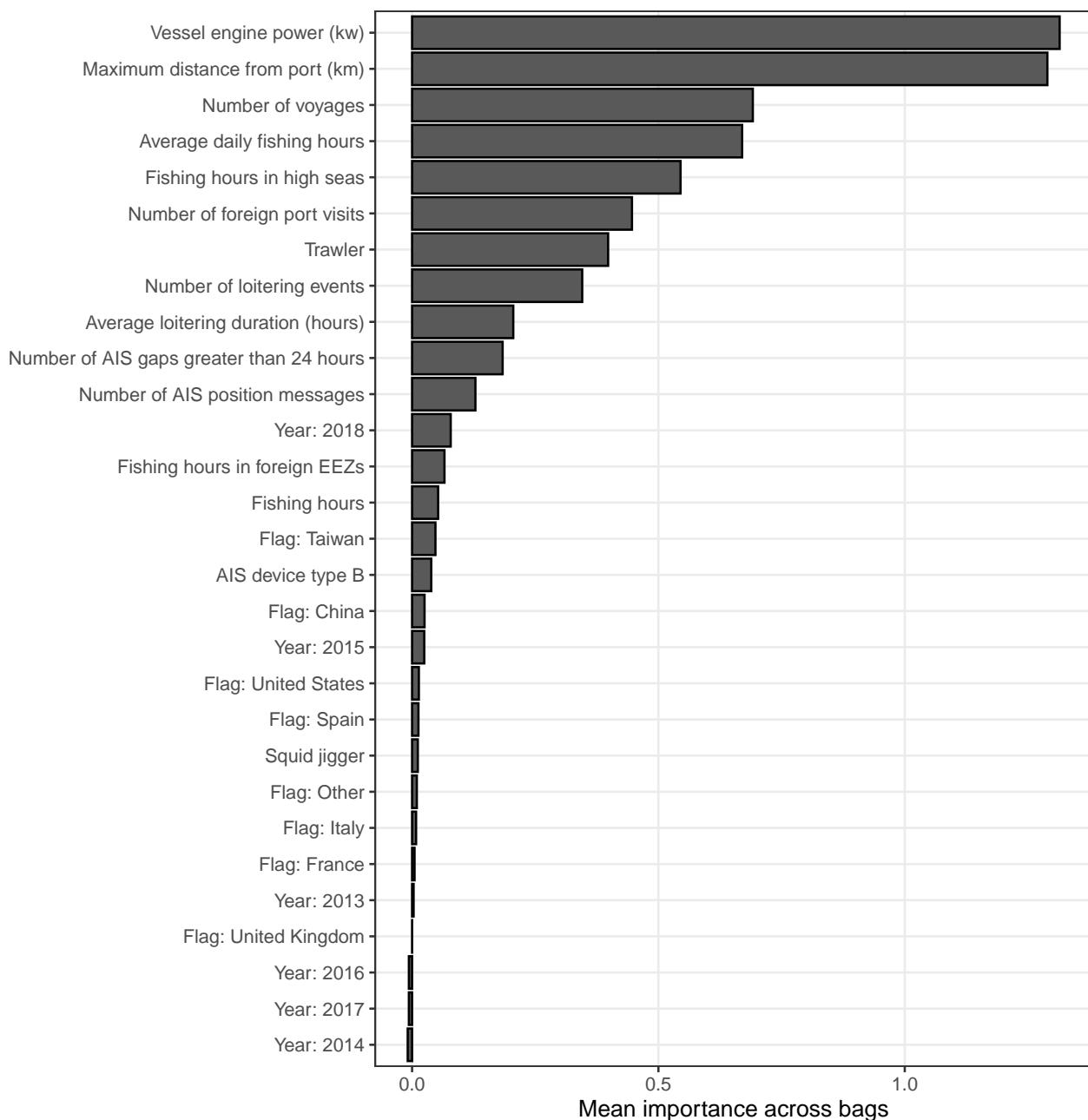


Figure S6 - Summary of known vessel registry information

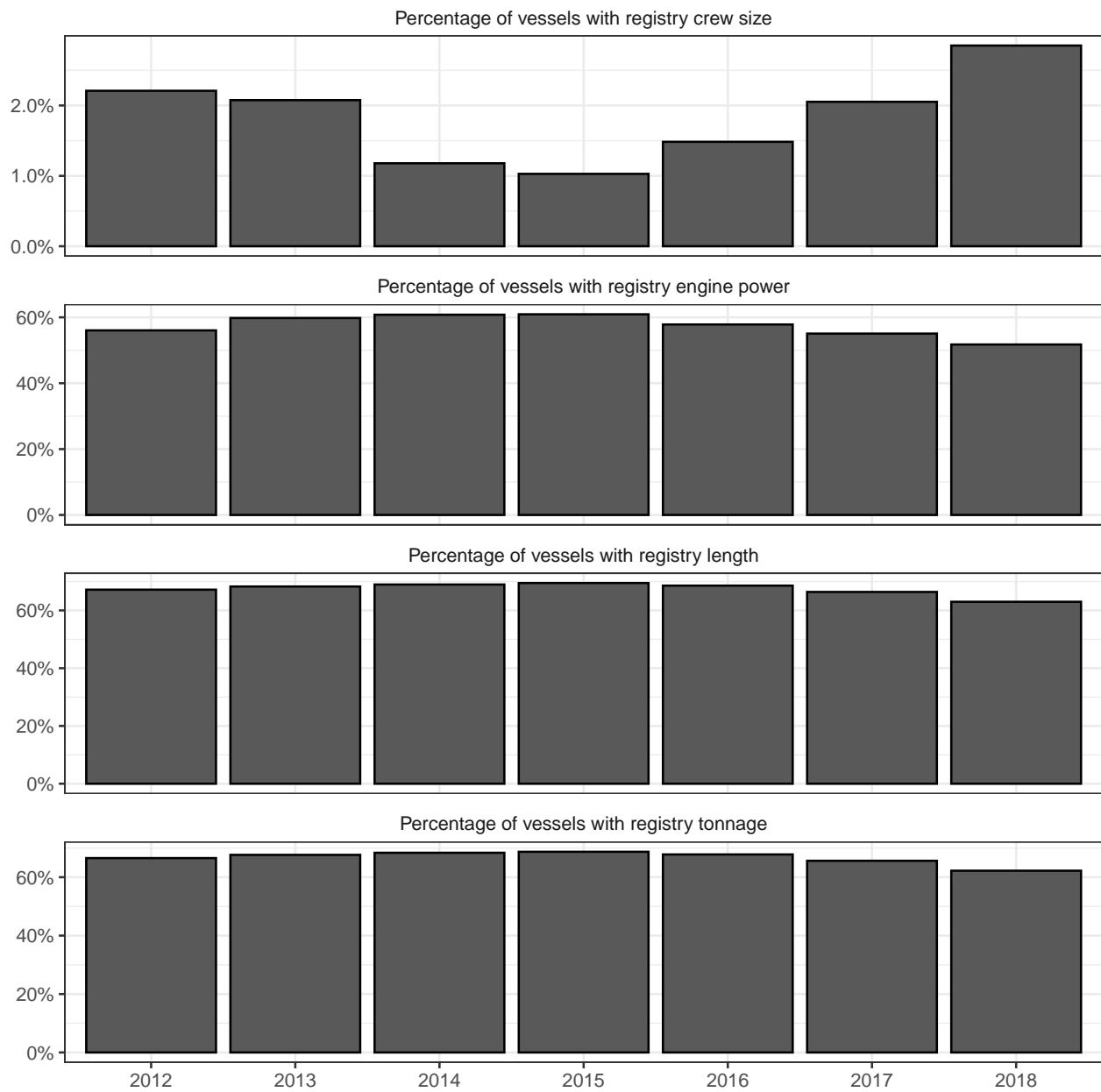


Figure S7 - Cross-validation results from robustness checks

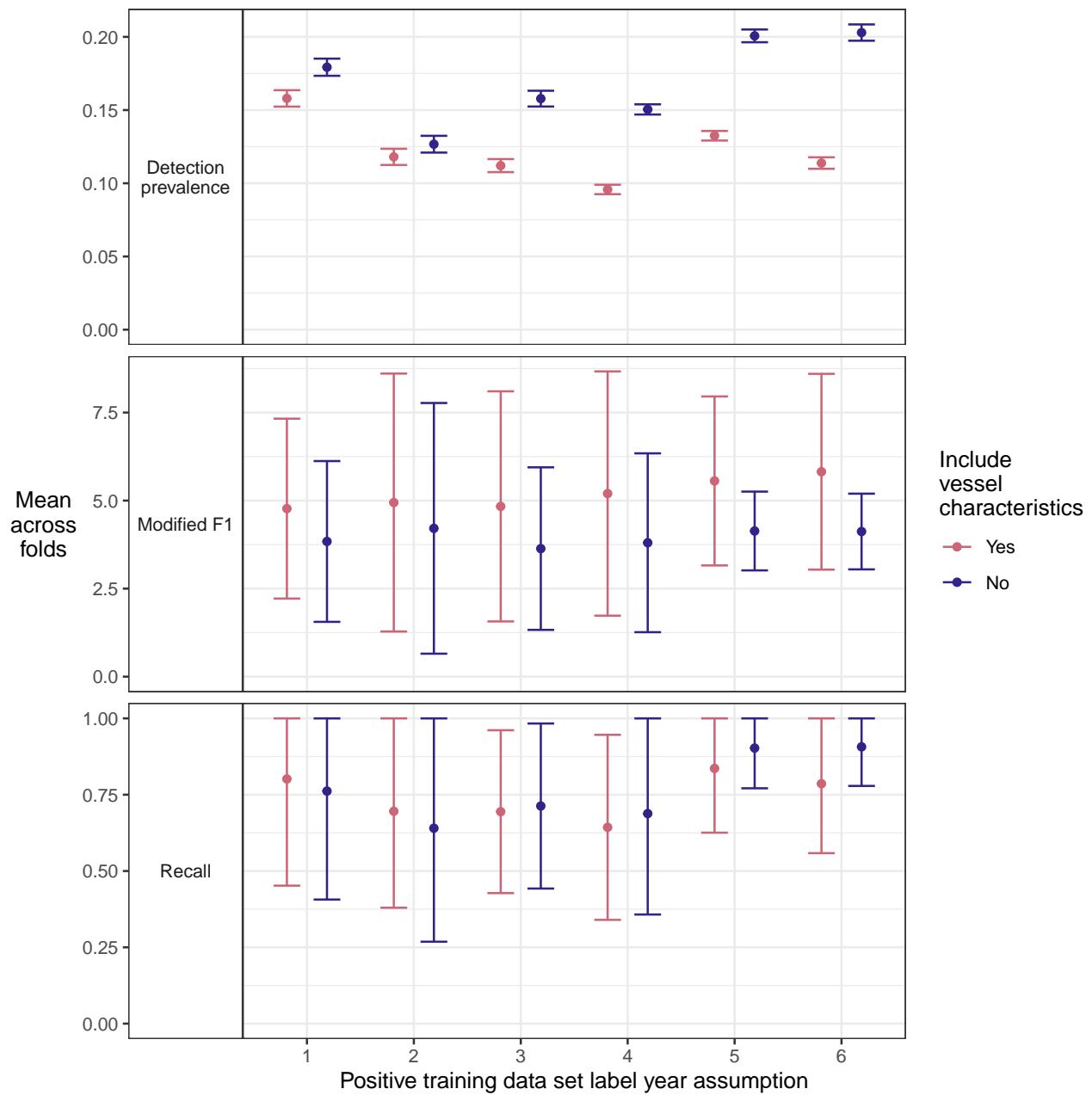
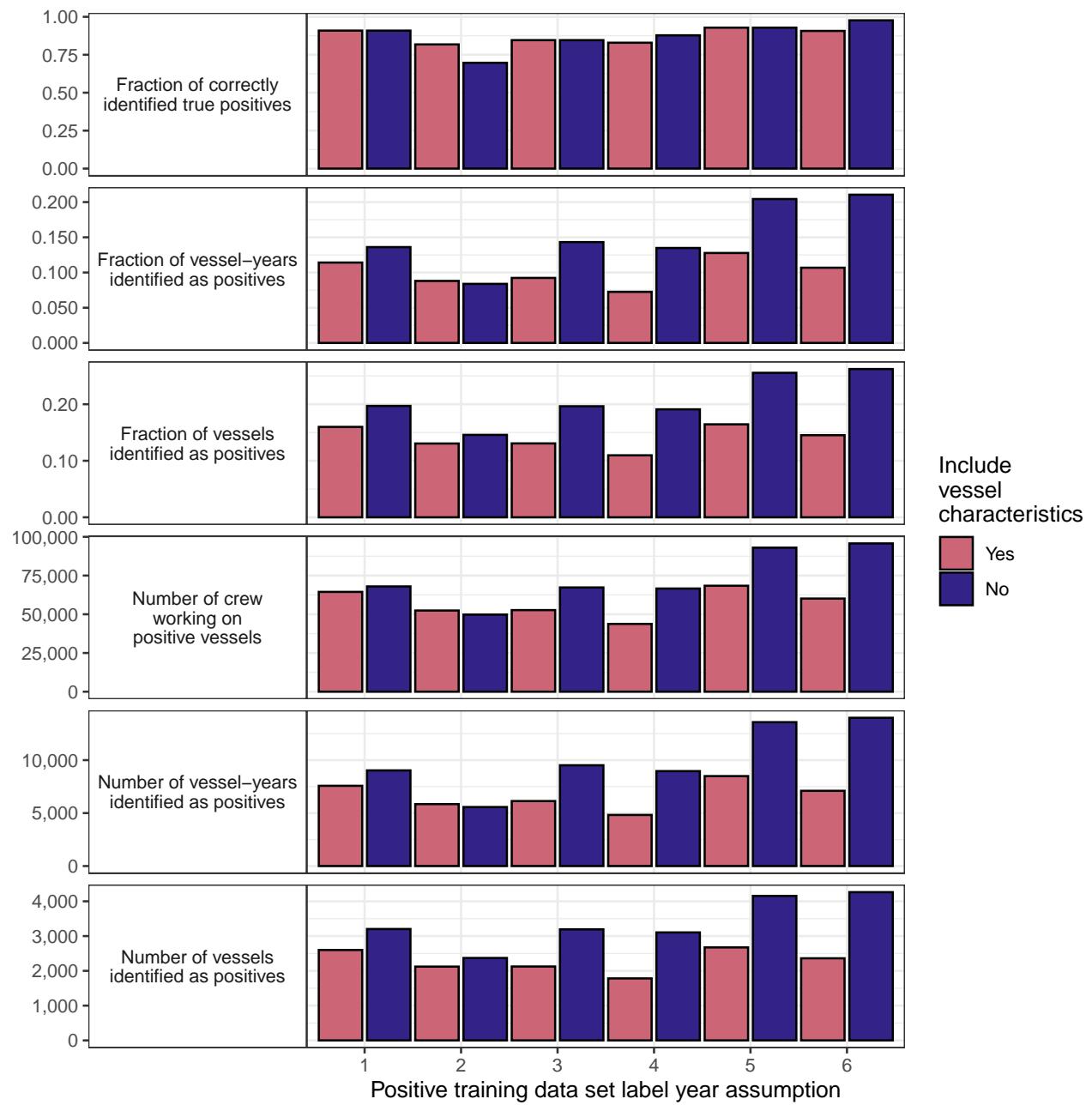


Figure S8 - Final model results from robustness checks



Statistics

We select our optimized model variation to be random forest with 100 bags and a downsampling ratio of 1. This specification has the following mean performance across the 10 folds: modified F1 score: 4.8; recall: 0.8; and detection prevalence: 0.16.

Using base model assumptions

The model identifies 7,556 new high-risk vessel-years that were previously unlabeled.

The model correctly identifies 91% of positive vessel-years as being high-risk, while also identifying 7,576 total high-risk vessel-years (11% of the total vessel-years).

2,600 unique vessels were high-risk during at least one year (16% of the total unique vessels).

64,458 crew members were working on these boats and thus potential victims of forced labor during at least one year.

Taiwan longliners, China squid jiggers, China longliners, Japan longliners, South Korea longliners represent the five fisheries with the largest number of unique high-risk vessels.

While longliners have the largest number of high-risk vessel-years across years, squid jiggers have the highest percentage of high-risk vessels across all years (44%), followed by longliners (39%) and trawlers (1%).

We also find that known positive vessels visited ports in 21 countries during the 2012-2018-time frame using our base model assumptions.

In 2018 alone, model-identified high-risk vessels visited ports across 53 developed and developing countries in 2018 (34% of all visited countries for these gear types), including 29 Parties to the Port State Measures Agreement (Figure 4). The visited ports are predominantly in Asia, Africa, and South America, with notable exceptions being Canada, United States, New Zealand, and several European countries.

35 of the countries visited by high-risk vessels in 2018 had not been visited by known positive vessels, which is reflective of our limited training data set but may also be reflective of the limited port oversight currently occurring in many countries.

In 2018, 14 of 24 (58%) vessels with reported forced labor were still operating, with 5 of those 14 vessels (36%) being classified as high-risk.

Table 1: Countries that were visited by high-risk vessels in 2018

country
American Samoa
Australia
Canada
Cape Verde
Chile
China
Congo - Brazzaville
Falkland Islands
Faroe Islands
Fiji
French Polynesia
Ghana
Guyana
Iceland
Indonesia

Japan
Kenya
Kiribati
Malaysia
Marshall Islands
Mauritania
Mauritius
Mayotte
Mexico
Micronesia (Federated States of)
Mozambique
Namibia
New Caledonia
New Zealand
Norway
Palau
Panama
Papua New Guinea
Peru
Philippines
Portugal
Samoa
Senegal
Seychelles
Singapore
Solomon Islands
South Africa
South Korea
Spain
Sri Lanka
Suriname
Taiwan
Tanzania
Thailand
Tonga
Trinidad & Tobago
United States
Uruguay

Table 2: Countries that had ratified the PSMA and were visited by high-risk vessels in 2018

country
Australia
Cape Verde
Chile
Faroe Islands
Ghana
Guyana
Iceland
Indonesia

Kenya
Mauritania
Mauritius
Namibia
New Zealand
Norway
Palau
Panama
Peru
Philippines
Portugal
Senegal
Singapore
South Africa
South Korea
Spain
Sri Lanka
Thailand
Tonga
United States
Uruguay

Using range of results from robustness checks

The model correctly identifies between 82% and 93% of positive vessel-years as being high-risk, while also identifying between 7576 total high-risk vessel-years (between 7% and 13% of the total vessel-years).

Between 1,783 and 2,674 unique vessels were high-risk during at least one year (between 11% and 16% of the total unique vessels).

Between 43,764 and 68,420 crew members were working on these boats and thus potential victims of forced labor during at least one year.

Looking across all model assumptions, Taiwan longliners, China squid jiggers, China longliners, Japan longliners, South Korea longliners represent the five fisheries with the largest number of unique high-risk vessels.

While longliners have the largest number of high-risk vessel-years across years, squid jiggers have the highest percentage of high-risk vessels across all years (between 20% and 60%), followed by longliners (between 29% and 40%) and trawlers (between 0% and 1%).