Supplementary Materials for

Tracking the global footprint of fisheries

David A. Kroodsma,* Juan Mayorga, Timothy Hochberg, Nathan A. Miller, Kristina Boerder, Francesco Ferretti, Alex Wilson, Bjorn Bergman, Timothy D. White, Barbara A. Block, Paul Woods, Brian Sullivan, Christopher Costello, Boris Worm

*Corresponding author. Email: david@globalfishingwatch.org

Published 23 February 2018, *Science* **359**, 904 (2018)
DOI: 10.1126/science.aao5646

This PDF file includes:

Table of Contents

Materials and Methods

1. Neural Network for Identifying Fishing Effort and Vessel Characteristics

1.1 Data Pipeline

We compiled all available AIS data provided by satellite service provider ORBCOMM from 2012-2016. According to International Maritime Organization (IMO) rules, AIS is required on all vessels >300 tons on international voyages, and many countries require smaller vessels to carry the device in their EEZ as well (*10*).

Raw NMEA encoded AIS messages, more than 20 million per day, are parsed using the Python library libais (*31*). For each positional message, we compute the distance to the nearest point on shore at 1km resolution. For each MMSI (the unique identifier for vessels in AIS), we calculate the distance and time between consecutive points, and if these imply an unrealistic speed between positions, or the time between points is greater than 24 hours, we create a new "segment," which is a physically possible track of a vessel with no gaps in time greater than 24 hours. The segment analysis allows us to eliminate incorrect positions, which result from noise in the GPS or incomplete transmission of AIS, and we eliminated short segments (<5 positions) for most analyses. When more than one vessel broadcasts with the same MMSI at the same time, we separate the messages into two or more segments that are spatially distinct. Fortunately, this behavior is not common. Less than 1% of the MMSI on our fishing vessel list are used by more than one vessel more than 5% of their active time.

1.2 Convolutional Neural Networks (CNN)

The raw AIS message data include a series of position messages for each vessel, consisting of a timestamp, a position, an instantaneous speed, and an instantaneous course. Although vessels transmit regularly, for crowded regions of the oceans, a considerable proportion of the messages can overlap in time and become corrupted and hence lost. As a result, the mean rate at which messages are collected by a satellite or terrestrial receiver varies considerably from region to region (fig. S6). The mean time delta between points in a given vessel history is, however, temporally correlated. When ships are in areas where the number of vessels is high, the mean time delta will be high; when they are in areas where the number of vessels is low, the mean time delta will be low. In general, boats take a significant period of time (days) to transition between areas of significantly different message rates, so the mean time between points varies only slowly as a series progresses. The exception is when vessels transition between satellite and terrestrial reception, in which case the message rate can change abruptly. Nevertheless, the mean rate remains temporally correlated.

Convolutional neural networks have shown great success in classifying images. For an image classification algorithm to be successful, it must be able to recognize the same object at arbitrary points in an image and at many different scales. Although it is unusual to use CNNs for non-uniformly sampled data, we argue that the AIS series is analogous (albeit in 1 dimension rather than 2 dimensions): Characteristic movement behaviors of vessels can be expected to appear at any point in a series and at different

apparent scales (as a result of the slow changes in mean sampling caused by variable levels of message loss).

We use two different convolutional neural networks (CNN) to aid in classifying vessel characteristics and activities:

1. Vessel characterization:
    - Classify vessels into one of eighteen classes
    - Predict vessel length, tonnage, and engine power
2. Fishing detection – classify each AIS point as either fishing or non-fishing.

These two CNNs use the same set of AIS derived features and share similar topologies. We have open sourced the code for these neural networks, and the code is available at https://github.com/GlobalFishingWatch/vessel-classification.

1.3 CNN Features

To generate features used by the CNN, AIS messages with missing or invalid fields are discarded and the messages are thinned so that there is at most one point every five minutes per MMSI. Twelve features are then derived at each time point:

- $\boldsymbol{log(1 + \delta_t)}$, where δt is the difference in the timestamps (s) of this point and the previous point.
- $\boldsymbol{log(1 + \delta_x)}$, where δx is the distance (m) between the current and previous point.
- $\boldsymbol{log(1 + s)}$, where s is the reported speed
- $\boldsymbol{\log\left(1 + s_{implied}\right)}$, where $s_{implied}$ is the speed implied by δt and δx
- $\boldsymbol{\delta COG/180}$, where δCOG is the change in course over ground (degrees) between the current and previous point
- $\boldsymbol{(t_{local} - 12) - 1}$, where t_{local} is the local time of day in 24-hour time
- $\boldsymbol{\frac{month-6}{6}}$, where $month$ is the month of the year
- $\boldsymbol{\delta COG_{implied}/180}$, where $\partial COG_{implied}$ is the change in course over ground as computed between the previous, current, and next position
- $\boldsymbol{\log(1 + d_{shore})}$, where d_{shore} is the distance to the nearest shore location
- $\boldsymbol{\log(1 + d_{anchorage})}$, $d_{anchorage}$ is the distance to the nearest anchorage. Anchorages are calculated by gridding the world with an S2 grid level 13 (mean area 1.27 km^2 per cell) using the s2-geometry-library (https://code.google.com/archive/p/s2-geometry-library/), and then finding all grid cells where at least 20 MMSI were stationary for at least 48 hours in a year
- $\boldsymbol{\log(1 + t_{anchorage})}$, where $t_{anchorage}$ is the time to the nearest anchorage visit
- $\boldsymbol{N_{neighbors}}$, where $N_{neighbors}$ is the number of vessels within 1 kilometer radius

1.4 Vessel Characterization CNN

The vessel characterization CNN has the structure shown in fig. S1A: a 12,000 time-points wide, by 12 deep array of features enters the net at the top, proceeds through a series of nine convolutional layers with a total of approximately 12 million learnable parameters. The output of the convolutional stage, 25 time-points wide by 768 features deep is passed on to four output stages as shown in fig. S1A.3: three regression stages for *length*, *tonnage*, and *engine power* and one classification stage for *vessel type*.

The nine convolutional layers shown share the structure shown in fig. S1A.2, which borrows heavily from Resnet (*32*,*33*) and Inception (*34*) neural net architectures. The left hand path in 1A is essentially a straight through path from the layer above, where the only transforms applied are average-pooling and zero-padding to match the next layer. The right hand path uses a simplified version of the inception architecture, where parallel filters are combined using concatenation and width-1 convolutions are used to control the feature depth. In this case, the two filters are a width-3, stride-2 convolution with depth-D, where D varies by layer, and a width-3, stride-2 MaxPool. The outputs of these two layers are combined by concatenating them along the feature axis and the feature depth then is reduced by performing a width-1 convolution with depth-D. Finally, the two paths are combined by summing the outputs of each path. The overall approach is similar to that used by the Inception-Resnet Architecture described in Szegedy et al. 2016 (*35*). All convolutional layers used batch-normalization (*36*) and relu nonlinearities unless otherwise specified.

The output stages consist of two, width-1 convolutional layers of depth 1024, a global pooling layer, and a final classification layer. The convolutional layers compute a set of features for each output stage at each of the 25 time points. For the classification output, this final layer consists of a 19-wide softmax layer, while the regression outputs have width-1 linear layers at the output. The objective functions are cross-entropy for classification out and mean-squared-error for the regression outputs.

For the regression outputs (length, tonnage, and engine power), we model the log of the value rather than the value itself. This method helps prevent the training being dominated by the larger vessels with their associated larger errors. However, this method also introduces the possibility of retransformation bias. To compensate for this bias we apply the so-called smearing correction (*37*,*38*) to inferred length, engine powers and tonnages. The correction for length is minimal (0.4%) because the residuals of the predicted lengths are relatively small. However, the correction for tonnages and engine powers are significant (11% and 21% respectively).

1.4 Fishing Detection CNN

The fishing detection CNN uses the same basic architecture as the vessel characterization CNN, although it is somewhat smaller, having approximately 4 million learnable parameters. The primary difference between the architecture relates to the objective of detecting whether the vessel is fishing at each time point. Note that the early layers on the network contain information that has high time-resolution, but does not contain longer time-scale information, while the later layers are in the opposite situation, having information at long time-scales, but coarse time-resolution. To take advantage of this information in different layers, we use an approach analogous to the layer fusion approach (*39*) and extract the output of each of the layers and concatenate them together

as shown in fig. S1B.2. Since the later layers have fewer time points, each time point is replicated so as to maintain a constant time width at each layer (by 2, 4, 8, 16, etc. as appropriate). These layers are then concatenated together and fed into two width-1 convolutions with depth 128. This has the effect of applying a 2-layer fully connected model *at each point*. Note that 128 time points are cropped from each edge to limit edge effects. Thus the model uses 1024 contiguous time points per application and predicts fishing for the middle 512 of them. To predict fishing at all time points, the model is scanned across the available points 512 points at a time; this scanning results in overlapping of the time points that are used as input, but the predicted time points are contiguous with no overlaps. A threshold of 0.5 is applied to the output of the CNN and the resulting series of predictions are converted into a series of time ranges for each MMSI during which fishing is predicted to be occurring. The model bases what it considers "fishing" on the training data (see next section).

1.5 Training

To obtain the identity of vessels, we matched the vessels to official registries using a vessel's name, call sign, IMO number, and flag state, accepting only vessels that had multiple matches on different characteristics. Many of these registries were obtained freely from the internet, including:

- EU's Community Fishing Fleet Register (http://ec.europa.eu/fisheries/fleet/index.cfm)
- International Telecommunications Union (http://www.itu.int/)
- Consolidated List of Authorized Vessels (http://www.tuna-org.org/)
- Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR https://www.ccamlr.org)
- Norwegian Directorate of Fisheries Vessel Registry (http://www.fiskeridir.no)
- the South Pacific Regional Fisheries Management Organization (SPRFMO) (https://www.sprfmo.org),
- Merchant Vessels of the United States (available at https://homeport.uscg.mil/mycg/portal/ep/home.do)
- Directorate of Fisheries, Iceland (http://www.fiskistofa.is)
- Pacific Island Forum Fisheries Agency (FFA) (http://www.ffa.int)

In addition, we manually reviewed the activity of a few thousand vessels to determine the gear type based on how the vessels moved, or by information obtained from the web as mentioned above. We also identified a few hundred gear or buoys based on information in the AIS messages and a manual review of the behavior. We matched or labeled 75,136 MMSI, identifying 16,689 fishing vessels, the gear types of 10,471 fishing vessels, the lengths of 14,384 vessels, the engine powers of 11,520, and the gross tonnages of 12,734. The results of these matches can be downloaded from globalfishingwatch.io.

Data for the fishing detection model primarily came from expert labelling of AIS fishing tracks, using the same method as described in de Souza et al. (2016) (*12*). Based on personal interviews with fishermen, experiences of our in-house former fisheries observers, and a review of literature, we developed judgements regarding whether points should be labeled as fishing or non-fishing based on the speed over ground, change of

direction within a defined area, spatio-temporal movement patterns, and operational time and duration of the fishing event. We used the training data from de Souza et al., and built on it by labeling another few hundred months' worth of vessel tracks from a broader variety of vessel classes. In all, more than 247,000 hours' worth of AIS tracks from 624 vessels, including over 569,000 points, were classified as fishing or not-fishing. The labeled data included 146 drifting longlines, 5 pole and line vessels, 36 purse seines, 9 set gillnets, 4 set longlines, 37 trawlers, and 3 trollers. Roughly 174,000 hours (503 MMSI) were used for training with the remainder used for testing.

The models were created in TensorFlow and trained asynchronously on Google's Cloud ML Engine using five GPU instances in parallel. Since TensorFlow does not support 1-D convolutions directly, 2-D convolutions were used instead, using kernel sizes and strides of 1 by W to perform a width-W convolution. Adam10 was used for optimization. The characterization model was trained for 600,000 examples, while the detection model was trained for 200,000 examples. The initial learning rate was 10^{-4} for characterization and 10^{-3} for detection and in both cases the learning rate was decayed by a factor of two every 50,000 examples. Otherwise, the default Adam parameters were used. Table S1 shows the vessel classes used for training the CNN.

1.6 Vessel Characterization Accuracy

Overall accuracy on the training set, when considering all classes, is about 88%, measured as the number of accurate classifications divided by total classifications. Most of the confusion of different classes occurs in cases where vessel movement behavior is likely to be similar. For example, cargo and tanker vessels are confused, as are cargo and reefer vessels (refrigerated cargo vessels) and these are both cases where there is an overlap in the boat function and, in the case of reefers, some overlap in the vessels themselves. Set longlines, set gillnets, and pots and traps were more likely to be confused with one another. We grouped the classes as shown in table S2.

Once grouped the overall accuracy improves to 95%. If the fishing and non-fishing classes are consolidated, the accuracy of distinguishing fishing from non-fishing vessels is 99%. Fig. S2A shows the confusion matrix.

The model performed the best at predicting vessel length (fig. S2C, R^2=0.9), and slightly less well at engine power (fig. S2D, R^2=0.83), and gross tonnage (fig. S2E, R^2=0.78).

1.7 Fishing Detection Effectiveness

The performance of the fishing detection model on each class of fishing vessel for which we have significant training data is shown in table S3 for various metrics. F1-Score is probably the most useful overall metric in our case since it balances precision and recall, and it is 0.93 for drifting longlines, 0.79 for purse seines, 0.9 for fixed gear, and 0.96 for trawlers.

We also compared our fishing detection to logbook data from a longliner operating in the Indian Ocean and a purse seiner in the Atlantic, and find a relatively good accuracy of identifying the location of fishing events (fig. S3).

2. Other Vessel Considerations

2.1 Assumption of One Maritime Mobile Service Identity (MMSI) to One Vessel

Ideally, each MMSI would correspond to exactly one vessel. Officially, MMSI numbers are rarely reassigned to a different vessel. Reviewing the International Telecommunications Union (ITU) list of MMSI numbers between 2013 and 2016, of 14,717 MMSIs that we identified as fishing and were also listed as fishing vessels by the ITU, we identified only one instance of an MMSI being officially reassigned to a different vessel. However, there are many examples of vessels unofficially using the wrong MMSI number. In our fishing list, about 3% of the MMSI are used by multiple vessels at the same time at some point during a year, which means our count of MMSI would undercount vessels. This potential undercounting, though, is offset somewhat by potential over counting due to the fact that some vessels change their MMSI over our five-year time period. A review of 3,643 vessels from the Consolidated List of Authorized Vessels (combined list form tuna Regional Fisheries Management Organizations) that we matched to MMSI numbers showed that 90 vessels (2.5%) changed their MMSI number at some point in the past five years. For this study, we assume that each MMSI corresponds to one vessel.

2.2 Identifying Fishing Vessels

Between 2012 and 2016, 336,221 unique MMSI were active enough to be classified by the neural network, meaning that they had at least 500 active positions over a six-month period at some point during these 5 years. Of these, 75,202 were classified in one of our fishing classes, which we call "inferred" fishing vessels. We used the best inferred label across the entire five-year period. Vessels also broadcast, in their AIS messages, the activity that they are engaged in. 81,442 MMSI consistently broadcasted that they were engaged in "fishing" in at least one year (>99% of their AIS type 5, 19, or 24 identity messages in that year said they were engaged in fishing). We call these "self-identified" fishing vessels.

14,406 MMSI were identified as fishing MMSI by the neural net ("inferred") but did not self-identify as fishing vessels, and 16,375 self-identified as fishing in at least one year, but were not on our inferred list. If our goal is to maximize precision, we would include the intersection of self-identified and inferred, but that would exclude a significant number of fishing vessels. To maximize recall, we would include the union of these two lists, but that would lower precision by including more non-fishing vessels. Many of the vessels which are on one list but not the other are nearly impossible for us to accurately identify, making precise characterization difficult. We manually reviewed 400 random MMSI that were inferred but not likely and 400 that were likely but not inferred. Of vessels we could identify, we found the inferred list had a slightly better accuracy (72% versus 64% of vessels). Also, we expect the self-reported list to be regionally biased, as we have found that different fleets have different consistency in how well they enter their identity information into the AIS device (for instance, we find that only about 30% of Taiwanese vessels consistently broadcast their names and call signs, while almost 80% of Japanese vessels do—it likely there are similar inconsistencies in how well they report other types of identity information in their AIS). The inferred list, on the other hand, relies on behavior and is the less likely to be regionally biased. For these reasons,

we used the inferred vessel list, even though it likely excludes a number of real fishing vessels.

A significant number of MMSI did not have enough activity to be classified by the neural net. The number of MMSI with at least 100 active positions in a year (positions with a speed above 0.1 knots) but not enough active positions to be categorized by the neural net, added up to 159,429 MMSI across the four years, of which about 43,445 consistently broadcast that they were fishing vessels. These very inactive self-reported fishing MMSI contributed very little to the total activity, accounting for less than 5% of the total hours of activity of all potential fishing vessels (inferred plus self-reported), and we excluded them from this study. This large number, though, suggests that many fishing vessels are inactive, a fact reflected by official statistics (see section 5.3).

Finally, we also excluded known research vessels from our list of fishing vessels, as many can look like fishing vessels to the neural net. We developed a list of a few hundred vessels that were listed as research vessels. We did not use these for training the model, but did remove them from the final vessel list.

2.3 Identifying Flag States

Each MMSI number is supposed to be nine digits long, and the first three digits, known as the MID Code (*40*), is supposed to correspond to the flag state of the vessel. A significant number of fishing vessels (17%) have fewer than nine digits or a MID code doesn't correspond to a flag state. Most of these vessels appear to be Chinese. In 2016, 61% of these MMSI operate predominantly in the Chinese EEZ, and most observe the summer moratorium and Chinese New Year. There are also other cases of vessels using incorrect MID codes, such as vessels claiming to be from Albania or Vatican City, but fishing in the Chinese EEZ. These vessels, ones that operate predominantly within the Chinese EEZ and were not claiming to be flagged neighboring countries (such as Taiwan or South Korea), were assigned to be Chinese vessels in our database. In addition, we reviewed another few hundred vessels with invalid MMSI numbers and assigned the flag state based on where the vessels operated and other identifiers. For instance, a dozen vessels that operated only with the Brazilian EEZ and had Portuguese names were assigned to Brazil. A remaining 161 MMSI we were not able to identify likely flag state. Table S4 shows the number of vessels by country by size.

2.4 Class A versus Class B

AIS has two main categories of devices, A and B. Class B devices are given lower priority in networks and broadcast a weaker signal (2 to 5 watts instead of 12.5) than class A. They also broadcast less frequently, about once every 5 to 30 seconds while moving, compared to once every 2 to 10 seconds for Class A devices. Class B devices have much poorer satellite coverage, but there are also far fewer vessels with class B that travel on the high seas. Use of class A and class B vary by fleet. Class B is most common in the Chinese fleet (>99% of MMSI), and less common in the EU (17% of MMSI).

2.5 AIS Offsets

We identified about 150 MMSI that broadcast continuous segments of activity where their positions were offset by hundreds of kilometers. These were identified by calculating the position of every ORBCOMM satellite (as well as the International Space

Station, which also provides AIS to the ORBCOMM database) every two minutes over the five years, and then calculating the distance from these satellites to every AIS point. Positions that were impossibly far from the satellite (>5,000 km) were identified as potential errors, and MMSI with a large number of errors were flagged for review. Also, MMSI whose messages were received by terrestrial stations, but which were broadcasting at more than 300 nautical miles from shore, were flagged. These MMSI with offset locations were eliminated from all spatial analyses.

3. Comparison with RFMO Hook Data and Drifting Longlines

To compare the fishing hours of drifting longline with the number of hooks set as reported by Regional Fisheries Management Organizations (RFMOs), we collected data from RFMOs for 2015 and re-gridded this data to 5 degrees. The Inter-American Tropical Tuna Commission (IATTC), the International Commission for the Conservation of Atlantic Tuna (ICCAT) and the Indian Ocean Tuna Commission (IOTC)'s online databases report the effort at this resolution. ICCAT reported fishing effort in variable resolutions: 5°x5° (90.5% of the data), 1°x1° (8.8%) and 10°x10° (0.7%). Similarly, IOTC reported at 5°x5° (75% of the data) and 1°x1° (25%). For cases where fishing effort was in a resolution different than 5°x5°, effort was transformed to that resolution by either aggregating the smaller pixels into the larger ones encompassing them (e.g. taking the sum of 25 pixels forming one 5° cell), or by dividing the effort of the larger pixels uniformly by the number of target pixels contained in it (i.e. 10x10 pixels were divided by 4). The Western Central Pacific Fisheries Commission (WCPFC) has not yet updated their 2015 statistics, and were excluded for this analysis.

To account for spatial correlation of residuals, we included a rational correlation structure of the variance. We tested several other correlation structures (exponential, Gaussian, linear and spherical) and the rational gave the best fit in terms of AIC (1694) against 1880 without correlation structure. The resulting GLS model explained 47% of the variance in the data. Globally, we found a fairly good correspondence between log-transformed hooks and log-transformed fishing hours (slope = 0.59, se: 0.03, $p<0.0001$). According to this model for each hour of fishing, we would predict about 447 hooks deployed (95% CI: 30,480 - 71,401 hooks). Deviations from the mean fit may be interpreted as regions where AIS coverage is poor (either due to vessels not having AIS or satellite reception), or regions where fishing is underreported to the RFMOs (fig. S5).

4. Fishing Effort Lost to Insufficient AIS Coverage

4.1 Limitations of Satellite and Terrestrial Coverage

A well-functioning AIS device on a moving vessel will broadcast tens of thousands of messages in a day. Only a fraction of these (~1 percent) are recorded in the global database because satellite receivers are not always overhead, and in regions with a high density of vessels, AIS signals interfere with one another, lowering reception. Terrestrial antennas, which generally have a range of tens of nautical miles, do not receive messages

from as many vessels in a given period of time as are received by a satellite, and thus are less affected by a high density of vessels.

The dataset has three main regions of poor reception: off the coast of Europe beyond the range of terrestrial receivers, in the Gulf of Mexico, and off the eastern and southeastern coast of Asia beyond the reach of terrestrial receivers. Class A devices perform better in satellite coverage than Class B. Coverage has improved as more satellites have been added to ORBCOMM's constellation, and more terrestrial receivers have been installed. In 2013, three satellites were recording data for the ORBCOMM database. In mid-2014, another six satellites were launched, and in December of 2016 another 11 were launched. Along many coastlines, terrestrial receivers provide relatively good coverage (fig. S6D).

4.2 Effect of Coverage on Fishing Hours

To aggregate fishing effort, we calculated the time between points and then assigned, to each position, half of the time to the next point and half of the time to the previous point. Because the maximum time between points before a new segment is created is 24 hours, 12 hours is added to the points at the start and end of a segment. If a position is labeled as fishing, the time associated with that position counts toward "fishing hours."

In regions where a vessel's AIS broadcast is less consistently received, we observed that the neural net identifies less fishing activity. To determine the sensitivity of fishing hours to poor AIS coverage, we took the data for 2016 and randomly removed points, creating 12 tables with 90% to 0.5% of the points as the full dataset. We then ran the neural net classifier on this data and calculated fishing hours. The accuracy of the neural net decreases as the number of AIS points per day is reduced. Instead of using points per day, we used "observed five minute intervals"—that is, the number of five minute intervals in which at least one position was observed. The reason for this interval is that the neural net thins data to once every five minutes before processing, so multiple positions within one five-minute interval does not affect our ability to identify behavior. We found that above 100 observed five minute intervals per day per vessel, the fraction of time estimated fishing by the neural net was unaffected by coverage. At a one degree pixel, we identified every cell that averaged more than 100 observed five minute intervals per vessel per day in 2016, and then measured how the fraction of time fishing for that cell decreased as data was removed. The fraction of time fishing decreased as points were removed from these cells. We found we could model the fraction of fishing effort missing, $F_{missing}$, as a function of observed five minute intervals, as:

$$F_{missing} = 1 - \beta * e^{(-\frac{x}{\gamma})} \qquad 1$$

where x is the number of observed five minute intervals per vessel per day, and beta and gamma are fit to the data and vary by gear type. The different gear types show different sensitivity to low point density. Longliners are the least sensitive, remaining fairly accurate down to eight observed five minute intervals per vessel per day, while purse seiners are the most sensitive, under predicting fishing by 20% below 44 observed five minute intervals per vessel per day (fig. S7).

Because coverage has improved over the past five years, the fraction of fishing lost to incomplete coverage has decreased in this time period. Table S6 estimates, using

equation 1, the fraction of missing fishing due to incomplete coverage each year by gear type.

We will provide at globalfishingwatch.io a layer that gives the number of observed five minute intervals per vessel per day per at 0.5 degree resolution averaged over the year, which can be used to identify areas with poor coverage.

4.3 Data Gaps Longer than 24 Hours

The previous section discussed fishing effort lost due to having a small number of positions per day, but it did not address the situation where there are gaps in AIS coverage over 24 hours in length. Such gaps can result when a vessel turns off its AIS or travels in a region with exceptionally poor satellite coverage.

To characterize the activity lost due to these gaps, we calculated every gap in transmission longer than 24 hours and the distance from shore at which the gap started and ended. We believe that during the majority of the time associated with these gaps the vessels are actually in port with their AIS turned off. If we instead assume that vessels are active during these gaps, we dramatically over count days at sea. For example, we compared the number of days at sea for vessels larger than 18 meters in our database with data from an official EU report from 2014 (*41*). Our number of days at sea was within 10% of the official statistics. The estimate from AIS was below the official estimate, but it is likely that this is in part because some EU vessels acquired AIS devices part way through the year, which means we likely see more than 90% of the activity of EU vessels equipped with AIS. If we assume that vessels were active during all gaps that start and end more than 10 nautical miles from shore, our count exceeds the official statistics by 47% and would exceed it by even more if not for the fact that vessels added AIS during the year. As a vessel can travel about 10 nautical miles in an hour, it is most likely that a lack of satellite or terrestrial coverage for an hour or two missed the vessel's transit to port, and then the vessel turned off its AIS in port. If this is the case, we would expect to see gaps be longer and more frequent the closer they start to shore, and we'd expect to see gaps generally short in distance. This is in fact the case; the average time in gaps is twice as long close to shore as it is for gaps that start 100 nautical miles to sea. Also, over 70 percent of the time in gaps is spent in gaps with a distance of less than 100 kilometers between the start and end of the gap in transmission.

As we lack good statistics on the estimated hours at sea of other fleets, it is difficult to do a similar comparison. The distribution of gaps – most are very short in distance and start closer to shore, similar to the EU – suggests that rampant disabling of AIS is not a global phenomenon. In specific regions and fleets, though, further analysis can assist with identifying if a significant fraction of the vessels are disabling their AIS.

5. Fraction of Global Fishing Effort in AIS Data

To estimate the fraction of global fishing effort in AIS data beyond 100 nautical miles from shore, which is half the distance to the edge of most countries' Exclusive Economic Zones, we calculated energy consumed by vessels of different sizes, observed the distance from shore fished by different size vessels, and then estimated the fraction of vessels with AIS by size class.

5.1 Energy Expended by Size

For each vessel, we estimate the energy expended in a given year (E) by multiplying its engine power (KW), as inferred by the neural net, by the number of hours the vessel is active, times the average load factor of a fishing vessels engine. The load factor accounts for the fact that the engine rarely operates at full capacity. According to unpublished research by the International Council on Clean Transportation, the average load factor of a fishing vessel 0.39. Independently, using AIS-implied vessel speed and characteristics, we estimated that gear-specific average load factors (LF) range between 0.25 and 0.37. These estimates do not account for higher engine loads when vessels drag gear in the water and thus are underestimates of the load factor for trawlers and dredgers. We used the average of these estimates, 0.35 for our estimates. This load factor of 0.35 was also used to estimate the total energy expended by fishing vessels in 2016. Other global estimates of the energy expended by fishing vessels (*8,42*) use engine power multiplied by active time without accounting for load factor, so to make our estimates directly comparable, one would have to divide by 0.35.

We examined the relationship between vessel length (L) and energy expended (E) using a log-log regression (fig. S4A) of the form:

$$\log(E) = a + b * \log(L) \qquad 2$$

The regression results: a = 4.85 (s.e. = 0.07, $p < 0.001$) and b = 2.34 (s.e. = 0.02, $p < 0.001$) can be rewritten as a power law relationship such that:

$$E = 90 * L_v^{2.34} \qquad 3$$

where E is in kilowatt hours per year and L is in meters.

5.2 Distance from Shore Fished by Size

Including only vessels that spent less than 10 percent of their time in gaps longer than 24 hours that start and end more than 20 nautical miles from shore, thus excluding vessels that might be missing activity, we charted the distance from shore fished by vessels 6-12 meters (2,551 vessels), 12-18m (10,323 vessels), 18-24m (16,297 vessels), 24-30m (9,339), 30-36m (4,940 vessels), and >36m (4,605 vessels). Vessels under 12 meters spend 0% of their time fishing >100 nautical miles from shore, while vessels over 36 meters spend more than 65% of their fishing time beyond this distance (fig. S4C).

5.3 Fraction of Active Vessels with AIS

The FAO State of the World Fisheries 2016 (*43*) estimates that there are 2.9 million motorized fishing vessels, with 2.5 million under 12 meters, 350,000 between 12 and 24 meters, and 64,000 larger than 24 meters. A second FAO report (*44*) from 2015 estimates that there are actually 2.5 million motorized vessels, with 2.1 million under 12 meters, 320,000 between 12 and 24 meters, and 43,700 larger than 24 meters. One reason for the high uncertainty might be that a large number of vessels are inactive. In the EU for instance, 25% of vessels under 12 meters, 7% percent between 12 and 24 meters, and 10% above 24 meters were inactive in 2014 (*41*).

We compiled our own dataset by collecting the most up-to-date official fleet size estimates for China, Japan, Korea, Taiwan, Norway and Chile, plus estimates of active vessels for all the European Union member states. Data gaps for remaining countries (excluding Thailand, Pakistan, Egypt, Guatemala, Guinea, Nicaragua, and Bahamas) were filled with FAO's vessel database, obtained from FAO staff. Our survey found far fewer vessels under 12m, likely due to different methodology for identifying these vessels. Table S5 shows the number of vessels by size class, globally, from the FAO high and low estimate, and from our survey, and table S4 shows, by country, the results of our survey for top countries.

We decided to use these three sources to estimate the high, low, and median guesses for the number of vessels in the world, assuming that the median value of these three was the best guess. So, for vessels under 12 meters, we estimated a high of 2.5 million, a low of 1 million, and a best guess of 2.1 million. For vessels 12 to 24, we estimate those figures are 312,000, 320,000, and 350,000, and for vessels over 24 meters, 43,700, 50,913, and 64,000. Because we are modeling active vessels, and not total vessels, we then reduced these estimates to account for inactive vessels, assuming the same number are inactive by size class as is found in the EU, similar to other global studies (*42*). It is likely that this fraction actually varies around the globe.

We also estimate the fraction of vessels above 300 tons (IMO regulation) that we see in AIS. Using a published relationship between vessel length and tonnage (*7*), the average length of 300 tonnage vessel is 37 meters. Using this threshold, we estimate that between 3847 and 4211 fishing vessels are above this limit and required to use AIS globally, excluding Russia and China, for which we were not able to obtain reliable information. We see between 77 and 84% of these vessels in the AIS data (excluding China, Russia, and vessels with invalid MMSI numbers).

5.4 Estimating Fraction of Fishing Effort Above 100 Nautical Miles

The energy expended fishing greater than a given distance from shore, E_d, over the course of one year would be

$$E_d = \sum_{l_{min}}^{l_{max}} N_{world,l} * F_{l,d} * E_l \qquad 4$$

where $N_{world,l}$ is the number of active vessels of a given length in the world, $F_{l,d}$ is the fraction of fishing effort by a vessel of given length beyond a given distance from shore, and E_l is the average energy expended by a vessel of size l, as calculated in equation 3. To calculate the fraction of energy expended by vessels with AIS, $FE_{ais,d}$,:

$$FE_{ais,d} = \sum_{l_{min}}^{l_{max}} \frac{N_{AIS,l}}{N_{world,l}} * F_{l,d} * E_l \qquad 5$$

For small lengths, our equation for E_l might be less accurate because we don't have many examples of small vessels in AIS (almost none shorter than six meters), making this number difficult to calculate for distances close to shore. However, small vessels stay close to shore and should not affect $FE_{ais,d}$ above 100 nautical miles (fig. S4B).

To apply equation 5, instead of using three size classes, we need to estimate a continuous distribution of vessels in the world. To do this, we fit a simple third degree

polynomial to the log(vessels larger than a given length) versus log(length), using only four points: the number of vessels larger than 0m, (V_0), the number larger than 12 (V_{12}), the number larger than 24 (V_{24}) and the number of vessels larger than the length above which all vessels in the world have AIS (V_{AIS}). Setting the distribution equal to the number of vessels with AIS, V_{AIS} is likely to be above 40 meters, the size at which most vessels exceed the IMO gross tonnage requirement (300 tons) to have AIS, and less than 150 meters, the size of the largest fishing vessels in our database. Varying this number does not change the total number of vessels larger than 24 meters, and we saw little difference in the results in varying the number between 50m and 140m. We used 70m, partly because it gave a slightly more conservative result.

We generated three distributions of vessels, using our low, medium, and high estimates of the number of active fishing vessels in the world. After generating these estimated distributions of vessels, we applied the empirically calculated $F_{l,d}$ (fraction of fishing above a given distance d from shore by a vessel of length l) from fig. S4B, using the bins of length size shown, and making the conservative estimate that all vessels under 12 meters have the same distribution from shore. We find that above 100 nautical miles, halfway to the EEZ boundary, vessels with AIS likely contribute 50 to 70 percent of the fishing effort, with a best guess of 60%. This estimate does not include gaps in transmission or losses due to incomplete coverage. This method also estimates that vessels with AIS contribute between 26% and 34% of the global fishing effort, by energy expended, of all vessels in the world (including those without AIS). This estimate is less certain than our estimate for vessels farther away from shore, as we have less confidence in using this method to estimate the activity of smaller vessels.

6. Area of the Ocean that is Fished

To estimate the area of the ocean fished, we used an equal area grid with a size of 0.5 degrees at the equator. While different fishing gears may impact marine ecosystems at different scales (e.g. trawlers impact the marine environment on a more localized scale than longliners) (*19,45*) we map impacts on a 0.5 degree grid to allow for most direct comparisons with previous estimates of global fishing effort (*14,15*). If we use a 0.25 grid at the equator instead of 0.5, the area fished by vessels with AIS drops from 55% of the ocean to 49%.

To account for areas of the ocean where vessels might not have AIS, we analyzed all the EEZs where the associated flag state had less than 50% of its vessels larger than 24 meters based on our survey of vessels with AIS (table S4, fig. S6) (see section 5.3). We also included all areas of the ocean where we had less than 20 positions per day per vessel for class A devices (fig. S6D), as our fishing algorithms deteriorated below this level (fig. S7). For every pixel in these regions where we did not see fishing activity, it is possible that fishing did indeed occur, and this unfished area in low-coverage areas equaled ~18% of the ocean. This likely significantly over counts the potential area in fishing, as it is unlikely that every bit of these exclusive economic zones is fished.

Another reason we do not believe that these areas are fished is that we see fishing vessels transiting through many regions of the ocean without engaging in fishing. About 15 percent of the ocean, at our same 0.5 degree grid, has fishing vessels transiting through the region, but no fishing activity. For instance, vessels cross the Southern

Atlantic between Argentina and South Africa, or transit across the northeast Pacific between Alaska and Seattle, but never exhibit fishing behavior (fig. S6A,B).

For the comparison with land area, global land cover surveys estimate that 38% of the ice-free land is covered in ice (*17*). Adjusting for the area of land covered by ice (10%) (*46*) gives 34% of the total land.

7. Price Elasticity of Fuel Demand Analysis

7.1 Data for Price Elasticity of Fuel Demand Analysis

We obtained global daily fuel prices of marine diesel oil ($/MT) from bunkerindex.com, averaged them for each month between 2014 and 2016, and adjusted prices for inflation using global inflation rates from data.worldbank.org. Estimates of hours at sea were aggregated monthly for every vessel that has been active at least 10 days every year since 2013 and that transmits more than 10 positions per day during at least two thirds of its time at sea. These restrictions select for vessels that have been exposed to all treatment levels (i.e., highest and lowest fuel prices) and filter out vessels that spend a large fraction of their time in areas with poor coverage. Lastly, we excluded MMSI that are used by multiple vessels. The resulting data set includes 6,015 distinct vessels from 79 flag states, between 7.2 to 127.8 meters in length (mean = 30 m) and across multiple gear types. Table S7 shows the number of vessels, by flag state, in this analysis.

7.2 Methods

We estimated a log-log linear panel regression model with multiple fixed effects, where MMSI is an individual vessel fixed effect to account for observed and unobserved time-invariant characteristics (e.g. flag state, gear type, length). Month and year fixed effects account for seasonal variation and time trend. Standard errors were adjusted with a heteroskedasticity-consistent covariance matrix of the coefficient estimates. These analyses were performed using the *plm*, *sandwich*, and *lmtest* R packages.

7.3 Results for Price Elasticity of Fuel Demand Analysis

The results of the panel regression suggest that the short-run price elasticity of fuel demand is small (-0.066, $p < 0.001$) (table S8). A different model specification with flag state and gear type fixed effects instead of individual vessel fixed effects gives a similar result (-0.050, $p < 0.001$). Similarly, removing the restriction on the number of positions per day had a small effect on the elasticity estimate (-0.055, $p < 0.001$); however, the resulting sample of vessels becomes heavily biased towards Chinese vessels.

8. Time Series of Fishing Effort

To analyze the temporal, global footprint of fishing, we constructed time series of daily fishing effort, as measured in hours of fishing per day, from late 2013 to early 2017 for non-Chinese vessels, only including vessels that were active all years (fig. S9). We used local time instead of UTC to better capture daily and weekly cycles.

From this time series, we derived additional time series for these vessels to investigate the seasonal, weekly, and yearly signals:

1. A long term, linear trend *T* (fig. S9A).
2. A weekly signal *W*, calculated by taking the original time series divided by a seven-day running average and harmonically averaging across weeks (fig. S9B).
3. A signal for yearly variation not including Christmas or weekdays *Y*, calculated by taking a seven-day running average, averaging across years, and then linearly interpolating values from the start to the end of the Christmas holiday (fig. S9D).
4. A signal for the Christmas holidays only *H*, which takes the seven-day running average for only the Christmas holiday and then assuming unchanging, flat fishing effort the rest of the year (essentially the complement of 2, fig. S9D).

To compare the contribution of Christmas, intra-week variation, and intra-year variation, we then computed the coefficient of partial determination (partial R^2) associated with each of the above signals. We found that the Christmas and intra-week variations each explained over nearly twice as much variance as intra-year variations with Christmas removed (fig. S9E). This implies that the time variation in non-Chinese fishing effort is dominated by weekly variations on short time scales and the winter holiday season on long time scales. In other words, at the global scale, non-Chinese fishing vessels take weekends and Christmas off and show little seasonality otherwise. Specific fleets do show seasonal fluctuations (as shown by Fig. 3), but when aggregated across all fleets, these fluctuations are small compared to weekends and holidays. Interestingly, it also appears that a significant number of fishers take Fridays off as well (fig. S9B). We should note that this dataset is heavily biased toward Europe and North America.

Chinese vessels show a very different temporal pattern. Although the fishing effort of Chinese vessels also shows strong temporal variation due to cultural and political factors, notably the drops in fishing due to the Chinese fishing moratoria and the Chinese New Year, this variation has added complexity due to its larger temporal extent and the timing variation of the Chinese New Year. In addition, the regular weekly component of the signal found in the non-Chinese fleet appears to be much weaker or missing altogether in the Chinese fleet.

9. Comparison of Net Primary Productivity with Fishing Effort

Primary production and sea surface temperature data were accessed from http://www.science.oregonstate.edu/ocean.productivity/index.php. To analyze the relationship between NPP and fishing effort, we fit a generalized additive model with Gaussian distribution of log(fishing effort) as a linear function of log(NPP) and a thin plate regression spline of latitude and longitude. Fishing hours and NPP had an equal area projection with grid size of 1 degree at the equator.

10. Response of Fishing Effort to Temperature and El Niño

10.1 Analysis of Effort in Indian Ocean

To assess the relationship between sea surface temperature and fishing effort by the longline fleet in the southern Indian Ocean we obtained Optimum Interpolation (OI) Sea Surface Temperature (SST) V2 from the NOAA Earth System Research Laboratory

(ESRL) for 2013 - 2016 (https://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.html). The position of the 19ºC isotherm in the southern Indian Ocean from 2013 to 2016 was found by analyzing ocean temperatures in the region between -37.5 and -27.5 latitude and 40 and 114 longitude. We then calculated the latitudinal center of mass of fishing effort by drifting longlines for this same region. Correlation coefficients between the 19ºC isotherm and center of mass of longline fishing were estimated by the cross correlation function (lags in weeks). The strongest correlation identified fishing effort lagging the 19ºC isotherm by 11 weeks ($R = 0.84$, $p < 0.05$).

10.2 El Nino Analysis

To estimate the effect of El Nino on the fishing fleet in the central Pacific Ocean we first identified the mean longitudinal position of the Pacific fishing fleet based upon vessel activity. As a metric, vessel activity is more uniform across gear types (in the central Pacific, purse seines and longliners) than fishing activity, which differs considerably between gears. The strength of El Niño was estimated from the Multivariate ENSO Index, obtained from https://www.esrl.noaa.gov/psd/enso/mei/index.html. Both variables exhibited strong overall trends as well as large seasonal components. Therefore, to estimate the effect of ENSO score on fleet dynamics we developed an autoregressive model with one-fold differencing (ma1) incorporating ENSO score and date as our predictors. Following a one-fold differencing model, the residuals exhibit no patterns or autocorrelation. Overall the model predicted a 3.5 +/-0.75 degree shift in the mean longitude of the central Pacific fleet for each one unit change in ENSO score ($Z=4.7$, $p <.001$). A similar model using individual fleets identified a longitudinal shift of 2.7 +/-1.1 ($Z=2.4$, $p = .02$) and 4.8 +/-1.2 ($Z=3.9$, $p < .001$) degrees for a one unit change in ENSO score, for the longline and purse seine fleets, respectively.

11. Data and Code Availability

Global raw AIS data is available from satellite providers ORBCOMM, Spire, and Exact Earth. These are commercial products, but it is possible to obtain a global dataset of fishing vessels through one of these providers for less than the cost of storage and processing of the data. Opportunities to obtain the raw AIS are also possible through research partnerships with one of Global Fishing Watch's many collaborators. Anonymized labeled AIS data is also available for training models at github.com/GlobalFishingWatch/vessel-classification, along with open source code we developed for the neural network.

The following datasets are available for research purposes at globalfishingwatch.io.:

- Fishing effort in fishing hours for each vessel for each day, gridded at 0.1 degrees.
- Fishing effort in fishing hours and vessel presence in hours grouped by flag state, gear type, and day, at 0.5, 0.1, and 0.01 degrees.
- A list of the ~70,000 identified fishing vessels (by MMSI number), as well as their identified flag state, gear type, flag state, length, gross tonnage, engine

power, years that the vessel is active, total fishing hours for that vessel for a given year, whether we have matched the vessel to an official registry, and the registries we have matched the vessel to if we have matched it. Length, tonnage, engine power, and gear type are inferred from the neural net unless the values are available from registries.

- Distance traveled, by month, of all fishing vessels used to generate the fuel elasticity calculations.

Supplemental Figures

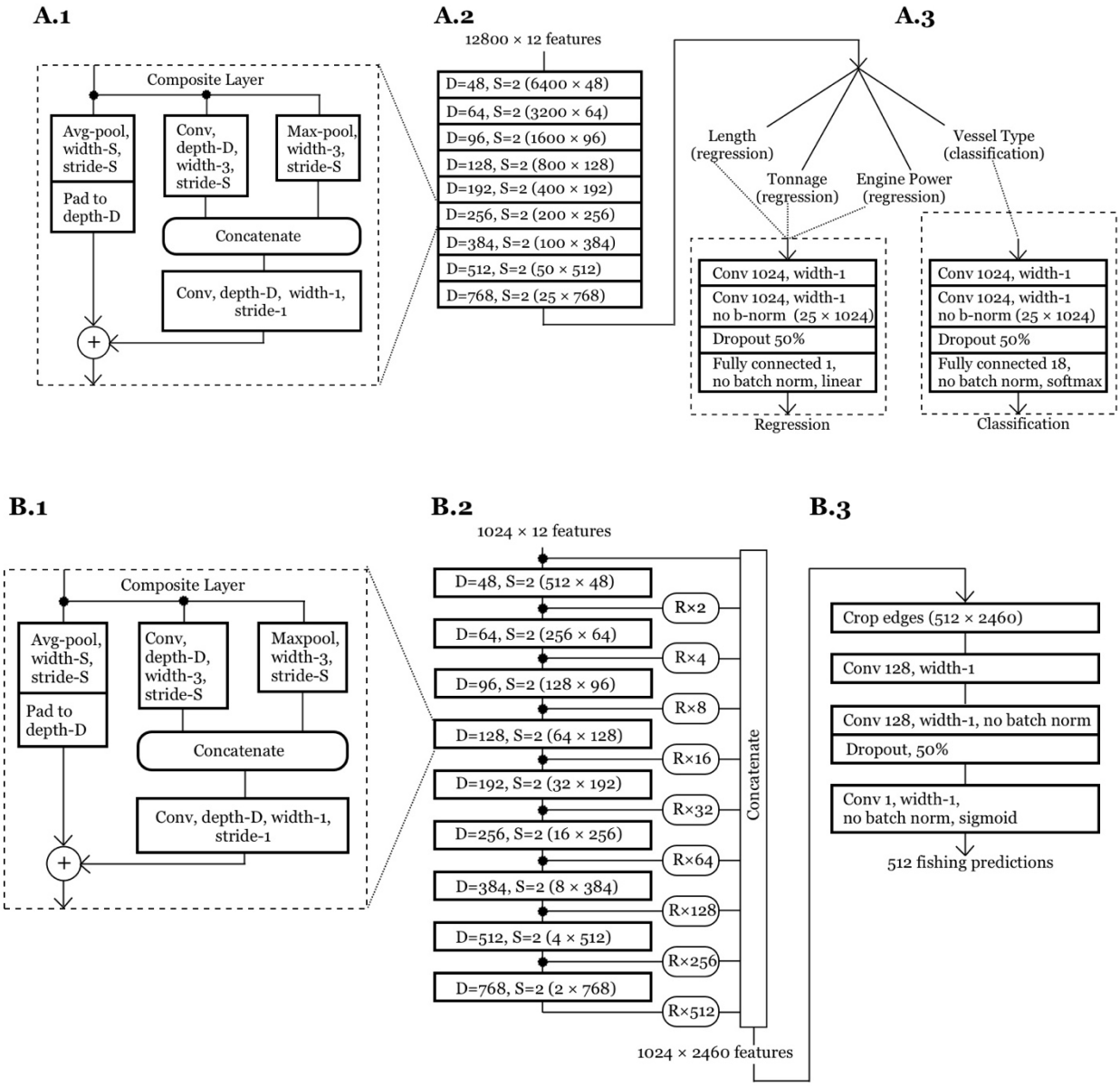


Fig. S1. Convolutional Neural Network (CNN) Architecture

(**A**) The neural network for classifying vessels consists of nine shared convolutional layers followed by output stages – each consisting of two, width-1 convolutional layers – for vessel classification and length, tonnage and engine-power regression. (**B**) The fishing detection CNN, which labels each point as fishing or non-fishing, uses a similar architecture, but we extract the output of each of the layers, replicating points to match sizes and concatenating the layers as shown in the figure. Two width-1 convolutions in the output stages combine the high resolution, local information, from the early layers, with the coarse, global information from the later layers to classify each time point.

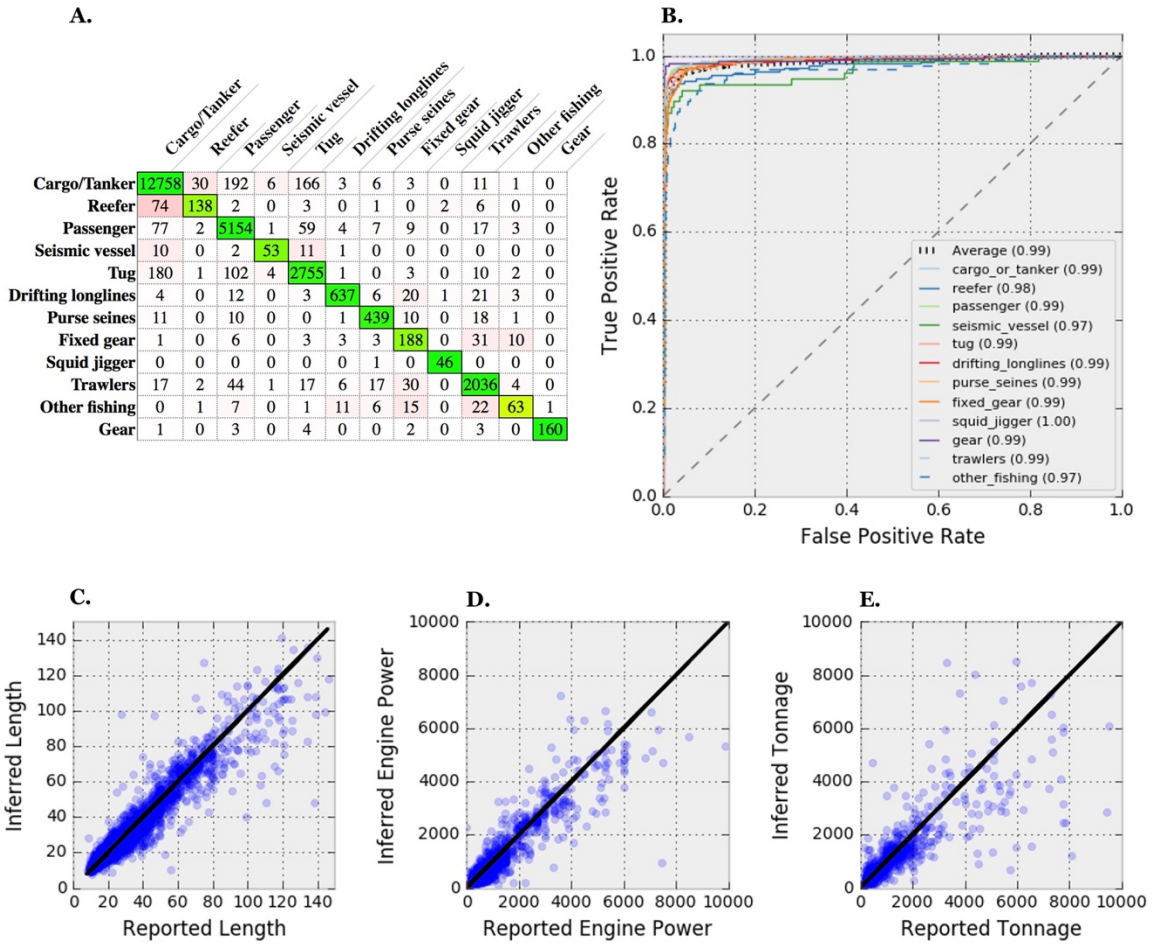
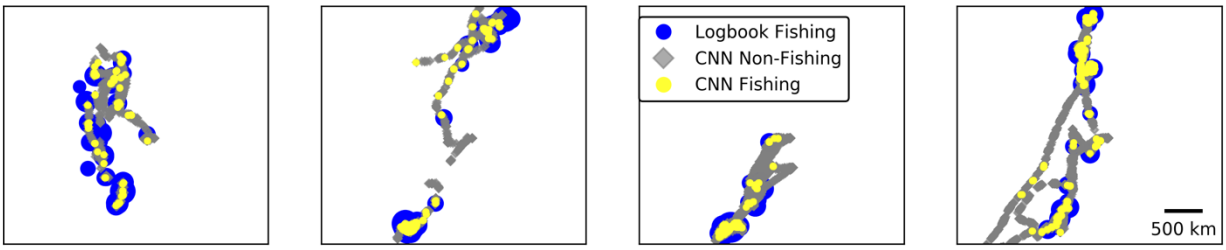


Fig. S2. Accuracy of CNN at Characterizing Vessels

(**A**) Confusion matrix showing the accuracy of CNN on test data. **(B**) Receiver Operator Characteristic (ROC) curves for the various vessel classes. (**C**) Reported length from vessel registries against inferred length from CNN (R^2 = 0.9). (**D**) Reported engine power against inferred engine power from the CNN (R^2 = 0.83). (**E**) Reported gross tonnage against inferred gross tonnage (R^2 = 0.77).

A.



B.

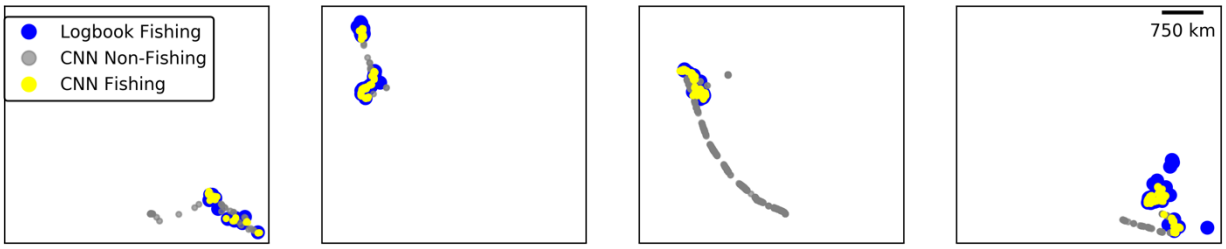


Fig. S3. Fishing Prediction Accuracy of CNN

(**A**) The locations of fishing indicated by logbook data for a French tuna purse seiner in the equatorial Atlantic in 2015 are shown in blue, with area of dot indicates hours of fishing. AIS data overlain, with yellow dots indicating where the CNN predicted fishing. Grey diamonds indicate positions from the AIS that were labeled as non-fishing. Each panel in the image covers 90 days of an approximately one year total span. To hide the identity of the vessel, the coastlines are removed and each panel is oriented differently with respect to north. (**B**) A comparison of logbook data for a Seychelles tuna longliner operating in the Indian Ocean in 2015 compared with the areas of fishing predicted by the CNN. As with (A), the coordinates are rotated in each panel to hide the identity of the vessel, and each panel represents 90 days of activity.

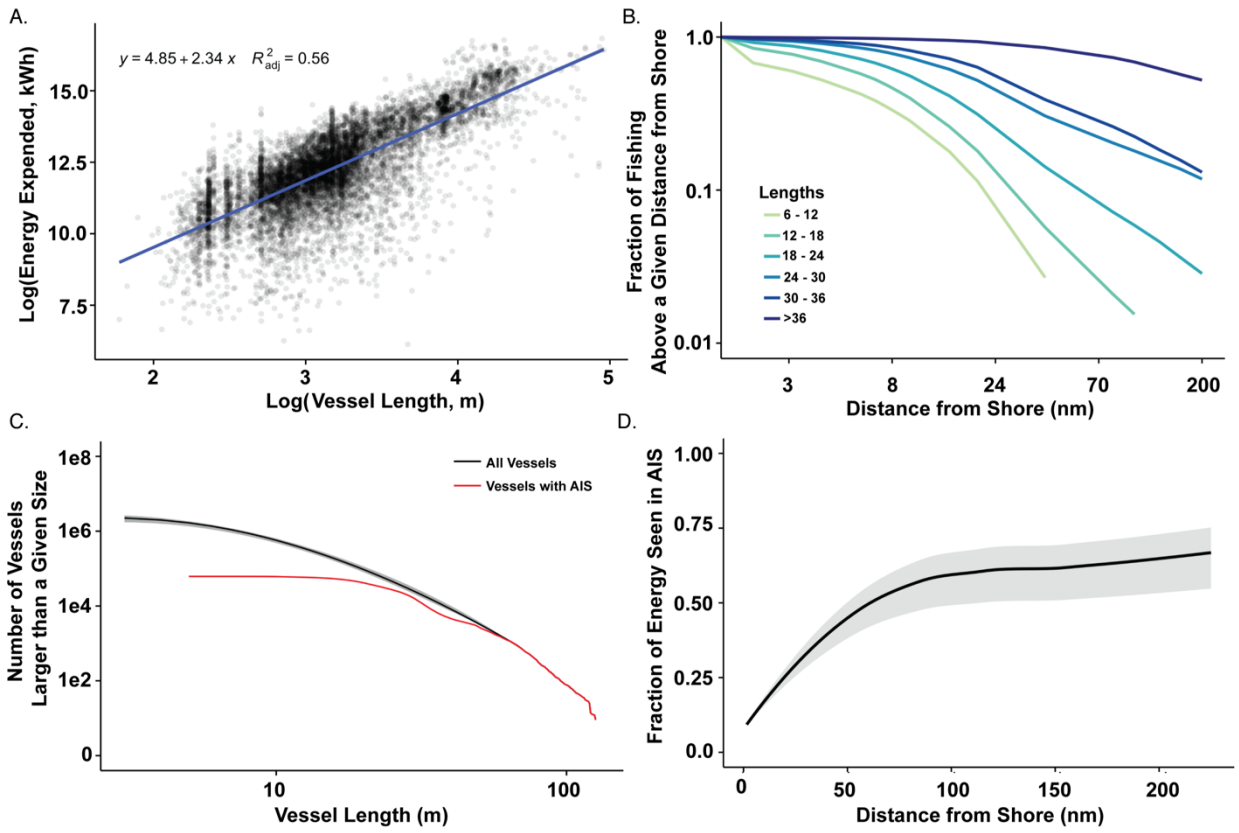
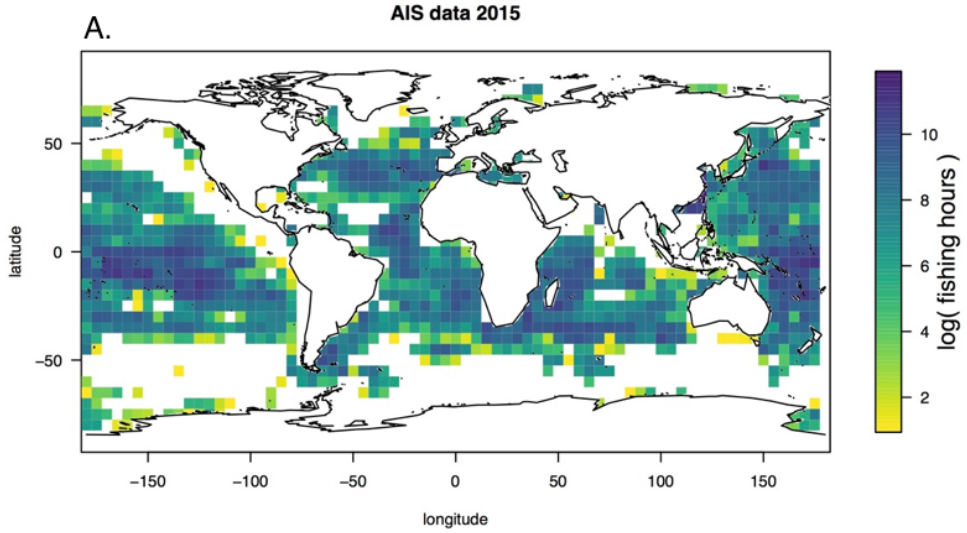
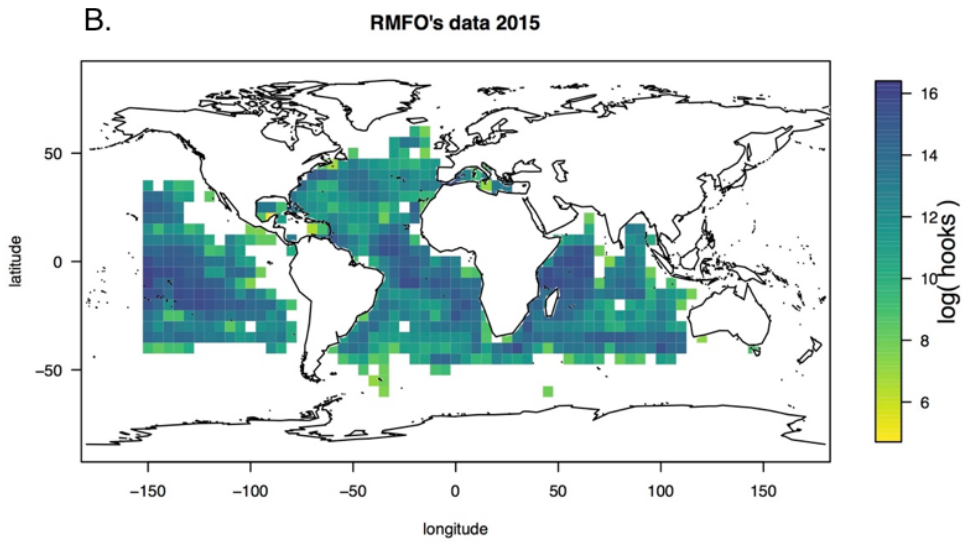


Fig. S4. Fraction of Fishing Effort in AIS Data

(**A**) The energy expended by vessels over the course of one year follows a power law distribution with exponent ~2.34 of length. (**B**) Larger vessels fish farther from shore than smaller vessels. (**C**) AIS covers the majority of large vessels, but only a small fraction of smaller ones. The red line is for vessels with AIS, while the grey line is the estimated number of vessels in the entire world. This approximation estimates that vessels with AIS comprise about 56% of vessels larger than 24m, 9% of vessels 12 to 24 meters, and 0.2% of vessels under 12m (**D**) Combining these relationships, we can estimate the fraction of fishing effort in AIS as a function of distance from shore, with about 50 to 70% of fishing effort above 100 nautical miles by vessels with AIS. The uncertainty (grey bar) is calculated by using an upper and lower bound on the number of vessels in the world (see section 5 of Methods and Materials).


A.
AIS data 2015
latitude
longitude
log(fishing hours)


B.
RMFO's data 2015
latitude
longitude
log(hooks)

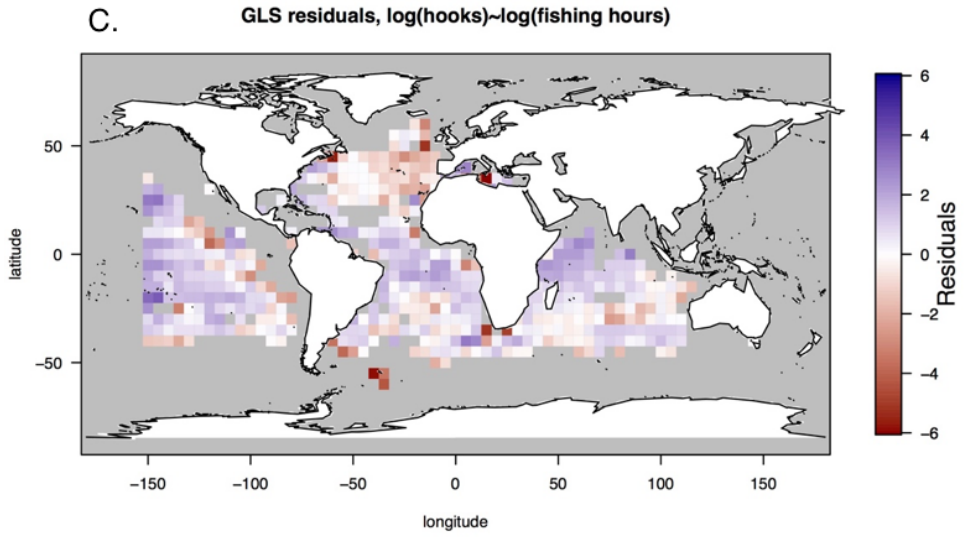

C.
GLS residuals, log(hooks)~log(fishing hours)
latitude
longitude
Residuals

Fig. S5. Comparison with Regional Fisheries Management Organizations

(**A**) Hours of fishing by drifting longlines in 2015 from the AIS data based on the neural net classification. (**B**) Hooks set by longlines according to Inter-American Tropical Tuna Commission, the International Commission for the Conservation of Atlantic Tunas, and the Indian Ocean Tuna Commission. Globally, we found a fairly good correspondence between log hooks and log hours (slope = 0.59, se: 0.03, $p<0.0001$), with 47% of the deviance in the AIS explained by the RFMO data (**C**), Residuals for a generalized least squares (GLS) regression between reported hooks set and fishing hours. Blue areas show where there is comparatively more fishing in the RFMO data than the AIS database, and red areas indicate where AIS predicts comparatively more fishing. The Western Central Pacific Fisheries Commission (WCPFC) has not yet reported its 2015 data, accounting for the lack of hook data in the Western Pacific.

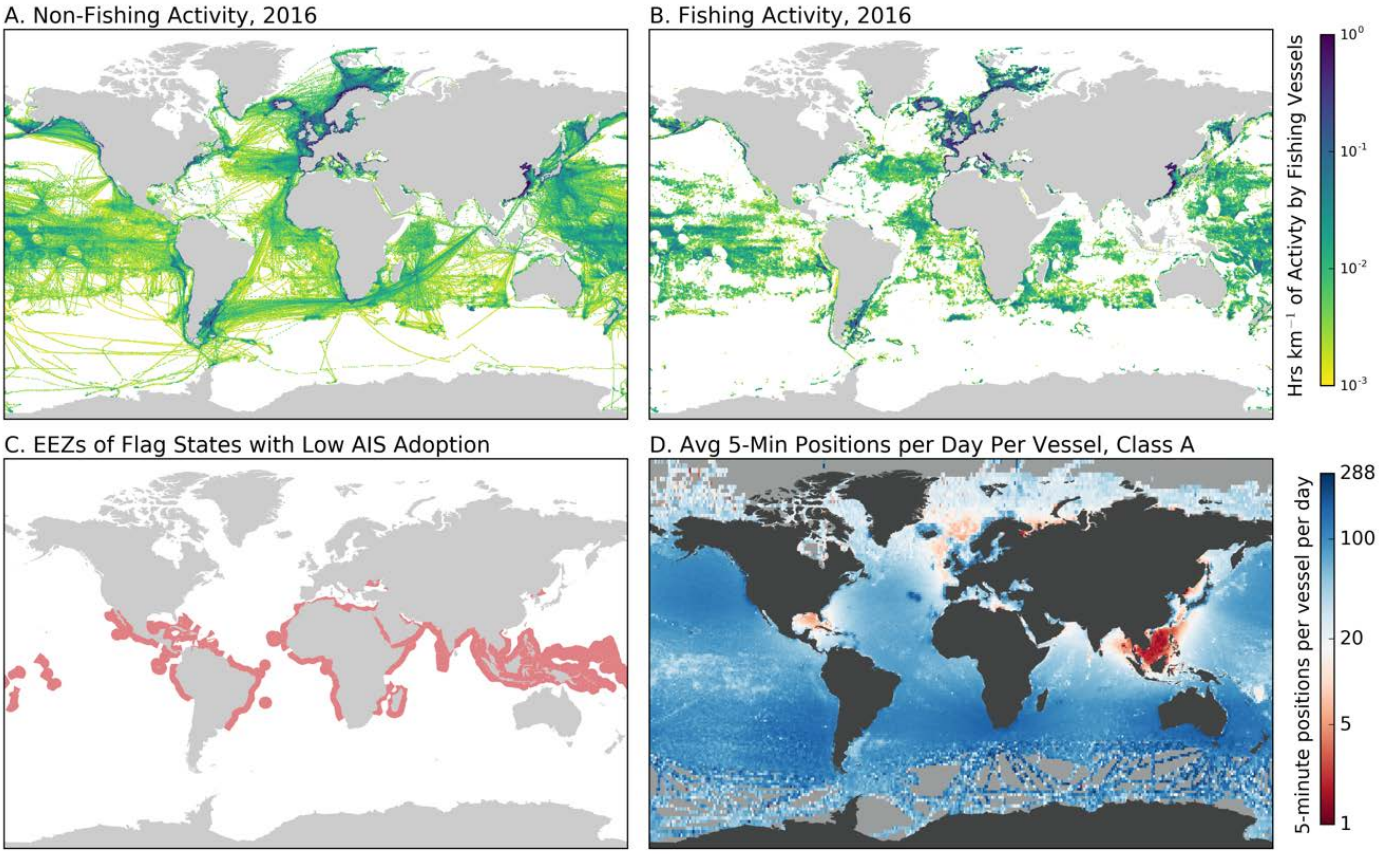


Fig. S6. AIS Coverage

(**A**) Non-fishing activity by fishing vessels, showing transits through many regions that are not fished in (**B**) Fishing Activity. (**C**) EEZs where the corresponding flag states have either an unknown number of vessels larger than 24 meters, or where less than half of the vessels larger than 24 meters have AIS, according to table S4. (**D**) Average observed five minute intervals per vessel per day in 2016 for Class A AIS devices (positions are thinned to one every five minutes, so the maximum number of positions in a day for a vessel is 288). Grey areas represent no data due to a lack of vessels with AIS being present in 2016.

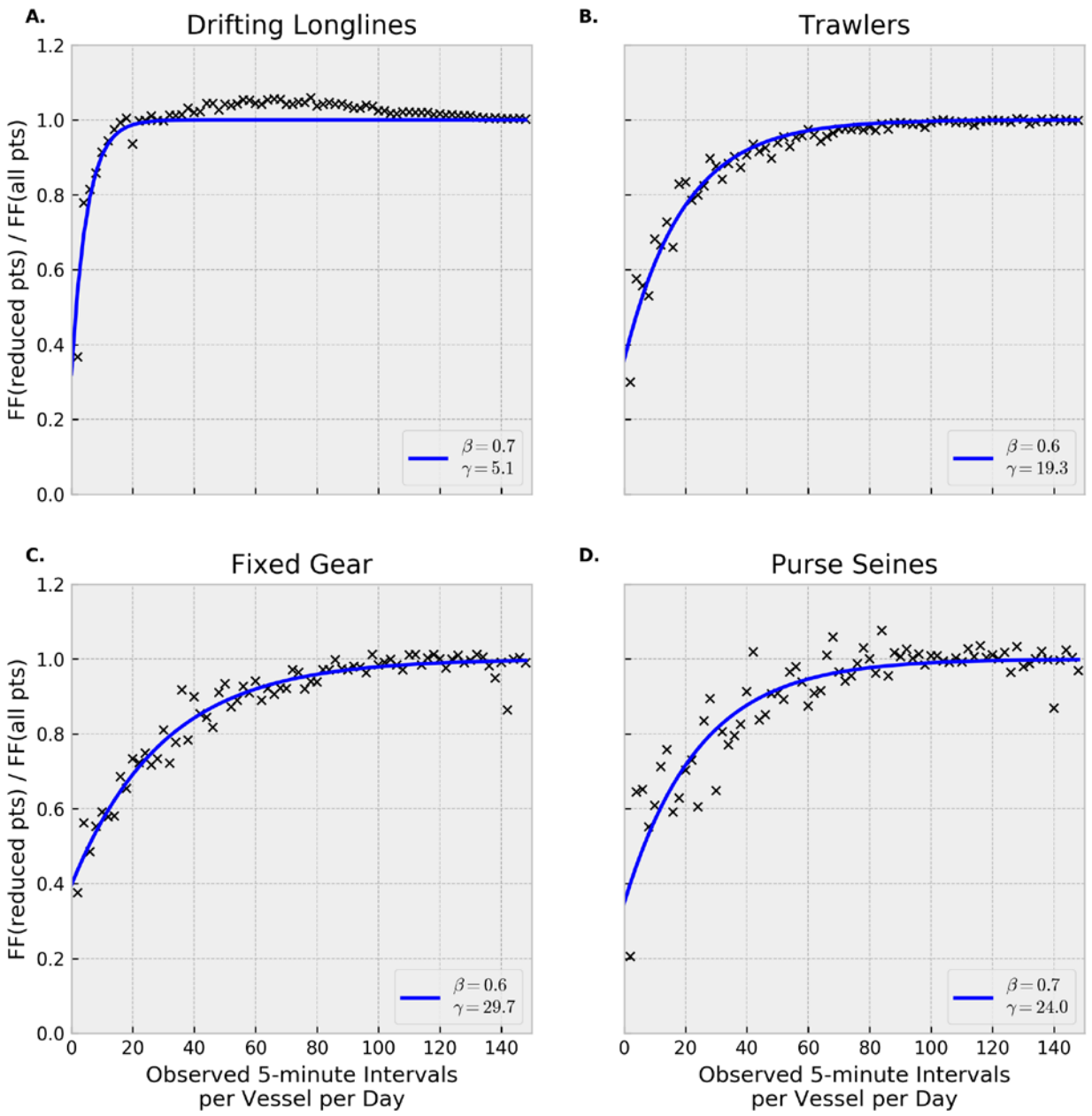


Fig. S7. Fishing Effort Not Observed Due to Poor Coverage

Taking 2016 data, we observed the estimated fishing effort from AIS in every grid cell with an average of more than 100 five-minute observed intervals per day per vessel. We then measured how the fraction of time fishing (FF) decreased as positions were randomly removed, where the fraction at 100 observed intervals per day = 1. An equation of form $FF = 1 - \beta * e^{(-\frac{x}{\gamma})}$ was fit to each gear type, where x is the observed number of five minute intervals per vessel in the grid cell. Purse seiners, (**A**)**,** were the most susceptible to decreasing the number of positions per day, while drifting longlines, (**D**)**,** were the least susceptible.

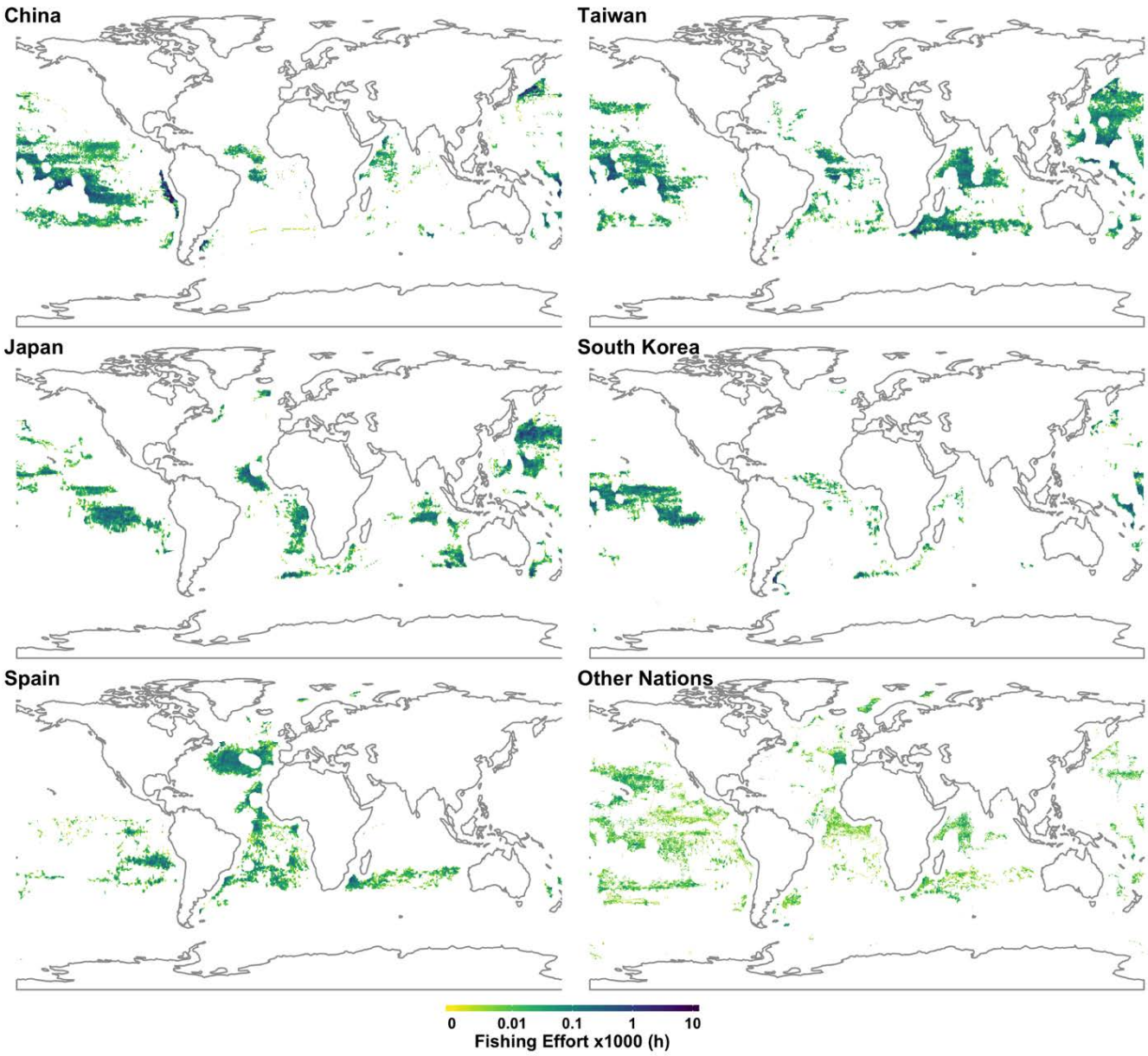


Fig. S8. High Seas Fishing

Five flag states accounted for 80% of the fishing hours in the high seas in 2016: China (33%, 2.3M fishing hours), Taiwan (21%, 1.5M fishing hours), Japan (15%, 1.1M fishing hours), South Korea (9%, 630k fishing hours), and Spain (6%, 460k fishing hours).

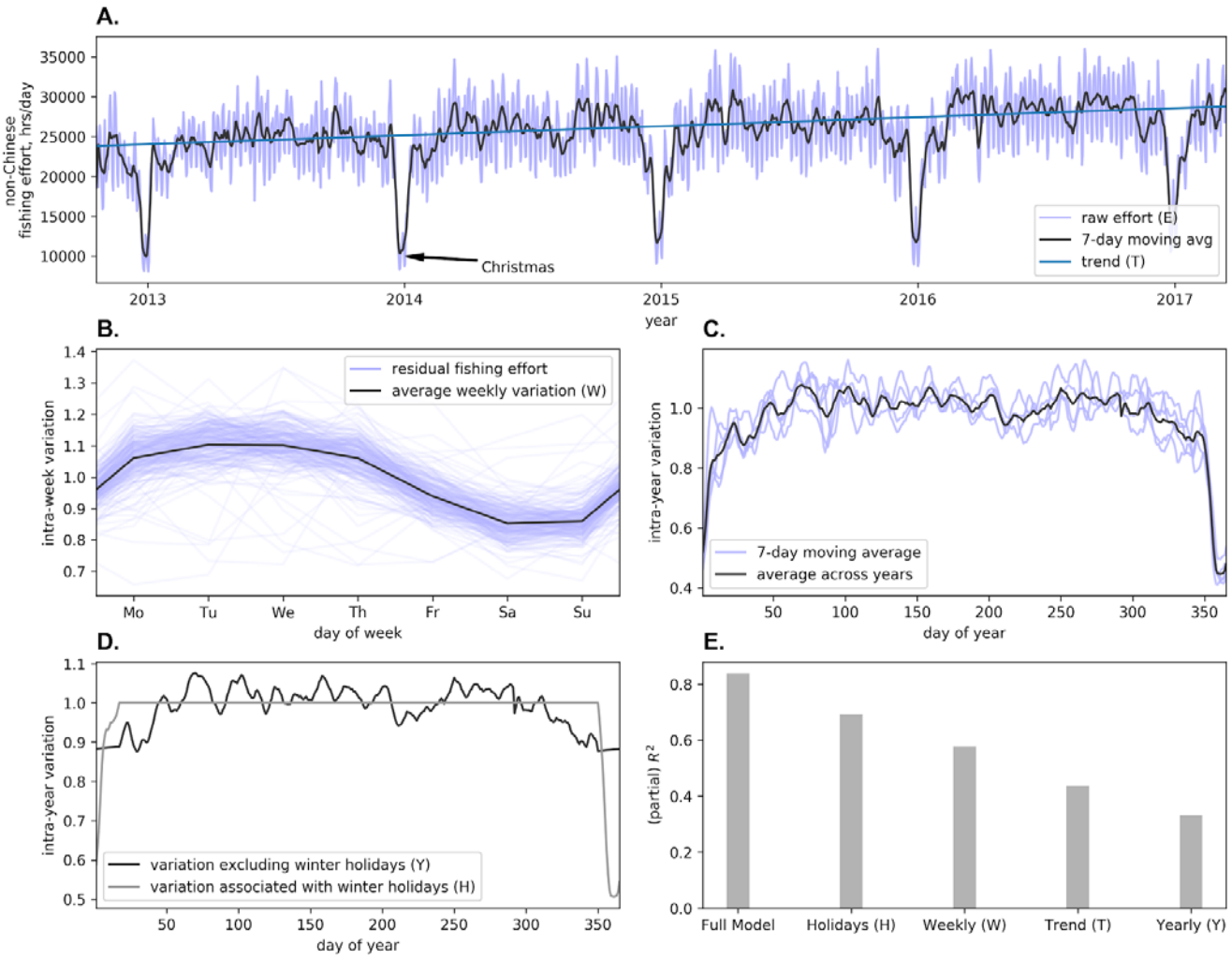


Fig. S9. Temporal Footprint of non-Chinese Fishing Vessels

(**A**) Fishing effort, in fishing hours per day, for non-Chinese vessels that were active from 2012 to 2016. The seven-day moving average removes any weekly effects (black lines). (**B**) Dividing the raw time series by the seven-day moving average gives the residual, weekly signal, and this can be averaged using the harmonic mean to obtain the average weekly variation. This panel shows that fishing, on average, decreases about two thirds as much on Friday as it does on Saturday and Sunday. (**C**) The 7-day moving average from (A) can be similarly averaged across years to obtain the average intra-year variation. The drop at the beginning and end of the year mark the end and start of Christmas holiday respectively. (**D**) The average intra-year variation from C is broken into two parts, one consisting of the month surrounding Christmas and the other the remaining variation. (**E**) Fishing effort is then modelled as E= T*W*Y*H or equivalently log(E)=log(T)+log(W)+log(Y)+log(H). The coefficient of determination (R^2) for the entire model and the partial coefficient of determination (partial R^2) for the components of the linear model. Holidays and weekly variation are the largest contributor to the accuracy of the model.

Supplemental Tables

Table S1. Vessels Classes Used for Training CNN

Methods and Materials section 1.5 describes how we identified these vessels. Parenthesis denote the number of unique vessels used for training the model.

Fishing Classes	Non-Fishing Classes
Drifting longlines (1,646)	Tanker (8,287)
Pole and line (173)	Cargo (18,348)
Set gillnets (330)	Refrigerated cargo vessel (468)
Set longlines (91)	Motor passenger (7924)
Squid jigger (106)	Sailing (4624)
Trawler (3293)	Seismic vessel (161)
Trollers (71)	Tug (6300)
“Other fishing” (70)	Gear/Buoys (1142)

Table S2. Merged Vessel Classes

Merged Class	Base Classes
Cargo or Tanker	Cargo or Tanker
Passenger	Motor Passenger or Sailing
Fixed Gear	Pots and Traps, Set Gillnets, or Set Longlines
Other Fishing	Pole and Line, Trollers, or Other Fishing

Table S3. Accuracy of Fishing Detection Neural Net on Test Data

Gear Type	**Precision**	**Recall**	**Accuracy**	**F1-Score**
Drifting Longlines	0.92	0.94	0.91	0.93
Purse Seines	0.78	0.81	0.95	0.79
Fixed Gear	0.95	0.88	0.97	0.9
Trawlers	0.98	0.94	0.96	0.96

Table S4. Countries by Number of Fishing Vessels > 24m, 12-14m, and <12m

Building on a database from the FAO, we compiled, and in some cases estimated, the number of vessels larger than 24 meters for each country, 12 to 24 meters, and under 12 meters. These are compared to the number of fishing vessels of these size classes in the AIS data that were active in 2016. The length of AIS vessels is estimated by the neural network or taken from registries. See supplemental materials section 5.3 for methods. Number of vessels from registries are from the FAO, except for China (Chinese Statistical Yearbook 2016), United States (Merchant Vessels of the United States), Russia (Russian Maritime Registry of Shipping, vessels under 24 meters are not listed), Chile (a registry obtained from the Secretaria Nacional de Pesca), all countries in the EU (the EU's Community Fishing Fleet Register), Japan (2013 Census of fisheries available at Portal Site of Official Statistics of Japan), Mexico (Mexican government's Anuario Estadístico de Acuacultura y Pesca, 2013), Norway (Norwegian Statistics Portal), Peru (Ministereo de la Produccion), South Korea (Korea Maritime Institute, 2015), and Taiwan (Taiwan Fisheries yearbook 2012). When possible, length overall is used. Countries with 0 or unknown number of vessels larger than 24 meters are not listed. Note that in some cases we identify more vessels in AIS than are in the listed registries or FAO, which could be due to misclassification of vessels as fishing vessels, or incorrect numbers from the registries. Accuracy of reporting to the FAO varies by country. Also, most figures from registries are for total vessels and not active vessels.

Country	Vessels > 24m		Vessels 12-24m		Vessels <12m	
	Registry	AIS	Registry	AIS	Registry	AIS
China	35,844	2,0311	49,697	17,975	186,781	678
Indonesia	2,577	18	182,996	19	194,197	1
United States	1,296	867	7,629	1,112	6,773	25
Russia	812	597	-	25	-	5
Spain	758	654	1,815	1,072	6,120	29
Taiwan	729	707	3,090	734	8,457	24
Myanmar	699	0	1,826	0	12,700	0
South Korea	657	813	2,183	460	63,394	14
Japan	654	587	8,227	230	72,766	7
Papua New Guinea	497	22	89	0	4	0
Thailand	362	4	17,256	14	15,454	0
Argentina	348	277	257	5	309	0
United Kingdom	348	243	798	680	3,659	206
Peru	345	100	1,415	1	5,269	0
Italy	331	257	3,263	1,209	7,961	3
Venezuela	310	13	11,211	0	36,365	0
Chile	297	76	1,411	2	11,085	0
Morocco	293	48	1,690	1	17,224	0

Norway	244	332	259	796	5,443	945
Tunisia	237	0	1,048	0	4,817	0
Greece	197	150	641	229	12,762	3
Bangladesh	196	0	43	0	32,859	0
Ecuador	190	22	238	0	19,639	0
Canada	185	110	4,760	401	9,473	16
France	178	229	721	643	4,859	103
Portugal	165	140	417	271	3,363	5
Mexico	164	31	1,576	18	74,356	1
Madagascar	160	0	71	0	7	0
Iceland	159	161	153	704	1,374	193
Sierra Leone	136	0	2	0	1	0
Netherlands	120	205	204	246	211	18
Algeria	104	1	1,361	1	3,312	0
Mozambique	95	12	239	3	1,064	0
Senegal	91	16	23	3	11,295	0
Panama	87	20	601	2	10,057	0
Croatia	87	87	388	113	2,241	1
Ireland	84	93	159	124	1,100	11
New Zealand	83	44	452	37	799	1
Tanzania	72	2	1	1	11,285	0
Brazil	70	23	34	41	0	0
Denmark	64	72	330	277	1,044	86
Vanuatu	56	56	55	2	8	0
Australia	55	76	233	84	18	2
Seychelles	55	36	116	0	132	0
Poland	50	23	140	62	606	10
Honduras	47	0	176	1	10,678	0
Iran	45	27	3,470	5	8,760	0
Latvia	45	49	12	22	221	0
Jamaica	45	0	126	0	170	0
Sweden	42	39	126	132	810	65
Albania	42	1	191	2	383	0
Côte d'Ivoire	42	2	1,577	0	200	0
Nicaragua	40	1	79	0	0	0
Germany	40	41	238	157	841	5
Uruguay	39	50	22	5	444	1
Saudi Arabia	37	0	1,013	1	10,145	0
Fiji	36	37	22	5	2,550	0
Iraq	35	0	194	0	53	0

Belgium	35	42	40	34	1	1
Suriname	33	0	358	0	292	0
Lithuania	32	28	2	9	63	0
Eritrea	31	0	55	0	127	0
Oman	31	1	710	0	17,841	0
Guam	30	0	211	0	0	0
Estonia	22	22	16	11	1,475	0
Finland	22	21	38	28	1,704	2
India	20	19	27	30	0	3
Congo	20	0	50	0	261	0
Mauritius	17	8	27	13	1,647	0
Colombia	16	12	158	3	0	1
Bahamas	16	0	135	0	833	0
Falkland Islands	14	15	0	0	0	0
Benin	13	0	125	0	182	0
Bulgaria	12	5	69	38	1,029	0
Montenegro	10	0	26	1	182	0
Cuba	10	1	462	0	227	0
Cook Islands	8	6	6	4	285	0
Tonga	8	1	17	0	791	0
Kuwait	8	0	217	0	630	0
Syria	7	0	67	0	1,589	0
Azerbaijan	6	0	0	0	790	0
Malta	6	9	43	33	660	2
Brunei Darussalam	5	0	34	0	2,143	0
French Polynesia	5	0	57	1	3,948	0
New Caledonia	5	2	13	16	155	0
El Salvador	5	5	47	0	6,663	0
Cyprus	5	6	22	14	827	0
Kazakhstan	3	0	33	0	1,723	0
Uganda	3	0	200	0	6,592	0
Guinea	3	0	0	0	0	0
Guatemala	3	2	48	0	6	0
Romania	2	0	10	3	111	0
Sri Lanka	2	0	2,367	10	28,781	0
Guinea-Bissau	1	0	0	0	750	0
French Guiana	1	14	30	1	134	0
Other / Unknown	0	983	1,968	805	34,566	12
Total	**51,546**	**28,984**	**324,052**	**28,981**	**1,012,906**	**2,479**

Table S5. Number of Fishing Vessels by Size Class, Globally

Vessel size	Our Survey	High FAO Estimate	Low FAO Estimate
< 12m	998,164	2,500,000	2,100,000
12-24m	312,531	350,000	320,000
>24m	50,913	64,000	43,700

Table S6. Fraction of Fishing Effort Lost to Incomplete Coverage

Year	Drifting Longlines	Trawlers	Purse Seines	Fixed Gear	All Fishing
2012	4%	37%	>60%	51%	27%
2013	7%	16%	33%	26%	15%
2014	6%	14%	34%	24%	13%
2015	4%	10%	27%	22%	10%
2016	2%	8%	20%	18%	8%

Table S7. Number of Vessels Included in Fuel Price Elasticity Estimation

Flag state	Number of Vessels	Hours at sea (10^6)
China	1015	5.3
Norway	517	3.3
United States	463	3.3
Spain	395	4.8
France	383	3.7
United Kingdom	340	2.8
Iceland	317	2.1
Italy	308	2.8
Netherlands	207	1.5
Denmark	196	1.3
Japan	155	2.7
South Korea	147	2.9
Ireland	115	0.8
Russia	111	1.7
Portugal	101	1.1
Sweden	94	0.5
Turkey	92	0.4
Canada	89	0.7
Argentina	88	1.1
Germany	81	0.5

Gear type	Number of vessels	Hours at sea (10^6)
trawlers	3601	29.8
fixed_gear	1100	6.5
purse_seines	607	4.6
drifting_longlines	486	8.8
other_fishing	184	1.0
squid_jigger	37	0.5

Table S8. Results of Price Elasticity of Fuel Demand Analysis

term	estimate	std.error	statistic	p.value
log(avg_fuel_price)	-0.066	0.021	-3.237	0.001
monthApr	0.026	0.014	1.773	0.076
monthAug	-0.044	0.014	-3.051	0.002
monthDec	-0.319	0.016	-19.362	0.000
monthFeb	-0.136	0.016	-8.670	0.000
monthJan	-0.152	0.016	-9.414	0.000
monthJul	-0.058	0.014	-4.243	0.000
monthMar	0.078	0.015	5.081	0.000
monthMay	0.069	0.014	4.836	0.000
monthNov	-0.018	0.015	-1.166	0.244
monthOct	0.092	0.015	6.116	0.000
monthSep	0.050	0.016	3.209	0.001
year2015	-0.046	0.011	-4.135	0.000
year2016	-0.047	0.016	-2.960	0.003

References and Notes

1. H. Haberl, K. H. Erb, F. Krausmann, V. Gaube, A. Bondeau, C. Plutzar, S. Gingrich, W. Lucht, M. Fischer-Kowalski, Quantifying and mapping the human appropriation of net primary production in earth's terrestrial ecosystems. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 12942–12947 (2007). doi:10.1073/pnas.0704243104 Medline

2. W. Swartz, E. Sala, S. Tracey, R. Watson, D. Pauly, The spatial expansion and ecological footprint of fisheries (1950 to present). *PLOS ONE* **5**, e15143 (2010). doi:10.1371/journal.pone.0015143 Medline

3. P. M. Vitousek, H. A. Mooney, J. Lubchenco, J. M. Melillo, Human domination of Earth's ecosystems. *Science* **277**, 494–499 (1997). doi:10.1126/science.277.5325.494

4. V. De Sy, M. Herold, F. Achard, G. P. Asner, A. Held, J. Kellndorfer, J. Verbesselt, Synergies of multiple remote sensing data sources for REDD+ monitoring. *Curr. Opin. Environ. Sustain.* **4**, 696–706 (2012). doi:10.1016/j.cosust.2012.09.013

5. M. Burke, D. B. Lobell, Satellite-based assessment of yield variation and its determinants in smallholder African systems. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 2189–2194 (2017). doi:10.1073/pnas.1616919114 Medline

6. W. Turner, S. Spector, N. Gardiner, M. Fladeland, E. Sterling, M. Steininger, Remote sensing for biodiversity science and conservation. *Trends Ecol. Evol.* **18**, 306–314 (2003). doi:10.1016/S0169-5347(03)00070-3

7. J. A. Anticamara, R. Watson, A. Gelchu, D. Pauly, Global fishing effort (1950–2010): Trends, gaps, and implications. *Fish. Res.* **107**, 131–136 (2011). doi:10.1016/j.fishres.2010.10.016

8. R. A. Watson, W. W. L. Cheung, J. A. Anticamara, R. U. Sumaila, D. Zeller, D. Pauly, Global marine yield halved as fishing intensity redoubles. *Fish Fish.* **14**, 493–503 (2013). doi:10.1111/j.1467-2979.2012.00483.x

9. International Telecommunication Union (ITU), "Technical characteristics for an automatic identification system using time division multiple access in the VHF maritime mobile frequency band" (Recommendation ITU-R M.1371-5, ITU, 2014); www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.1371-5-201402-I!!PDF-E.pdf.

10. D. J. McCauley, P. Woods, B. Sullivan, B. Bergman, C. Jablonicky, A. Roan, M. Hirshfield, K. Boerder, B. Worm, Ending hide and seek at sea. *Science* **351**, 1148–1150 (2016). doi:10.1126/science.aad5686 Medline

11. F. Natale, M. Gibin, A. Alessandrini, M. Vespe, A. Paulrud, Mapping fishing effort through AIS data. *PLOS ONE* **10**, e0130746 (2015). doi:10.1371/journal.pone.0130746 Medline

12. E. N. de Souza, K. Boerder, S. Matwin, B. Worm, Improving fishing pattern detection from satellite AIS using data mining and machine learning. *PLOS ONE* **11**, e0158248 (2016). doi:10.1371/journal.pone.0158248 Medline

13. T. D. White, A. B. Carlisle, D. A. Kroodsma, B. A. Block, R. Casagrandi, G. A. De Leo, M. Gatto, F. Micheli, D. J. McCauley, Assessing the effectiveness of a large marine protected area for reef shark conservation. *Biol. Conserv.* **207**, 64–71 (2017). doi:10.1016/j.biocon.2017.01.009

14. D. Pauly, D. Zeller, Eds., Sea Around Us concepts, design, and data (2015); www.seaaroundus.org.

15. R. A. Watson, A database of global marine commercial, small-scale, illegal and unreported fisheries catch 1950–2014. *Sci. Data* **4**, 170039 (2017). doi:10.1038/sdata.2017.39 Medline

16. Food and Agriculture Organization of the United Nations (FAO), *The State of World Fisheries and Aquaculture 2016* (FAO, 2016).

17. J. A. Foley, N. Ramankutty, K. A. Brauman, E. S. Cassidy, J. S. Gerber, M. Johnston, N. D. Mueller, C. O'Connell, D. K. Ray, P. C. West, C. Balzer, E. M. Bennett, S. R. Carpenter, J. Hill, C. Monfreda, S. Polasky, J. Rockström, J. Sheehan, S. Siebert, D. Tilman, D. P. M. Zaks, Solutions for a cultivated planet. *Nature* **478**, 337–342 (2011). doi:10.1038/nature10452 Medline

18. L. Guanter, Y. Zhang, M. Jung, J. Joiner, M. Voigt, J. A. Berry, C. Frankenberg, A. R. Huete, P. Zarco-Tejada, J.-E. Lee, M. S. Moran, G. Ponce-Campos, C. Beer, G. Camps-Valls, N. Buchmann, D. Gianelle, K. Klumpp, A. Cescatti, J. M. Baker, T. J. Griffis, Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E1327–E1333 (2014). doi:10.1073/pnas.1320008111 Medline

19. E. Chassot, S. Bonhommeau, N. K. Dulvy, F. Mélin, R. Watson, D. Gascuel, O. Le Pape, Global marine primary production constrains fisheries catches. *Ecol. Lett.* **13**, 495–505 (2010). doi:10.1111/j.1461-0248.2010.01443.x Medline

20. P. L. Barnard, D. Hoover, D. M. Hubbard, A. Snyder, B. C. Ludka, J. Allan, G. M. Kaminsky, P. Ruggiero, T. W. Gallien, L. Gabel, D. McCandless, H. M. Weiner, N. Cohn, D. L. Anderson, K. A. Serafin, Extreme oceanographic forcing and coastal response due to the 2015–2016 El Niño. *Nat. Commun.* **8**, 14365 (2017). doi:10.1038/ncomms14365

21. P. Lehodey, M. Bertignac, J. Hampton, A. Lewis, J. Picaut, El Niño Southern Oscillation and tuna in the western Pacific. *Nature* **389**, 715–718 (1997). doi:10.1038/39575

22. World Bank, *The Sunken Billions Revisited: Progress and Challenges in Global Marine Fisheries* (World Bank, 2017).

23. H. Stouten, K. Van Craeynest, A. Heene, X. Gellynck, H. Polet, "The effect of fuel price scenarios on Belgian fishing fleet dynamics," paper presented at the International Council for the Exploration of the Sea (ICES) Annual Science Conference, Halifax, Canada, 22 to 26 September 2008 (CM 2007/M:04); www.vliz.be/imisdocs/publications/135276.pdf.

24. A. Cheilari, J. Guillen, D. Damalas, T. Barbas, Effects of the fuel price crisis on the energy efficiency and the economic performance of the European Union fishing fleets. *Mar. Policy* **40**, 18–24 (2013). doi:10.1016/j.marpol.2012.12.006

25. D. J. Beare, M. Machiels, Beam trawlermen take feet off gas in response to oil price hikes. *ICES J. Mar. Sci.* **69**, 1064–1068 (2012). doi:10.1093/icesjms/fss057

26. J. E. Hughes, C. R. Knittel, D. Sperling, "Evidence of a shift in the short-run price elasticity of gasoline demand" (Working paper 12530, National Bureau of Economic Research, 2006).

27. J. J. Winebrake, E. H. Green, B. Comer, C. Li, S. Froman, M. Shelby, Fuel price elasticities in the U.S. combination trucking sector. *Transp. Res. D* **38**, 166–177 (2015). doi:10.1016/j.trd.2015.04.006

28. T. Havranek, Z. Irsova, K. Janda, Demand for gasoline is more price-inelastic than commonly thought. *Energy Econ.* **34**, 201–207 (2012). doi:10.1016/j.eneco.2011.09.003

29. FAO, *FAO Statistical Pocketbook 2015* (FAO, 2016).

30. R. Lewison, A. J. Hobday, S. Maxwell, E. Hazen, J. R. Hartog, D. C. Dunn, D. Briscoe, S. Fossette, C. E. O'Keefe, M. Barnes, M. Abecassis, S. Bograd, N. D. Bethoney, H. Bailey, D. Wiley, S. Andrews, L. Hazen, L. B. Crowder, Dynamic ocean management: Identifying the critical ingredients of dynamic approaches to ocean resource management. *Bioscience* **65**, 486–498 (2015). doi:10.1093/biosci/biv018

31. K. Schwehr, C++ decoder for Automatic Identification System for tracking ships and decoding maritime information, version 0.15 (2015); https://github.com/schwehr/libais.

32. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition. arXiv:1512.03385 [cs.CV] (10 Dec 2015).

33. K. He, X. Zhang, S. Ren, J. Sun, Identity mappings in deep residual networks. arXiv:1603.05027 [cs.CV] (16 Mar 2016).

34. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, "Going deeper with convolutions," *2015 IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, 7 to 12 June 2015, pp. 1–9; www.infona.pl//resource/bwmeta1.element.ieee-art-000007298594.

35. C. Szegedy, S. Ioffe, V. Vanhoucke, A. Alemi, Inception-v4, Inception-ResNet and the impact of residual connections on learning. arXiv:1602.07261 [cs.CV] (23 Feb 2016).

36. S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv:1502.03167 [cs.LG] (11 Feb 2015).

37. C. Costello, O. Deschênes, A. Larsen, S. Gaines, Removing biases in forecasts of fishery status. *J. Bioeconomics* **16**, 213–219 (2014). doi:10.1007/s10818-013-9158-4

38. N. Duan, Smearing Estimate: A nonparametric retransformation method. *J. Am. Stat. Assoc.* **78**, 605–610 (1983). doi:10.1080/01621459.1983.10478017

39. E. Shelhamer, J. Long, T. Darrell, Fully convolutional networks for semantic segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**, 640–651 (2017). doi:10.1109/TPAMI.2016.2572683 Medline

40. Z. Ou, J. Zhu, AIS database powered by GIS technology for maritime safety and security. *J. Navig.* **61**, 655–665 (2008). doi:10.1017/S0373463308004888

41. Scientific, Technical and Economic Committee for Fisheries, "The 2015 annual economic report on the EU fishing fleet," A. Paulrud, N. Carvalho, A. Borrello, A. Motova, Eds. (STECF 15-07, Publications Office of the European Union, 2015).

42. J. D. Bell, R. A. Watson, Y. Ye, Global fishing capacity and fishing effort from 1950 to 2012. *Fish Fish.* **18**, 489–505 (2017). doi:10.1111/faf.12187

43. FAO, *FAO Statistical Yearbook 2016* (FAO, 2016).

44. J. F. Muir, "Fuel and energy use in the fisheries sector—approaches, inventories and strategic implications" (FAO Fish. Aquac. Circ. no. C1080, FAO, 2015).

45. N. Queiroz, N. E. Humphries, G. Mucientes, N. Hammerschlag, F. P. Lima, K. L. Scales, P. I. Miller, L. L. Sousa, R. Seabra, D. W. Sims, Ocean-wide tracking of pelagic sharks reveals extent of overlap with longline fishing hotspots. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 1582–1587 (2016). doi:10.1073/pnas.1510090113 Medline

46. S. Solomon, *Climate Change 2007—The Physical Science Basis: Working Group I Contribution to the Fourth Assessment Report of the IPCC* (Cambridge Univ. Press, 2007).