The GFW Fishing Score

Background

The Global Fishing Watch (GFW) fishing score model computes the probability that a vessel is fishing based on its AIS track data. The combined fishing score of all vessels is used to estimate the fishing effort worldwide as shown in Figure 1 below. *Fishing* can be defined several ways, however, for the purposes of the model described in this document, we define *fishing* as the period when a vessel has fishing gear in the water.

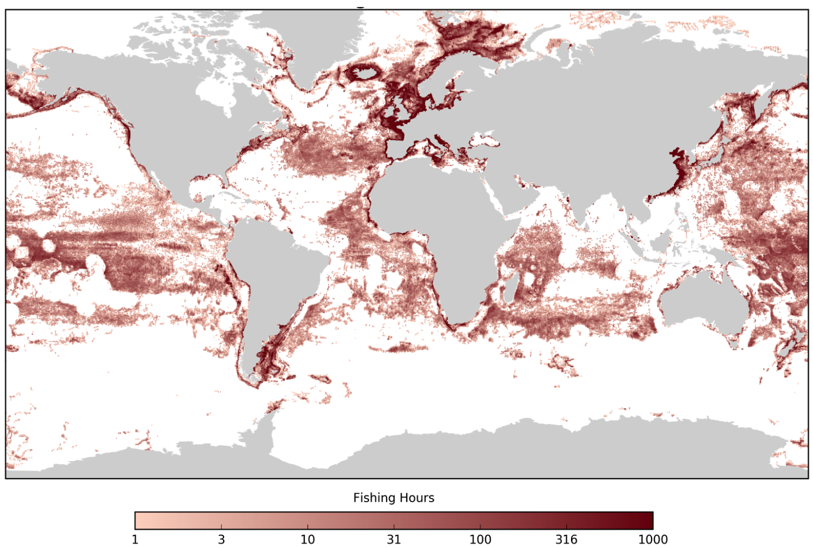

Figure 1: Fishing effort in 2015

An example of an AIS vessel track is shown in Figure 2 below, with the points where the vessel is fishing are shown in red. The job of the fishing score model is to estimate the probability that a vessel is fishing at each of the points along the AIS track. The fishing score for a vessel at a given point is computed using a model based on three primary features: average speed, the amount of variation in speed, and the amount of variation in course. The "variation" in speed and course is technically called the *standard deviation*, and we shall refer to these features as the *speed deviation* and *course deviation*. In addition, points within 3 nautical miles of shore are uniformly considered non-fishing. Ordinary vessel behavior near shore is easily confused with fishing and using this 3 nautical mile cut off avoids a number of false positive values.

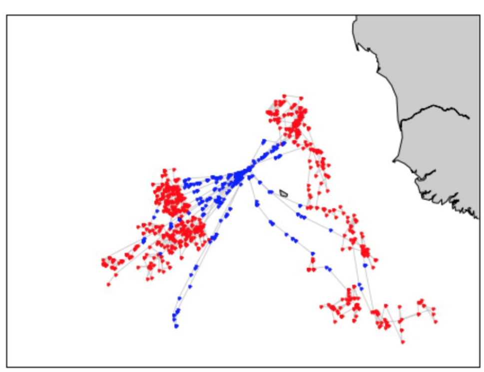

Figure 2: Example fishing vessel AIS track with fishing shown in red.

To see how this works, examine the scatter plot shown in Figure 3, which shows how these three features, computed over a six hour window, relate to whether a vessel is fishing. It is apparent from Figure 3 that fishing activity tends to be be clustered in certain regions of average-speed, speed-deviation and course-deviation.

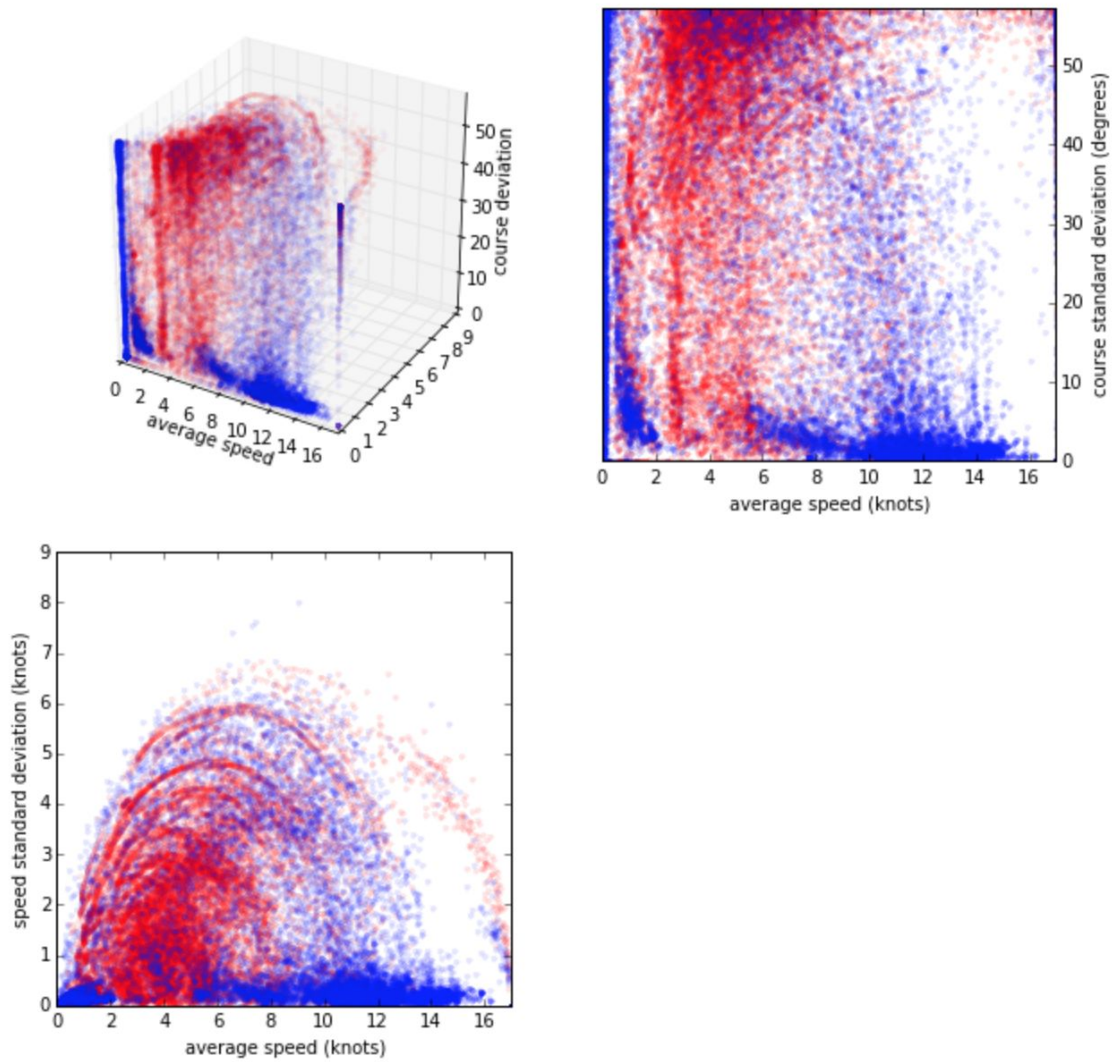

Figure 3: Fishing activity, shown in red, is most common for speeds in the range of 2 to 5 knots and is also associated with larger deviations in speed and course.

The Model

We use a logistic regression model to predict the likelihood that a vessel is fishing based on its average-speed, speed-deviation and course-deviation. In addition to the six hour window discussed above, these features were computed over one-half, one, three, twelve and twenty-four hour windows. Using multiple windows increases the overall accuracy of the model.

We also use *feature augmentation* to increase the expressiveness of the model. A simple logistic regression would effectively divide the regions of the graph above with a flat surface, where values on one side of the surface are classified as fishing. However, dividing the space shown in Figure 3 with a *flat* surface would not split the fishing and non-fishing regions effectively. To overcome this we add powers of the base features. For example, in addition to the *average speed*, we also include the square of the *average speed*, as well as the cube all the way up to the sixth power. This gives the logistic regression model the ability to divide the space along a curved surface and thus identify fishing more accurately.

The model is trained using data that has been hand-labeled as fishing or non-fishing by Kristina Boerder at Dalhousie University. To train the model, we provide it with three-quarters of the labeled data, and have it find the optimal way to divide the feature space into fishing and non-fishing regions as discussed above. We then validate the model by inputting the remaining data to see how accurate it is at determining fishing versus non-fishing behaviour. Table 1, below, shows how the model performs on various gear types. The table's three columns show:

- **Precision:** the percentage of points that the model predicts as fishing which are in fact fishing.
- **Recall:** the percentage of all of the fishing that the model classifies as fishing.
- **False Positive Rate:** the percentage of non-fishing points mistakenly classified as fishing.

The table shows that the model does well predicting fishing for longliners and trawlers, with high precision, high recall and low false positive rates. However, the model performs poorly on purse seiners, as indicated by the the low precision.

	Predicted fishing classified correctly (precision)	Fishing captured (recall)	Non-fishing classified as fishing (false positive rate)
Longliner	97%	78%	9%
Trawler	93%	91%	10%
Purse Seine	11%	73%	22%

Table 1:precision, recall and false positive rate for various gear types.

The low precision for purse seiners is a result of two factors. The first is simply that fishing for purse seiners is difficult to predict based on AIS tracks. Purse seiners spend only a small amount of time with their gear in the water, and because there are gaps in AIS coverage, periods of time where no AIS signal is received due to lack of satellite coverage or other factors, the resulting track may skip over and miss these short fishing events. This makes detecting fishing events reliably difficult for purse seiners. These short fishing events are also more difficult to distinguish from other, non-fishing, activities than the longer fishing events of trawlers and longliners, resulting in more incorrect classifications.The second factor is that the small amount of actual fishing means that even a small fraction of non-fishing classified as fishing (false positive rate) can overwhelm the true fishing, hence the low precision. In addition, we are currently using a single model for all three gear types, which limits the accuracy on any given gear type. Based on early experiments, we expect the precision for purse seiners to at least double and the false positive rate to drop by at least a factor of two as we move to more sophisticated models.

Training Data

The training data from Dalhousie consists of hand-classified AIS data for 29 unique vessels with complete tracks classified as *fishing* (gear in the water) or *non-fishing* over long periods. These vessels are divided between the different gear types as shown in the table below. In addition, data from two vessels performing slow transits is added to the training data to help the model learn to avoid classifying these transits as fishing.

	Vessels	Points	Training Points	Validation Points
Longliner	16	569,504	15,000	5000
Trawler	6	828,162	15,000	5000
Purse Seine	7	398,897	15,000	5000
Slow Transits	2	9,038	6,514	

Table 2: Number of vessels and AIS points available for model training and validation.

The vessels for each class are divided between the training and test sets so that roughly three-quarters of the points from each class are in the training set, with the remainder in the test set for that class. Fifteen thousand points are randomly sampled from the training points for each gear type. The training data for each class along with with the slow transit data is combined to create a training set of 49,430 points. Five thousand randomly selected points from the test set for each gear type are used for testing.

See Also

- The [Model-Descriptions notebook] has more details for this model as well as earlier and upcoming models.
- Wikipedia has a nice illustration of [precision and recall].