Observing unobserved fishing characteristics in the drift gillnet fishery for swordfish

Jenny Suter[1], Rob Ames[1], Brett Holycross[1], Jordan Watson[2]
[1]Pacific States Marine Fisheries Commission
[2]NOAA Alaska Fisheries Science Center

Introduction

Observer programs are a valuable source of information for the management and compliance of regulations in many fisheries globally. Fishery observer duties may vary from fishery-to-fishery but in many cases, they record information on fishing locations and gear, enumerate catches of target and non-target species (including bycatch), collect biological information about certain catches, and in some cases, they may also collect physical samples like otoliths, scales, stomach contents, or tissue samples. Such fishery observer programs can be extremely expensive, as well as logistically complicated. Thus, a common pursuit in the design of many fishery observer programs is to identify the minimum observer coverage necessary to provide a representative sample of fishery behaviors and catches. However, biases in observer data can easily be introduced and affect the potential inference regarding catch of target, non-target (including bycatch) species across a fishery.

Biases in observer data may result from several sources, some of which may be intentional behaviors while others may be inadvertent (Babcock & Pikitch, 2003; Benoit & Allard 2009). A major concern is bias associated with an “observer effect,” where fishers behave differently (e.g., different fishing locations, depths, times of day, bycatch avoidance, trip durations, or other operational characteristics) during observed trips than during unobserved trips (Vølstad and Fogarty 2006; Benoit & Allard 2009). Another potential form of bias, which we call the “unobservable vessel effect,” may occur for trips made by vessels that are never observed and whose behaviors may thus be unrepresented by trips of vessels that are observed. Any such biases, may lead to skewed estimates of fishing effort, catch compositions, bycatch impacts, trip- and set-level metrics or more. These so-called unobservable vessels may result from a suite of factors, further described by Babcock and Pikitch (2003), including logistical constraints (e.g., vessels are too small or unsafe to carry observers), inappropriate stratification or allocation of sampling effort.

In recent years, the requirement of vessel monitoring systems (VMS) for many fleets has provided a new tool by which to assess fleet dynamics and potential biases in some observer data. VMS transmit vessel locations at regular time intervals, often every 30 - 60 min, regardless of whether the vessel was ever observed. When combined with fishery landing records, the combination of vessel spatial, temporal, and movement characteristics derived from the VMS data can facilitate new analyses to quantify operational differences across vessels, trips, or even fishing sets.

One fishery requiring VMS on all vessels (since 2013) is the large-mesh drift gillnet (DGN) fishery off the U.S. West Coast. The DGN fishery has targeted swordfish (*Xiphias gladius*) and other highly migratory species (HMS) off the U.S. West Coast for more than four decades (Urbisci et al., 2016; Eguchi et al., 2017). Throughout its history, operational characteristics, targeting behaviors, and regulations have evolved considerably. The fishery has seen a suite of time-area closures aimed at bycatch reduction (e.g., leatherback turtles), leading to the current distribution of the fishery, which is largely centered in the Southern California Bight (SCB). For much of the fishery's existence, fishery observers have collected critical information, facilitating validation of fisher logbooks, standardization of CPUE, and documentation of interactions between the fleet and marine mammals (Urbisci et al., 2016; Carretta 2020). Such observer data has also been critical for the continued management of the fishery and development of more advanced dynamic ocean management tools for bycatch avoidance (e.g., Eguchi et al., 2017; Brodie et al., 2018; Hazen et al. 2018).

The DGN fleet is required to carry fisheries observers on a portion of their trips each season. The NOAA West Coast Region Observer Program (WCROP) works with the fleet to cover a portion of the total fishing effort each fishing season. This typically results in 20 – 30 percent observer coverage. In the most recent Biological Opinion on the DGN fishery (NMFS 2013), questions about the representative quality of observer coverage and the overall reliability of bycatch estimates produced by observer data in the DGN fishery were considered. These questions ultimately led to the 2013 implementation of a VMS and a pre-trip notification requirement, providing expanded information about the spatiotemporal distributions of the entire fleet during all observed and unobserved trips in the fishery.

We hypothesized that the integration of observer, logbooks, and landings data with VMS data was sufficient to identify individual set and trip metrics (e.g., season, location, trip duration, trip distance, average depth, sea-surface temperature), and to determine whether differences existed between observed and unobserved fishing. First, we developed a machine learning approach to quantify trip and set metrics for DGN activity. Second, we compared observed and unobserved trips to assess potential observer effects. Finally, we compared trip and set metrics for those vessels that were periodically observed with those vessels that were never observed, which we refer to as unobservable.

Methods

This data intensive study proceeded via several discrete steps: (1) Integrating data and engineering features; (2) Differentiating fishing and non-fishing trips; (3) Developing classification models to identify fishing sets; (4) Identifying true DGN trips; (5) Exploring observer coverage rates; (6) Analyzing trip and set level differences in observed and unobserved fishing behaviors.

Integrating Data and Engineering Features

The foundation of this study was the integration of disparate types of data, many of which are confidential. These data included fishery-dependent (observer, logbook, VMS, fisheries landings, and permit data) and fishery-independent (environmental, geospatial, and U.S. Coast Guard (USCG) vessel registries) data for six fishing seasons, September 2013 – January 2019. The requisite data were accessed from the Pacific Fisheries Information Network (PacFIN), which functions as an intermediary network that consolidates State and Federal fisheries data within a centralized data warehouse. Specifically, these data include:

- **Vessel Monitoring System (VMS) data**: Since September 2013, DGN vessels have been required[1] to transmit their locations (latitude, longitude) hourly via VMS for the duration of all observed and unobserved fishing trips. VMS data do not explicitly identify vessel behaviors (e.g., fishing, transiting, drifting, searching), but a suite of studies have used VMS data to infer vessel

[1] 4 Sept 2013 – 31 Jan 31 2014 (78 FR 54548); 22 May 2014 – 5 Aug 2014 (79 FR 29377); 9 Jul 2015 – present (80 FR 10392, 80 FR 32465)

activities (e.g., Russo et al., 2011; Muench et al., 2018; Watson et al., 2018). We engineered a suite of additional features within the VMS data (e.g., vessel speed, distance traveled between VMS records, distance to the coastline, bottom depth) to support the development of a behavioral inference model.

- **Vessel Characteristics**: Vessels > 5 mt are required to register with the USCG and supply vessel characteristics such as vessel length and net and gross tonnage, along with other information including the owner, year built, and the ship builder.
- **Observer**: The data used in this study are a subset of the information collected by WCROP observers on DGN vessels including the location, date, and time of set deployment and retrieval (aka haul).
- **Vessel Logbook**: During the study timeframe, DGN vessel captains were required to submit their logbooks to the California Department of Fish and Wildlife (CDFW) and the data were keypunched into the NOAA Southwest Fisheries Science Center (SWFSC) logbook database. Vessel logbook information used in this study were the haul date (without time), set type, gear configuration, and area of catch.
- **Fish Landing Receipts**: Upon completion of a commercial fishing trip, the sales transaction (aka fish ticket) between fishing vessels and dealers are submitted to the state of landing. The state landings data for the West coast are aggregated in the PacFIN data warehouse. Fish tickets include information such as the species, quantity in pounds, price paid, fishing gear type, and port of landing. For HMS fisheries like the DGN fishery, the state-reported gear types and species are translated into HMS fishery codes (D'Angelo, *unpub*), which aided in identifying DGN landings. All DGN landings used in this analysis were from the CDFW.
- **Geospatial features**: We implemented a suite of geospatial analyses (ArcGIS Pro 2.4, Esri) to engineer additional features for behavioral comparison and fishing detection, such as bottom depth (GEBCO 2019), distance from shore, and distance to the nearest port. For each port visited by the DGN fleet during this study, we iteratively defined port areas polygons by first using simplified estuary polygons (buffered by 100m) from the Pacific Marine and Estuarine Fish Habitat Partnership (PMEP) West Coast USA Current and Historical Estuary Extent layer (2017). For ports without matching PMEP polygons, we created polygons manually around VMS records for which vessels were not moving and appeared to be inside a known fishing port. VMS activity "event types" were calculated based on a VMS record's presence inside ('in port') or outside ('at sea') a port polygon (see example in Figure 2). Events were further distinguished for VMS records that occurred prior to entering ('arrive port') and subsequent to leaving ('depart port') a port polygon.
- **Environmental data**: DGN gear is deployed along the ocean surface, typically on ocean temperature gradients, where target species are known to occur. To incorporate such behaviors, daily satellite records of sea surface temperatures (SST) (JPL MUR MEaSUREs Project 2015) were obtained through an ERDDAP server at the NOAA CoastWatch Program (Simons 2020). Daily SST records were matched with the spatial coordinates for each VMS record.

The approach we used to integrate these data into a Comprehensive DGN database table was stepwise (Figure 1, steps 1 – 6), and followed a hybrid approach of the cross-industry standard process for data mining (CRISP-DM) (Chapman et al. 2000). CRISP-DM steps were not always linear, and lessons learned during one-step often triggered the improvement in logic for previous steps. Data integration steps were done in Oracle and ArcGIS Pro. The ArcGIS-created features were merged into the Comprehensive DGN table via an ArcGIS-Oracle database link. Likewise, data preparation for the classification models, as well as for the observer bias analysis were performed outside the database in Python, and the classification model output features were merged into the Comprehensive DGN table using the cx_Oracle library in Python.

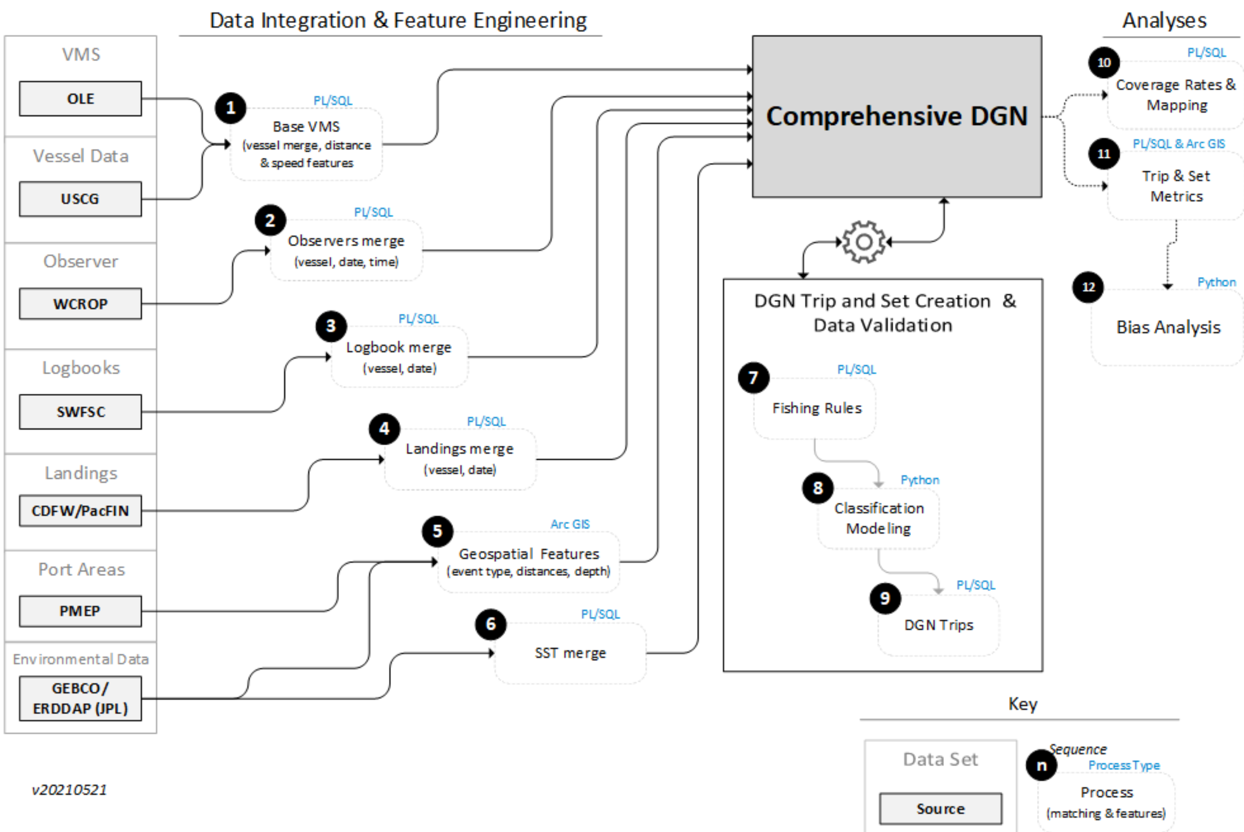

Figure 1. Stepwise approach to integrate datasets, create sets and trips, validate the data, and analyze the data. Data source names include OLE – NOAAs Office of Law Enforcement; USCG – U.S. Coast Guard; WCROP – West Coast Region Observer Program; SWFSC – NOAA Southwest Fisheries Science Center; CDFW – California Department of Fish and Wildlife; PacFIN – Pacific Fisheries Information Network; PMEP - Pacific Marine and Estuarine Fish Habitat Partnership. GEBCO and ERDDAP are commonly used names for a public bathymetric repository and a data server, respectively. Jet Propulsion Lab (JPL) satellite data were accessed via the ERDDAP server.

Trip and Set Creation, Fishing Rules, and Data Validation

Initially, we used the 'depart port' and 'arrive port' events created in the Geospatial Features step to delineate individual vessel trips (Watson & Haynie 2016). However, vessels do not always fish when they leave port. Depending on the configuration of a particular port, for example, a vessel may leave a defined port polygon for fueling, gear maintenance, long-term vessel storage, or other activities after which no fish were landed or sold. Such activities may trigger a new "trip" based on VMS data even if no fishing occurs. Most non-fishing trips were relatively brief, typically lasting less than 12 hours. Differentiating between fishing and non-fishing trips represents a first step in the process of modeling fishing activity (Watson and Haynie 2016).

We used a stepwise approach to create and validate true DGN trips and sets from the individual vessel trips (Figure 1, steps 7 – 9). We first created features to denote known parameters of DGN fishing, employed classification models to identify fishing and non-fishing pings, and then reviewed the outputs of models with the known fishing parameters to determine true DGN trips and sets. We validated the data and checked for errors before and after each step.

We created more engineered features for known parameters of DGN fishing (Figure 1, step 7) such as fishing time (3:00 PM to 8:00 AM) and the speed of the vessel during a fishing set, estimated to be < 1.25 knots (Sippel & Stohs, 2017). Combining fishing time (is_during_fishing_time) and fishing speed (is_fishing_speed) with the 'at sea' event type, we created another variable called 'is_fishing_sql'. From Observer data, we calculated that DGN vessels fished their gear for about 11-12 hours, slightly longer than estimated in previous studies (Stohs and Sippel, 2017, Urbisci, Stohs, and Piner 2016). Thus, with hourly transmission rates, we considered 10-12 consecutive VMS points to be indicative of fishing sets (is_fishing_sql).We created the variable 'is_observed_ping' to denote that an observer was on board during the matched date and time, 'is_driftgillnet_fishing' to denote the presence of observer, logbook, or DGN landing data matched to VMS data by the date and time, and 'season' to denote the season of trip. Due to the closure of the northern portion of the fishing grounds in the Pacific Leatherback Closure Area (PLCA) (66 FR 44549; August 24, 2001) until November 15 of each year, we split the fishing year into two seasons (Season 1, 1 May – 15 Nov and Season 2, 16 Nov – 31 Jan). If the trip overlapped two seasons, it was classified as the season that included the majority of fishing days.

Classification Modeling to Identify Fishing Sets

We evaluated multiple machine learning classification models to identify fishing versus non-fishing records (Figure 1, step 8). These models included: (1) naïve Bayes, (2) support-vector machine, (3) decision tree, (4) random forest, and (5) gradient boosting. All records from the Comprehensive DGN dataset that were defined as (a) DGN fishing (is_driftgillnet_fishing) and (b) event_type = 'at sea' were exported to Python. An exploratory analysis for all observed trips was conducted to evaluate the relationship between the features and the target variable, which was the observer record that identified fishing versus non-fishing records (is_observed_ping). This approach along with domain knowledge guided model feature selection. All features used in the classification models were numeric and were scaled using min-max normalization and missing values in the SST data were filled using nearest-neighbor imputation. The data from the observed trips were randomly split into training (75% or 14342 records) and testing (25% or 4781 records) datasets.

Each classification model was fitted to the training data using the default hyperparameter settings from the scikit-learn software libraries for machine learning in Python. The performance of each model was evaluated for prediction accuracy (i.e., fishing vs non-fishing) along with the receiver operating characteristic (ROC) curve and the area under the curve (AUC). Based on these criteria, rigorous model fitting was pursued for random forests and gradient boosting models, both of which are ensemble decision tree approaches. Random forest and gradient boosting models were optimized using k-fold cross-validation and a grid-search for hyperparameter tuning and selection based on the highest mean predictive accuracy. Models with the best hyperparameters were fitted to the training data and models were evaluated on the test data. The final models were applied to the out-of-sample (unobserved) data and the resulting predicted fishing records were merged into the Comprehensive DGN dataset (Appendix 1 - Table 1: is_fishing_rfm and is_fishing_gbc).

By combining the results of the classification modeling with the fishing rules, we made the final determination of DGN fishing sets (is_fishing_set). Non-observed sets were considered valid fishing sets if, (a) both models agreed that the vessel was fishing, (b) the set occurred between 3:00 PM and 8:00 AM (is_during_fishing_time), and (c) the set was >= 6 hours (determined either by observer data soak times or estimated by the number of pings >= 6). We manually reviewed and validated all of the observed and unobserved sets that did not meet these criteria.

Identifying DGN Trips

We developed a set of rules to characterize *true* fishing trips (aka DGN trips) from GIS-defined VMS trips based on the 'depart port' and 'arrive port' events. Trips were classified as DGN trips if they met at least one of these criteria:

- There was positive DGN effort (is_driftgillnet_fishing)
- There was a landing (i.e., at least one fish ticket) of DGN target species at the end of a trip ('arrive port' date) or when the vessel was 'in port' and a landing was made within five days of the 'arrive port' date.
- The linked landings had the HMS fishery code = DGNLM.

We flagged trips for review that did not easily link to landings, but had positive DGN fishing effort (is_driftgillnet_fishing). Various reasons caused problems with the trip-landing link process including miscoded gear or fishery (CDFW gear code or HMS fishery code), date errors on fish tickets, or erratic VMS ping rates. We reviewed and verified the unlinked DGN trip and landings data and made some manual links that fell into these categories:

- Landings that occur within a trip ('at sea')
- Landings with a date that exceeded five days after the 'arrive port' date.
- Landings with HMS fishery code = DGNLM.

The definition of a DGN trip is different from how the WCROP defined trips. Observers were assigned to cover a portion of fishing effort for a vessel and thus, their assignment period may have covered more than one fishing trip, so observer trip identifiers were not sufficient to differentiate trips.

We created a sequential trip number (trip_num) and set number (set_num) for each verified trip and set.

Exploring Observer Coverage

Our ultimate goal was to evaluate bias in observer data, which first required summarizing the distribution of observer coverage rates throughout a suite of strata within the fishery. To investigate the relative representativeness of observer coverage spatially and temporally, we compared model-estimated fishing sets to observed fishing sets within and between weeks, seasons, and vessel categories. Figure 2 shows the distribution of estimated fishing effort (left panel) and observer coverage (right panel) for the six fishing years combined, September 2013 to January 2019.

We categorized vessels as *periodically observed vessels, unobservable vessels,* and *excluded vessels*. A few vessels with very low observer coverage (only one observed trip in six fishing years) were grouped with the *unobservable vessels* category. The WCROP classifies some vessels in the fishery as *unobservable* for logistical reasons such as that they are too small or not safe enough to accommodate an observer. Excluded vessels included those with erratic transmission rates or that only participated in the fishery for the first year of the study timeframe.

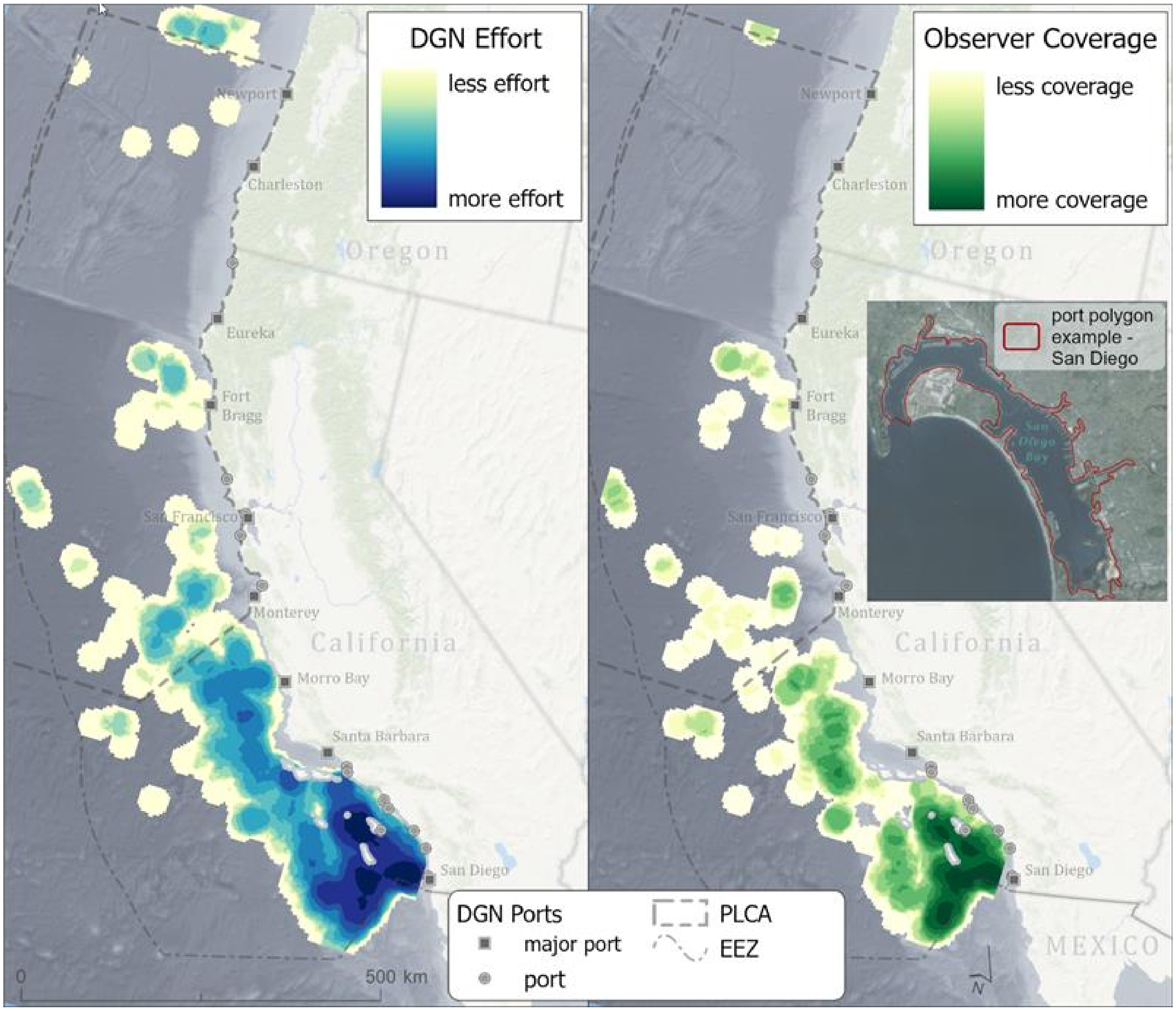

Figure 2: Map of the extent of the DGN fishery (total DGN effort (left panel) and observer coverage (right panel)) and ports visited by DGN fishing vessels for the six fishing years combined (September 2013 to January 2019). The DGN trips only land to California ports, but may stop in Oregon and Washington during northern DGN trips (or while participating in other west coast fisheries). The inset shows an example of the port polygon for San Diego. PLCA is the Pacific Leatherback Closure Area (66 FR 44549) and EEZ is the U.S. Exclusive Economic Zone.

Observer Bias Analysis

In order to test if vessels fish differently when an observer was onboard, we compared metrics at the trip and set levels for the different vessel groups and seasons. The two comparisons were between: (1) *Unobserved trips and sets* versus *observed trips and sets* by vessels that periodically carry observers to test the observer effect; and (2) *Trips and sets on vessels that periodically carry observers* versus *trips and sets by unobservable vessels* that are exempted from observer coverage or that had very low observer coverage (only one observed trip in five fishing seasons) to test the unobservable vessel effect.

We fit linear mixed-effect models (LME) to quantify the percent change (i.e., resulting model coefficients) in a number of log-transformed fishery metrics (i.e., response variables) between the vessel groups (Watson et al., 2018). The LME models were fit to the fishery metrics individually and for each season (Season 1, 1 May – 15 Nov; Season 2, 16 Nov – 31 Jan):

$$Y_{t,v}=(\beta_0+b_{0v})+\beta_1*is\_observed\_ping_{t,v}+\beta_2*Vessel+\varepsilon_{t,v}$$

where, for example, for the observed and unobserved trip or set comparison, the subscripts *t* and *v* represent trip and vessel, and 'is_observed_ping' is a binary, fixed-effect indicating if the trip was observed and individual 'vessel' was the random effect (to account for differences between individual vessels), which allows for broad level inferences about the fixed effects, i.e., observed and unobserved trips or sets. The random intercept for vessel (b_{0v}) and residual ($\varepsilon_{t,v}$) were assumed to be independent and normally distributed with means zero and variances α^2_v and α^2, respectively. Similarly, the 'is_observed_ping' binary variable was replaced with another binary variable, 'vessel_group', for the comparison between vessels that periodically carry observers versus unobservable vessels.
The fixed coefficient on 'is_observed_ping' and 'vessel_group' measured the change in the response between observed and unobserved trips on periodically observed vessels or between these observed vessels and the unobservable vessels.

The fishery metrics to which we fit models represented a suite of behavioral and operational aspects of fishing at the trip- and set-level (Table 1) that we hypothesized could result in different catch or bycatch compositions (Brodie et al., 2018; Hazen et al., 2018; Mason et al., 2019). For example, during different seasons, sea surface temperatures have been used to indicate a greater probability of interaction between leatherback sea turtles and fishing fleets (Eguchi et al., 2017). If fishers sought to avoid bycatch to a greater extent when observers were onboard, then we might expect temperature profiles to be different for observed versus unobserved trips.

Table 1. Description of trip- and set-level metrics that were fit to models to compare observed versus unobserved fishing behaviors.

Scale	Name	Description
Trip	Sea Surface Temperature (SST)	Average satellite-derived SST value from the final VMS location for each set per trip
Trip	Trip Duration	Trip duration (hours) calculated between the first and last VMS record timestamps of each trip
Trip	Trip Distance	Distance (nm) traveled during the trip, calculated from the sum of the distances between each VMS record.
Trip	Proportion of Trip Fishing	Duration of the combined fishing sets (predicted or observed) divided by the trip duration
Trip	Distance from Shore	Average distance (nm) from shore of the final VMS location of each set per trip
Trip	Depth	Average depth of final VMS record of each set per trip
Trip	Catch per Unit Effort	Total weight (lbs) of DGN target species (swordfish, thresher sharks, mako sharks, opah, and tunas) sold for each trip divided by the count of fishing sets per trip
Trip	Ex-Vessel per Unit Effort	Nominal value (USD) of target fish sold per trip divided by the number of sets per trip
Set	Sea Surface Temperature (SST)	Average SST extracted from gridded satellite SST data for each VMS position
Set	Soak Time	Duration (min) between the begin set and begin haul, calculated from observer data when available or from the model estimated set data
Set	Distance from Shore	Average distance (nm) from shore for each fishing set
Set	Depth	Average bottom depth (m) for each fishing set

The data from the 2013/2014 fishing year (Sept 2013 to Jan 2014) were not included in the observer bias analysis because all offshore sets deeper than the 1100 fathoms (2012 m) depth contour were intentionally observed during that season as required by federal regulation (78 FR 54548).

Results

Integrating Data and Engineering Features

A major highlight of this project is that we successfully built a repeatable process to integrate the fishery-dependent and fishery-independent datasets. Along each step of the way, we learned something new and updated the process to be a bit more agile. The power of this integrated fishery-dependent dataset made it much easier to recognize and rectify the mistakes when the datasets did not agree. We explored the datasets visually, which aided in developing some of the engineered features and stepwise process for determining valid DGN trips and sets.

Figure 3 shows the VMS, observer, logbook, and landings data for a vessel on a timeline during a typical fishing season. Using the combination of fishery-dependent data with engineered features, fishery rules, and machine learning, we could confidently determine fishing versus non-fishing activities and positive DGN fishing.

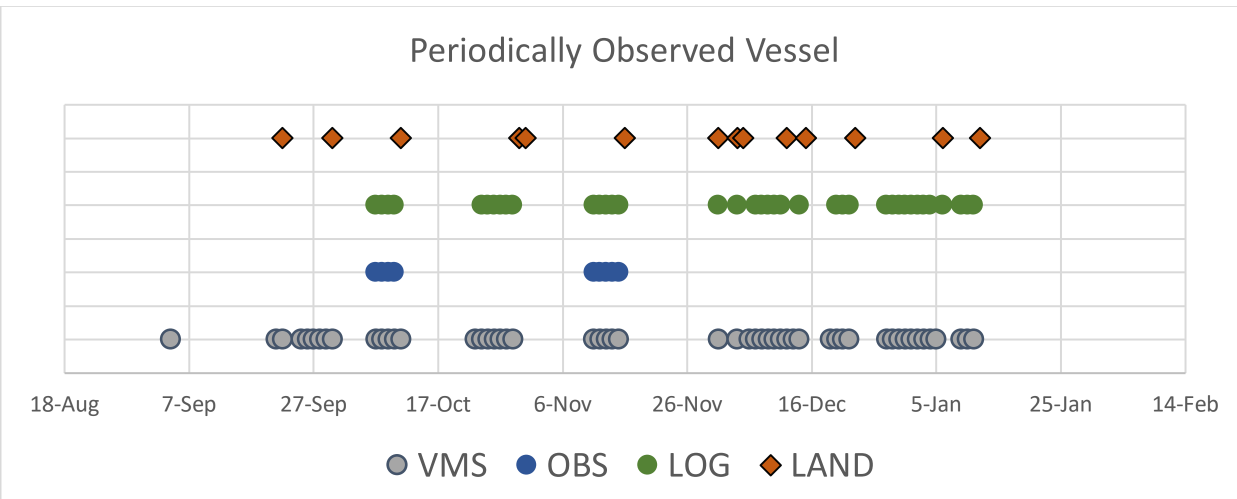

Figure 3: Timeline of VMS data matched by date to observer (OBS), logbook (LOG), and landings data (LAND) for a periodically observed vessel.

Plotting the recorded dates and positions from the datasets together on a map highlights how the data align in time and space. Figure 4 shows a portion of a typical observed trip. The VMS points are color-coded to reflect the average vessel speed that we used to determine potential fishing activity (gray > 1.25 knots, blue < 1.25 knots, 'is_fishing_speed'). The observer set and haul positions overlay the VMS pings. We calculated the number of pings in each CDFW fishing block to determine primary (black outline) and secondary (gray outline) fishing areas on a given trip and matched that to what was recorded on the logbook (green) and fish ticket (gray hatch marks). We did not try to quantify the rate at which the recorded CDFW block data matched the actual fishing activity as that was beyond the scope of this study, however such explorations of the data are possible.

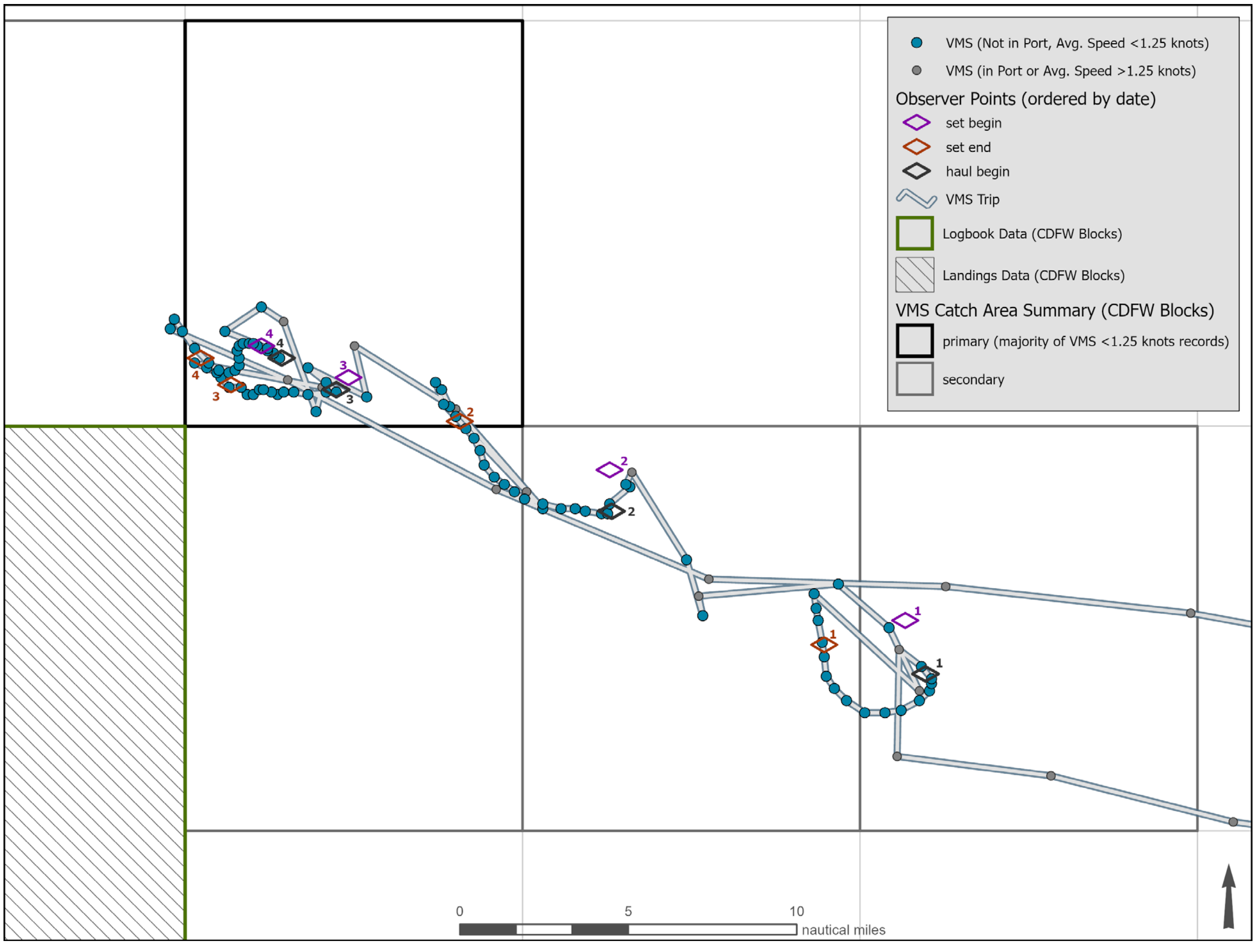

Figure 4: Comparison of VMS, observer, logbook, and landings data for a portion of an observed fishing trip.

Trip and Set Creation, Fishing Rules, and Data Validation

Prior to validating trips and sets, we scrutinized vessel transmission rates along with other vessel characteristics. Of the 22 vessels that participated in the DGN fishery for the six fishing years, six of the vessels were *unobservable* and 18 were *periodically observed*. However, three of the periodically observed vessels had only one observed trip during the study period, and were thus treated as unobservable. At most, 20 vessels participated in the DGN fishery in a given year (Table 2).

A few vessels had chronic VMS issues (unstable transmission rates, VMS turned off for long stretches of time, etc.) and could not be included in many of the analyses except for some basic trip and set metrics. The problem with unstable transmission rates is that trips or sets were initially created, but were flagged for review when they did not meet certain criteria. For instance, if a vessel's VMS did not transmit for a week and it arrived and departed port during that time, initially we may have only estimated one long trip by stringing the 'depart port' and 'arrive port' events together. However, if landing occurred in the middle of a trip, the trip would be flagged for review. During the review process, we would look at the vessel track and the logbook or observer data, if available, and could determine whether the vessel made two distinct trips instead of one.

Table 2: Number of vessels participating in the DGN fishing each year. Some vessels had DGN landings, but were not matched to trips due to chronic problems with their VMS transmission rate.

Metric	2013	2014	2015	2016	2017	2018	2019*	All
Vessels with landings	17	19	18	20	18	19	10	22
Vessels with matched trips	17	17	14	19	14	15	9	19
*2019 includes January only								

Erroneous transmission rates could also hamper estimating the total number of sets made in a given trip, which could inflate the catch per unit of effort. For example, if a vessel's VMS did not transmit a signal for a few days while it was out at sea, or there were less than six pings during the night when vessels tend to be actively setting, it was challenging to estimate fishing versus non-fishing.

Classification Models to Identify Fishing Sets

The random forest and gradient boosting models performed better than the other modeling frameworks at identifying fishing versus non-fishing behaviors based on VMS transmitted pings. A ROC curve illustrates a model's trade-off between its true positive rate (y-axis) and its false positive rate (x-axis), and the area under this curve (AUC) is a numeric measure of model performance (with an optimal value of 1). The AUC for both of our modeling frameworks was 0.96, representing highly accurate predictions with few false positives (Figure II-a). Additional metrics (Table 3) derived from a confusion matrix further illustrate the similar performance of both models.

Table 3: Summary of fishing and non-fishing classification performance on an out-of-sample testing dataset.

Model Performance	Random Forest Classification Model (RFM)	Gradient Boosting Classification Model (GBC)	Metrics Derived from Confusion Matrix*
Accuracy Rate	0.90	0.90	AR = (TP + TN)/n
Error Rate	0.10	0.10	ER = (FP +FN)/n
Precision	0.85	0.86	P = TP / (TP + FP)
Recall	0.93	0.91	R = TP / (TP + FN)
F1 Score	0.89	0.89	F1 = 2TP / (2TP + FP +FN)

* TP = true positive, TN = true negative, n = samples (i.e., VMS pings), FP = false positive, FN = false negative.

For the six fishing years, we estimated 2,448 total DGN sets from observer data and model predictions (Table 4) for all 22 vessels.

Table 4: Estimated number of sets by year resulting from observer data and classification model predictions.

Metric	2013	2014	2015	2016	2017	2018	2019*	All
Estimated count of sets	355	257	167	693	491	383	102	2448
*2019 includes January only								

Identifying DGN Trips

For the six fishing years, we estimated 571 total DGN trips for which we were able to match to landings (HMS Fishery code = 'DGNLM') to trips created from VMS data (for all 22 vessels). The 571 DGN trips represented approximately 81% of the total estimated landings across years, from a low of 58% in 2015 to a high of 94% in 2016 (Figure 3, Table 4). DGN trip landings (lbs) were calculated as the sum of the target and non-target species that are typically landed by the fleet (swordfish, mako shark (*Isurus oxyrinchus*), thresher sharks (*Alopias spp.*), tunas (*Thunnus spp.*), and opah (*Lampris guttatus*)), which may not match official landing records.

Nineteen percent of the landings (lbs) were not linked to trips due to gaps in VMS data either due to the implementation of the rules or errors in transmission rates. The lower rates of trip-link matching during the first three fishing years (2013 – 2015) were caused by data gaps due to the implementation of the rules that required VMS transmission. In 2013, vessels were not required to start using VMS until the fishing season was underway in the fall (78 FR 54548). In 2014 and 2015, VMS usage lapsed for most vessels between the expiration of the emergency rule (79 FR 29377) and the implementation of the final rule that required vessels to install VMS units and transmit locations (80 FR 10392, 80 FR 32465). There were fewer issues with VMS data and determining DGN trips for 2016 – 2019. There were a few cases where there was no matching fish tickets for an observed trip (e.g., the fish ticket may have been lost).

The process for matching trips to landings was iterative. For example, there were a few cases where positive DGN effort (is_driftgillnet_fishing) was determined by the presence of observer or logbook data, or the classification model estimated fishing effort, but the landing did not have the HMS fishery code = 'DGNLM'. We found that most of these cases were due to errors with CDFW gear codes or HMS fishery codes and would manually update the HMS fishery code in the database (and report the error to CDFW).

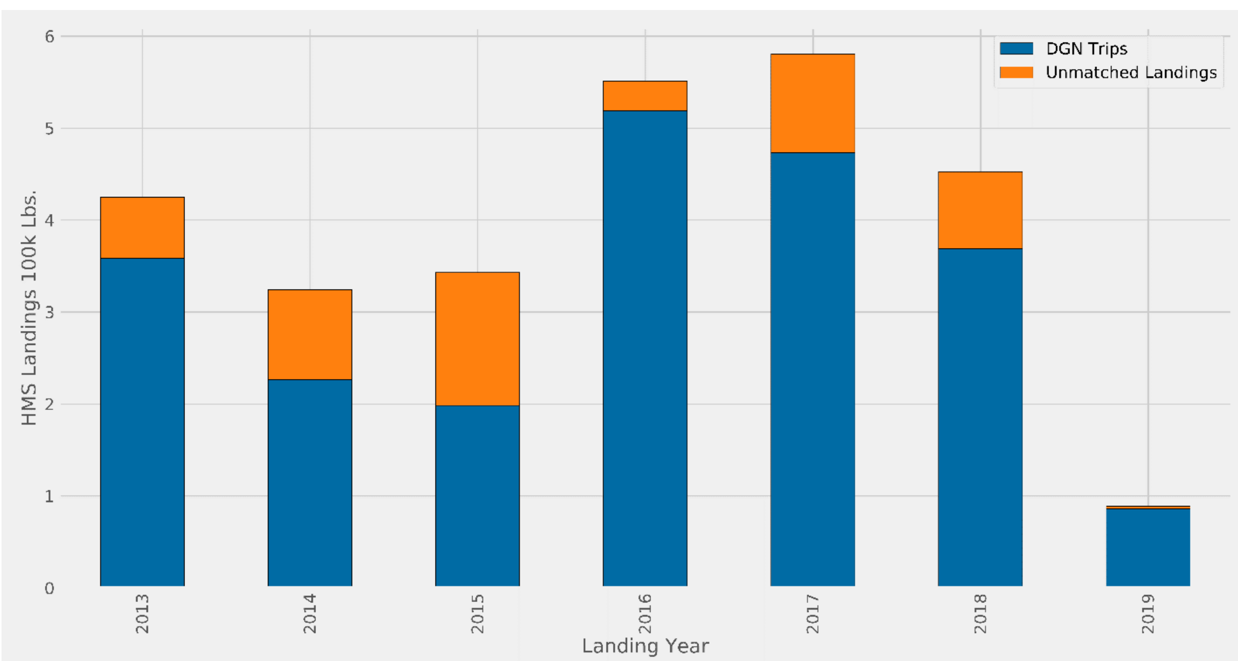

Figure 5: Total pounds landed of DGN target species by year and proportion of landings that verified DGN trips (VMS-landings linked) covered.

Table 5: Trip-link results for landings (lbs) and number of distinct fish tickets and estimated number of trips by year.

Metric	2013	2014	2015	2016	2017	2018	2019*	All
Trip-linked Landings (lbs)	358,014	226,656	198,302	520,370	474,737	374,117	84,915	2,237,110
Unmatched Landings (lbs)	66,204	97,380	144,614	32,603	107,207	83,647	2,869	534,524
Total Landings (lbs)	424,218	324,036	342,916	552,973	581,944	457,763	87,784	2,771,634
% Trip-linked Landings	84%	70%	58%	94%	82%	82%	97%	81%
Trip-linked Trips	76	53	54	146	89	92	20	529
Unmatched Trips	0	2	6	18	9	6	0	42
Count of estimated trips	76	55	60	164	98	98	20	571
*2019 includes January only								

Exploring Observer Coverage

The WCROP provides observer coverage on 20 – 30% of all sets in a given fishing year for the DGN fleet (NMFS 2013). Table 6 shows the estimated number of unobserved and observed trips and sets by vessel group for all six fishing years combined. For unobserved trips with missing or erratic VMS transmission rates, we estimated the total number of sets by calculating the average number of sets per trip for each vessel separately, and then substituted those averages for the problem trips. Periodically observed vessels tended to have two or three observed trips per year, which equates to seven to 13 sets per year. The six unobservable vessels represent the three vessels that were unobservable with the three vessels with low observer coverage (only one observed trip in the six fishing years).

Table 6: The estimated number of unobserved and observed trips and sets and coverage rates (percent) by vessel group.

Vessel Group	Number of Vessels	Unobserved Trips	Observed Trips	Percent Observed Trips	Unobserved Sets	Observed Sets	Percent Observed Sets
Excluded	5	47	4	9%	156	17	11%
Unobservable	6	145	3	2%	528	11	2%
Observable	11	296	121	41%	1200	558	47%
Totals	22	443	128	22%	1956	586	23%

Approximately 40% of the fleet was made up of unobservable vessels (six unobservable plus three excluded), an estimated 47% of sets made by periodically observed portion of the fleet was necessary to achieve the goal of 20 – 30% of observer coverage during this study.

We examined the number of trips and sets observed versus unobserved by week within season. We found that there was a systematic gap in observer coverage during parts of season 1 when fishing rates are low. Observers do not regularly begin covering the DGN fishery until September, although a few vessels fish before that in May through August.

We explored patterns in the estimated effort and observer coverage spatially. Figure 1 shows a the comparison between total estimated DGN effort (left panel) and observer coverage (right panel)) for the six fishing years combined. Overall, there were no detectable areas with effort and no coverage. We mapped and grouped the data in various ways and calculated differences in coverage in smaller 15km-hexbins to larger areas (i.e. PLCA, SCB, etc), but never found any evident differences in the distribution in estimated effort compared to coverage.

Observer Bias Analysis

We plotted the distributions of the trip and set level metrics prior to their inclusion in the LME models. Of the estimated 571 DGN trips and 2448 DGN sets, 364 trips (64%) and 1784 sets (73%) were included in the LME models. The various reasons trips and sets were not included in the LME models are listed in Table 7. We did not include data from the 2013/2014 fishing season in the observer bias analysis since the emergency rule (78 FR 54548) required all sets made deeper than 1100 fathoms to be observed.

Table 7: Count of estimated number of valid DGN trips and sets by year and the number of records removed from the observer bias analysis by issue and metric type.

Analysis Affected	Issue (count)	2013	2014	2015	2016	2017	2018	2019	Total
Trip	Estimated trips	76	55	60	164	98	98	20	571
Trip	Trips excluded due to missing fish ticket or erratic VMS transmission	0	2	6	17	9	8	0	43
Trip	Trips excluded due to potential bias (2013/2014 fishing year) or data quality issues	76	39	6	8	2	7	0	137
Trip	Trips with landed HMS pounds or ex-vessel revenue > 2 Std. Dev.	0	1	2	8	9	5	2	27
Trip	**Trips used in Bias Analysis**	**0**	**13**	**46**	**131**	**78**	**78**	**18**	**364**
Set	Estimated sets	355	257	167	693	491	383	102	2448
Set	Sets excluded due to potential bias (2013/2014 fishing year) or data quality issues	355	196	6	19	1	1	0	578
Set	Sets with set durations > 2 Std. Dev.	0	6	5	31	22	18	4	86
Set	**Sets used in Bias Analysis**	**0**	**55**	**156**	**643**	**468**	**364**	**98**	**1784**

Trip and set data from five of the 22 vessels (three periodically observed and two unobservable) were excluded from the observer bias analysis, but some data were used for trip and set metrics (Table 8). Of those five, two of the 22 DGN vessels only fished during that year and three had unstable or unreliable VMS data where ping transmissions were missing for long periods and therefore, characterizing fishing behavior was impossible.

Of the 17 remaining vessels, two vessels were missing VMS data for entire fishing seasons, but had DGN landings. Several other vessels had intermittent problems with their VMS transmissions and trips were omitted because unobserved fishing could not be sufficiently characterized.

The unobservable vessel group tended to be slightly smaller in median length (48 ft) than the periodically observed vessels (51 ft), however the median engine horsepower and net tonnage were 16% and 18% greater, respectively, for periodically observed vessels. The range of vessels metrics was quite wide both the unobservable and periodically observed vessels groups contain vessels with the smallest and largest capacities and engine horsepower.

Table 8: Vessel metrics by vessel group.

Vessel Group	Number of Vessels	Capacity (net tons)		Engine (hp)		Length (ft)	
		Median	Range	Median	Range	Median	Range
Excluded	5	14	8-16	300	120-471	38	35-40
Unobservable	6	23	5-30	213	165-335	48	42-51
Observable	11	27	5-65	250	180-1000	51	27-65
All	22	16	5-65	244	120-1000	45	27-65

We fit linear mixed-effect models (LME) to quantify the percent difference (i.e., resulting model coefficients) in a number of log-transformed fishery metrics (i.e., response variables) between the vessel groups (Watson et al., 2018).

The results of the LME models show the percent difference in response variables by two vessel groupings and the trip level by season (Figure 11 and Appendix II) for (1) unobservable vessels versus periodically observed vessels (left panel), and (2) observed versus unobserved trips by periodically observed vessels.

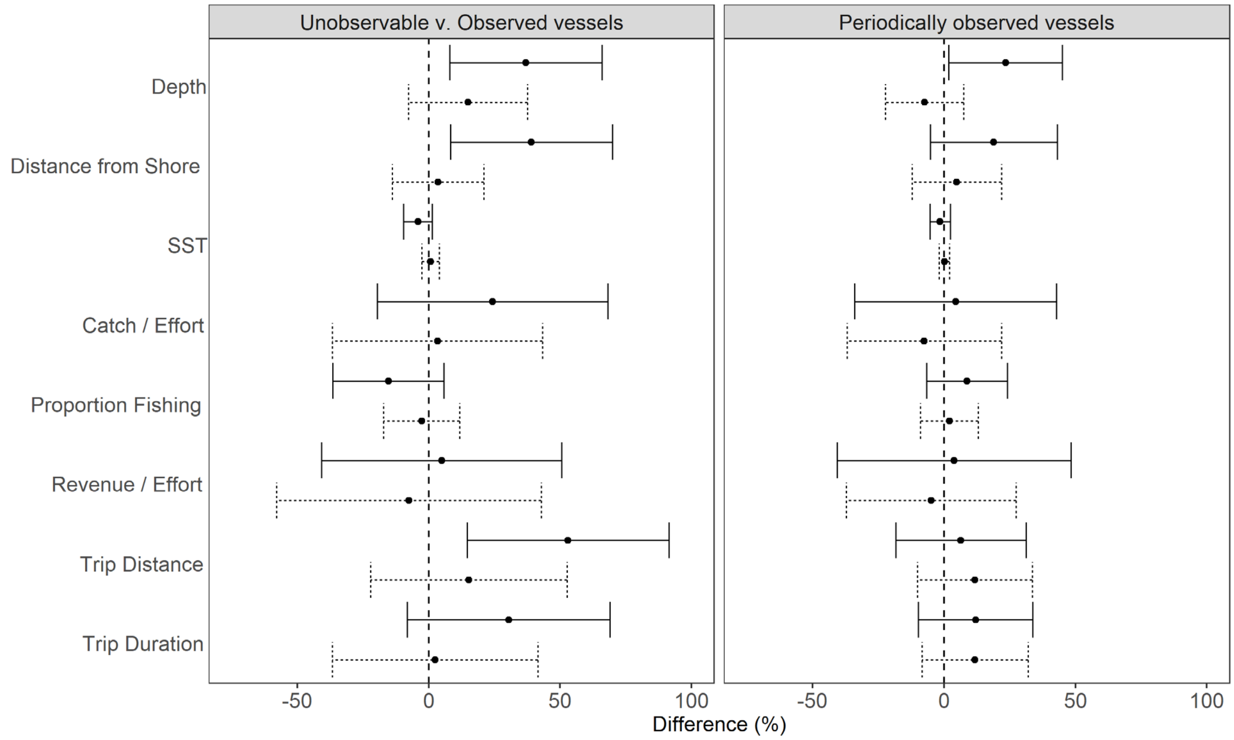

Figure 11: Percent difference between observed and unobservable (left), and observed and unobserved (right) trips. Points represent mean difference error bars bound the 95% confidence intervals. Error bars that span zero illustrate no significant difference. For each metric, solid symbols designate Season 1 (May 1 - Nov 15) and dotted symbols designate Season 2 (November 15 - January 31).

Overall, there were few statistically significant differences in fishing metrics between observed and unobserved trips on periodically observed vessels, or between unobservable and periodically observed vessels. Generally, the trip-level metrics showed more variability over the set-level metrics, probably since trip-level metrics were averaged over all of the sets in a trip and set-level metrics were only averaged over all of the data points for each individual set. We included many trip-level metrics that are similar, such as trip distance and trip duration, or distance from shore and depth and the results show those pairs of metrics had similar results. All of the model results are in Appendix 2.

Notably, the periodically observed vessels fished about 45% deeper and farther from shore than the unobservable vessels at the trip- and set-level during Season 1. However, there was no significant difference during Season 2. Likewise, the observed versus unobserved trips for the periodically observed vessels showed a similar pattern, where observed trips fished 23% deeper when an observer was onboard during Season 1. This pattern was less evident in the distance-from-shore metrics. The pattern reversed for both depth and distance from shore in Season 2 for the set-level metrics and for depth in the trip-level metrics. There were slight differences between seasons for unobservable vessels versus periodically observed vessels. Catch and effort (number of sets) by unobservable vessels tended to be greater in Season 1 compared to the Season 2); however, the opposite was true for periodically observed vessels. The periodically observed vessels tended to have longer trip distances, fish at greater depths, and fished farther offshore than unobservable vessels. This may be explained by the fact that there are more periodically observed vessels than unobservable vessels and more of those vessels tend to have greater capacities and more engine horsepower. The difference in SST at the trip and set level between the vessel groups and seasons had very low variability.

Discussion

A foundational aspect of partial coverage fishery observer programs is the assumption that sampling is sufficient to represent the full landscape of behaviors in a system (e.g., Pella & Geiger 2009). In traditional statistical designs, sample size power analyses typically benefit from the assumption that processes or individuals being sampled are either unaware or unconcerned with the fact that they are part of a study; this leads to little concern that subjects may bias the outcome of the results. However, in fisheries, the subjects (fishermen) being sampled may have perverse incentives to behave differently when they are observed versus unobserved. This is especially true when putative measures may result for some behaviors (e.g., exceeding bycatch quotas). We undertook a comprehensive and novel data-driven approach to examine potential differences between observed and unobserved fishing. While previous studies have explored biases by comparing, for example, landed catch compositions from observed and unobserved trips (e.g., Babcock & Pikitch 2003; Benoit & Allard 2009), ours is the first study (to our knowledge) to quantify potential biases by coupling trip- and set-level catches and revenues with spatial, environmental, and operational characteristics of observed and unobserved trips. In doing so for the DGN fishery, we observed little statistical evidence of differences between observed and unobserved trips, neither for vessels that periodically carried observers nor for vessels that were never observed.

The DGN fishery has been the target of numerous studies focused on environmentally-driven distributions of target and bycatch species. Soykan et al. (2014) used two decades of landings and ocean condition data to demonstrate the strong association between fishing locations and the expected spatial distributions of target species. Such strong relationships have led the DGN fishery to be a seminal case study for dynamic ocean management and environmentally-explicit species distribution models (e.g., Brodie et al., 2018; Eguchi et al., 2017; Hazen et al., 2018). The strength of these associations makes it unsurprising that we found few statistical differences between observed and unobserved behaviors. Moreover, because the target species for this fishery are pelagic and are influenced by more ephemeral conditions like water temperatures at the surface, species movements are more likely to cover a broad spatial range. The fidelity of target species to certain environmental conditions is illustrated by Figure 11, which demonstrates a small variance in the SST differences for observed and unobserved trips, but a high variance for spatial metrics like trip distance or distance from shore, suggesting that vessels traveled variable distances in search of less variable environments. These observations support the notion that the drivers of selecting fishing locations may have been relatively persistent, regardless of observer presence.

While our particular results apply to only one fishery, we applied a framework for large-scale integrations of fishery-dependent and fishery-independent datasets within an operational database backend that facilitate an understanding of a broad landscape of potential set- and trip-level fishing behaviors. Moreover, while several studies have used VMS data to classify fishing and non-fishing behaviors using machine learning approaches (e.g., Russo et al., 2011; Joo et al., 2013; O'Farrell et al., 2017), this may be the first study to do so using boosted regression trees, expanding the set of tools with which to detect fishing events. However, despite having a robust database infrastructure with which to efficiently merge datasets from multiple sources (e.g., VMS, an observer program, landings records) for model development and behavioral inference, a great deal of manual quality control was still required before the machine learning vessel behavior models could be trained.

Several data challenges typified the quality control and data management for this study. First, despite a requirement for vessels to transmit location data via VMS, several vessels' VMS data were insufficient for analyses. Gaps and inconsistencies in VMS data have been reported previously to affect inference (e.g., Watson & Haynie 2016; Thoya et al., 2021) and in this case, some vessels lacked enough VMS data to match to any fishing activity (e.g., observer or landings data). A potential and unknown sampling bias may occur for such vessels. One theory would be that vessels that are less likely to reliably transmit data despite the VMS requirement might also be more likely to behave differently when unobserved. However, such vessels represented a small component of our dataset (one to three vessels per year) and as previously noted, environmental conditions may serve as a more predictable driver of fishery behavior than observer or VMS coverage. An additional challenge in data integration was matching VMS data to human-recorded data fields like dates reported in logbooks or landing records. This challenge, also reported elsewhere (e.g., Bastardie et al., 2010; Watson et al., 2018) typically occurs when there are no landings records whose dates reasonably align with the automated time stamps generated by VMS data or the generally more meticulously-recorded time stamps from fishery observers.

Implications and Conclusions

We have developed a system within an integrated database backend that would facilitate the broad implementation of multiple different aspects of our approach for fisheries with partial observer coverage and VMS data for the West Coast, Alaska, or Pacific Islands regions. We have presented a comprehensive analysis of trip- and set-level characteristics to explore potential differences between observed and unobserved trips. Benoit & Allard (2009) assert that to detect an "observer effect," analysts should examine statistical differences in the amounts of target catch, bycatch, and fishing effort. However, many key species of concern (e.g. leatherback sea turtles) are only rarely caught in the DGN fishery. Thus, the amounts of bycatch or bycatch to target catch ratios were not a meaningful analysis to explore. Our approach however quantified a broader set of fishery metrics that enabled a more comprehensive assessment of the behaviors that characterized the observed and unobserved trips and vessels. Thus, while no approach can detect whether rare bycatch species were discarded at sea, our approach is at least more robust to understand whether the behaviors and spatial footprint of unobserved fishing are consistent with those that were observed.

References

Babcock & Pikitch. 2003. How much observer coverage is enough to adequately estimate bycatch?. Miami, FL: Pew Institute of Ocean Science, 2003.

Bastardie, F., Nielsen, J.R., Ulrich, C., Egekvist, J., and Degel, H. 2010. Detailed mapping of fishing effort and landings by coupling fishing logbooks with satellite-recorded vessel geo-location. Fisheries Research 106:41-53.

Benoît, H.P. and Allard, J., 2009. Can the data from at-sea observer surveys be used to make general inferences about catch composition and discards? Canadian Journal of Fisheries and Aquatic Sciences, 66(12), pp.2025-2039.

Brodie, S., Jacox, M. G., Bograd, S. J., Welch, H., Dewar, H., Scales, K. L., Maxwell, S. M., Briscoe, D. M., Edwards, C. A., Crowder, L. B., Lewison, R. L., & Hazen, E. L. 2018. Integrating dynamic subsurface habitat metrics into species distribution models. Frontiers in Marine Science, 5(JUN), 1–13. https://doi.org/10.3389/fmars.2018.00219

Carretta, J.V.. 2020. Estimates of marine mammal, sea turtle, and seabird bycatch in the California large-mesh drift gillnet fishery:1990-2018. U.S. Department of Commerce, NOAA Technical Memorandum NMFS-SWFSC-632.

Chapman, P., Clinton, J., Kerber R., Khabaza T., Reinartz T., Shearer C., and Wirth R. 2000. CRISP-DM step-by-step data mining guide. The Modeling Agency Website https://www.the-modeling-agency.com/crisp-dm.pdf

D'Angelo, C.D. 2014 (unpublished). An Alternative Method for Estimating West Coast HMS Landings

Eguchi, T., Benson, S. R., Foley, D. G., & Forney, K. A. (2017). Predicting overlap between drift gillnet fishing and leatherback turtle habitat in the California Current Ecosystem. Fisheries Oceanography, 26(1), 17-33.

GEBCO Compilation Group. 2019 GEBCO Grid. (doi:10.5285/a29c5465-b138-234d-e053-6c86abc040b9) Website: https://www.gebco.net/data_and_products/gridded_bathymetry_data/gebco_2019/gebco_2019_info.html

Hazen, E. L., Scales, K. L., Maxwell, S. M., Briscoe, D. K., Welch, H., Bograd, S. J., Bailey, H., Benson, S. R., Eguchi, T., Dewar, H., Kohin, S., Costa, D. P., Crowder, L. B., & Lewison, R. L. 2018. A dynamic ocean management tool to reduce bycatch and support sustainable fisheries. Science Advances, 4(5), 1–8. https://doi.org/10.1126/sciadv.aar3001

Joo, R., Salcedo, O., Gutierrez, M. Falbet, R. Bertrand, S. 2015. Definiting fishing spatial strategies from VMS data: Insights from the world's largest monospecific fishery. Fisheries Research. 164:223-230. https://doi.org/10.1016/j.fishres.2014.12.004

JPL MUR MEaSUREs Project. 2015. GHRSST Level 4 MUR Global Foundation Sea Surface Temperature Analysis. Ver. 4.1. PO.DAAC, CA, USA. Dataset accessed [7 May 2021] at https://doi.org/10.5067/GHGMR-4FJ04

Mason, J. G., Hazen, E. L., Bograd, S. J., Dewar, H., & Crowder, L. B. 2019. Community-level effects of spatial management in the California drift gillnet Fishery. Fisheries Research, 214(September), 175–182. https://doi.org/10.1016/j.fishres.2019.02.010

Muench, A., Depiper, G. S., & Demarest, C. (2018). On the precision of predicting fishing location using data from the vessel monitoring system (Vms). Canadian Journal of Fisheries and Aquatic Sciences, 75(7), 1036–1047. https://doi.org/10.1139/cjfas-2016-0446

NMFS. 2013. Biological Opinion on the continued management of the drift gillnet fishery under the Fishery Management Plan for U.S. West Coast Fisheries for Highly Migratory Species. U.S. Dep. Commer., NOAA, NMFS, SW Reg., Sustain. Fish. Div., 2 May 2013. westcoast.fisheries.noaa.gov/publications/protected_species/marine_mammals/memo_signed_dgn_biop_050213.pdf.

Shay O'Farrell, James N. Sanchirico, Iliana Chollett, Marcy Cockrell, Steven A. Murawski, Jordan T. Watson, Alan Haynie, Andrew Strelcheck, Larry Perruso, Improving detection of short-duration fishing behaviour in vessel tracks by feature engineering of training data, ICES Journal of Marine Science, Volume 74, Issue 5, May-June 2017, Pages 1428–1436, https://doi.org/10.1093/icesjms/fsw244

Pacific Marine and Estuarine Fish Habitat Partnership (PMEP) West Coast USA Current and Historical Estuary Extent (2017) https://www.pacificfishhabitat.org/data/estuary-extents

Pella, J. J., and H. J. Geiger. 2009. Sampling considerations for estimating geographic origins of Chinook salmon bycatch in the Bering Sea pollock fishery. Alaska Department of Fish and Game, Special Publication No. 09-08, Anchorage.

Russo, T., Parisi, A., Prorgi, M., Boccoli, F., Cignini, I., Tordoni, M., & Cataudella, S. (2011). When behaviour reveals activity: Assigning fishing effort to métiers based on VMS data using artificial neural networks. Fisheries Research, 111(1–2), 53–64 (https://doi.org/10.1016/j.fishres.2011.06.011)

Simons, R.A. 2020. ERDDAP. https://coastwatch.pfeg.noaa.gov/erddap. Monterey, CA: NOAA/NMFS/SWFSC/ERD. Last Accessed: 7 May 2021.

Soykan, C.U., Eguchi, T., Kohin, S., Dewar, H. 2014. Prediction of fishing effort distributions using boosted regression trees. Ecological Applications. 24:71-83. https://doi.org/10.1890/12-0826.1

Stohs, S., Sippel, T. 2017. Analysis of Increasing the Required VMS Ping Rate for the California Drift Gillnet Fishery. NOAA Technical Memorandum NMFS. (doi:10.7289/V5/TM-SWFSC-570)

Thoya, P.; Maina, J.; Möllmann, C.; Schiele, K.S. AIS and VMS Ensemble Can Address Data Gaps on Fisheries for Marine Spatial Planning. Sustainability 2021, 13, 3769. https://doi.org/10.3390/su13073769

Urbisci, L.C., Stohs, S.M., Piner, K.R. 2016. From Sunrise to Sunset in the California Drift Gillnet Fishery: An Examination of the Effects of Time and Area Closures on the Catch and Catch Rates of Pelagic Species. Marine Fisheries Review.MFR 78(3-4), (doi: dx.doi.org/10.7755/MFR.78.3–4.1)

Vølstad, J.H. and M. Fogarty. 2006. Report on the National Observer Program Vessel Selection Bias Workshop, Woods Hole, MA, May 17-19, 2006.

Watson, J.T., Haynie, A.C., Sullivan, P.J., Perruso, L., O'Farrell, S., Sanchirico, J.N., Mueter, F.J. 2018. Vessel monitoring systems (VMS) reveal an increase in fishing efficiency following regulatory changes in a demersal longline fishery. Fisheries Research. 207:85-94

Watson JT, Haynie AC 2016. Using Vessel Monitoring System Data to Identify and Characterize Trips Made by Fishing Vessels in the United States North Pacific. PLoS ONE 11(10):e0165173. (doi:10.1371/journal.pone.0165173)

Appendix I – Comprehensive DGN

The stepwise process for creating the Comprehensive DGN VMS integrated dataset creating engineered features (Table I-a) is as follows:

1. Create a database table with VMS data for vessels participating in the DGN fishery from September 2013 to January 2019.
    1. Create engineered fields such as average vessel speed, distance between points, average distance between current point, previous point, and subsequent point, distance from shore, etc. (Watson et al., 2018).
    2. Create value-added fields such as 'is_fishing_speed', 'is_during_fishing_time', 'is_fishing_sql'.
    3. Merge vessel characteristics from the USCG data (net tonnage, length, year built, etc).
2. Merge observer data into the table by matching on vessel, date, and time. The merged observer data (daily set-by-set records) included date, time, and position (latitude and longitude) of the set start, set end, and begin haul events.
3. Merge logbook data into the table by matching on vessel and date. The merged logbook data included date, location (CDFW block), net mesh size, target species, set type (drift or set).
4. Merge landings data in the table by matching on vessel and date and create 'is_driftgillnet_fishery'.
5. Add Geospatial Features with ArcGIS Pro
    1. Define port polygons.
    2. Label each VMS record with vessel activity (in port, at sea, depart port, arrive port). The 'in port' activity was derived by assessing whether the vessel was within the defined port polygon.
    3. Merge bottom depth data for each position.
    4. Using ArcGIS Pro, create trips by linking the sequential 'depart port' to 'arrive port' activities by vessel, referred to as 'GIS-defined trips'.
6. Merge SST for each position using a PL/SQL procedure, which retrieved the SST data through the ERDDAP web service.
    1. Create Trip and Set data, apply fishing rules, validate data.
7. Determine fishing and non-fishing events using classification models with data fields described in steps 2 - 9. Merge model results into 'is_fishing_rfm' and 'is_fishing_gbc', then verify sets (is_fishing_set).
8. Identify true fishing trips (aka DGN trips) from GIS-defined VMS trips based on the 'depart port' and 'arrive port' events. Assign a sequential trip number (trip_num) and set number (set_num) for each verified trip and set.

Table I-a: Engineered features defined for each VMS record

Variable	Description	Procedure	Use
event_type	In port, at sea, depart port, arrive port	ArcGIS	Determine fishing effort Mapping
distance_from_shore	Distance to nearest point on coast (m)	ArgGIS	Trip & set metrics
port_distance	Distance to nearest port (m) (Fig. 1)	ArgGIS	Determine event types (depart port, arrive port, in port)
distance_m	Distance between current and previous vessel position (m)	PL/SQL	Classification model input
distance_tplus1	Difference between subsequent 'distance_m' and current 'distance_m' (m)	PL/SQL	Classification model input
distance_tneg1	Difference between current 'distance_m' and previous 'distance_m' (m)	PL/SQL	Classification model input
distance_t5	Average of a 5-record moving window (current 'distance_m', two subsequent 'distance_m', and two previous 'distance_m') (m)	PL/SQL	Classification model input
speed	Speed between current and previous vessel position (m/min)	PL/SQL	Classification model input
speed_tplus1	Difference in vessel speed between current and subsequent record (m/min)	PL/SQL	Classification model input
speed_tneg1	Difference in vessel speed between current and previous record (m/min)	PL/SQL	Classification model input
speed_t5	Average of a 5-record moving window (current speed, two subsequent speeds, and two previous speeds) (m/min)	PL/SQL	Classification model input
depth	Bottom depth (m)	ArgGIS	Classification model input
daily_temperature	Daily sea surface temperature	PL/SQL	Classification model input

is_fishing_logbook_day	Boolean variable denoting if VMS record occurred on a date recorded in the vessel logbook	PL/SQL	Determine fishing effort
is_observed_ping	Boolean variable denoting if VMS record occurred during an observed set retrieval. VMS records were matched to observer records on date, time, and vessel	PL/SQL	Determine fishing effort
is_observed_trip	Boolean variable denoting if an observer was onboard during the trip. VMS records were matched to observer records on date, time, vessel.	PL/SQL	Determine fishing effort
is_fishing_speed	Boolean denoting if speed < 38.58 (m/min) or 1.25 knots[1]	PL/SQL	Determine fishing effort
is_during_fishing_time	Boolean variable denoting pings from 3:00 PM to 8:00 AM	PL/SQL	Determine fishing effort
is_fishing_sql	is_during_fishing_time + is_fishing_speed = 2 AND event_type = ‘at sea’	PL/SQL	Determine fishing effort Classification model input
is_driftgillnet_fishery	Boolean variable denoting positive DGN fishing trip (is_observed_ping=1) OR (landing gear type = DGN) OR (is_fishing_logbook_day = 1)	PL/SQL	Trip & set metrics
is_fishing_rfm	Boolean variable denoting random forest classification model output (fishing or not fishing) for all defined ‘at sea’ event_types for DGN trips	Python	Classification model output
is_fishing_gbc	Boolean variable denoting gradient boosting classification model output (fishing or not fishing) for all defined ‘at sea’ event_types for DGN trips	Python	Classification model output
is_fishing_set	Boolean variable denoting a set is verified (is_during_fishing time = 1, event_type = ‘at sea, >=6 pings from both is_fishing_rfm AND is_fishing_gbc) OR where ping rate was erratic (is_during_fishing time = 1, event_type = ‘at sea, >3 pings	PL/SQL	Determine fishing effort (ultimate determination)

	AND >6 hours between first and last is_fishing_rfm/gbc		
is_set_end_point	The last ping of is_fishing_set equivalent to when a vessel would begin to haul their net	PL/SQL	Mapping Trip & set metrics Bias analysis
trip_num	Sequential trip number used to identify for individual trips	PL/SQL	Trip & set metrics Bias analysis
set_num	Sequential set number based on trip number and order of sets as defined by the observer or as defined by classification models	PL/SQL	Trip & set metrics Bias analysis
season	Denotes season of trip (Season 1, 1 May – 15 Nov or Season 2, 16 Nov – 31 Jan). If the trip overlapped two seasons, season was based on the season that included the majority of fishing days.	PL/SQL	Bias analysis
vessel_group	Code denoting the vessel group that records belong to (excluded, unobservable, or observable)	PL/SQL	Bias analysis
[1]Stohs & Sippel, 2017			

Appendix II: Classification and Mixed-linear Effects Model Results

Classification Model Results

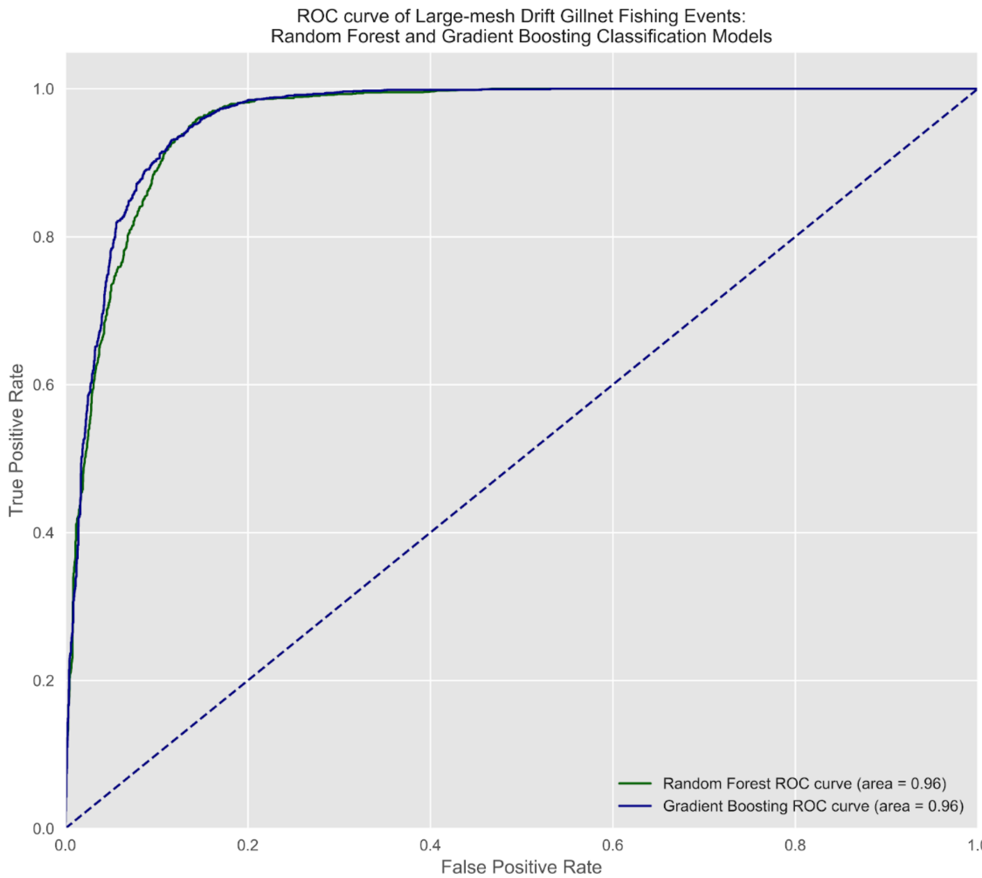

Figure II-a: ROC curve for the random forest and gradient boosting classification models on the testing data.

LME Model results

Table II-b. Trip summary statistics from the LME models for Season 1, May 1 - Nov 15 comparing unobservable and periodically observed vessels.

LME	Coefficient	P-value	Lower CI	Upper CI	Trips
Average Set Longitude	0.41	0.340	-0.43	1.25	136
Average Set Latitude	0.85	0.573	-2.10	3.80	136
Average Set Sea Surface Temperature	-4.12	0.141	-9.60	1.36	136
Trip Duration	30.40	0.122	-8.14	68.93	136
Trip Distance	53.04	0.007	14.59	91.49	136
Effort	18.67	0.383	-23.30	60.64	136
Proportion of Trip Fishing	-15.36	0.155	-36.53	5.80	136

Average Set Distance from Shore	39.05	0.013	8.27	69.83	136
Average Set Depth	36.98	0.012	8.03	65.93	136
Ex-vessel Revenue	18.37	0.622	-54.59	91.32	136
RPUE	4.93	0.833	-40.77	50.62	136
Catch	40.62	0.249	-28.38	109.62	136
CPUE	24.25	0.278	-19.58	68.09	136

Table II-c. Trip summary statistics from the LME models for season 2, November 15 - January 31, comparing unobservable and periodically observed vessels.

LME	Coefficient	P-value	Lower CI	Upper CI	Trips
Average Set Longitude	-0.07	0.791	-0.57	0.43	228
Average Set Latitude	-0.19	0.724	-1.25	0.87	228
Average Set Sea Surface Temperature	0.67	0.695	-2.68	4.03	228
Trip Duration	2.48	0.901	-36.63	41.59	228
Trip Distance	15.29	0.423	-22.14	52.72	228
Effort	7.02	0.693	-27.80	41.84	228
Proportion of Trip Fishing	-2.66	0.719	-17.14	11.83	228
Average Set Distance from Shore	3.54	0.690	-13.87	20.94	228
Average Set Depth	14.99	0.194	-7.63	37.60	228
Ex-vessel Revenue	-0.82	0.974	-50.43	48.78	228
RPUE	-7.52	0.770	-57.82	42.78	228
Catch	7.24	0.784	-44.50	58.99	228
CPUE	3.37	0.869	-36.62	43.35	228

Table II-d. Trip summary statistics from the LME models for Season 1, May 1 - Nov 15, comparing unobserved trips versus observed trips from periodically observed vessels.

LME	Coefficient	P-value	Lower CI	Upper C	Trips
Average Set Longitude	0.12	0.755	-0.66	0.91	79
Average Set Latitude	-1.54	0.391	-5.06	1.98	79
Average Set Sea Surface Temperature	-1.51	0.447	-5.41	2.38	79
Trip Duration	11.99	0.280	-9.76	33.74	79

Trip Distance	6.42	0.612	-18.37	31.20	79
Effort	14.66	0.222	-8.86	38.17	79
Proportion of Trip Fishing	8.67	0.269	-6.71	24.06	79
Average Set Distance from Shore	18.89	0.125	-5.24	43.02	79
Average Set Depth	23.42	0.034	1.82	45.02	79
Ex-vessel Revenue	14.42	0.561	-34.24	63.08	79
RPUE	3.76	0.869	-40.75	48.26	79
Catch	17.91	0.424	-26.03	61.86	79
CPUE	4.38	0.823	-33.97	42.73	79

Table II-e. Trip summary statistics from the LME models for Season 2, November 15 - January 31, comparing unobserved trips versus observed trips from periodically observed vessels.

LME	Coefficient	P-value	Lower CI	Upper CI	Trips
Longitude	-0.22	0.230	-0.57	0.14	173
Latitude	-0.39	0.419	-1.35	0.56	173
Sea Surface Temperature	0.14	0.888	-1.83	2.11	173
Trip Duration	11.76	0.254	-8.45	31.96	173
Trip Distance	11.72	0.293	-10.10	33.54	173
Effort	14.27	0.184	-6.77	35.32	173
Proportion of Trip Fishing	2.01	0.720	-8.97	12.99	173
Distance from Shore	4.83	0.578	-12.18	21.84	173
Depth	-7.44	0.329	-22.36	7.49	173
Ex-vessel Revenue	10.32	0.578	-26.02	46.65	173
RPUE	-4.93	0.765	-37.21	27.35	173
Catch	7.982806	0.645215	-25.999251	41.964863	173
CPUE	-7.537897	0.614912	-36.905354	21.82956	173

Table II-f. Trip summary statistics from the LME models for Season 1, May 1 - November 15, comparing sets made by unobservable versus periodically observed vessels.

LME	Coefficient	P-value	Lower CI	Upper CI	Sets
Longitude	0.53	0.156	-0.20	1.27	654

Latitude	0.79	0.522	-1.62	3.19	654
Sea Surface Temperature	-4.70	0.068	-9.75	0.36	654
Soak Time	9.13	0.219	-5.44	23.69	654
Distance from Shore	38.17	0.026	4.52	71.83	654
Depth	50.57	0.002	18.92	82.21	654

Table II-g. Trip summary statistics from the LME models for Season 2, November 15 - January 31, comparing sets made by unobservable versus periodically observed vessels.

LME	Coefficient	P-value	Lower CI	Upper CI	Sets
Longitude	0.21	0.556	-0.48	0.89	1130
Latitude	0.28	0.720	-1.23	1.79	1130
Sea Surface Temperature	-0.80	0.660	-4.38	2.77	1130
Soak Time	-5.70	0.453	-20.58	9.18	1130
Distance from Shore	15.10	0.122	-4.05	34.24	1130
Depth	28.10	0.045	0.65	55.55	1130

Table II-h. Trip summary statistics from the LME models for Season 1, May 1 - November 15, comparing unobserved and observed sets made by periodically observed vessels.

LME	Coefficient	P-value	Lower CI	Upper C	Sets
Longitude	0.04	0.824	-0.32	0.40	396
Latitude	-1.12	0.189	-2.79	0.55	396
Sea Surface Temperature	-1.02	0.286	-2.90	0.86	396
Soak Time	11.85	0.014	2.42	21.28	396
Distance from Shore	10.26	0.095	-1.77	22.30	396
Depth	16.61	0.010	3.90	29.33	396

Table II-i. Trip summary statistics from the LME models for Season 2, November 15 - January 31, comparing unobserved and observed sets made by periodically observed vessels.

LME	Coefficient	P-value	Lower CI	Upper CI	Sets
Longitude	-0.17	0.095	-0.37	0.03	859
Latitude	-0.08	0.751	-0.60	0.44	859
Sea Surface Temperature	0.98	0.021	0.15	1.82	859

Soak Time	15.39	<<0.001	9.25	21.53	859
Distance from Shore	-8.68	0.025	-16.27	-1.09	859
Depth	-4.79	0.243	-12.81	3.24	859