INTER-AMERICAN TROPICAL TUNA COMMISSION

2nd WORKSHOP ON DATA IMPROVEMENT (C-03-05): PURSE-SEINE VESSELS ≤363 t

(*by videoconference*)
18-20 February 2025

DOCUMENT DAT-02-02

ASSESSING OBSERVER COVERAGE LEVELS FOR ESTIMATING TOTAL BYCATCH IN CLASS 1-5 PURSE-SEINE VESSELS

Dan Ovando, Leanne Fuller, Shane Griffiths and Jon Lopez

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This document provides advice on representativeness and coverage levels potentially needed from observer programs on Class 1-5 purse seine vessels (<364 metric tons) in order to reliably estimate species-specific total catches. Currently, species-level bycatch data are not consistently recorded in logbook data, meaning that estimates of total bycatch removals for individual species must be made by extrapolating rates from the observed subset of Class 1-5 vessels sets to the larger unobserved sets. In recent years observer coverage levels reached 30-40% for Class 1-5 sets and trips. Using a range of analyses, this report assesses how accurate this process may be and provides recommendations on what levels may be needed to accurately estimate total bycatch, with different associated error rates, from this fleet segment. Our results suggest that existing coverage levels (i.e., 30-40%) may be adequate to estimate total catches for some commonly encountered and representatively observed species (e.g., common

dolphinfish, wahoo, silky shark), but that much higher observer rates (greater than 80%), whether through human or electronic monitoring, spread more representatively over the fishing operations are likely to be needed for rarer species (e.g., oceanic whitetip shark, mobulid rays, bigeye thresher shark).

1. INTRODUCTION

Tropical tuna fisheries in the eastern Pacific Ocean (EPO) interact with many different species besides the target species, resulting in incidental catch (bycatch) of species such as common dolphinfish, wahoo, billfish, sharks, rays, and sea turtles, among others. Obtaining accurate estimates of the amount of bycatch of individual species is important for better understanding the impacts of the fishery on the species and the ecosystem, and generating annual reports (e.g., EB-02-01), indicator-based management (e.g., EB-02-02), parameterization of ecological risk assessments (e.g., EASI-Fish) and updates of the IATTC's ecosystem model (DAT-02-01). This is especially true for key bycatch species, such as those with specific resolutions or listed under the purview of the IATTC (e.g., see Reports and Provision of Data in the IATTC website, under the section "Bycatch Management").

The IATTC maintains 100% observer coverage on Class 6 purse-seine vessels operating within the IATTC convention area per the Agreement on the International Dolphin Conservation Program (AIDCP). These observers collect data on both tropical tuna catches and incidentally captured species herein termed "bycatch". However, observer coverage for Class 1-5 purse-seine vessels, (which is produced by a mix of national, IATTC, and the voluntary TUNACONS observer programs, see DAT-02-01), is lower, covering roughly 38% of sets (34% of trips) by this fleet segment during the study period of 2022-2024. Logbook data exist for unobserved purse-seine sets in the EPO (DAT-02-01), but these logbook data rarely include information on bycatch.

As a result, total bycatch of individual species by the Class 1-5 purse seine vessels are poorly understood in the EPO. This document provides an evaluation of the potential implications of different observer coverage levels, and its spatiotemporal distribution, for the Class 1-5 purse seine vessels with regards to estimating total bycatch levels. It also provides the foundation to draft recommendations on potential observer coverage levels needed to monitor and estimate total catch of key bycatch species by this segment of the EPO tropical tuna purse seine fleet.

2. METHODS

The present work is based on 2022-2024 observer and logbook data from Class 1-5 purse seine vessels (see DAT-02-01), along with observer data from Class 6 purse seine vessels. In order to be more representative and provide guidance on species of interests for the Commission (i.e., species of special conservation interest, species with existing Resolutions), the analysis focused on target species, as well as common (e.g., dorado, silky shark) and rare (e.g., oceanic whitetip shark, mobulid rays, bigeye thresher sharks, sea turtles) bycatch species.

2.1 Exploratory Analysis

We first conducted an exploratory analysis of the available data. Observer and logbook data were combined to examine observer coverage of Class 1-5 purse seine vessels through a range of lenses, including coverage rates in total as well as in space, time, flag, and vessel size class. We also explored the nature of variation of species level bycatch in the Class 6 data as a point of reference for Class 1-5 purse seine vessels.

2.2 Representativeness of Class 1-5 Observer Data for Tropical Tunas

We followed a similar methodology to BYC-10 INF-D, focused on longline data, as one means of evaluating the ability of the current Class 1-5 purse seine vessels observer coverage to accurately estimate total annual retained catches of bycatch species. While the logbooks for unobserved Class 1-5 sets do not report bycatch totals, they do record the volume of tropical tuna retained, by species. Taking these data as truth, an estimate of the total retained catch of tropical tuna species by the Class 1-5 vessels can be produced by summing the logbook and observer data from those vessels.

If the Class 1-5 purse seine vessels observer data are representative of unobserved Class 1-5 sets, the total retained tropical tuna catches reported in the logbook data could be reconstructed using only the observer data. If the observer data are not representative, estimates of total logbook catch based on the observer data should be inaccurate and/or biased. Subsequently, if the Class 1-5 observer data are not representative for tropical tunas, it is unlikely they would be representative for bycatch species. However, representativeness for the tropical tuna species does not guarantee representativeness for bycatch species.

We attempted to estimate the total retained Class 1-5 purse seine vessels tropical tuna logbook catch by first fitting a random forest model predicting catch per set (catch per unit effort, CPUE) in the observed Class 1-5 sets as a function of vessel carrying capacity, fishing location, and fishing month. This model was then used to predict the average catch per set per tropical tuna species of the unobserved Class 1-5 sets and then summing across the number of unobserved Class 1-5 sets. We then compared the reported total logbook catch of tropical tuna against the total tropical tuna logbook catches obtained by predicting the unobserved tuna catch rates using the observed tuna catch rates.

2.3 Comparison of Class 6 and Class 1-6 Bycatch Rates

While observer coverage is limited for Class 1-5 purse seine vessel trips, 100% observer coverage exists for the Class 6 purse seine vessels. Given that there is spatial overlap between the Class 6 and Class 1-5 purse seine vessels, we evaluated whether bycatch dynamics, particularly variability in bycatch rates, were sufficiently similar between the Class 6 vessels and observed segments of the Class 1-5 vessels such that Class 6 data could be used to provide insights on potential required observer coverage levels for the Class 1-5 fleet segment.

To accomplish this, observer-based bycatch data for the Class 1-5 vessels was extracted. Note that this restricts the resulting dataset to only Class 3-5 vessels since no observer data were recorded for Class 1-2 vessels in our database. We then created a database of bycatch data from a subset of the Class 6 vessels observer data selected to more closely match the Class 1-5 vessel sizes and fishing grounds. First, a convex hull covering the spatial distribution of the Class 1-5 sets recorded by either logbook or observer was estimated. We then filtered the Class 6 observer data to only include sets within the Class 1-5 operations polygon and corresponding only to Class 6 vessels with less than 600 metric tons of carrying capacity to more closely match the sizes of the Class 1-5 vessels.

We then evaluated whether bycatch rates in this subset of the Class 6 vessels are sufficiently similar in their dynamics (i.e., bycatch rates) to the observed Class 1-5 sets so as to be useful for this study (i.e., provide guidance on the impact of observer coverage levels on our ability to estimate total catch of individual bycatch species reliably). Total catch can be calculated as:

$$C_s = CPUE_s \times E_s$$

where *C* is total catch, $CPUE_s$ is mean catch per unit effort, *E* is effort, and *s* denotes species. In this case, *E* is known, defined as the number of applicable Class 1-5 sets.

However, when observer coverage is less than 100%, the true average CPUE across the population is unknown and therefore needs to be estimated, which is denoted as $\widehat{CPUE_s}$. Given number of sets N_s and individual observations *i*, the estimated average CPUE ($\widehat{CPUE_s}$) can be calculated as:

$$\widehat{CPUE_s} = \frac{\sum_i^N C\,PUE_{i,s}}{N_s}$$

The question then is how many sets (*N)* need to be sampled (in this case, a proxy for observer coverage), are required for our estimate $\widehat{CPUE_s}$ to become sufficiently close to $CPUE_s$. The answer to this depends on the variability of $CPUE_s$. If $CPUE_s$ is constant, then one observation would suffice for $\widehat{CPUE_s}$ to equal $CPUE_s$. However, the more variable $CPUE_s$ is, the greater the number of observations will be needed in order for $\widehat{CPUE_s}$ to approach $CPUE_s$.

We quantified the variability in CPUE between the Class 6 and Class 1-5 vessels through three approaches. First, the coefficient of variation (CV) of set-level CPUE per bycatch species between the Class 6 and Class 1-5 observer programs were calculated and compared. This provides a measure of the similarity in total variability between the two programs. Second, quantile-quantile plots were then created comparing the set-level CPUE per bycatch species (scaled relative to the maximum CPUE) between the Class 6 and Class 1-5 observer programs. This provides a measure of the similarity in the shape of the distributions of set level CPUE between the Class 6 and Class 1-5 observer programs. However, distributions can be similar in shape but differently distributed in space, which could impact estimates of total catch based on observer data if observer coverage is systemically different in space between the Class 6 and Class 1-5 programs. Third, to account for this, we also calculated the Hellinger Distance of the spatial distributions of CPUE per bycatch species between the Class 6 and Class 1-5 programs. Hellinger Distance is a measure of the similarity of the spatial distribution between the two datasets. Hellinger distances are bounded between 0 and 1, where 0 means two spatial distributions are identical, and 1 means they are mutually exclusive (Wilson, 2011).

2.4 Simulated Alternative Observer Coverage Levels

Because the steps conducted in the previous sections showed the nature of variability in bycatch rates are roughly comparable between the Class 6 and Class 1-5 fleet segments, Class 6 observer data was used to simulate the effect of different observer coverage rates in space and time on the accuracy of estimated total catch of bycatch species. To accomplish this, two separate simulations were run. In the first, we randomly subsampled a given proportion (“random sampling scenario”), ranging from 5-95%, of Class 6 trips and classified these as “observed”. For each simulated set of observed trips, we calculated the mean catch per set $\widehat{CPUE_s}$ of the bycatch species from the observed subset, and then estimated the total catch of the bycatch species by multiplying $\widehat{CPUE_s}$ by the number of unobserved sets and summing with the total observed catch to provide an estimate of total catch which we then compared to the actual reported catch. This process was repeated 42 times per simulated observer coverage level (the percent of trips with an observer) to generate a Monte-Carlo distribution of total catch error rates per bycatch species.

However, observer coverage is not randomly assigned in the Class 1-5 data. Instead, observer coverage is concentrated among specific segments of the fleet, specifically being most common on Ecuadorian flagged vessels as a result of the voluntary TUNACONS observer program. To mimic this dynamic using the Class 6 data for the second experiment we limited the trips that could be sampled to only Ecuadorian flagged vessels (“Ecuador-only scenario”).

3. RESULTS

3.1 Exploratory Analyses

Over the study period sets were dominated by vessels in Classes 5 and 4, with some Class 3 sets and a very small number of sets by Class 1 and 2 vessels (Figure 1). Observer records were available for 38% of Class 1-5 sets (34% of trips) over the study period and were entirely from Ecuadorian flagged vessels in size classes 3-5. Among the Ecuadorian-flagged Class 1-5 vessels with an observer, observers recorded data on roughly 75% of Class 5 sets, 30% of Class 4 sets, and less than 5% of Class 3 vessels (Figure 2). No observer data were available for Class 1-2 vessels from any flag.

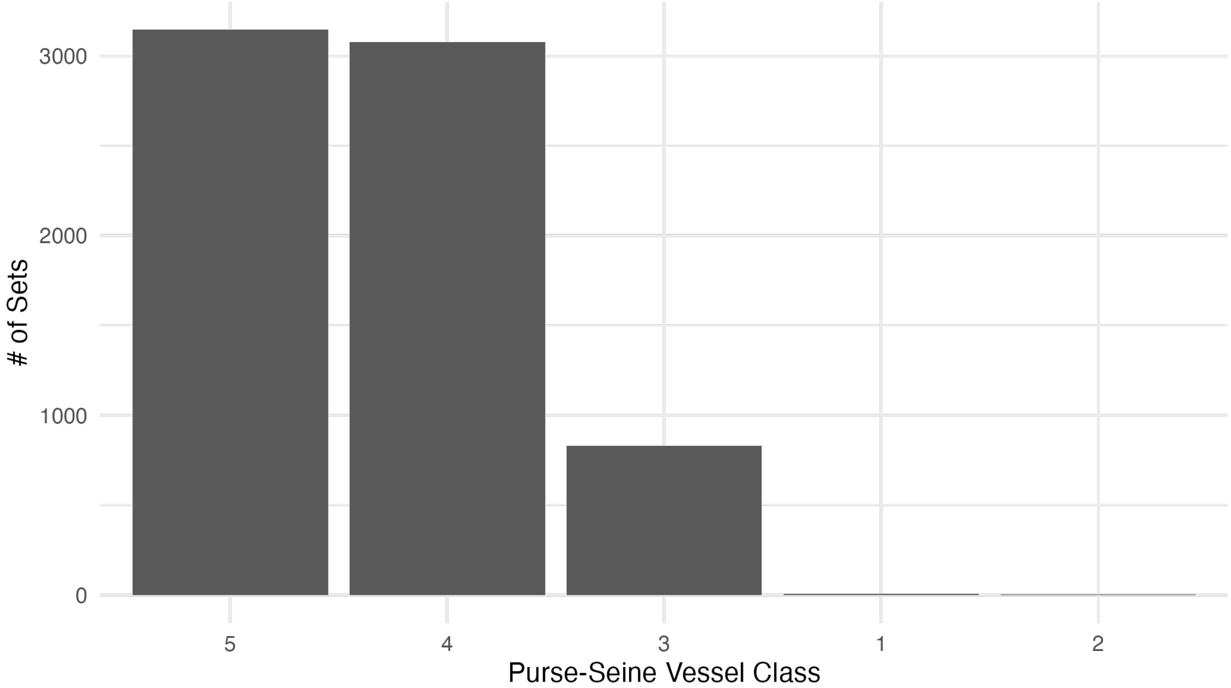

FIGURE 1. Total number of observed sets by vessel class category during the study period (2022-2024).
FIGURA 1. Número total de lances observados, por categoría de clase de capacidad de buque, durante el periodo de estudio (2022-2024).

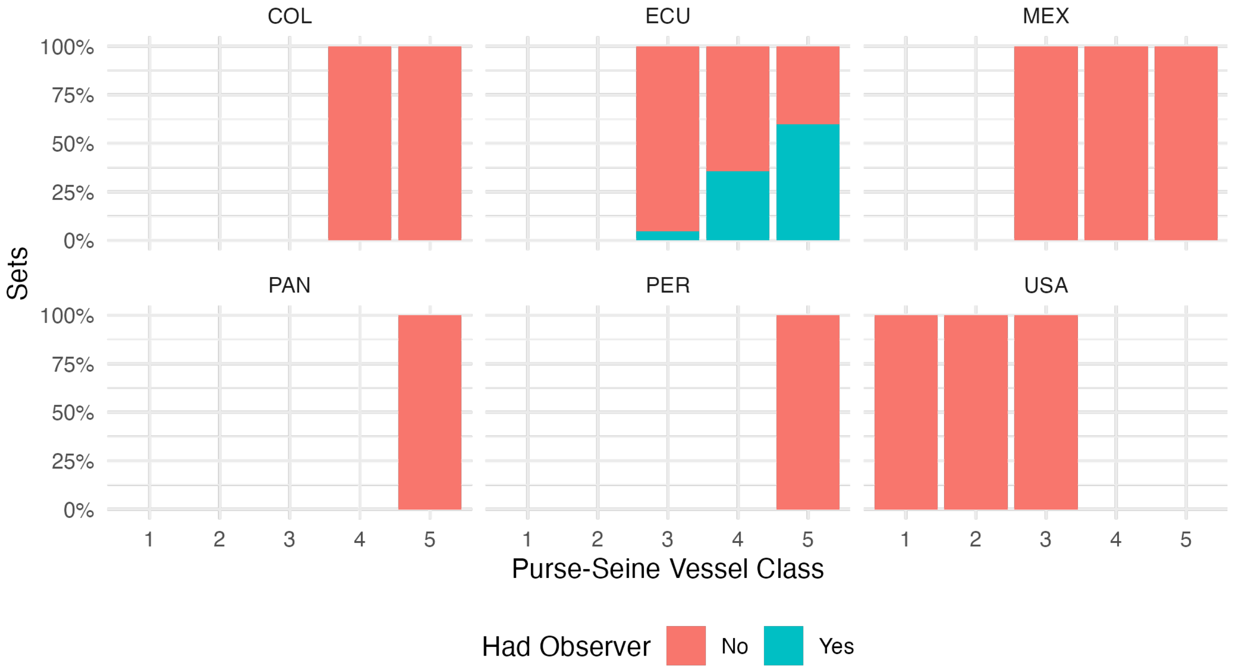

FIGURE 2. Proportion of Class 1-5 sets with observer coverage broken out by purse-seine vessel class and flag.
FIGURA 2. Proporción de lances de buques de clases 1-5 con cobertura por observadores, desglosada por clase de buque cerquero y pabellón.

Observer coverage was concentrated in the south and east of the Class 1-5 vessels fishing grounds. Coverage was in general lower closer to shore and was non-existent in the northwestern parts of the Class 1-5 vessels fishing grounds (note that relatively few sets occurred in this region). Class 1-5 observer and logbook data were contained within a polygon roughly north of 20 degrees south and 30 degrees north, and 75-120 degrees West (Figure 3).

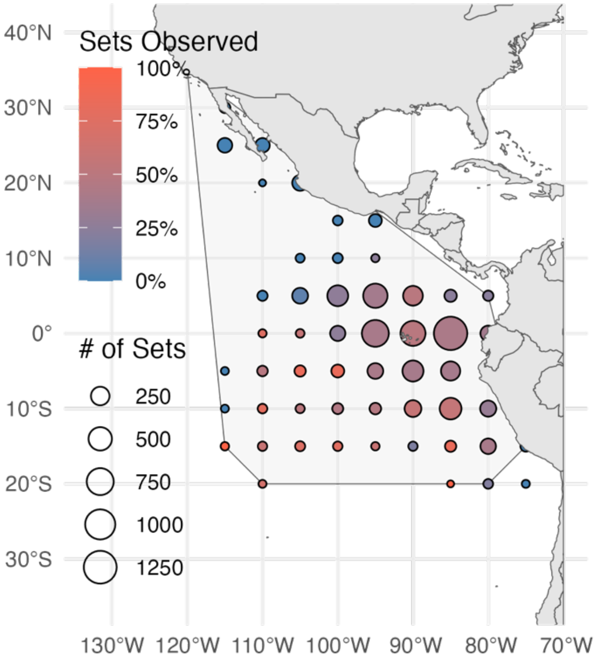

FIGURE 3. Proportion of Class 1-5 sets with observer data by 5-degree cell. Color indicates percent of sets within that cell with observer data, circle size is proportional to the number of sets in that cell over 2022-2024. Grey polygon indicates convex hull of locations with more than one set by a Class 1-5 vessel over the study period.

FIGURA 3. Proporción de lances de buques de clases 1-5 con datos de observadores por celda de 5 grados. El color indica el porcentaje de lances dentro de esa celda con datos de observadores, el tamaño del círculo es proporcional al número de lances en esa celda durante 2022-2024. El polígono gris indica la envolvente convexa de las ubicaciones con más de un lance por un buque de clases 1-5 durante el periodo de estudio.

3.2 Representativeness of Class 3-5 Observer Data for Tropical Tunas

Unlike BYC-10 INF-D, the Class 1-5 observer data was sufficiently representative of the catch rates of the tropical tunas to reconstruct the total retained logbook catches of tropical tuna in the Class 1-5 vessels (Figure 4). The mean absolute percent error in total annual catch across all species and vessel classes was 27%. If observed sets were systemically different than unobserved sets, higher error rates would be expected. This is somewhat evident in the case of the Bullet tuna (*Auxis rochei*); based on the observer data, the model predicted much lower catch rates of this species in the logbook data than were reported, in particular for the Class 5 vessels. This indicates that the observed Class 1-5 sets that reported catches of Bullet tuna are in some way unrepresentative of the logbook Class 1-5 sets that caught Bullet tuna.

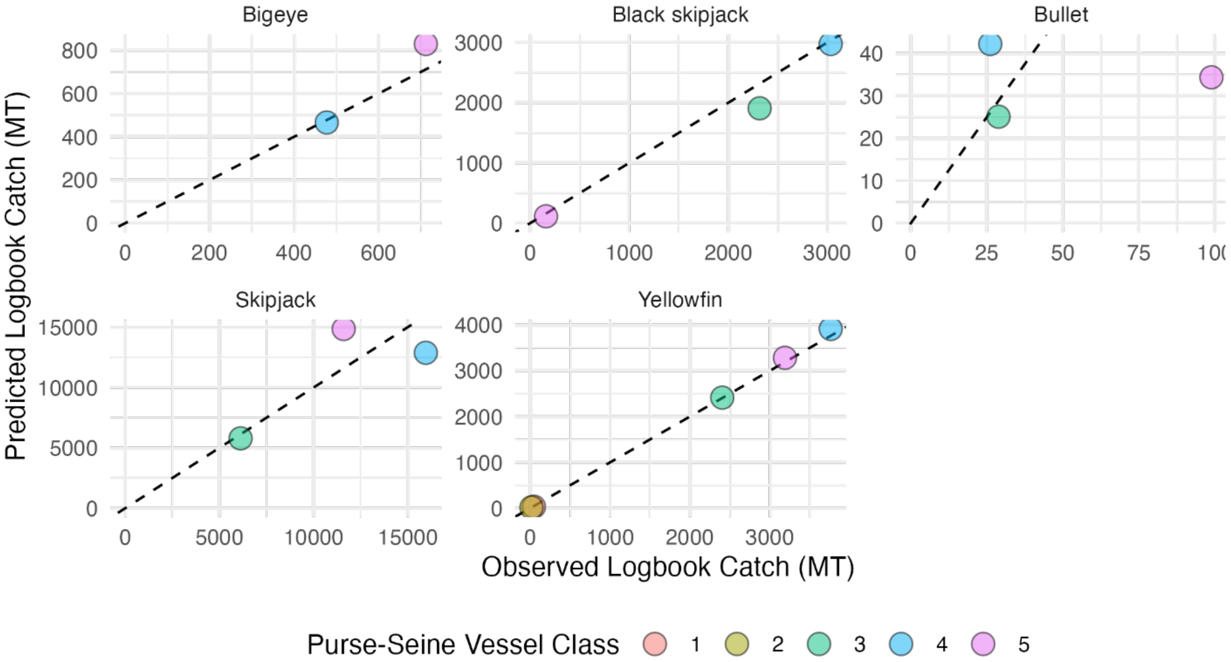

FIGURE 4. Total observed (x-axis) and predicted (y-axis) retained logbook catch of selected tuna species for Class 1-5 vessels disaggregated by vessel class.

FIGURA 4. Captura retenida total registrada en las bitácoras, observada (eje 'x') y predicha (eje 'y'), de especies seleccionadas de atunes para buques de clases 1-5, desglosada por clase de buque.

The potential uncertainty in the true CPUE of bycatch rates in the Class 1-5 fleet segment was quantified by bootstrapping the bycatch observations from the observed segment of the Class 1-5 vessels, and then calculating the mean absolute percent error (MAPE) between the bootstrapped database and the held-out samples (repeating the process 200 times).

The resulting MAPE values largely mirror the encounter rates of the species in question. For example, common dolphinfish was relatively commonly encountered and the bootstrapped MAPE values were low (less than 5%). Conversely, pelagic thresher sharks were only rarely observed in the observed Class 1-5 sets, and therefore the bootstrapped MAPE values were higher than 75%. While we cannot know if these error rates would truly apply to the unobserved Class 1-5 sets, this analysis does suggest that for many rarer bycatch species estimates of CPUE and subsequently total catch based on current Class 1-5 observer coverage may be somewhat imprecise (Figure 5).

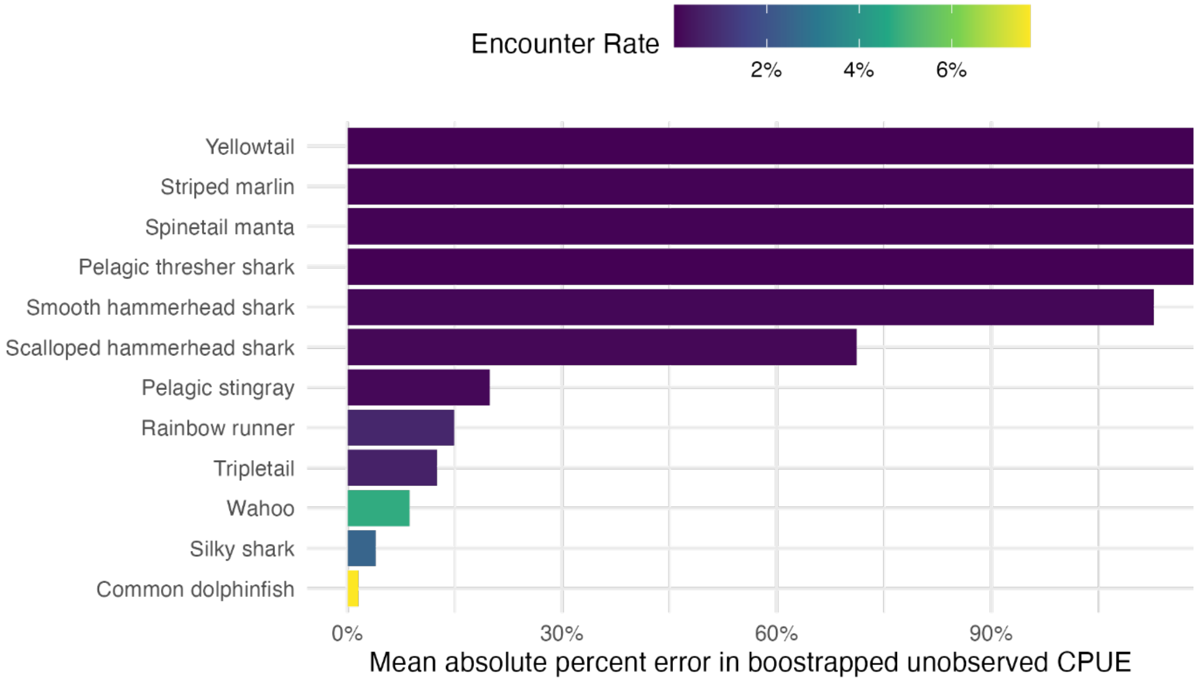

FIGURE 5. Mean absolute percent error in estimated unobserved CPUE resulting from applying CPUE from a simulated observed segment of sets to a simulated unobserved segment of sets.
FIGURA 5. Error porcentual absoluto medio en la CPUE no observada estimada resultante de aplicar la CPUE de un segmento de lances observado simulado a un segmento de lances no observado simulado.

Another basic consideration in determining observer coverage is the rarity of the species in question. Achieving accurate estimates of catches of rarely encountered species will require extremely high observer coverage levels. In the case of the species considered here, for the Class 6 observer data, the number of encounters is highly variable, ranging from thousands per year in the case of common dolphinfish, to ten or fewer encounters per year across all Class 6 sets in the case of bigeye thresher sharks (Figure 6).

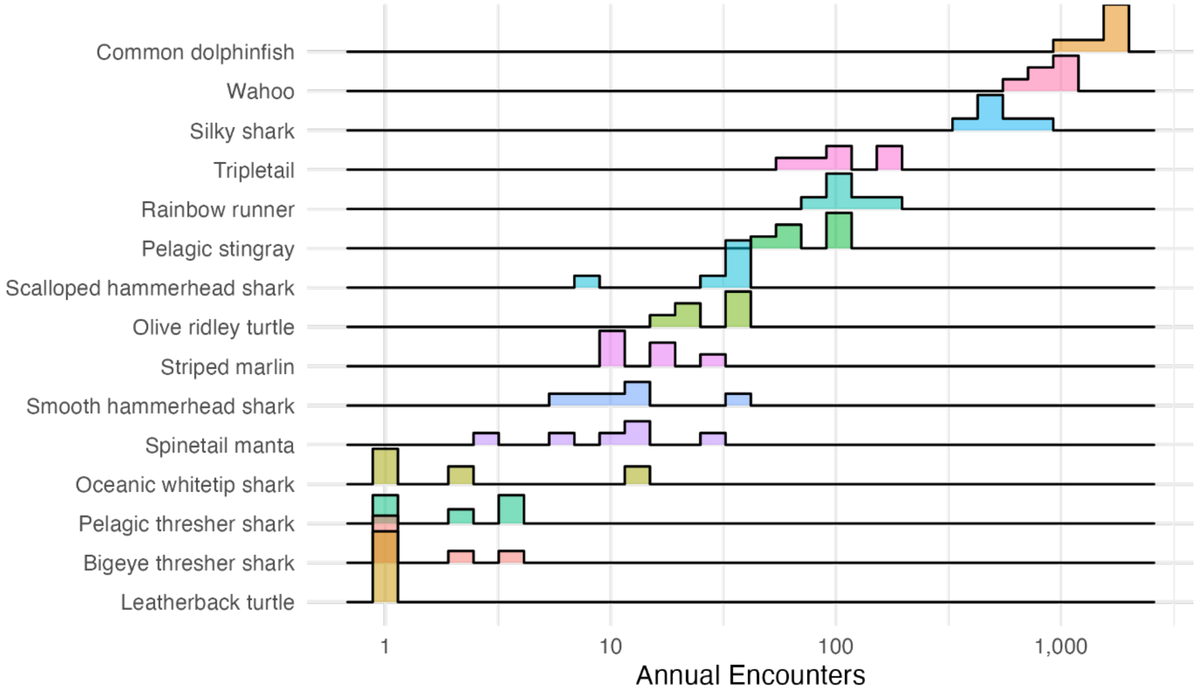

FIGURE 6. Distribution of number of encounter events (noting that an "encounter" simple means more than zero individuals were recorded) per year per species across Class 6 vessels. Note that the x-axis is on $\log_{10}$ scale.

FIGURA 6. Distribución del número de eventos de encuentro (teniendo en cuenta que un "encuentro" significa simplemente que se registraron más de cero individuos) por año y por especie en buques de clase 6. Nótese que el eje 'x' está en escala $\log_{10}$.

3.3 Comparison of Class 6 and Class 1-5 Bycatch Rates

Across all key bycatch species shared between the Class 1-5 and Class 6 vessels with observer data, the average CV of set-and-species-level CPUE was 10.3 for Class 6 vessels, and 11.31 for class 1-5 vessels. Across all species considered in the analysis, the correlation between the Class 6 and Class 1-5 CV was 0.97; while the species-level variation was somewhat different in some cases, in general they were reasonably similar between the two datasets (except for Class 3 vessels and billfish and sharks) (Figure 7).

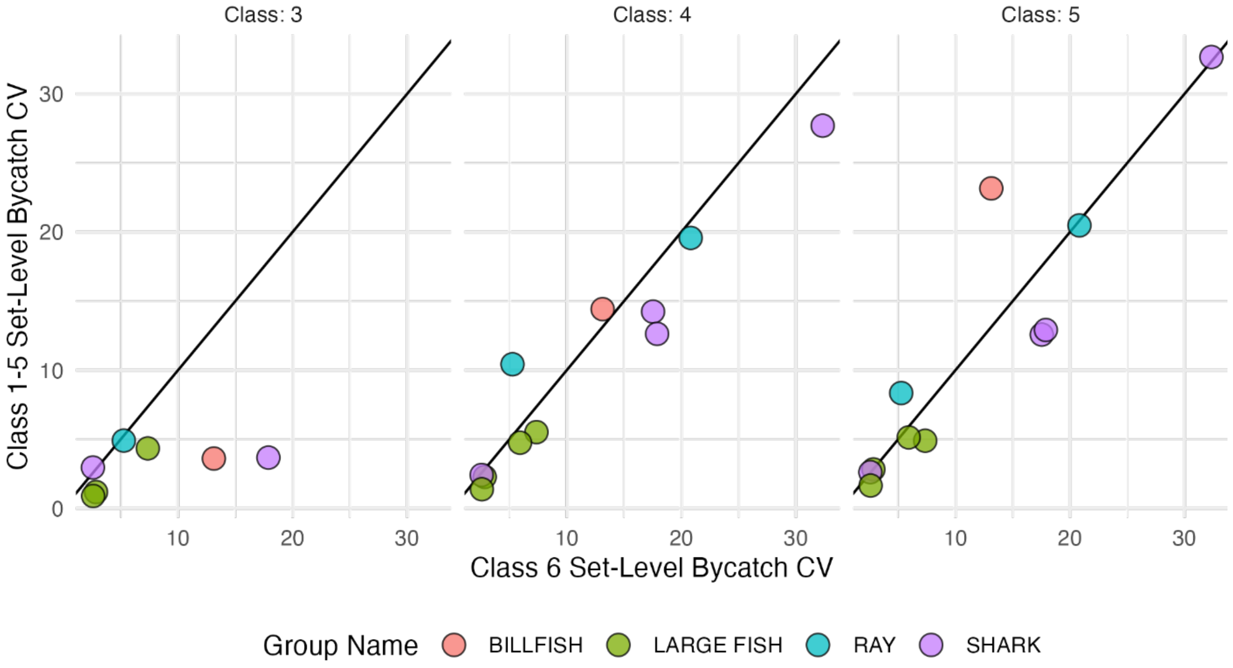

FIGURE 7. Set-level coefficient of variation (CV) in CPUE of individual bycatch species groups by vessel class size between the Class 6 and Class 1-5 observer programs. Note that turtles are omitted as they are not reported in the Class 1-5 observer data used in this report.

FIGURA 7. Coeficiente de variación (CV) por lance en la CPUE de grupos de especies de captura incidental, por clase de capacidad de los buques, entre los programas de observadores de buques de clase 6 y de clases 1-5. Cabe señalar que no se incluyen las tortugas, ya que no figuran en los datos de observadores de buques de clases 1-5 utilizados en este informe.

The distribution of CPUEs per species were relatively similar between the Class 6 and Class 1-5 observer programs in most cases (Figure 8). Examining spatial differences, Hellinger distances ranged from roughly 0.25 (pelagic stingray) to slightly over 0.5 (tripletail). Hellinger distance values above 0.5 indicate that the two distributions are more dissimilar in space than they are similar. These results indicate that for many species the CPUE intensity is similarly distributed in space, but not all (Figure 9), a result which mirrors that of the quantile-quantile plot analysis (Figure 8).

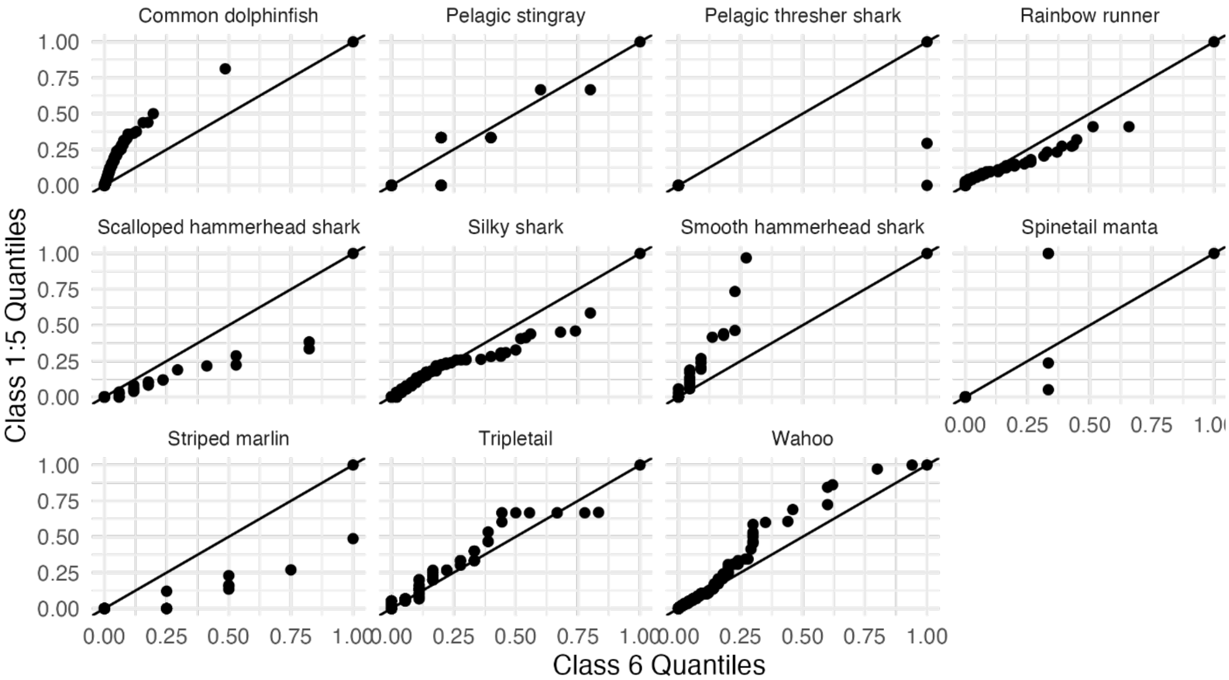

FIGURE 8. Quantile-quantile plot of Class 6 and Class 1-5 bycatch CPUE, each normalized to a zero to one interval. Note that turtles are omitted as they are not reported in the Class 1-5 observer data used in this report.

FIGURA 8. Gráfica cuantil-cuantil de la CPUE de captura incidental de buques de clase 6 y de clases 1-5, cada una normalizada a un intervalo de cero a uno. Cabe señalar que no se incluyen las tortugas, ya que no figuran en los datos de observadores de buques de clases 1-5 utilizados en este informe.

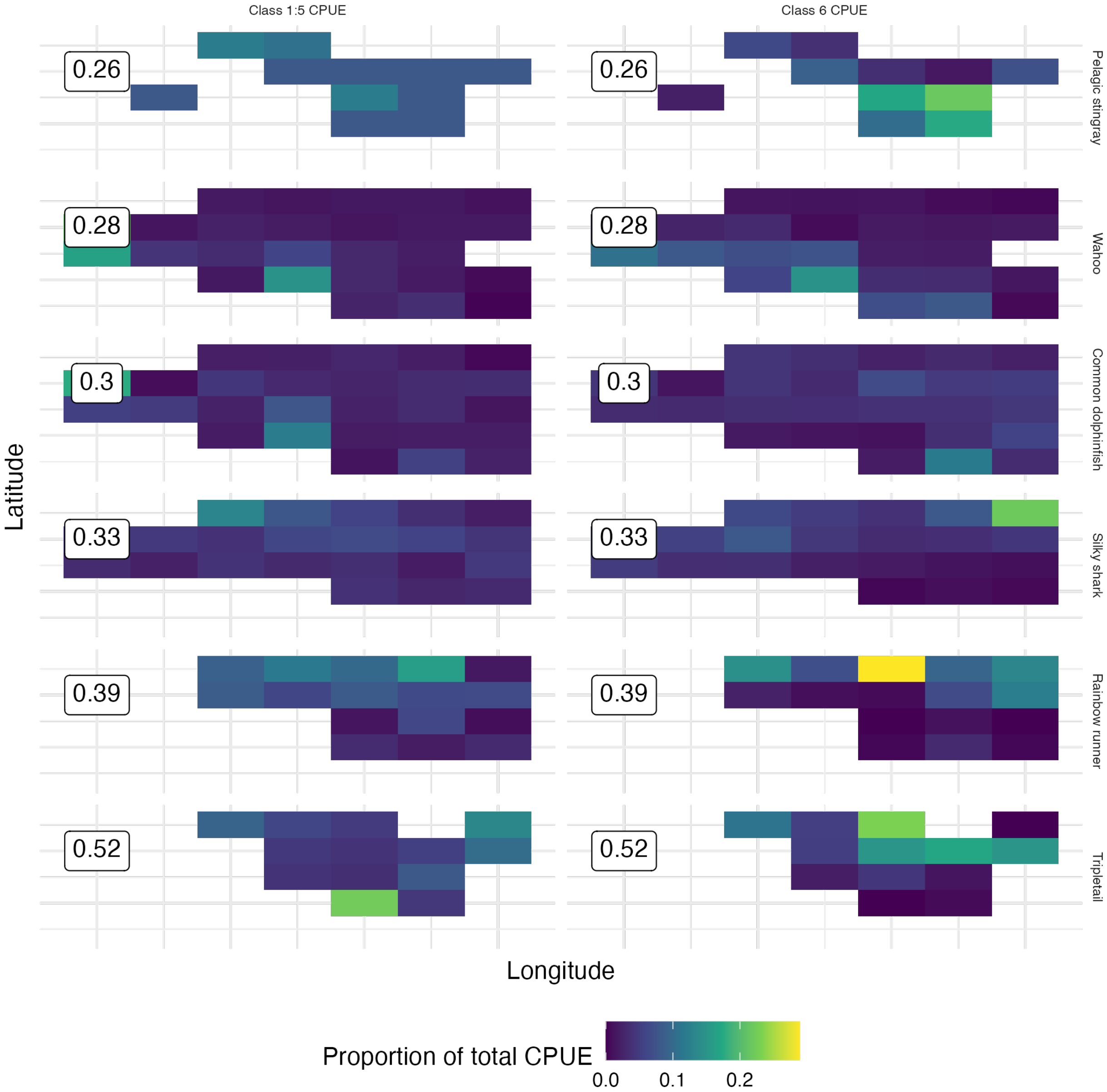

FIGURE 9. Distribution of CPUE in space for individual species broken out by Class 6 (right column) and Class 1-5 (left columns) vessels. Text indicates Hellinger distance in the spatial distribution of CPUE between the Class 6 and Class 1-5 vessels. Plotted species represent pairs for which Hellinger distances could be estimated for both vessel class groups.

FIGURA 9. Distribución de la CPUE en el espacio para especies individuales, desglosada por buques de clase 6 (columna derecha) y de clases 1-5 (columnas izquierdas). El texto indica la distancia de Hellinger en la distribución espacial de la CPUE entre los buques de clase 6 y de clases 1-5. Las especies representadas en la gráfica son pares para los que se pudieron estimar las distancias de Hellinger para ambos grupos de clases de buques.

In total, the key question in determining whether the Class 6 observer data can provide insights on observer coverage requirements for Class 1-5 vessels is to what extent they have similar variability. Our analysis reveals that variability of bycatch rates share many similarities between the Class 1-5 and Class 6 observer programs. While some species have different distributions between the two datasets, these differences are in general not severe enough to prevent broad insights gained from evaluation of the Class 6 observer data to be potentially applicable to the Class 1-5 cases.

3.4 Simulated Alternative Observer Coverage Rates

The analysis in Section 3.2 assessing the representativeness of Class 1-5 observer data for tropical tunas as a proxy for bycatch species was inconclusive as to the utility of Class 1-5 observer coverage for bycatch estimation. While the results shown in Figure 7 indicate that the total unobserved Class 1-5 logbook catch of tropical tunas can be estimated reasonably well based on the catch rates of the Class 1-5 observer data, that does not mean that the current observer coverage for the Class 1-5 vessels would be suitable for all key bycatch species. This is because individual bycatch species may have different spatio-temporal dynamics and interactions with fishing gear than the tropical tunas, meaning that a program that may be representative for tropical tunas may not be representative for some bycatch species. As such and given the relatively similar nature of variability seen between the Class 6 and Class 1-5 vessels bycatch CPUE, we turned to the Class 6 data with 100% observer coverage as a proxy, testing the effects of simulated reductions in observer coverage on the accuracy of total catch of various species of interest.

The results of this exercise were visualized by plotting the distribution of Monte Carlo draws for each species. Observing 34% of the Class 6 trips produced a median MAPE of roughly 25% across all considered bycatch species, with some species having MAPE values greater than 100% (e.g., thresher sharks) and others as low as roughly 10% (e.g., silky shark). Increasing observer coverage beyond 34% of trips decreased the median and maximum MAPE, but slowly, with observer coverage of 85% needed for all species to have a MAPE less than 50% (Figure 10).

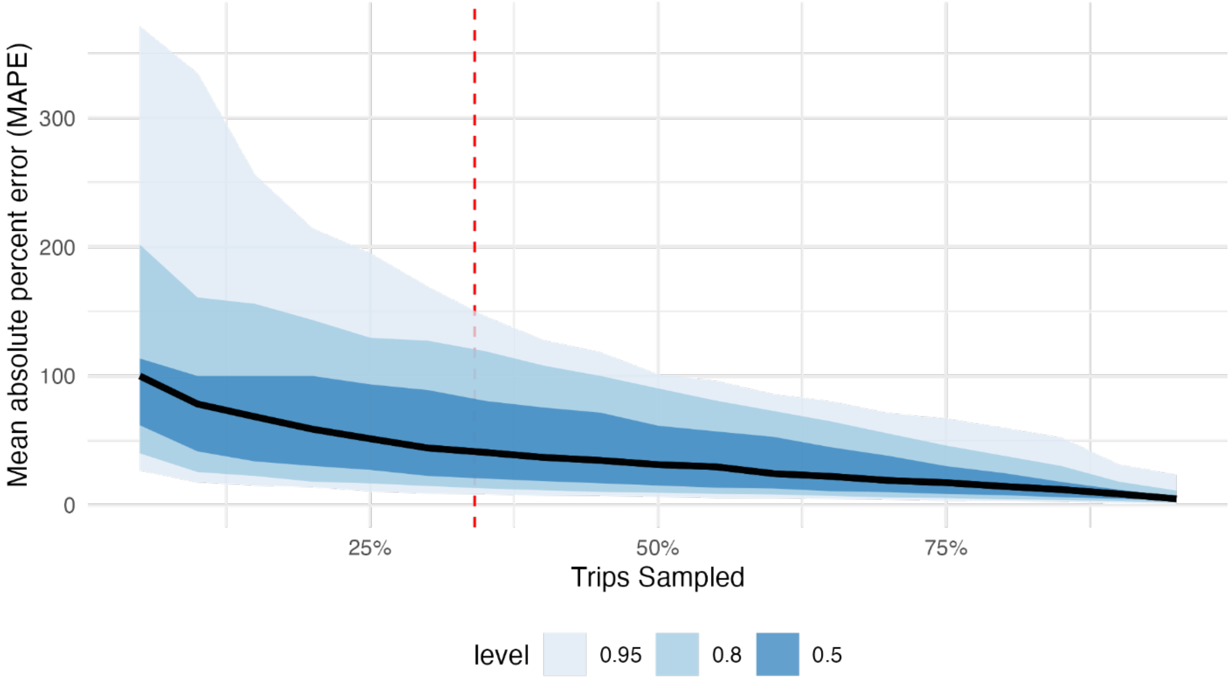

FIGURE 10. Distribution of estimated mean absolute percent error (MAPE) in selected Class 6 vessels catch across all Monte Carlo draws, study years, and species included in the study. Vertical dashed line is the percent of Class 1-5 trips with observer data in the subset of data used in this study. The simulation assumes all sets are observed on observed trips.
FIGURA 10. Distribución del error porcentual absoluto medio (MAPE) estimado en capturas de buques de clase 6 seleccionados en todos los sorteos de Montecarlo, años de estudio y especies incluidas en el estudio. La línea vertical discontinua es el porcentaje de viajes de buques de clases 1-5 con datos de observadores en el subconjunto de datos utilizados en este estudio. La simulación supone que todos los lances se observan en viajes observados.

Breaking the totals out by the species considered in this study shows a wide range of MAPE values (Figure 11). Some common species such as common dolphinfish, wahoo or silky sharks were well represented even with 34% observer coverage of trips, whereas other rarer species like bigeye thresher sharks required coverage levels over 80% for MAPE to approach 25%. Total catch estimates were on average unbiased, consistent with random sampling design (Figure 12).

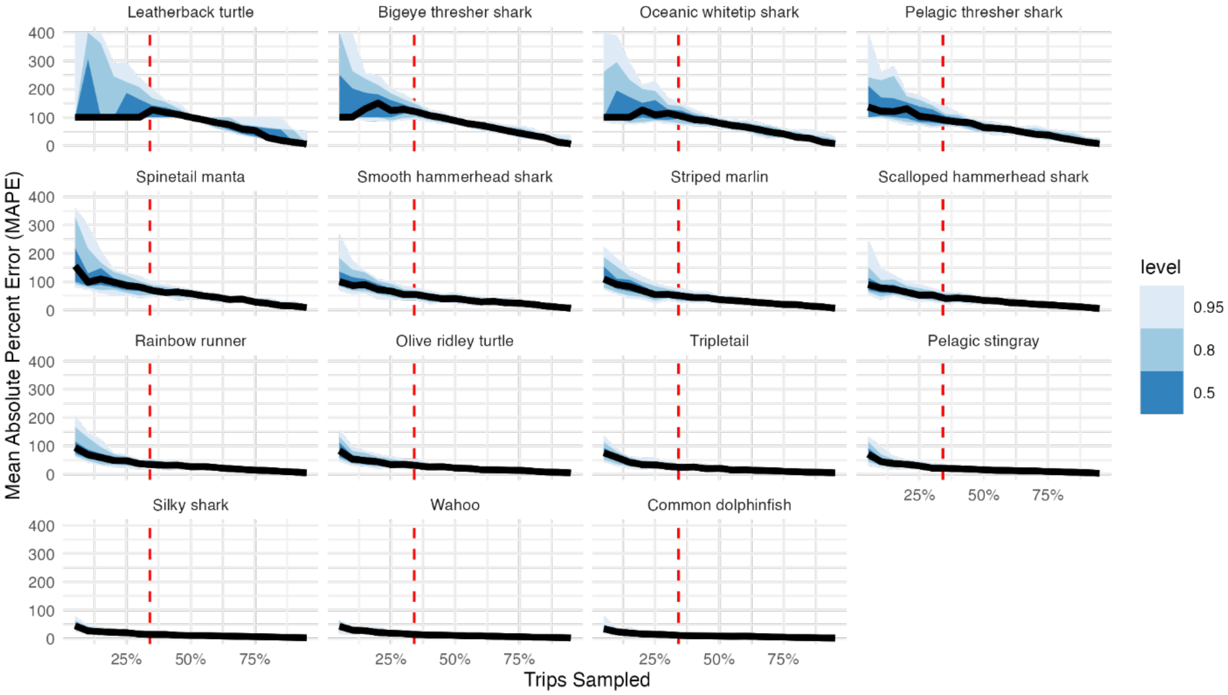

FIGURE 11. Distribution of mean absolute percent error (MAPE) in total catch across all study years across Monte Carlo draws and species, disaggregated by individual species. Y-axis values capped at 400%. Vertical dashed line is the percent of Class 1-5 trips with observer data in subset of data used in this study. Simulation assumes all sets are observed on observed trips.

FIGURA 11. Distribución del error porcentual absoluto medio (MAPE) en la captura total a lo largo de todos los años de estudio en todos los sorteos de Montecarlo y especies, desglosado por especies individuales. Los valores del eje 'y' tienen un límite de 400%. La línea vertical discontinua es el porcentaje de viajes de buques de clases 1-5 con datos de observadores en el subconjunto de datos utilizados en este estudio. La simulación supone que todos los lances se observan en viajes observados.

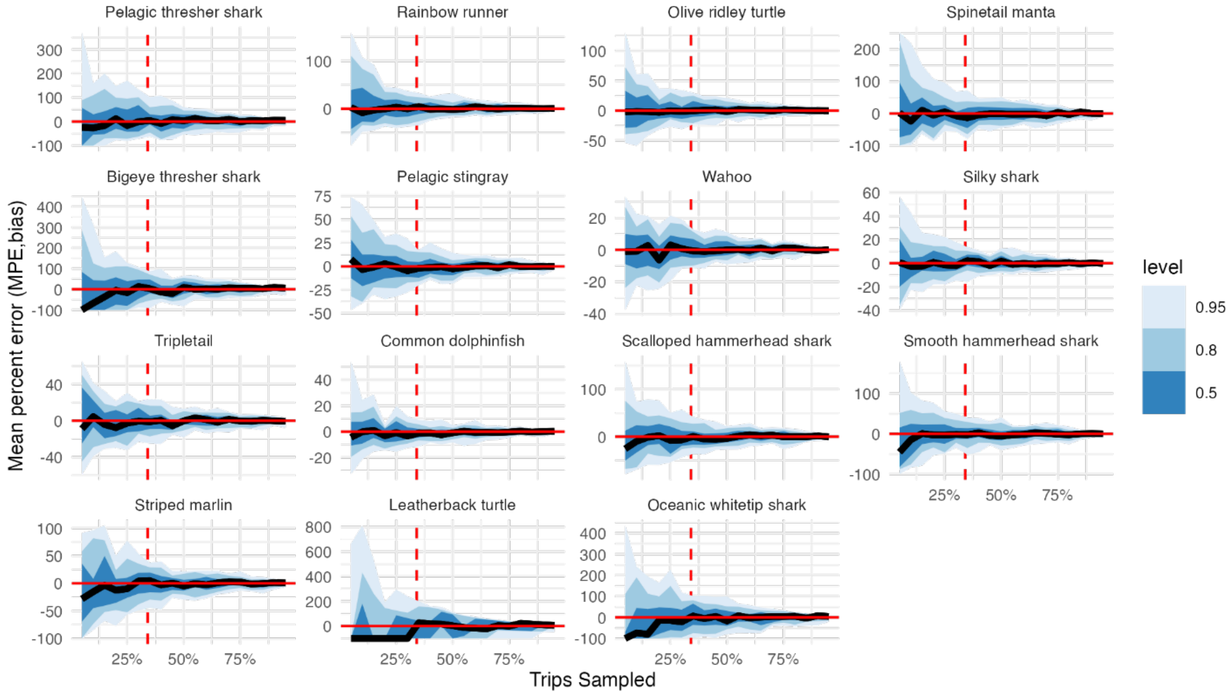

FIGURE 12. Distribution of mean percent error (bias) in total catch across all study years across Monte Carlo draws and species, disaggregated by individual species. Vertical dashed line is the percent of Class 1-5 trips with observer data in subset of data used in this study. Simulation assumes all sets are observed on observed trips.

FIGURA 12. Distribución del error porcentual medio (sesgo) en la captura total a lo largo de todos los años de estudio en todos los sorteos de Montecarlo y especies, desglosado por especies individuales. La línea vertical discontinua es el porcentaje de viajes de buques de clases 1-5 con datos de observadores en el subconjunto de datos utilizados en este estudio. La simulación supone que todos los lances se observan en viajes observados.

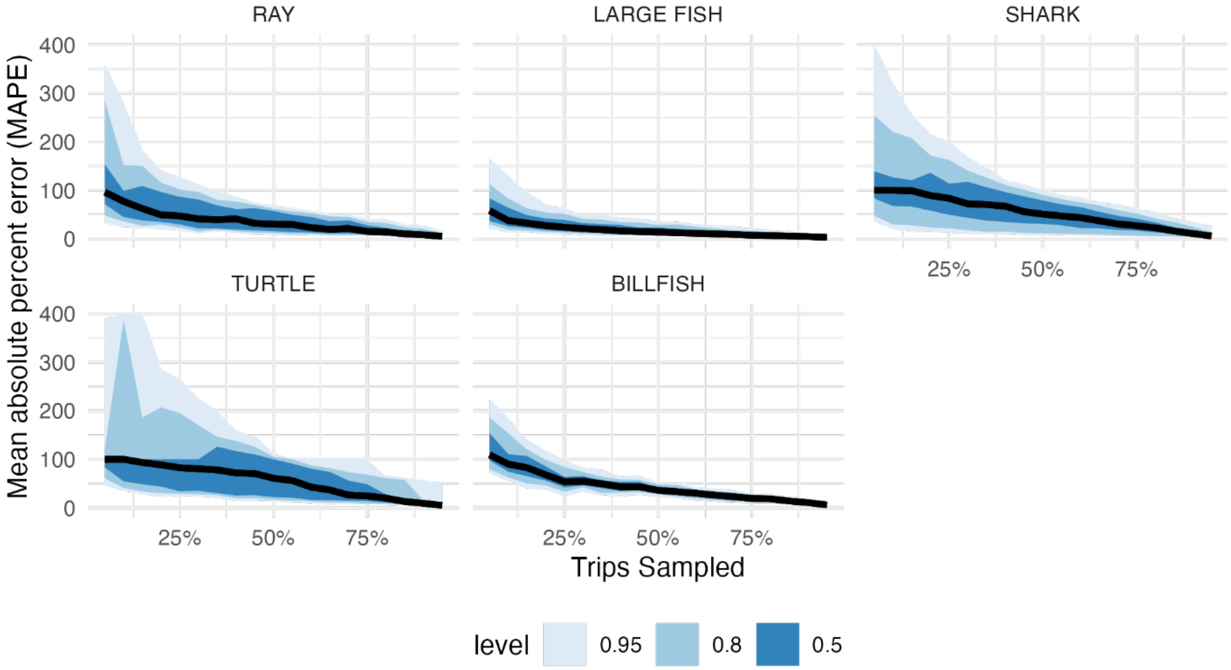

FIGURE 13. Distribution of mean absolute percent error (MAPE) in total catch across all study years for all Monte Carlo draws and species, separated by broad taxonomic group. Y-axis capped at 400%.
FIGURA 13. Distribución del error porcentual absoluto medio (MAPE) en la captura total a lo largo de todos los años de estudio en todos los sorteos de Montecarlo y especies, separado por grupo taxonómico amplio. Los valores del eje 'y' tienen un límite de 400%.

3.5 Ecuador-Flagged Vessels Only Case Study

The previous simulation study presents an optimistic case where Class 6 vessels trips were randomly selected for observation. As an alternative, we repeated our simulated observer coverage exercise but limited observed trips to only Class-6 Ecuadorian flagged vessels, to mimic the dynamics of the current Class 1-5 observer program.

Observing only Ecuadorian vessels increased the MAPE and bias for some species. Median MAPE values were similar across the two exercises (random sampling of all trips versus random sampling of Ecuadorian vessels trips only), reaching quite low values at 34% of trips observed. However MAPE values in total catch for some species remained at or above 50% even in the case where the simulation observed nearly 100% of Ecuadorian Class-6 vessels trips in the database (Figure 14). This occurred due to the introduction of systemic sampling biases between the observed and unobserved trips (e.g., lack of spatial overlap between the observed and unobserved trips). For example, even sampling 95% of Ecuadorian vessels trips resulted in underestimating total silky shark catch by over 10%, or bigeye thresher catch by 75% (Figure 16).

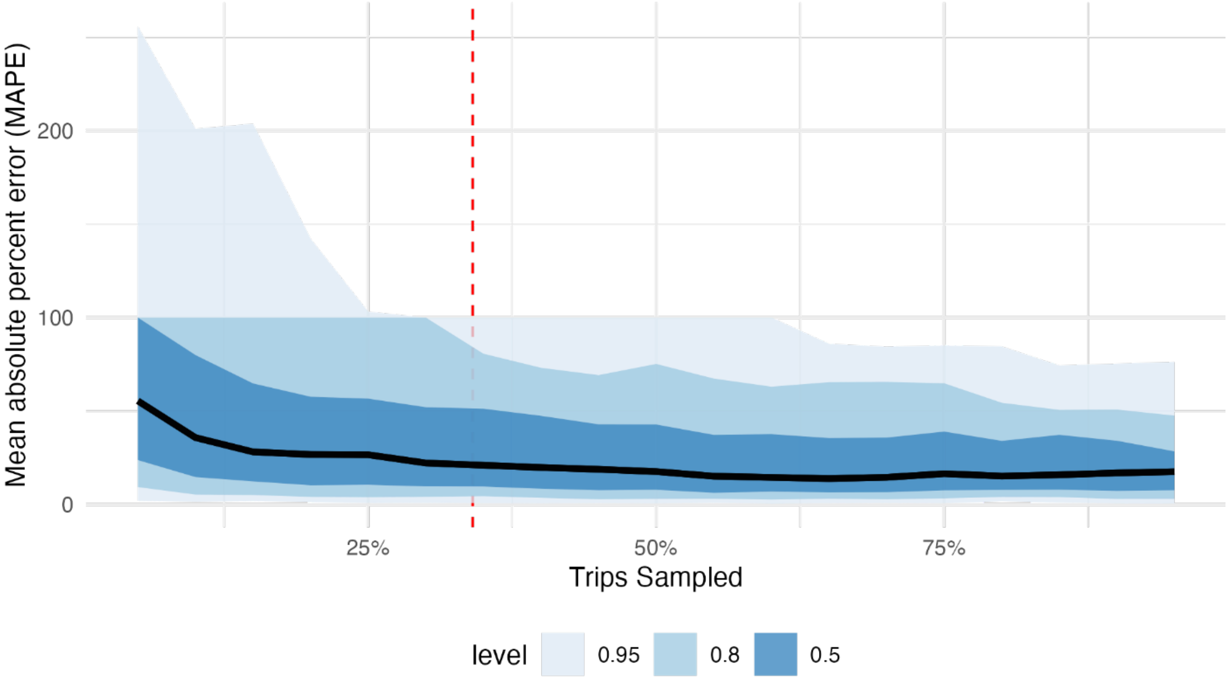

FIGURE 14. Distribution of mean absolute percent error (MAPE) in total catch across all study years for all Monte Carlo draws and species for a simulated case study where only Ecuadorian flagged vessels trips are observed. Vertical dashed line is the percent of Class 1-5 trips with observer data in subset of data used in this study. Simulation assumes all sets are observed on observed trips.

FIGURA 14. Distribución del error porcentual absoluto medio (MAPE) en la captura total a lo largo de todos los años de estudio en todos los sorteos de Montecarlo para un estudio de caso simulado en el que solo se observan viajes de buques de pabellón ecuatoriano. La línea vertical discontinua es el porcentaje de viajes de buques de clases 1-5 con datos de observadores en el subconjunto de datos utilizados en este estudio. La simulación supone que todos los lances se observan en viajes observados.

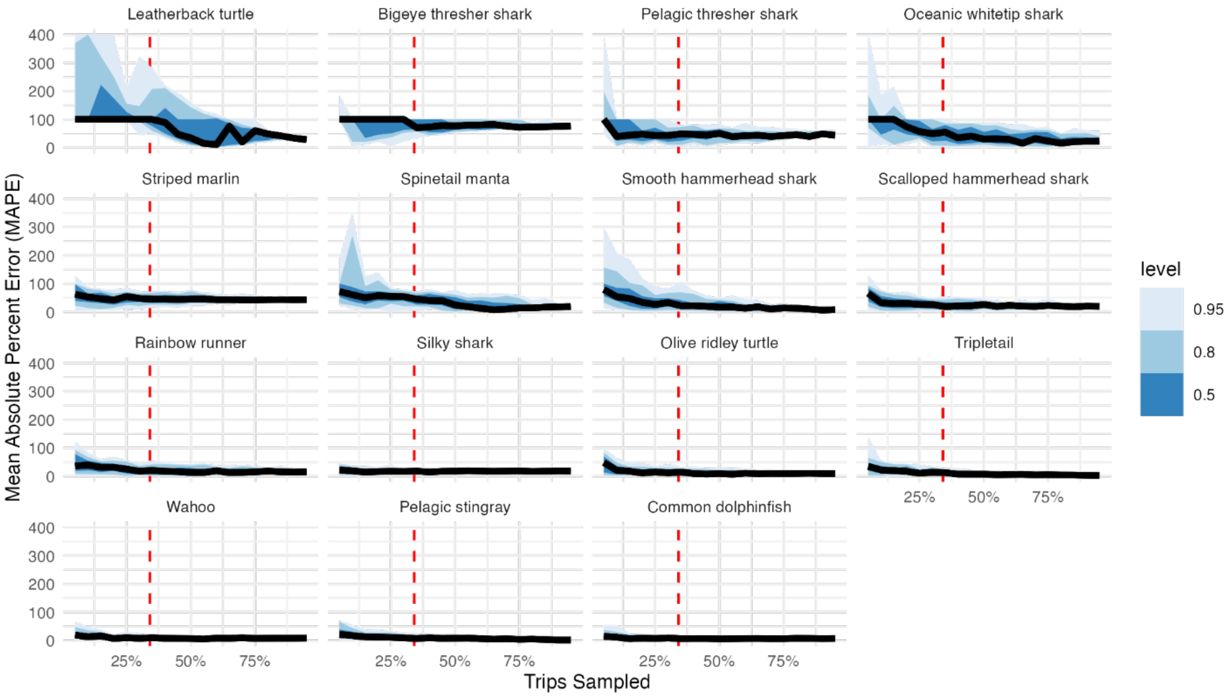

FIGURE 15. Distribution of mean absolute percent error (MAPE) in total catch across all study years for all Monte Carlo draws and species, disaggregated by individual species for a simulated case study where only Ecuadorian flagged vessel trips are observed. Y-axis values capped at 400%. Vertical dashed line is the percent of Class 1-5 trips with observer data in subset of data used in this study. Simulation assumes all sets are observed on observed trips.

FIGURA 15. Distribución del error porcentual absoluto medio (MAPE) en la captura total a lo largo de todos los años de estudio en todos los sorteos de Montecarlo y especies, desglosado por especies individuales, para un estudio de caso simulado en el que solo se observan viajes de buques de pabellón ecuatoriano. Los valores del eje 'y' tienen un límite de 400%. La línea vertical discontinua es el porcentaje de viajes de buques de clases 1-5 con datos de observadores en el subconjunto de datos utilizados en este estudio. La simulación supone que todos los lances se observan en viajes observados.

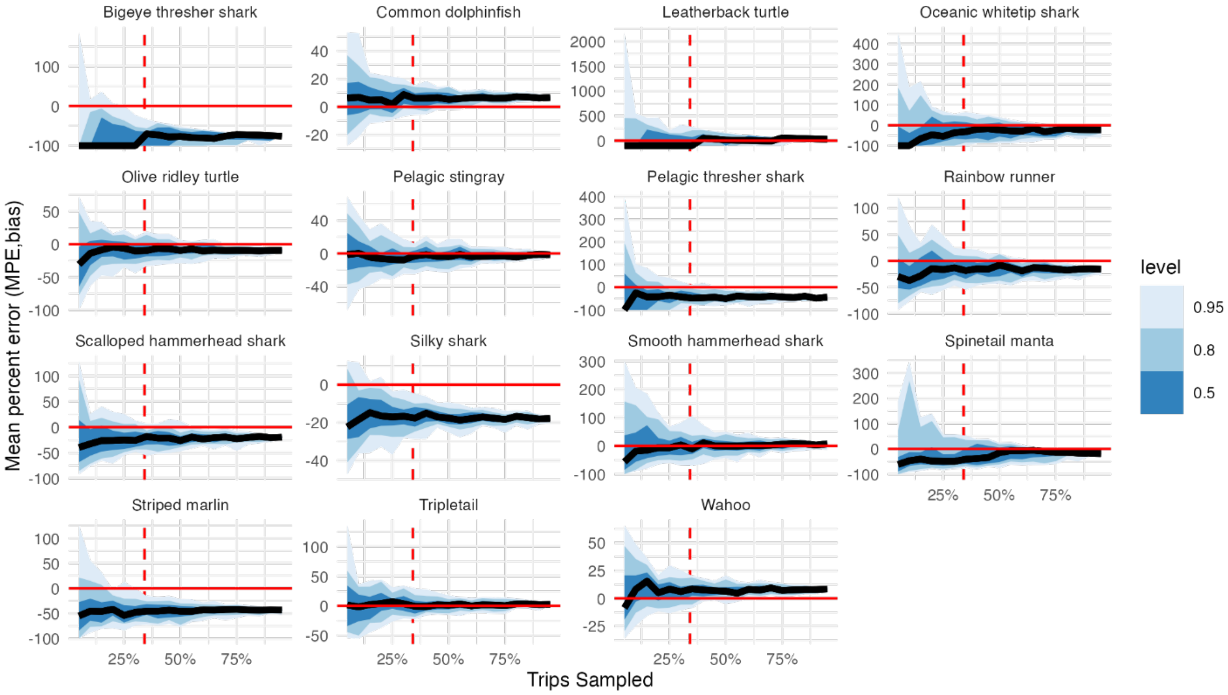

FIGURE 16. Distribution of mean percent error (bias) in total catch across all study years for all Monte Carlo draws and species, disaggregated by individual species for a simulated case study where only Ecuadorian flagged vessels trips are observed. Vertical dashed line is the percent of Class 1-5 trips with observer data in the subset of data used in this study. Simulation assumes all sets are observed on observed trips.

FIGURA 16. Distribución del error porcentual medio (sesgo) en la captura total a lo largo de todos los años de estudio en todos los sorteos de Montecarlo y especies, desglosado por especies individuales, para un estudio de caso simulado en el que solo se observan viajes de buques de pabellón ecuatoriano. La línea vertical discontinua es el porcentaje de viajes de buques de clases 1-5 con datos de observadores en el subconjunto de datos utilizados en este estudio. La simulación supone que todos los lances se observan en viajes observados.

4. DISCUSSION

Efficient sampling designs seek to maximize accuracy and precision of desired outcomes through the implementation of logistically practical yet statistically rigorous protocols, all while minimizing costs. The severity of the tradeoff between these objectives will depend in large part on the costs of sampling and the variability of the process being measured, as well as the extent of fleet involvement in trip selection process. This study sought to provide guidance on observer coverage levels and sampling design requirements for the Class 1-5 vessels in order to estimate total catch of key bycatch species from Class 1-5 vessels.

Unlike the logbook analysis conducted in BYC-10 INF-D, our analysis suggests that the current observer coverage in the Class 1-5 purse seine vessels may be sufficiently representative to estimate the retained tuna catches reported in the logbooks from this group. However, this does not imply that the same would apply for bycatch species. Bycatch species may have very different spatio-temporal distributions than tropical tunas, and it is possible that, for example, a given bycatch species may be mostly caught by Class 1-2 vessels despite those vessels contributing only a small number of sets to the total in the region (Figure 6).

Since logbook data on bycatch from unobserved Class 1-5 sets are not available, is not possible to directly quantify differences between observed and logbook bycatch rates among this fleet segment. Moreover, simulating higher Class 1-5 observer coverage levels than currently exist cannot easily be done. However, Class 6 observer program data with 100% observer coverage can be used as a proxy to address these questions. Based on Class 6 observer data, for many species, observer rates near 30% are sufficient to produce relatively low (25% or less) MAPE in total catch (Figure 11). However, obtaining low MAPE values for all species requires much higher observer coverage rates, and in cases where systemic differences exist between the nature of observed and unobserved fishing operations (i.e., the "Ecuador-flagged vessels only scenario", Figure 15), high error and biases in total catch persist even at high coverage levels. In fact, it is reasonable to believe that the current nature of observer coverage in the Class 1-5 vessels is closer to the Ecuador-only scenario than the randomly-assigned scenario.

As mentioned above, the results of the present study suggest that current observer coverage rates of the Class 1-5 vessels may be sufficient for estimating total catch for many species, but substantial increases in both amount and representativeness may be needed if the goal is to have accurate estimates for all of the bycatch species considered in this study, particularly for species that are sometimes only encountered a few times per year. Taking the Ecuador-only scenario as being most representative of nature of observer coverage in the Class 1-5 vessels, current rates of roughly 34% of trips (38% of sets) may be sufficient to provide accurate estimates of total catch for many important and common bycatch species (e.g., dorado, silky shark, olive ridley turtle). However, higher observer coverage levels may be needed for rare key bycatch species such as oceanic whitetip shark; and for other rare species such as bigeye thresher shark, our analysis suggests that improved estimates of total catch would have to come from expanding the representativeness of the Class 1-5 observer programs, as observing even 100% of the Class 6 Ecuadorian trips in the database still produced imprecise and biased estimates of total catch for this species. Given the clear potential for biased estimates of total bycatch removals when observer coverage is biased to a particular segment of the fleet, efforts to increase observer coverage into areas of the Class 1-5 vessels fishing grounds not currently well represented by the existing Class 1-5 observer programs are desirable.

It is also important to note that the results presented in this study stem from very optimistic scenarios in which all sets on a sampled trip are observed, and within the subset of candidate trips (in this case, all trips or Ecuador-only trips) trips are selected at random. Deviations from these optimistic assumptions would likely increase the error and bias in bycatch estimates from a given level of observer coverage. The purpose of this study was not to precisely estimate the error rates of existing programs, but to provide general recommendations on whether the current program is likely to be sufficiently representative to estimate bycatch of selected species, and an initial estimate of the potential observer coverage levels that might be needed to improve estimates for these bycatch species.

It is important to note that the definition of "representativeness" depends on the objectives of the observer program. A sampling program could be considered "representative" if the relevant attributes (e.g. catch per set) of the sampled observations are sufficiently similar to those same attributes in an unsampled set of observations. Taking the "Ecuador-Flagged Vessels Only" experiment as an example, such a program would be unrepresentative of the full spatial and temporal extent of Class 1-5 fishing operations, but regardless may be sufficiently representative of some species (e.g. Smooth hammerhead sharks), while being unrepresentative for other species (e.g. Bigeye thresher shark) (Figure 16). As such, if the goal is for an observer program to provide an unbiased estimate of total captures for all the species considered in this report, then a baseline definition of "representative" might be "distributed randomly across the full spatial and temporal extent of Class 1-5 fishing operations", simulated results of which are shown Figure 12. However, different definitions of representivity might exist for different objectives of the observer program.

Purely random sampling of all possible sets is the more reliable way of ensuring a representative sample, but is not always possible or efficient. Future simulation work could test various options for more realistic probability sampling designs to determine how the necessary coverage might be achieved through protocols that are adapted to logistical constraints that might make random selection of trips challenging. For example, sampling protocols that involved systematic sampling of trips might be considered. Additionally, if bycatch estimates may be required for areas where species are particularly vulnerable or occurring, or different coverage levels are required for different vessel size classes, advantages and disadvantages of stratified sampling designs, in terms of statistical performance and logistical complexity, could be evaluated.

The primary information needed to provide more specific guidance on potential coverage levels would be a list of acceptable error rates for specific species or groups of species. For example, clarity on whether the goal of the program is that no individual species have a catch MAPE greater than 25%, 50%, or any other value, or whether the goal should apply to taxonomic groups rather than individual species. These guidelines should also consider the realities of sampling rare species, some of which are of interest for the Commission as per existing Resolutions (e.g., C-19-04, C-24-05). Some bycatch species considered here were quite commonly in present in the Class 6 vessels observer data, for example dorado. However, many other species were encountered extremely rarely. For example, bigeye thresher sharks were often reported less than 10 times in a year throughout the Class 6 vessels observer data (Figure 6). The reality is that obtaining accurate total catch numbers for these rarely encountered species in the Class 1-5 vessels observer data would likely require near 100% observer coverage of some kind (i.e., a mix of human and electronic monitoring), which may deem impractical and logistically challenging.

5. REFERENCES

Wilson, P.D., 2011. Distance-based methods for the analysis of maps produced by species distribution models. Methods in Ecology and Evolution 2, 623–633. https://doi.org/10.1111/j.2041-210X.2011.00115.x