INTER-AMERICAN TROPICAL TUNA COMMISSION

SCIENTIFIC ADVISORY COMMITTEE

16TH MEETING

La Jolla, California (USA)
02-06 June 2025

DOCUMENT SAC-16 INF-J

AN EVALUATION OF THE PROPOSED IPSP PROBABILITY SAMPLING PROTOCOL FOR COLLECTION OF PURSE-SEINE PORT-SAMPLING DATA

Cleridy E. Lennert-Cody, Mark N. Maunder, Cristina De La Cadena, Luis Chompoy and Alexandre Aires-da-Silva

SUMMARY

A simulation study using 2022 - 2023 observer data was conducted to evaluate the performance of the proposed port-sampling protocol of the Integrated Port-Sampling Program (IPSP) for the Class-6 purse-seine vessel fleet (SAC-16-05). The proposed IPSP sampling protocol involves sampling 450 trips and 1,350 wells, annually. For 2022 – 2023, on average, this level of sampling would have produced 71% coverage of sampleable trips and 18% coverage of sampleable wells. Results of the simulation indicated that about 50% or more of the stock assessment fisheries for the two dominant set types, sets on dolphin-associated tunas (DEL) and sets on floating-object associated tunas (OBJ), would have had 10 or more wells sampled under the IPSP sampling protocol, and over 80% of fisheries for those two set types would have had at least 2 or more wells sampled. For the stock assessment fisheries for sets on unassociated schools of tunas (NOA), on average, 13% of fisheries would have had 10 or more wells sampled, with coverage increasing to about 50% to 60% at 2 or more wells sampled per fishery.

In the simulation, the relative bias and the coefficient of variation (CV) on fleet-level species catch estimates for the primary tropical tuna species caught in OBJ sets and NOA sets, were computed by stock assessment fisheries, for the IPSP sampling protocol[1]. (Simulations to estimate relative bias and CVs were not run for the DEL fishery, which generates almost exclusively YFT catch, nor for BET in the NOA fishery because it makes up only a very small fraction of the total purse-seine catch of BET.) For OBJ fisheries with at least 10 sampled wells, and for species that were relatively common in the catch (true species proportion > 0.1), the median of relative bias values tended to be centered at or near zero, with 50% of the values falling within about +/-5% or less of the true species proportion. For species that were less common, the relative bias was larger. In terms of CVs, the performance of the proposed IPSP sampling protocol for OBJ fisheries was best for SKJ, with almost all CV values below 0.20. For YFT, about 75% of the CV values tended to be at or below 0.3, as long as the true proportion of YFT in the fishery was > 0.1. For BET, in two of the three OBJ fishery areas that have generated most of the BET catch in recent years, 75% of the CV values were at or below 0.30, and for the third area, the median of the CV values was about 0.30. For the two more inshore OBJ areas, where the true proportion of BET in the simulation was around

[1] The within-well sampling component was adapted to the observer data format.

0.05, the CVs were considerably higher. Overall, for the OBJ fishery these results suggest that BET should be the species that determines the within-well coverage of the IPSP sampling protocol for OBJ-set wells.

For SKJ and YFT (true species proportion > 0.1) in the few NOA fisheries that had at least 10 sampled wells, the median of relative bias values tended to be centered at or near zero, with 50% of the values falling within about +/-5% or less of the true species proportion. In terms of CVs for SKJ, 75% of the CV values were well below 0.3 2022 and below 0.4 in 2023. For YFT, the median of CV values was below 0.3 in the area where the species proportion was > 0.1, but considerably higher in the other area where the true species proportion was < 0.05.

For BET in OBJ fisheries, additional simulations were run with alternate protocols that had higher sampling coverage of trips and wells than the proposed IPSP protocol, but lower within-well sampling coverage. In terms of CVs, those alternate protocols generally did not perform as well as the simulated IPSP sampling protocol. This suggests that the higher within-well sampling coverage of OBJ-set wells proposed for the IPSP sampling protocol is beneficial. (The proposed IPSP sampling protocol has a lower within-well coverage for DEL-set and NOA-set wells, as compared to that for OBJ-set wells.)

Based on the results of the simulation study, it is concluded that the proposed IPSP sampling protocol for Class-6 vessels would be expected to:

a) Provide sufficient data with which to make reliable design-based estimates of species catch, and their variances, for relatively common species in the dominant stock assessment fisheries for OBJ and NOA. Although not tested in the simulation, the same is expected for DEL, which is largely monospecific, given the percentage of DEL stock assessment fisheries with at least 10 wells sampled under the simulated IPSP sampling protocol.

b) Generate improved port-sampling data that can be used as part of model-based estimation methods for fleet-level species catch, using multiple data sources.

c) Provide a sampling framework that can be adapted, if necessary, to: address specific data needs for model-based estimation methods, such as might be beneficial for NOA-set wells; and, address any further data needs for estimation of catch size composition.

BACKGROUND

Traditional port-sampling for species composition and length-frequency distributions of the purse-seine tuna catches

Since 2000, IATTC traditional port-sampling (TPS) data have played a central role in the methodology used to estimate the purse-seine fleet-level tuna species catch. That methodology[2], which is based on a ratio estimator approach, is applied to the TPS data to estimate the species and size composition of the total catch of tropical tunas[3] (sum of catches of yellowfin (YFT), bigeye (BET) and skipjack (SKJ) tunas), by area (Figure 1), month of fishing, purse-seine set type and IATTC vessel size class category (Classes 1-5; Class-6). For tropical tuna catch associated with combinations of these factors for which there is no port-sampling data, species and size composition estimates are based on TPS data from a 'neighboring' area, month and/or vessel size class category, where the 'best' neighbor is determined through a set of ad hoc

[2] A description of the statistical methodology used since 2000 to estimate the purse-seine tuna catch composition for the three target tuna species can be found in several documents: starting on page 339 of Stock Assessment Report 2; starting on page 311 of Stock Assessment Report 4; and, for both time periods (1975-1999 and 2000 onwards), in Document BET-02-06 .

[3] Based on catches reported by canneries (also referred to as 'processors'), observers, and in logbooks.

hierarchical rules that were largely established prior to adoption of the current TPS in 2000 (Tomlinson et al. 1992; Suter 2010; SAC-13 INF-L).

The TPS data are collected by IATTC field office staff when purse-seine vessels unload their catch in port. The sampling has always followed a multi-stage protocol, where first a fishing trip and then a well (or wells) of that trip are selected for sampling, and lastly fish within the well are selected. The selection of trips and wells to sample, from among those wells of a trip that meet the same area/month/set type criteria, is largely opportunistic due to logistical constraints. The sampling of fish within a well resembles a single cluster sample, but not initiated from a random starting point. The data collected include length measurements from a sample of fish, and, separately from the fish measured, counts of species from another sample of fish. The protocol requires samplers to alternate between measuring and counting fish (typically in groups of 25 fish for measurements, and 50 to several hundred fish for counts) (see appendix of Suter 2010 for more details). This procedure extends each type of sampling over a larger fraction of the unloading, as recommended by Wild (1994), than would occur if the counts and measurements were each collected from a contiguous number of fish. The total number of fish in a sample depends on the set type and the number of species believed to be in the well catch, as determined from observer or logbook data just prior to sampling. For example, for a well with catch from sets on tunas associated with floating objects (OBJ sets) that was thought to contain catch of all three tuna species, the sample would typically consist of about 550 fish (150 length measurements, 50 for each species, and counts of species from 400 fish); if one of the species is rare in the catch, the sample would consist of up to 750 fish (150 length measurements, 50 for each species, and counts of species from 600 fish).

TPS data coverage characteristics, Class-6 fleet component

Defining sampling coverage under the TPS sampling protocol requires information on operational characteristics associated with the catch of each well, so that the number of wells that would have been considered 'sampleable' under the protocol can be determined. For the TPS sampling protocol, sampleable refers to wells that contained catch from the same market measurement area (Figure 1), the same month and set type, and were unloaded in a port where sampling occurred at some point during the year (additional details are provided in the Appendix).

For IATTC Class-6 vessels during 2015 – 2023, excluding the pandemic years of 2020 – 2021[4], on average, approximately 68% of all sampleable trips (those with at least one sampleable well), and 15% of all sampleable wells, were sampled by the TPS (Table 1). Almost all sampling occurred in the ports of Ecuador (ECU) and Mexico (MEX)[5]. The level of unloading activity differed between ECU and MEX ports, with typically many more trips (and wells) unloaded in ECU ports (Figure 2). On average, one or more wells were sampled for approximately 87% of vessels that had at least one sampleable trip (Table 2).

At the well level, typically 1-2 wells were sampled per trip for trips unloading in ECU ports and 1-3 wells per trip for trips unloading in MEX ports (Figure 3). The majority of sampleable trips unloading in those ports had more than 3 sampleable wells per trip (Figure 3). On average, approximately 76% of samples came from the upper half of the well and approximately 6% from the bottom quarter of the well, with similar distributions for ECU and MEX ports (Table 3). Sampleable wells with catch from sets on dolphin-associated tunas (DEL) and OBJ-set wells typically contained catch from more than one set, whereas, in

[4] Collection of port-sampling was limited in some ports (SAC-13 INF-L; SAC-14 INF-D) during the pandemic years.

[5] Very limited sampling occasionally occurred in ports of Colombia, Costa Rica, Guatemala, Panama, Peru and Venezuela. Typically, only a few wells per year were sampled in any of these ports, and sampling did not occur in all years in all of these countries.

some years, sampleable wells with catch from unassociated schools of tunas (NOA) were somewhat more likely to contain catch from a single set (Figure 4).

On average, approximately 48% - 59% of tropical tuna catch occurred in sampleable wells, depending on the set type (Table 4). Those values increased by approximately 7% to 11%, on average, when all ports were considered, not just ports where TPS data were collected.

Potential for improvements **to data collection and species catch estimation**

Scientific research related to the development of the trip-level sampling protocol used by the IATTC Enhanced Monitoring Program (EMP) identified several potential areas of improvement relevant for fleet-level catch estimation. Those improvements included minimizing the opportunistic aspects of the TPS sampling protocol, while increasing flexibility in trip and well selection, and increasing within-well coverage of OBJ-set wells. These points are covered in more detail below.

a) Minimization of opportunistic aspects

Currently, there are opportunistic aspects of data collection which arise because the timing of sampling is driven by the work schedules of field office staff, given their other duties (e.g. observer placement and debriefing) and their scheduled work hours (SAC-16-05). Adopting a probability sampling protocol for collection of port-sampling data for fleet-level catch estimation would minimize the opportunistic aspects of the TPS sampling protocol, with several potential benefits. First, under a probability sampling protocol, trips, wells and sections of a well to be sampled would be selected at random, using specific statistical procedures. This would minimize the possibility for bias to be accidentally introduced into the data collection process. Second, under a probability sampling protocol, the probability that trips, wells and fish are included in a sample is known, and thus, can be used to improve the sample weighting in the species catch estimation methodology, for example, when using a design-based estimator. This would reduce the chance that underrepresentation of some trips (vessels) in the sample data, relative to the Class-6 vessel population (e.g. Figure 2), would negatively impact catch estimation (e.g. as illustrated in SAC-15 INF-J). Finally, adopting a probability sampling protocol would allow estimation of the variance on the estimated species catch for design-based estimators, based on the sampling probabilities, which is not possible under the TPS sampling protocol.

b) Increase flexibility in trip and well selection

Increasing flexibility of trip and well selection in the sampling protocol is important to ensure that the protocol does not unnecessarily constrain the fishery definitions used in the stock assessments and catch estimation for those fisheries. Under the TPS sampling protocol, wells of a trip can only be considered for sampling if all the catch in the well came from the same month, set type, and market measurement area (Figure 1) (appendix of Suter 2010). The market measurement areas are no longer used in the tuna stock assessments, and thus, the TPS sampling protocol imposes a structure on the data collection that is no longer useful for species catch estimation. Removing the area constraint, as well as the month constraint, on well selection would increase the number of wells considered sampleable. Less restrictive spatial and temporal sampling protocols are already used in some tuna purse-seine fisheries (Duparc et al. 2018, 2020).

c) Increasing within-well coverage for OBJ-set wells

Increasing the extent of within-well sampling for wells with OBJ-set catch, to sample the entire well catch, will lead to fleet-level estimates of species composition with lower estimated variance. The TPS protocol limits the extent of within-well sampling through the number of fish to be counted or measured for a sample and its compact structure, relative to the duration of the unloading; data are almost always collected from within a single quarter of the well (Table 3). Given that catch from several OBJ sets is typically loaded into the same well (Figure 4), and that OBJ sets catch all three species (Fishery Status Report, Table A-7), differences in species composition among sets can lead to unrepresentative estimates of the species composition for a well if only a subsample of the well catch is collected. This was demonstrated for BET in OBJ-set wells by the EMP pilot study, where trends in BET species composition during uploading of individual OBJ-set wells were identified (SAC-14-10; Lennert-Cody et al. 2024). Moreover, for a fixed coverage level of trips and wells, subsampling of OBJ-set well catch resulted in higher estimated variance on the fleet-level estimates of BET catch for the OBJ-set fishery, as compared to estimates based on data from wells where the entire well catch was sampled (Lennert-Cody et al. 2024).

Purpose of this document

The purpose of this document is to provide an evaluation of the proposed IPSP sampling protocol for collection of purse-seine port-sampling data. The IATTC staff is recommending the implementation of the IPSP as a replacement for the TPS data collection and the EMP. The IPSP will have its main focus on data collection for scientific research while also maintaining some of the features previously provided by the EMP (SAC-16-05). The evaluation presented in this document focuses on the IATTC Class-6 fleet component. To evaluate the protocol performance, simulation studies were conducted using observer data for 2022 – 2023. Anticipated coverage of stock assessment fisheries under the proposed IPSP sampling protocol was computed. Relative bias and coefficient of variation (CV) were computed and compared for several levels of coverage of trips, wells per trips, and within-well sample size. Although the proposed IPSP sampling protocol is for the collection of port-sampling data, well-level observer data were considered the best data source to evaluate the performance of the protocol because of the 100% observer coverage of the Class-6 fleet.

IPSP PORT-SAMPLING PROTOCOL

Assumptions and purpose

The IPSP sampling protocol was developed assuming the following:

1) The priority for estimation of annual fleet-level species catch is by stock assessment fisheries. In addition, the definition of the stock assessment fisheries, especially their spatial structure, is expected to evolve over time as new data and scientific understanding become available. (It is for this reason that the proposed IPSP sampling protocol adopts a dynamic approach, in contrast with the static structure of the TPS sampling protocol, which has been in place since 2000.)

2) The sampling protocol must minimize any interference with catch unloading (as was required for data collection by the TPS and the EMP).

3) Prior to sampling, the database information on well-level catch characteristics relevant for sampling decisions will be limited to the weekly at-sea reports made by AIDCP observers, with more detailed data, i.e. the observer Set Summary/Resumen De Lances data, only available in paper form, a few days before a vessel arrives to port.

4) The annual amount of fishing activity, by set type, month and area, is not knowable in advance of sampling, and the possible combinations of these factors (e.g., 3 set types x 12 months x number of areas), hereafter referred to as stock assessment 'domains', will be too large to make a stratified sampling design logistically feasible.

5) Other existing data sources (i.e., observer, logbook, cannery) are either not universally available and/or their species composition is not sufficiently reliable to be used as the sole source of species composition information for species catch estimation.

The IPSP sampling protocol is a probability sampling protocol. The purpose of this sampling protocol is to:

1) Minimize the need for subjective decision making (i.e. minimize selection bias) during data collection.

2) Generate data with which the variance on species composition estimates can be computed using the sampling protocol, i.e., based on the selection probability associated with each well in the sample. (This is not possible with the TPS sampling protocol because it is not a probability sampling protocol.)

3) Generate port-sampling data that will produce unbiased estimates of species composition at the well level, so that the data can be used, in combination with other well-level data sources (e.g., observer data), to develop methods for estimation of fleet-level species catch using model-based approaches. (For OBJ-set wells, the TPS protocol can lead to biased estimates for BET at the well level.)

The IPSP port-sampling protocol

The IPSP port-sampling protocol is to apply to the entire IATTC Class-6 vessel purse-seine fleet. Logistical details are provided in SAC-16-05. The following describes the three stages of the proposed protocol, which will generate data from about 1,350 wells of 450 trips, annually.

1) *Trips*. To make random sampling of trips logistically feasible, groups of trips are selected using a cluster sampling protocol, instead of selecting individual trips by simple random sampling (SRS). To implement the IPSP sampling protocol, trips will be grouped into clusters of the same number of trips, where the trips in each cluster form a temporally ordered sequence. The ordering of trips will be by approximate arrival date to port. In this way, each cluster will contain trips from the entire year, from the beginning of the year to the end of the year.

   The following is an example to illustrate how clusters of trips would be formed. Suppose there is expected to be a total of 600 sampeable trips[6] for the year, and that a cluster size of 100 trips is chosen. Then there will be a total of 6 cluster samples of sampleable trips for the year (600/100 = 6). The table below shows a list of those 6 cluster samples, where the numbers in the table are the ordered arrival of the trips to port. The top row lists the first 6 trips arriving to port, from left to right, the second row lists the next 6 trips arriving to port, and so on. The table is static in the sense that, once there is a mechanism in place to order trips, and the total number of clusters has been determined, every trip that occurs can be assigned to a cluster (to a column of the table) using information on when it will arrive to port, i.e., based on its place in the sequence of arrivals for all sampleable trips at sea. This does not require that all arrival dates for the year are known at the start of the year, but rather that a list of consecutive arrivals is maintained and continually updated in near-real time. For Class-6 purse-seine vessels, this can be made possible through the use of data contained in the observers' at-sea reports.

[6] To be sampleable under the IPSP sampling protocol, a trip must have: a) tropical tuna catch from the EPO; b) be unloading in one of the four ports where sampling will occur (Manta, Manzanillo, Mazatlán, Posorja); and, c) have at least one well with catch entirely from the year of interest and from a single purse-seine set type (i.e. have at least one sampleable well).

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
1st	2nd	3rd	4th	5th	6th
7th	8th	9th	10th	11th	12th
13th	14th	15th	16th	17th	18th
….	….	….	….	….	….
589th	590th	591st	592nd	593rd	594th
595th	596th	597th	598th	599th	600th

Now, suppose we want to sample three clusters of trips during a year. At the start of the year, we would select three of the 6 possible clusters of trips at random, by randomly selecting (without replacement) three numbers from the sequence: 1, 2, 3, 4, 5, 6. Suppose those randomly selected numbers were: 1, 3, 4. Then the three cluster samples, each of 100 trips, would be: Cluster 1, with trips 1, 7, 13, …, 595; Cluster 3, with trips 3, 9, 15, …, 597; and, Cluster 4, with trips 4, 10, 16, …, 598. In the case of cluster sampling, it is the clusters that are selected at random, not the individual trips. Once a cluster is selected, every trip in the cluster would be sampled. Creating clusters by ordering trips by the date they arrive to port is a way to create clusters that span the full year of fishing. Selecting at least two cluster of trips for sampling will allow for estimation of the among-trip variation in species composition.

For the IPSP sampling protocol, the following information led to the specifics of the cluster sampling of trips. First, based on observer data for 2022 and 2023, it was anticipated that at least 600 sampleable trips would be made annually by the Class-6 fleet component. Second, with the operating budget of the IPSP (SAC-16-05), it is expected that a total of about 450 trips could be sampled annually. And, lastly, in order to estimate the among-trip variation in species composition, and to capture seasonality of the fleet activity in each cluster of trips, it was decided to sample three clusters of 150 sampleable trips each. With these parameters, and assuming a total of 600 sampleable trips, there would be a total of 4 clusters possible (600/150 = 4). Identical to the example above, to form the clusters of trips, trips would be ordered from first to last based on their estimated arrival date to port, which allows a table similar to the one above to be created. The difference relative to the table above is that the clusters of trips each contain 150 trips, instead of 100 trips. At 150 trips per cluster, if trips were uniformly distributed over the year, about 12- 13 trips per month would be sampled, with about 2 - 3 trips sampled per week.

In practice, two important considerations are worth mentioning. First, the total number of sampleable trips for the year will not be known in advance, but the clusters of trips can still be created. If there were to be fewer total trips than expected, then the actual clusters would contain fewer trips at year's end (i.e., the last few rows of the table would not be used). If there were to be more trips than expected, then the number of trips sampled per cluster could be extended, resources permitting (i.e., more rows could be added to the table). Adjustments to the protocol, in terms of number of clusters and number trips per cluster, can be made from year to year, based on realized fishing activity of the previous year (or years), and taking into consideration available resources. If there are more trips than expected and the number of trips sampled per cluster cannot be increased to accommodate the excess, the trips not sampled per cluster will not likely occur at random. The best option to account for this in the species catch estimation will be discussed at the planned External Review of the IPSP sampling protocol (see Discussion section). Second, the exact arrival date of each trip will not be known in advance. However, a proxy for arrival date can be obtained from data in the observer at-sea reports (Informes Desde el Mar or IDM), which are submitted by observers on a weekly basis. Using those data, an estimate of how much catch a vessel has on board, relative to its capacity (i.e., percent full) can be computed, and trips can be ordered according to the date at which they were estimated to be 'nearly' full (e.g. above a threshold of 90% full).

It is important to note that the definition of sampleable under the proposed IPSP sampling protocol is different from the definition under the TPS sampling protocol. To be sampleable under the IPSP sampling protocol, a trip must have: a) tropical tuna catch from the EPO; b) be unloading in one of the four ports where sampling will occur (Manta, Manzanillo, Mazatlán, Posorja); and, c) have at least one well with catch entirely from the year of interest and from a single purse-seine set type (i.e. have at least one sampleable well).

2) *Wells*. Select 3 wells per trip from each trip selected in (1), by SRS, without replacement, for an annual total of 1,350 wells (3 clusters of trips x 150 trips per cluster x 3 wells/trip = 1,350 wells).

3) *Fish*. Collect 1 cluster sample, or subsample, of 'units' of fish from each well selected in (2). This within-well component of the IPSP sampling protocol differs by set type. For all OBJ-set wells, a unit refers to a container of fish[7]. The cluster sample, which will be a sequential group of containers of fish, will cover the entire unloading of the well (i.e., identical to the EMP within-well sampling protocol; SAC-14 INF-I). For all DEL-set wells, a subsample of a single cluster sample will be collected. In the case of DEL-set wells, a 'unit' refers to an individual fish, and a single cluster sample is a sequential group of individual fish, created by sampling one out of every 10 fish unloaded from the well (in the case of catch dominated by small fish, the interval may be one out of every 30 fish). A subsample from a single cluster sample is, therefore, a subsample from this sequential group of individual fish. In practice, quarters of the well will be used to define subsamples; that is, there will be 4 possible subsamples that could be collected from a single well. A quarter of the well will be selected at random, and the part of the (full) single cluster sample that falls in that quarter of the well is the subsample that will be collected. For all NOA-set wells, a unit will refer to a container of fish (identical to OBJ-set wells), however, only a subsample of the single cluster sample of containers of fish will be collected. The subsample will be the sequential group of containers of fish that were unloaded from a randomly selected quarter of the well.

The collection of data from every unit of the single cluster sample, or subsample, of fish from the well will be as follows. All fish in each unit will be identified to species (BET; YFT; SKJ; Other). For OBJ-set and NOA-set catch, all tropical tunas in each unit will be individually weighed to the nearest 0.02 kg, or for the largest fish[8], measured to the nearest millimeter (mm). For DEL-set catch, instead of collecting weight data, every tropical tuna will be measured for length to the nearest mm. Lengths can be converted to weight using length-weight relationships that will be developed from the morphometric sampling being conducted by the EMP in 2025 (SAC-15 INF-H).

[7] In the case that physical containers are not used during the unloading, virtual containers of fish will be created, where a virtual container refers to a contiguous group of 50 fish. This is identical to the EMP within-well sampling protocol. Additional details are provided in SAC-14 INF-I.

[8] Due to the upper weight limitation of the portable scales, which was 30 kg, tunas larger than 28 kg were measured for length (SAC-14 INF-I).

EVALUATION OF THE IPSP PORT-SAMPLING PROTOCOL

Data

Observer Set Summary[9] (Resumen De Lances (RDL), in Spanish) data collected by AIDCP observers for Class-6 vessel trips were used in this study because they contain set-level species composition information, by well, for all Class-6 trips and vessels due to the 100% coverage of the AIDCP observer program. Data for the most recent years, 2022 – 2023, were used. The RDL data had to be reconstructed from different observer data tables[10] because, although RDL data are recorded on data forms by all AIDCP observers, those data forms are only keypunched for some trips[11]. To reconstruct the RDL data set for the entire Class-6 vessel fleet, different types of observer data were linked using trip numbers and fishing locations and dates (set numbers were not available in all required data tables). About 2% of wells could not be linked to specific sets, and data of those wells were excluded from the simulation. The RDL data contain information on set type, dates and locations of fishing, and catch amounts, by species, for the catch from every set loaded into each well of a trip.

Methods

The simulation study to evaluate the performance of the proposed IPSP sampling protocol is outlined in this section. For the simulation, a modification had to be made to the IPSP sampling protocol presented above to adapt it to the RDL data format. Specifically, the third stage of the simulation sampling protocol was formulated in terms of sampling sets from the well, instead units of fish unloaded from the well. This was necessary because the RDL data do not have fine-scale within-well resolution. Under the proposed IPSP sampling protocol, a number of units of fish would be sampled from the catch of each set in a well, and thus, the data actually generated by the IPSP sampling protocol will be more detailed at the well level than the RDL data. It is noted that, in practice, selection of catch from different sets in the well would be logistically very difficult because the physical boundaries between catch from different sets in the well are not known and unlikely to be distinct.

Domain spatial definitions and sampling parameter combinations used in the simulation

Because different areas are used in the stock assessments for each of the three tropical tuna species, the simulation was run by set type x species, for those set types that contributed significantly to the EPO species catch (Figure 5). For BET, the OBJ-set spatial fishery definitions from the 2024 BET benchmark assessment (SAC-15-02) were used. The simulation was not run for BET in NOA-set wells or DEL-set wells because OBJ sets produce the vast majority of the BET purse-seine catch in the EPO (Fishery Status Report). For YFT, separate simulations were run for OBJ-set wells and NOA-set wells using the set type-specific areas of the 2025 YFT benchmark assessment (SAC-16-03). Only the sampling coverage-related aspects of the simulation were run for YFT in DEL-set wells because DEL sets produce almost exclusively YFT (Fishery Status Report), however, future research will evaluate the IPSP sampling protocol with respect to estimation of length composition. For SKJ, simulations were run for OBJ-set wells and NOA-set wells, using the 2024 SKJ benchmark assessment set type-specific spatial fishery definitions (SAC-15-04).

There were three sampling parameter combinations considered in the simulation, which represent three different sampling protocols. The first parameter combination approximates the proposed IPSP sampling protocol. The other two parameter combinations were run to evaluate the effect of increased sampling

[9] This observer data type is also referred to as observer 'well plan' data.
[10] Permanent observer database and Tuna Tracking form database.
[11] RDL data are considered a preliminary data source and are ultimately replaced with fully-edited observer data once those data become available.

effort, in terms of wells and trips, on the performance of the two within-well sampling options (described below). The three parameter combinations, i.e. three simulated sampling protocols, were:

1) IPSP protocol: 3 wells/trip; 3 clusters each of 150 trips, for an annual total of 450 trips and 1,350 wells;
2) Double the wells sampled per trip: 6 wells/trip; 3 clusters each of 150 trips, for a total of 450 trips and 2,700 wells;
3) Sample almost all trips: 3 wells/trip, 3 clusters each of 206 trips (2022) or 217 trips (2023), for a total of 618 trips and 1,854 wells for 2022, and 651 trips and 1,953 wells for 2023.

It is noted that since there were more than 600 sampleable trips in both 2022 and 2023, simulated sampling protocol (1) above is an approximation to the implementation of the proposed IPSP protocol, in practice, in terms of sampling coverage. For example, the total number of sampleable trips for 2022 in the simulation data set was 619 trips. The number of trips per cluster, for 4 clusters, would be 154 (= 619/4, rounded to the smaller integer). Thus, in practice, resources permitting, 3 clusters of 154 trips per cluster might have been sampled by the IPSP. Implementing this in the simulation would have led to different sample sizes in 2022 and 2023, since the number of sampleable trips differed between the years (there were a total of 653 sampleable trips in the 2023 data set). A decision was made to fix the sample sizes for the simulation at 150 trips per cluster for simulated sampling protocols (1) - (2), despite the differences in coverage this would produce in the simulation, because extending the sampling beyond the 150 trips per cluster will, in practice, be dependent on resource availability (see also Discussion section). The relatively small differences in coverage between the two years may contribute to differences in performance between the two years in the simulation.

Simulation procedure

The simulation, which was programmed using the statistical freeware R (R Core Team 2024), had the following steps:

i. Creating synthetic full data sets
    a. Generate 100 synthetic data sets by resampling sampleable trips, with replacement, from the original data, to the total number of sampleable trips in the original data.
    b. Order trips in each synthetic data set by unloading date (first well); trips with the same unloading date appear in a random order.
ii. Generating samples from each synthetic full data set
    a. For each synthetic full data set, 30 different sample data sets were generated as follows:
        i. Three cluster samples of trips were selected. Given the number of trips per cluster in simulated protocols (1) and (2) (i.e., 150 trips), and the total number of sampleable trips of each population (619 for 2022, 653 for 2023), it was assumed that there was a total of 4 possible clusters in each population under simulated sampling protocols (1) and (2). Out of these 4 clusters, the 3-cluster samples that could be drawn were: (1,2,3); (1,3,4); (1,2,4); and (2,3,4). For simulated sampling protocol (3), the number of trips per cluster was increased to about 1/3 of the total number of sampleable trips, and thus, there were only three possible cluster samples in the population. That is, under simulation sampling protocol (3), all clusters were sampled, and nearly all possible trips. On the other hand, under simulation sampling protocols (1)-(2), 450 trips were sampled out of an assumed

600 trips. For all three simulation sampling protocols, however, performance was evaluated against the population values (see below), which were based on the total number of trips in the synthetic full data set.

ii. For each trip of a cluster, 3 sampleable wells were drawn per trip, by SRS without replacement, from among all sampleable wells of the trip, for simulated sampling protocols (1) and (3), and 6 wells per trip for simulated sampling protocol (2).

iii. For each well selected, two set-selection options were used: 1) 'all sets': all the catch of all sets in the well, i.e., a census of the catch in the well; and, 2) '1 set at random': all the catch of one set in the well, where the set was selected at random from among those sets in the well. The two options are intended to approximate the two types of within-well sampling under the proposed IPSP sampling protocol: complete sampling of the well catch, i.e. the IPSP within-well sampling protocol for OBJ-set wells, which is approximated in the simulation by the 'all sets' option; and, subsampling of the well catch, i.e. the IPSP within-well sampling protocol for DEL-set and NOA-set wells, which is approximated in the simulation by the '1 set at random' option. For selected wells that contain catch from only one set, the two set-selection options in the simulation will produce the same well-level result. Note that the RDL catch amounts are in weight of fish (not numbers of fish).

iii. Estimating species proportions by domain (i.e., by a specific combination of set type x month x stock assessment fishery area) for each of the 30 samples from a synthetic full data set

a. For poststratification of the sample data to domains, sampled wells were assigned to set type x month x area, based on the information associated with the set(s) from which catch was loaded into the well. For those wells with catch from more than one set, and for which the sets were not from the same month and/or area, the month and area assigned to the well was that of the set with the majority of the catch in the well. (The set type of all sets in the well was the same, as required by the criteria used to select sampleable wells.)

b. Species proportions, by domain, were estimated according to eq. 4, shown below, for those domains that were represented by at least 10 sampled wells.

c. This produced up to 30 estimates of species proportions, for each domain with sufficient data, for each of the 100 synthetic full data sets.

iv. Computing metrics of performance for different parameter combinations for each synthetic full data set

a. Coverage of domains, only computed from the simulation output for the first synthetic data set, was summarized by the number (and percent) of domains associated with a particular set type, that had: 2 or more wells sampled; and, 10 or more wells sampled.

b. Performance of a parameter option for a domain and species was summarized by:

i. Relative Bias = $\frac{\left(\overline{\hat{p}_{d\_spp}} - {p^*}_{d\_spp}\right)}{{p^*}_{d\_spp}}$

ii. CV = $\frac{\sqrt{Variance}}{{p^*}_{d\_spp}} = \frac{\sqrt{\left(\frac{1}{(number\ of\ samples-1)}\right)\sum_{samples}\left(\hat{p}_{d\_spp} - \overline{\hat{p}_{d\_spp}}\right)^2}}{{p^*}_{d\_spp}}$

where $\hat{p}_{d\_spp}$ is the estimated fleet-level proportion of the species for the domain based on a single sample from the synthetic full data set, according to eq. 4 below;

$\overline{\hat{p}_{d\_spp}}$ is the average of the sample estimates for the particular synthetic full data set; and, $p^*{}_{d\_spp}$ is the true fleet-level proportion of the species for the domain of the synthetic full data set (computed as the sum of all of the species catch divided by the sum of all tropical tuna catch, from the same full data set).

c. Summarize the relative bias and CV estimates from the 100 synthetic full data sets, for a particular combination of sampling parameters, by area and set type, using box-and-whisker plots. Note that variability in these box-and-whisker plots for a single area and set type is due to both variability among months for the same population and variability across populations. Also to note is that for some populations, some combinations of sampling parameters, areas and set types might not have had sufficient sample data to make an estimate (i.e. at least 10 sampled wells) and hence the summaries for that particular combination would be based on fewer than all 100 synthetic full data sets.

Estimation of $\hat{p}_{d\_spp}$

Following the general estimation methodology of Lohr (2022) for probability sampling, a design-based estimator of the species proportion can be obtained from the simulation probability sampling protocols described above. For those sampling protocols, the probability that an 'observation' from a well (one set, all sets) would be included in a sample (the 'inclusion' probability), is given by the following equation, assuming that all trips in each cluster are sampled:

$$P(observation\ k\ of\ well\ j\ of\ trip\ i\ of\ cluster\ c) = \left(\frac{s_c}{S_c}\right)\left(\frac{m_{ci}}{M_{ci}}\right)\left(\frac{l_{cij}}{L_{cij}}\right) \qquad \text{eq. 1}$$

where, s_c clusters of trips were selected at random from a total of S_c possible clusters (all n_c trips in cluster s_c are sampled), m_{ci} wells of M_{ci} for trip i are selected by SRS without replacement, and l_{cij} sets out of L_{cij} sets for the well selected from each of the m_{ci} wells of trip i selected by SRS without replacement. For the '1 set at random' option, l_{cij} = 1, and for the 'all sets' option, l_{cij} = L_{cij}. Note that Eq. 1 assumes that the S_c clusters, combined, contain the total number of trips in the population, which for simulation sampling protocols (1) – (2) is an approximation.

Given the inclusion probability of eq. 1, the sampling weight for an observation from a well (one set, all sets), which is used to raise the data to the fleet level, is the inverse of the probability it was sampled:

$$\omega_{cijk} = \frac{1}{P(observation\ k\ of\ well\ j\ of\ trip\ i\ of\ cluster\ c)} = \left(\frac{S_c}{s_c}\right)\left(\frac{M_{ci}}{m_{ci}}\right)\left(\frac{L_{cij}}{l_{cij}}\right) \qquad \text{eq. 2}$$

It is noted that, identical to the proposed IPSP sampling protocol, the simulation protocol is not species-specific. For simulation sampling protocols (1) – (2), the raising factor $\left(\frac{S_c}{s_c}\right)$ of eq. 2 does not account for the trips beyond 600 that were not included for in the 4 clusters of 150 trips. In this sense, the performance metrics, computed below, reflect an outcome of the situation where the 450 trips and 3 wells per trip of the proposed IPSP sampling protocol could not be extended to cover all trips in the 4 clusters. As an aside, it is also noted that, if the level of sampling coverage at all stages of the protocol were to be very low, raising those sample data, as is done with the sample weights of eq. 2, will not necessarily correct for a lack of representativeness of the sample, and might therefore result in a poor estimate of the fleet-level species catch in the domain. This would apply to the proposed IPSP sampling protocol, were the sampling coverage to be very low.

An estimator of the proportion of species *spp* in the fleet-level catch from the sample data, for the entire sampleable population of trips, i.e. without poststratification, would be given by:

$$\hat{p}_{spp} = \frac{\hat{C}_{spp}}{\hat{C}_{tropical\_tuna}} = \frac{\sum_c \sum_i \sum_j \sum_k \omega_{cijk} y_{spp\_cijk}}{\sum_c \sum_i \sum_j \sum_k \omega_{cijk} y_{tropical\_tuna\_cijk}} \quad \text{eq. 3}$$

where $\hat{C}_{...}$ is the estimated catch (in weight), and *y* is the sample amount (in weight) of the species or all tropical tuna. It is noted that, in future research, a revision to eq. 3 could be considered under the assumption that the total catch of all three species in the well is known.

With the proposed IPSP sampling protocol, poststratification of the sample data is required to obtain domain-level estimates because the IPSP sampling protocol is not a stratified sampling protocol. Estimation therefore requires adjusted sampling weights, which, per Lohr (2022 page 330 eq. 8.12), account for the fact that the proportion of observations in each domain (poststratum) in the sample data may not equal that of the population, due to taking a sample rather than obtaining a census. A design-based estimator of the proportion of species *spp* in the fleet-level catch of domain *d*, using adjusted sampling weights, is given by:

$$\hat{p}_{d\_spp} = \frac{\hat{C}_{d\_spp}}{\hat{C}_{d\_trop\_tun}} = \frac{\sum_c \sum_i \sum_j \sum_k \omega^*_{cijk\_d}\, y_{spp\_cijk} x_{cijk\_d}}{\sum_c \sum_i \sum_j \sum_k \omega^*_{cijk\_d}\, y_{tropical\_tuna\_cijk}\, x_{cijk\_d}} \quad \text{eq. 4}$$

where $\omega^*_{cijk\_d} = \omega_{cijk} \frac{O_d}{\sum_B \omega_{cijk}\, x_{cijk\_d}}$, O_d is the total number of sampleable wells of the population of sampleable trips in domain *d*, x_{cijk_d} is an indicator variable that takes the value 1 if the well sample is from domain cell *d* and 0 otherwise, and *B* is the collection of all wells in the sample data. The purpose of the adjustment is to adjust the sample weights to represent the total number of sampleable wells in the domain.

Results

Coverage

Under the IPSP sampling protocol (450 trips and 1,350 wells sampled annually), the percent coverage of IPSP-sampleable trips and wells in 2022 and 2023 (Table 5) would have been similar to the actual percent coverage by the TPS of TPS-sampleable trips and wells in those years (Table 1). At 1,350 wells to sample, the IPSP sampling protocol would sample considerably more wells compared to the average number of wells sampled by the TPS sampling protocol, approximately 847 wells. The percent coverage shows little difference, however, because the flexibility of the IPSP protocol with respect to area and month means that there are more wells in the IPSP's sampleable population (i.e. more wells that could be sampled).

The number of stock assessment domains with sample data under the proposed IPSP sampling protocol varied by set type (Table 6). About 50% to 60% of the OBJ-set domains had at least 10 wells sampled per domain, and over 80% had at least 2 or more wells sampled per domain, regardless of the species-specific area configuration of the domains. For the DEL-set domains, just under 40% of the domains had at least 10 wells sampled per domain, and about 80% had at least 2 wells sampled per domain. For NOA-set domains, the coverage was lower and varied with the area configuration (i.e. YFT assessment areas *versus* SKJ assessment areas). For the SKJ assessment NOA areas, 27% in 2022 and 12% in 2023 had 10 or more sampled wells per domain. For the YFT assessment NOA areas, 7% of the areas in both years had 10 or more sampled wells per domain. Coverage increased to about 50% to 60% when considering 2 or more wells sampled per domain.

General simulation output characteristics

In general, across set types and species there was an increasing relationship between relative bias for the two types of within-well sampling options, and between the CVs for those two options (Figure 6). For a species, on a domain by domain basis, the relative bias of the 'all sets' option could be larger or smaller than that of the '1 set at random' option, but the CV of the 'all sets' option was almost always smaller than the CV of the '1 set at random' option. Of course, if there were catch from only one set in every sampled well for every sample from a synthetic full data set, the relative bias and CV for a domain and species would be the same for both within-well sampling options because both within-well sampling protocols would produce exactly the same sample data. However, as shown in Figure 4, this is unlikely to be the case.

IPSP sampling protocol performance for OBJ-set domains

In general, the simulated IPSP sampling protocol performance was better for those domains where the species of interest was relatively common in the catch, from which it follows that in those OBJ domains with at least 10 wells or more per sample, the performance was largely quite good for SKJ, reasonable for YFT, in almost all such domains, and reasonable for BET in those domains that were based on areas where the true proportion of BET in the catch for the simulations was typically above 0.1 (Figures 7 – 12). For relative bias on the estimated proportion of SKJ, across months and populations in all 5 OBJ-SKJ areas (Figures 9, 12), the median of the simulation values was centered on zero and the interquartile range (IQR) was at or less than +/- 5% of the true population proportion value. For relative bias on the estimated proportion of YFT, across months and populations, the median relative bias was generally close to zero and the IQR was at about +/- 5% of the true population proportion value (Figures 8, 11). For relative bias on the estimated proportion of BET, across months and populations, the median relative bias was generally close to zero and the IQR was at about +/- 5% of the true population proportion value in Areas 1-2 (Figure 5), the two areas with the greatest catches of BET in recent years (SAC-15-02), and less than about +/-10% in Areas 3 -5 (Figures 7, 10). For all three species, the 'all sets' option generally led to a smaller IQR and smaller overall range of relative bias than the '1 set at random' option.

The performance of the IPSP protocol with respect to estimated CVs, and the differences among species and domains, was generally similar to that for relative bias (Figures 7-12), but the improvement of the 'all sets' option over the '1 set at random' option was greater for YFT, and especially for BET, than for SKJ. For SKJ in OBJ-SKJ areas, the maximum estimated CVs, for both the 'all sets' and the '1 set at random' options, were largely below 0.2, and the upper quartile of CV values was below 0.15 (Figures 9, 12). For YFT in OBJ-YFT areas, the upper quartile on the estimated CV values for the 'all sets' option was 0.30 or less for all but Area 5 in 2022 (Figures 8, 11). For BET in OBJ-BET domains, the upper quartile of CV values for the 'all sets' option was less than 0.30 in Areas 1-2, and the median CV value was about 0.3 in Area 4. For Areas 3 and 5, where the true population proportion BET was less than 0.1, the lower quartile of CV values for the 'all sets' option was greater than 0.3 (Figures 7, 10).

IPSP sampling protocol performance for NOA-set domains

There were fewer NOA domains with at least 10 sampled wells with which to evaluate the performance of the IPSP protocol (Table 6), which is related to the smaller numbers of NOA sets in 2022 and 2023, among other factors. In 2022 – 2023, there were between 17,000 – 18,000 OBJ sets made, as compared to 4,000 – 5,000 NOA sets (all vessel size classes; Fishery Status Report, Table 7). For YFT in NOA-YFT Area 4 (Figure 13), under the '1 set at random' option the IQR of relative bias was within +/- 5% of the true population proportion value, and the median CV value was below 0.3, with an upper quartile value of about 0.4. However, in Area 5 where the mean population proportion value was less than 0.1, for the '1 set at random' option the IQR of relative bias values was about +/- 10% or greater and the lower quartile of CV values was above 0.3. For SKJ in NOA-SKJ Area 3 (Figure 14), the only area with at least 10 wells per sample, the IQR of relative bias was within -/+ 5% of the true population value for the '1 set at random' option, and the upper quartile of CV values was well below 0.3 in 2022 but about 0.4 in 2023.

Comparison of IPSP protocol to those sampling more wells, trips

Given the results presented above, the comparison between the simulated IPSP sampling protocol and the simulated alternate sampling protocols focused on performance for BET in the OBJ-BET stock assessment domains. The comparison was made between the 'all sets' option under the simulated IPSP sampling protocol, i.e., simulation sampling protocol (1), and the '1 set at random' option under the alternate simulation sampling protocols (2) and (3), with their higher coverage of trips and/or wells.

Overall, the IPSP sampling protocol with the 'all sets' option often had similar or better performance than the '1 set at random' option under alternate sampling protocols (2) and (3), especially as regards the CV (Figures 15 - 22). For relative bias, the IPSP sampling protocol with the 'all sets' option gave very similar results as compared to the two alternate sampling protocols under the '1 set at random' option, except for Area 3, where protocol (3) that sampled more trips (and hence also more wells) led to slightly less negative bias. For the CV, the IPSP sampling protocol with the 'all sets' option had similar or better CVs than the CVs obtained from either of the alternate sampling protocols with the '1 set at random' option, except for sampling protocol (2), that sampled more wells, in Area 3 for 2023.

DISCUSSION

The proposed IPSP sampling protocol for collection of port-sampling data from the EPO purse-seine Class-6 fleet was presented and its performance was evaluated through simulations that used observer RDL data for 2022 – 2023. The proposed IPSP sampling protocol is a three-stage protocol, with cluster sampling of trips, simple random sampling (without replacement) of wells, and cluster sampling of units of fish within a well. Assuming fishing dynamics remain similar, on average, the IPSP sampling protocol would be expected to achieve about 71% coverage of sampleable trips and 18% coverage of sampleable wells, annually. Under the IPSP sampling protocol, about 50% or more of the stock assessment fisheries ('domains') for the two dominant set types (DEL and OBJ) would be expected to have 10 or more wells sampled per year. However, on average, only 13% of NOA domains had 10 or more sampled wells per year.

Simulation results on the performance of the IPSP sampling protocol, summarized in terms of the relative bias and the CV on the estimated species proportions, by domain, were generally similar for the two years[12]. For the relatively common tropical tuna species (true species proportion > 0.1) in OBJ domains with at least 10 sampled wells, the median of relative bias values tended to be centered at or near zero, with an IQR around +/-5% or less of the true proportion value. For species that were less common, the relative bias was larger. In terms of CVs, the performance of the IPSP sampling protocol in OBJ domains was best for SKJ, with almost all CV values below 0.20. For YFT, the upper quartile of CV values tended to be at or below 0.3, as long as the true proportion of YFT in the domain was > 0.1. For BET, the upper quartile of CV values was at or below 0.30 only in those domains that generated the most BET catch (Figure 5, BET-OBJ Areas 1-2). In BET-OBJ Area 4, the median of CV values was about 0.30. For the other two areas, where the true proportion of BET in the simulation was around 0.05, the CVs were considerably higher. Overall, for the OBJ fishery, these results suggest that BET should be the species that determines the within-well coverage of the IPSP sampling protocol for OBJ-set wells. For SKJ and YFT in the few NOA domains that had at least 10 sampled wells and where the true species proportions > 0.1 (SKJ-NOA Area 3 and YFT-NOA Area 4), the median of relative bias values tended to be centered at or near zero, with an

[12] Simulations were not run for the DEL fishery, which generates almost exclusively YFT catch, nor for BET in the NOA fishery because it makes up a very small fraction of the NOA catch (Fishery Status Report).

IQR around +/-5% or less of the true proportion value. In terms of CVs, for SKJ in SKJ-NOA Area 3 the upper quartile of CV values was well below 0.3 in 2022 and around 0.4 in 2023. For YFT, the median of CV values was below 0.3, and the upper quartile was about 0.4, in YFT-NOA Area 4.

The 'all sets' within-well sampling option of the simulated IPSP sampling protocol for OBJ-set wells generally had similar or better performance for CVs for BET in OBJ domains than the '1 set at random' option under the two alternate simulation sampling protocols that doubled the number of wells sampled or sampled nearly all sampleable trips. Given that OBJ-set wells typically have catch from more than one set, when the entire well catch is sampled, the catch from all sets in the well will be sampled. The improved CVs with the 'all sets' option is consistent with results of previous studies with EMP pilot study port-sampling data and simulations with observer RDL data, which suggested that precision would be better when the entire catch of OBJ-set wells is sampled (Lennert-Cody et al. 2024). At a coverage of trips and wells below that of the proposed IPSP sampling protocol (Table 6), such as might occur if it were not possible for the IPSP sampling protocol to be implemented by dedicated samplers, neither within-well sampling option would be expected to yield satisfactory results in terms of relative bias and CVs (Figure 23).

A further advantage of the IPSP sampling protocol's higher within-well coverage for OBJ-set wells is that it will have more reliable species estimates at the well level, opening up more methods options for fleet-level species catch estimation (see below). Examples of well-level bias for BET due to the limited within-well sampling protocol used by the TPS can be seen in a comparison of the EMP well-level estimates and TPS well-level estimates, for the same wells (Figures 24-25). Although the number of wells sampled under both protocols is small, and not a random sample of wells, it can be seen that, for the estimates that are most different, the TPS protocol likely led to an overestimation of the proportion of BET in the well because not enough of the well catch was sampled. This bias will lead to increased prediction error for a port-sampling – observer well-level model (e.g. SAC-16 INF-I), complicating any effort to use a model-based approach for estimation of fleet-level species catch that incorporates both observer data and port-sampling data.

It is noted that, despite several important improvements of the proposed IPSP sampling protocol over the TPS sampling protocol, the low level of sampling coverage for some domains (Table 6) is an outcome common to both the IPSP and the TPS sampling protocols, due to finite resources. For the TPS, this was noted previously (Tomlinson et al. 1992; Suter 2010; SAC-13 INF-L). With the IPSP, several steps will be taken to try to address this challenge within the approved budget. These steps include exploring options for additional methodological improvements to the sampling protocol, as well as the species catch estimation methodology for both design-based estimators and model-based estimators. In addition, as noted in SAC-16-05, collaboration with universities to establish student internships to help with sampling is also being explored.

As regards improving the proposed IPSP sampling protocol, simulation studies with the data collected under Project C.1.b (SAC-16 INF-H) will be conducted to further evaluate the within-well sampling component of the IPSP sampling protocol for DEL and NOA wells, which might lead to modifications of the IPSP sampling protocol that could make more efficient use of available resources. Project C.1.b generated high-frequency species and length composition data for DEL-set and NOA-set wells, similar to the data collected during the EMP pilot study for OBJ-set wells. Simulation studies will evaluate bias and precision of well-level estimates under various types of within-well subsampling, for both species and size composition.

In addition, an external review of the proposed IPSP sampling protocol and species catch estimation methodology is expected to be held later this year. It is anticipated that the review panel will provide suggestions for improving the IPSP sampling protocol, such as how best to handle the estimation when there are unsampled trips in the selected clusters, e.g. because more trips occurred per cluster than could be sampled with available resources. The fact that the poststratification of the port-sampling data includes month as one of the poststratification variables may be an advantage in this regard. The external review may also produce suggestions for alternate design-based estimators, beyond eq. 4 adapted to the proposed IPSP sampling protocol, as well as model-based estimators related to 'small area estimation' methods (e.g. Porter et al. 2014; Erciulescu et al. 2019).

As part of development of an estimator of the total fleet-level species catch by domain, it is helpful to first consider a simple conceptual formulation. The total estimated catch of species *spp* in stock assessment domain *d*, $\hat{C}_{d\_spp}$, can be expressed as the sum of two estimates:

$$
\begin{aligned}
\hat{C}_{d_{spp}} &= \hat{p}_{d_{spp}} C_{d_{tropica\_tuna}} \\
&= \hat{p}_{d\_spp\_sampleable}\, C_{d\_tropical\_tuna\_sampleable} \\
&\qquad + \qquad \hat{p}_{d\_spp\_UNsampleable}\, C_{d\_tropical\_tuna\_UNsampleable}
\end{aligned}
$$

where $C_{d_{tropical\_tuna\_...}}$ is the total tropical tuna catch for the fleet in domain *d* (obtained using cannery, observer and logbook information) and is assumed to be known. For domains with sufficient port-sampling data, the first term can be estimated from port-sampling data alone, using a design-based estimator for the species proportion. The estimator for the species proportion in the simulation study was a simplistic illustration of this, which could be adapted to the proposed IPSP sampling protocol. In the case of domains with little to no port-sampling data, the first term would be estimated from a model-based approach. The second term, by definition, would have to be estimated from a model-based approach. It is worth noting that estimation for unsampleable trips and wells requires the assumption that processes associated with the unsampleable catch are the same as those associated with the sampleable catch. Examples of such processes include vessel specific effects, where some vessels may never be sampled because of the ports they use to unload their catch, and processes that lead to small-catch sets and hence mixed set type wells. In the case of the current estimation methodology with the TPS data for domains with little to no port-sampling data and unsampleable data (Tomlinson et al. 1992; Suter 2010; SAC-13 INF-L), a type of model-based approach is used, based solely on port-sampling data. However, the model structure of the approach is ad hoc and not informed by analysis of recent data.

To develop new model-based approaches to species catch estimation, either for all domains or only those with little or no port-sampling data, there are three data-driven improvements for species catch estimation presently being considered: 1) expanded use of covariates; 2) use of both port-sampling and observer data; and 3) inclusion of fine scale spatio-temporal correlations.

Covariates could include trip, vessel, well-level covariates (e.g., summary of fishing dates for sets in the well, latitude centroid of sets in the well, longitude centroid of sets in the well), and vessel-level operational characteristics covariates (e.g. hopper use). Including well-level covariates may improve model performance for trips and vessels never sampled. The error on model predictions, and hence the error on the estimated fleet-level species catch, could be estimated. The mixed-effects modelling study presented in SAC-16 INF-I uses this approach by predicting the well-level port-sampling species composition as a function of the observer species composition and covariates. Alternative approaches could fit to both port sampling and observer data simultaneously as a function of covariates, with the observation model for the fit to the observer data representing the relationship between the port sampling estimates of species composition and the observer estimates.

Species composition is expected to be correlated in space and time. This correlation is addressed above through the large-scale definitions of the domains. However, modelling finer scale correlations using spatial-temporal models may improve the estimates. This approach is used in the analysis of CPUE data to develop indices of abundance for tunas in the EPO (e.g. Xu et al. 2019).

For a design-based approach to species catch estimation, such as that shown above in eq. 4 but adapted to the proposed IPSP sampling protocol, an estimator of the variance on the estimated species proportions can be developed using the sampling weights of the proposed IPSP sampling protocol, because that protocol is a probability sampling protocol. The variance on the species catch estimates has three components: among trips; among wells of a trip; and within wells. Estimation of the among-trip variance is made possible because three cluster samples of trips would be obtained (2 cluster samples would be the minimum to be able to estimate variance). Estimation of the among-well variance component is possible because three wells per trip will be sampled; when those three wells do not fall in the same domain, the variance will be estimated from a model developed with port-sampling data for those trips with multiple sampled wells in the same domain. For the within-well variance, since only one cluster sample will be collected per well, it will either be assumed that the IPSP within-well sampling protocol had sufficient within-well coverage that this component of the variance can be assumed to be small (Lennert-Cody et al. 2024), which is a practice sometimes adopted in complex surveys, or an approximation to the within-well variance will be obtained through modelling. It is noted that the design-based approach to estimating variance on the fleet-level species catch, by domain, would apply to the sampleable population of wells, which for 2022 – 2023 was, on average, about 66% of all Class-6 vessel wells.

For purse-seine Class 1-5 vessels, which are not currently included in the proposed IPSP sampling protocol, but may be sampleable with additional cooperation from vessels (see SAC-16-05 for details), species catch estimates for domains with insufficient sample data might be estimated with model-based approaches similar to those described above. However, any well-level model would largely have to be developed between port-sampling and logbook data.

CONCLUSIONS

Based on the results of the simulation study presented in this document, it is concluded that the proposed IPSP sampling protocol would be expected to:

a) Provide sufficient data with which to make reliable design-based estimates of species catch, and their variances, for relatively common species in the dominant domains of OBJ and NOA fisheries. Although not tested in the simulation, the same is expected for the DEL fishery, which is largely monospecific, given the percentage of DEL domains with at least 10 wells sampled under the simulated IPSP protocol.

b) Generate improved port-sampling data that can be used as part of model-based estimation methods for fleet-level species catch, using multiple data sources.

c) Provide a sampling framework that can be adapted, if necessary, to: address specific data need for model-based estimation methods, such as might be beneficial for NOA-set wells; and, address any further data needs for estimation of catch size composition.

ACKNOWLEDGEMENTS

The authors extend a special thank you to Marti McCracken for discussions on survey sampling.

REFERENCES

Duparc, A., Cauquil, P., Depetris, M., Dewals, P., Floch, L., Gaertner, D., Hervé, A., Lebranchu, J., Marsac, F., Bach, P. 2018. Assessment of accuracy in processing purse seine tropical tuna catches with the T3 methodology using French fleet data. IOTC-2018-WPTT20-16.

Duparc, A., Aragno, V., Depetris, M., Floch, L., Cauquil, P., Lebranchu, J., Gaertner, D., Marsac, F., Bach, P., 2020. Assessment of the species composition of major tropical tunas in purse seine catches: a new modelling approach for the tropical tuna treatment processing (Case of the French fleet in Atlantic Ocean). Collect. Vol. Sci. Pap. ICCAT 76 (6), 951–982. SCRS/2019/189.

Erciulescu, A.L., Cruze, N.B., Nandram, B. 2019. Model-based county level crop estimates incorporating auxiliary sources of information. Journal of the Royal Statistical Society Series A 182, Part 1, 283 – 303.

Lennert-Cody, C.E., De La Cadena, C., McCracken, M., Chompoy, L., Vogel, N.W., Maunder, M.N., Wiley, B.A., Altamirano Nieto, E., Aires-da-Silva, A. 2024. Within-well patterns in bigeye tuna catch composition and implications for purse-seine port-sampling and catch estimation in the Eastern Pacific Ocean. Fisheries Research https://doi.org/10.1016/j.fishres.2024.107079

Lohr, S.L. 2022. Sampling: Design and Analysis, 3rd Edition. CRC Press, Boca Raton. 650 pp.

Porter, A.T., Holan, S.H., Wikle, C.K., Cressie, N. 2014. Spatial Fay-Herriot models for small area estimation with functional covariates. Spatial Statistics 10: 27 – 42.

R Core Team 2024. _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Wood, S.N. 2017. Generalized Additive Models: An Introduction with R, 2nd Edition. Chapman and Hall/CRC. 496 pp.

Xu, H., Lennert-Cody, C.E., Maunder, M.N. and Minte-Vera, C. 2019. Spatiotemporal dynamics of the dolphin-associated purse-seine fishery for yellowfin tuna (Thunnus albacares) in the eastern Pacific Ocean. Fisheries Research 213: 121-131.

TABLE 1. Approximate coverage of sampleable trips and wells under the TPS protocol, for Class-6 vessel trips making any sets in the EPO. The number of sampleable wells, and by extension, trips, does not include those wells that could not be set-linked to observer RDL data but were actually sampled. Details of the definition of 'sampleable' can be found in the appendix.

Year	Number trips sampled	Number sampleable trips	% Sampled trips of Sampleable	Total number trips	% Sampled trips of Total	% Sampleable trips of Total	Number wells sampled	Number sampleable wells	% Sampled wells of Sampleable	Total wells	% Sampled wells of Total	% Sampleable wells of Total
2023	416	720	58%	857	49%	84%	734	6802	11%	11581	6%	59%
2022	477	687	69%	815	59%	84%	889	5650	16%	10323	9%	55%
2019	515	748	69%	830	62%	90%	902	6205	15%	10102	9%	61%
2018	465	671	69%	789	59%	85%	765	4905	16%	9190	8%	53%
2017	471	676	70%	840	56%	80%	870	5047	17%	9250	9%	55%
2016	486	669	73%	851	57%	79%	866	5304	16%	9959	9%	53%
2015	524	733	71%	849	62%	86%	903	5648	16%	9793	9%	58%
Average	479	701	68%	833	58%	84%	847	5652	15%	10028	8%	56%

TABLE 2. Number of IATTC Class-6 vessels that were sampled, had any sampleable trips under the TPS protocol, and total. Details of the definition of 'sampleable' can be found in the appendix.

Year	Number vessels sampled	Number sampleable vessels	% Sampled (of sampleable)	Total number of vessels	% Sampled of Total
2023	127	156	81%	182	70%
2022	143	161	89%	176	81%
2019	154	170	91%	180	86%
2018	139	164	85%	180	77%
2017	141	162	87%	179	79%
2016	148	166	89%	183	81%
2015	147	165	89%	172	85%

TABLE 3. Percentage of well sampled by the TPS, by quarter of the well, using the well quarter that corresponded to the length data. Using the quarter corresponding to the count data, results were similar and are not shown; on average (over years), for 2% of the sampled wells, the quarter from which the length data were collected differed from that of the count data by 1 (maximum percent, over years, was 4% in 2023, with 4 of the 7 years ≤ 1%). By row, percent values may not sum to exactly 100% due to rounding of column-specific values to integers. In some years, the table for 'ALL' sampled ports may include a few sampled wells from ports of countries other than Ecuador (ECU) and Mexico (MEX). The data shown for Ecuadorian ports does not include Monteverde, but Monteverde is included under 'ALL'.

Year	% 1st quarter of well	% 2nd quarter of well	% 3rd quarter of well	% 4th quarter of well	% missing quarter
ALL sampled ports					
2023	40%	33%	19%	8%	< 1%
2022	45%	34%	15%	5%	0%
2019	45%	33%	15%	7%	0%
2018	41%	37%	13%	9%	< 1%
2017	42%	38%	14%	6%	0%
2016	43%	32%	17%	0%	< 1%
2015	39%	33%	18%	10%	0%
ECU Ports					
2023	34%	36%	21%	9%	< 1%
2022	49%	37%	12%	2%	0%
2019	45%	37%	14%	4%	0%
2018	39%	42%	12%	6%	< 1%
2017	38%	42%	13%	6%	0%
2016	39%	35%	17%	9%	< 1%
2015	38%	35%	19%	8%	0%
MEX Ports					
2023	48%	28%	17%	7%	0%
2022	39%	30%	20%	11%	0%
2019	46%	25%	19%	11%	0%
2018	46%	27%	16%	11%	0%
2017	50%	29%	14%	7%	0%
2016	53%	26%	14%	7%	0%
2015	41%	29%	17%	13%	0%

TABLE 4. Percentage of tropical tuna catch (YFT + SKJ + BET) for Class-6 vessels from EPO sets that was in sampleable wells under the TPS protocol, and into single-cell wells (i.e. single area x month x set type) without regard for port of unloading (shown in parentheses), by set type and year.

Year	% Sampleable of Total DEL-wells	% Sampleable of Total NOA-wells	% Sampleable of Total OBJ-wells
2023	55% (69%)	54% (62%)	60% (68%)
2022	49% (63%)	50% (55%)	57% (65%)
2019	58% (61%)	62% (67%)	63% (66%)
2018	52% (60%)	45% (52%)	54% (63%)
2017	36% (50%)	60% (66%)	58% (68%)
2016	37% (54%)	49% (61%)	62% (71%)
2015	50% (58%)	58% (63%)	61% (66%)

TABLE 5. Approximate sampling coverage of the sampleable population under the 450 trips/1350 wells IPSP protocol, that would have occurred in 2022 and 2023, based on the 4 ports where the IPSP program would operate (Manta, Posorja, Mazatlán, Manzanillo).

	Percent of Trips	Percent of wells
2023	69%	17%
2022	73%	20%

TABLE 6. Approximate coverage of domains, per simulation results (average, over samples, for 1st population), at two levels of number of wells sampled per domain: at last 10 wells; at least 2 wells. By set type, the total number of possible domains, which may not actually have any tropical tuna catch in a given year, is equal to 12 months x number of areas; OBJ-BET, OBJ-YFT, NOA-YFT, OBJ-SKJ: 60 possible domains; NOA-SKJ, DEL-YFT: 48 possible domains. Dashed lines indicate simulations not run because the general outcomes can be easily inferred from other simulation runs shown in the table.

	Number (%) of domains, out of those with tropical tuna catch, with ≥ 10 wells sampled (2022; 2023)	Number (%) of domains, out of those with tropical tuna catch, with ≥ 2 wells sampled (2022; 2023)
OBJ domains for BET assessment		
450 trips (3 clusters); 3 wells/trip (1350 wells)	32 of 59 (54%); 30 of 60 (50%)	56 of 59 (95%); 54 of 60 (90%)
Almost all trips (3 clusters); 3 wells/trip	43 of 59 (73%); 39 of 60 (65%)	---------
450 trips (3 clusters); 6 wells/trip (2700 wells)	46 of 59 (78%); 45 of 60 (75%)	---------
200 trips (2 clusters); 2 wells/trip (400 wells)	7 of 59 (12%); 6 of 60 (10%)	---------
OBJ domains for YFT assessment		
450 trips (3 clusters); 3 wells/trip (1350 wells)	21 of 42 (50%); 25 of 43 (58%);	35 of 42 (83%); 37 of 43 (86%)
OBJ domains for SKJ assessment		
450 trips (3 clusters); 3 wells/trip (1350 wells)	35 of 52 (67%); 30 of 53 (57%)	48 of 52 (92%); 46 of 53 (87%)
NOA domains for SKJ assessment		
450 trips (3 clusters); 3 wells/trip (1350 wells)	7 of 26 (27%); 4 of 34 (12%)	16 of 26 (62%); 20 of 34 (59%)
NOA domains for YFT assessment		
450 trips (3 clusters); 3 wells/trip (1350 wells)	3 of 41 (7%); 3 of 43 (7%)	24 of 41 (58%); 22 of 43 (51%)
DEL domains for YFT assessment		
450 trips (3 clusters); 3 wells/trip (1350 wells)	17 of 46 (37%); 18 of 46 (39%)	36 of 46 (78%); 37 of 46 (80%)

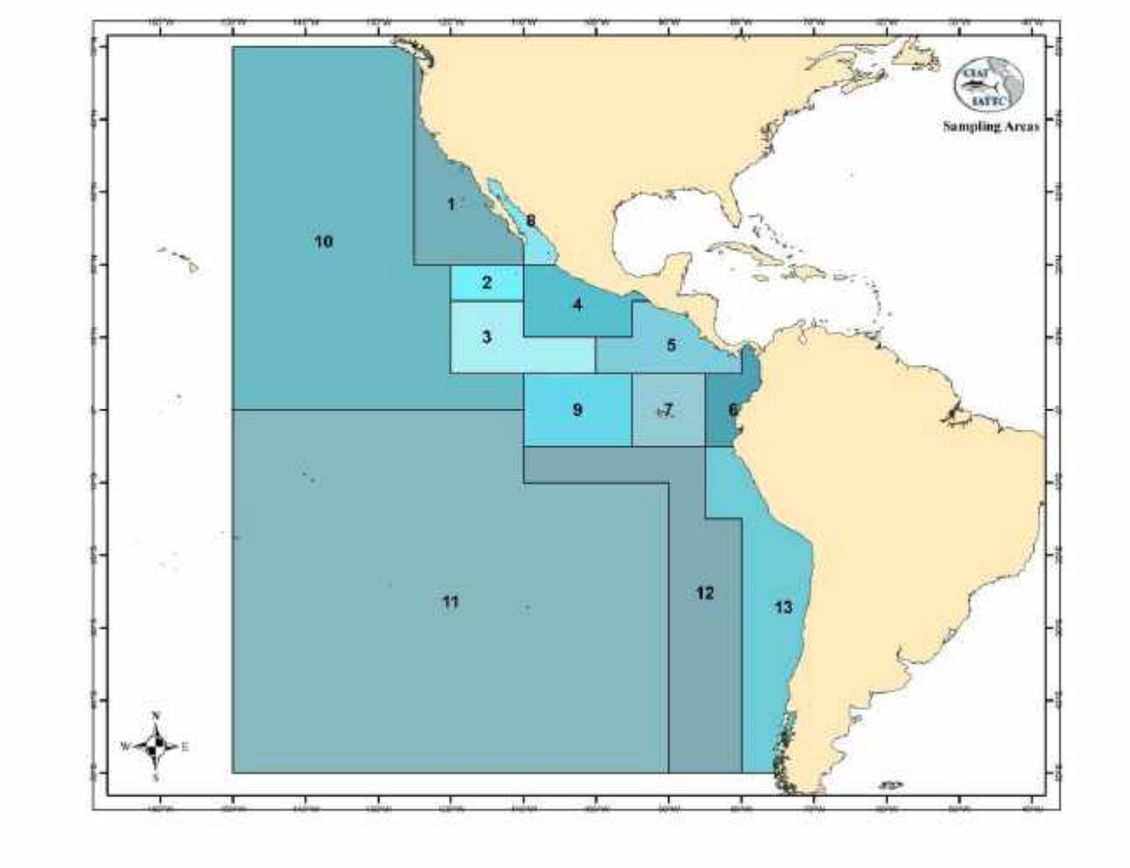

FIGURE 1. The 13 market measurement areas used in the TPS protocol.

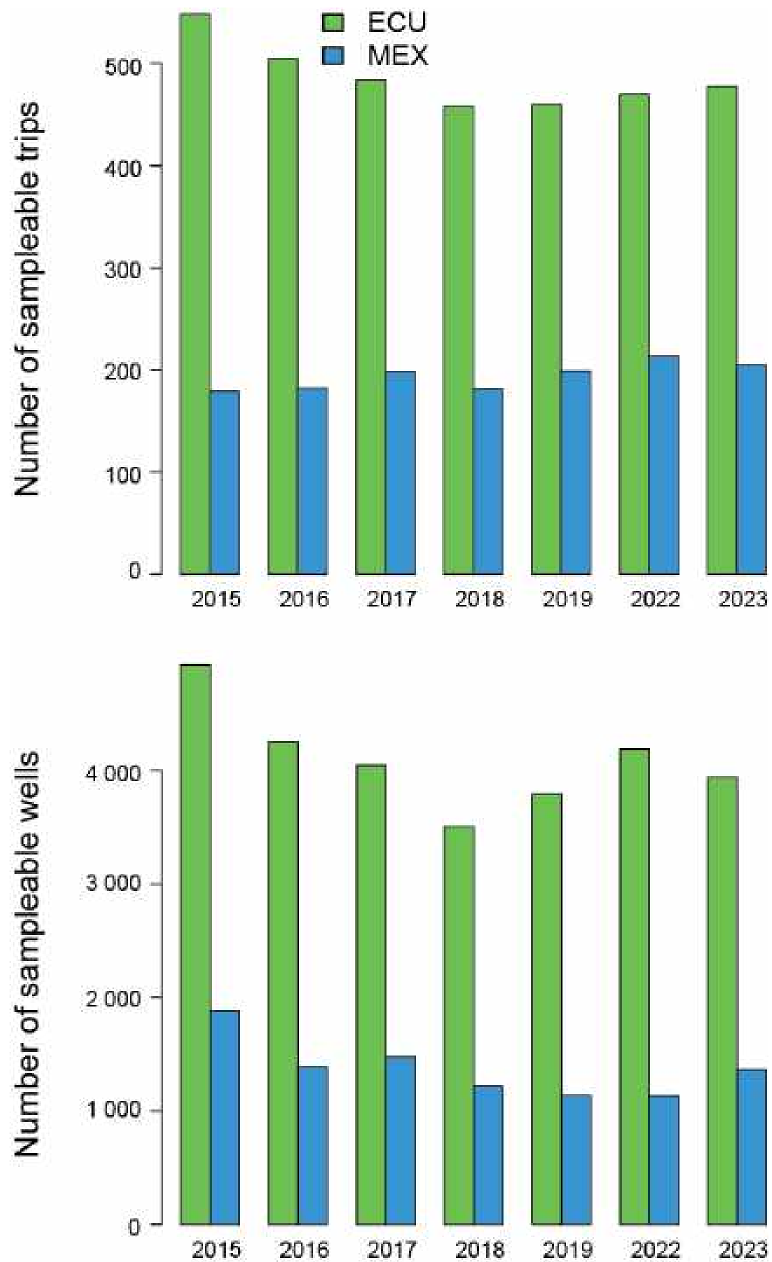

FIGURE 2. Number of sampleable trips and wells, under the TPS protocol, by year and country. "ECU": Ecuadorian ports where TPS data were collected (excluding Monte Verde); "MEX": Mexican ports were TPS were collected.

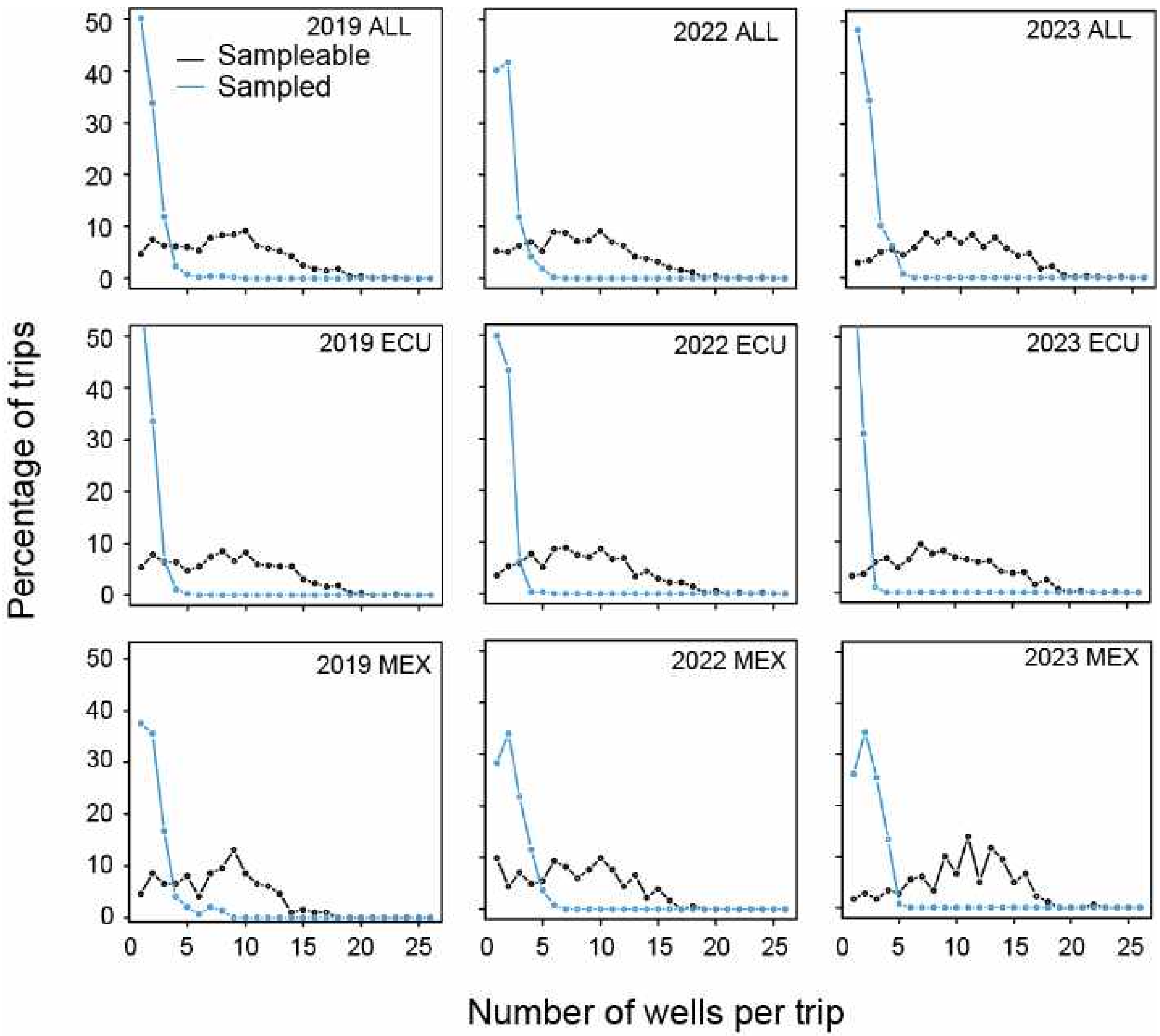

FIGURE 3. Frequency distributions of the number of 'sampleable' (black circles) under the TPS protocol, and sampled (blue circles) wells per trip, Class-6 EPO all set types, for 2019, 2022 and 2023 (results for other years are similar and not shown). ALL: all sampled ports (1st row); ECU: sampled ports in Ecuador (2nd row; excluding Monte Verde); MEX: sampled ports in Mexico (3rd row). In some years, ALL includes a few sampled wells from ports of countries other than Ecuador (including Monte Verde) and Mexico. For some panels, the y-axis range was trimmed to show detail.

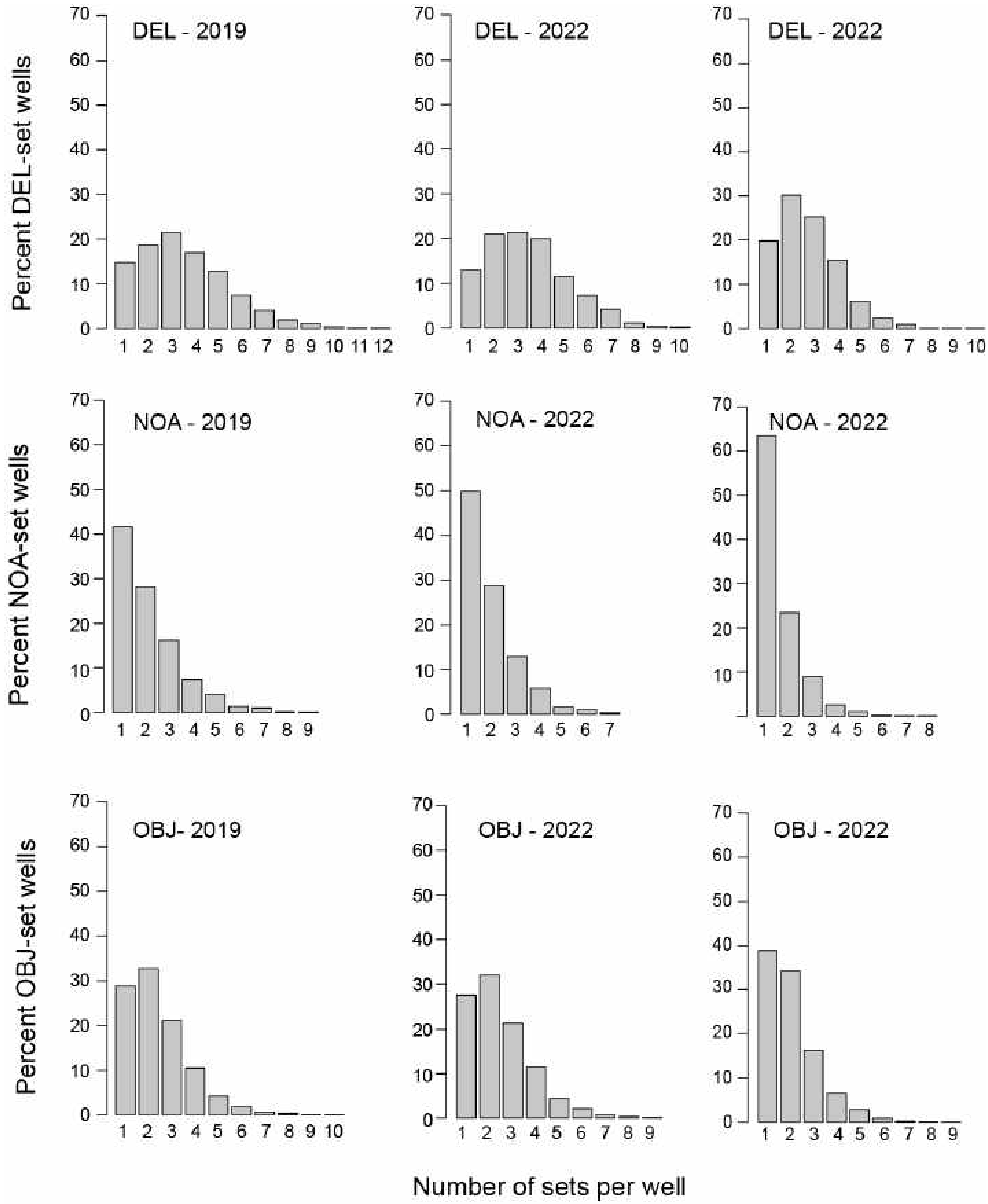

FIGURE 4. Frequency distributions of the number of sets per well, by set type, for sampleable wells, under the TPS protocol, unloaded in all sampled ports, 2019, 2022 and 2023 (Class-6, EPO) (results for other years are similar and not shown).

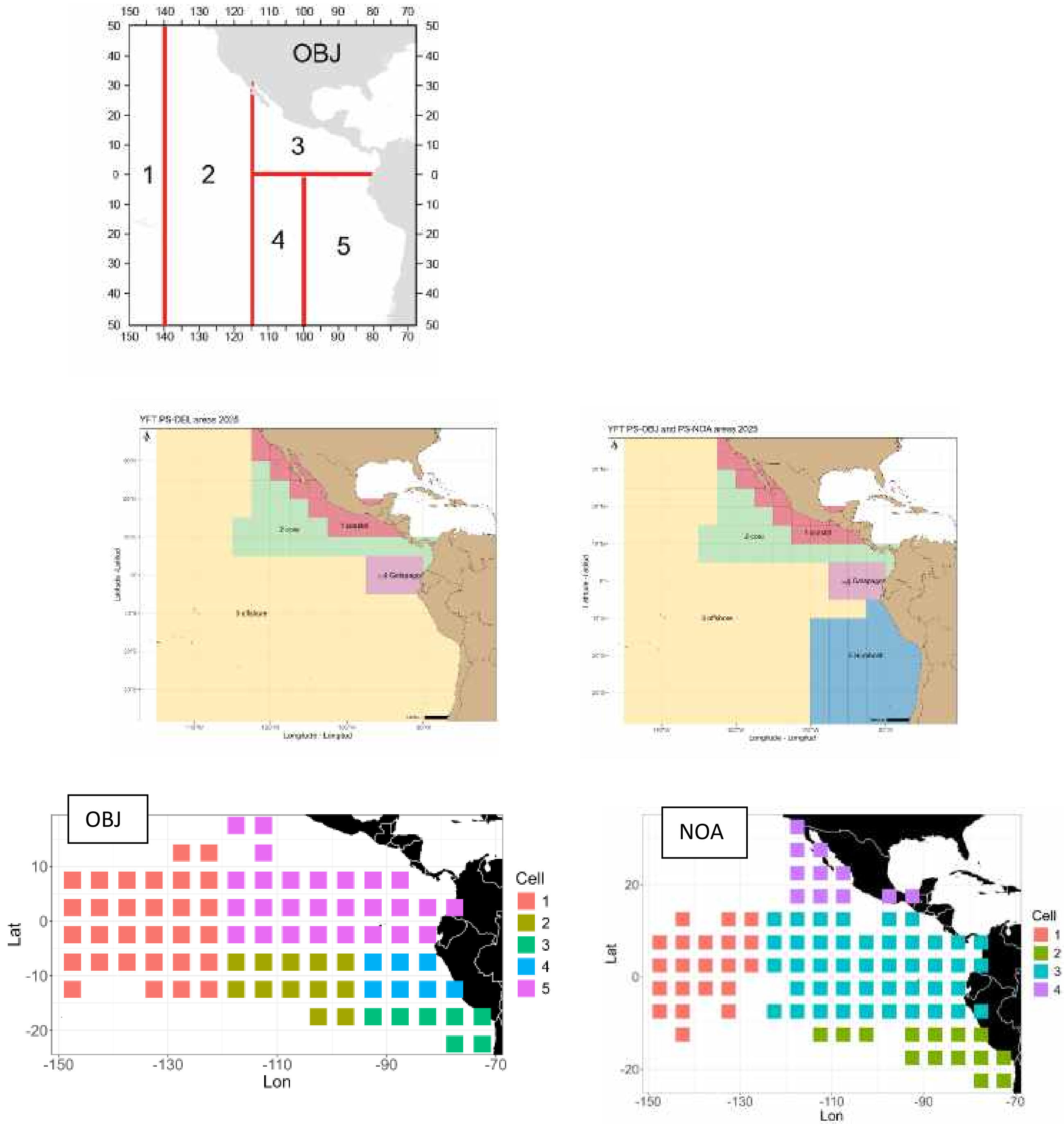

FIGURE 5. Spatial structure for tuna stock assessment fisheries. Top row: OBJ fisheries in the BET benchmark assessment [(SAC-15-02)](); middle row: DEL and NOA/OBJ areas in YFT assessment (SAC-16-03); bottom row: OBJ and NOA fisheries in the SKJ benchmark assessment ([SAC-15-04]()).

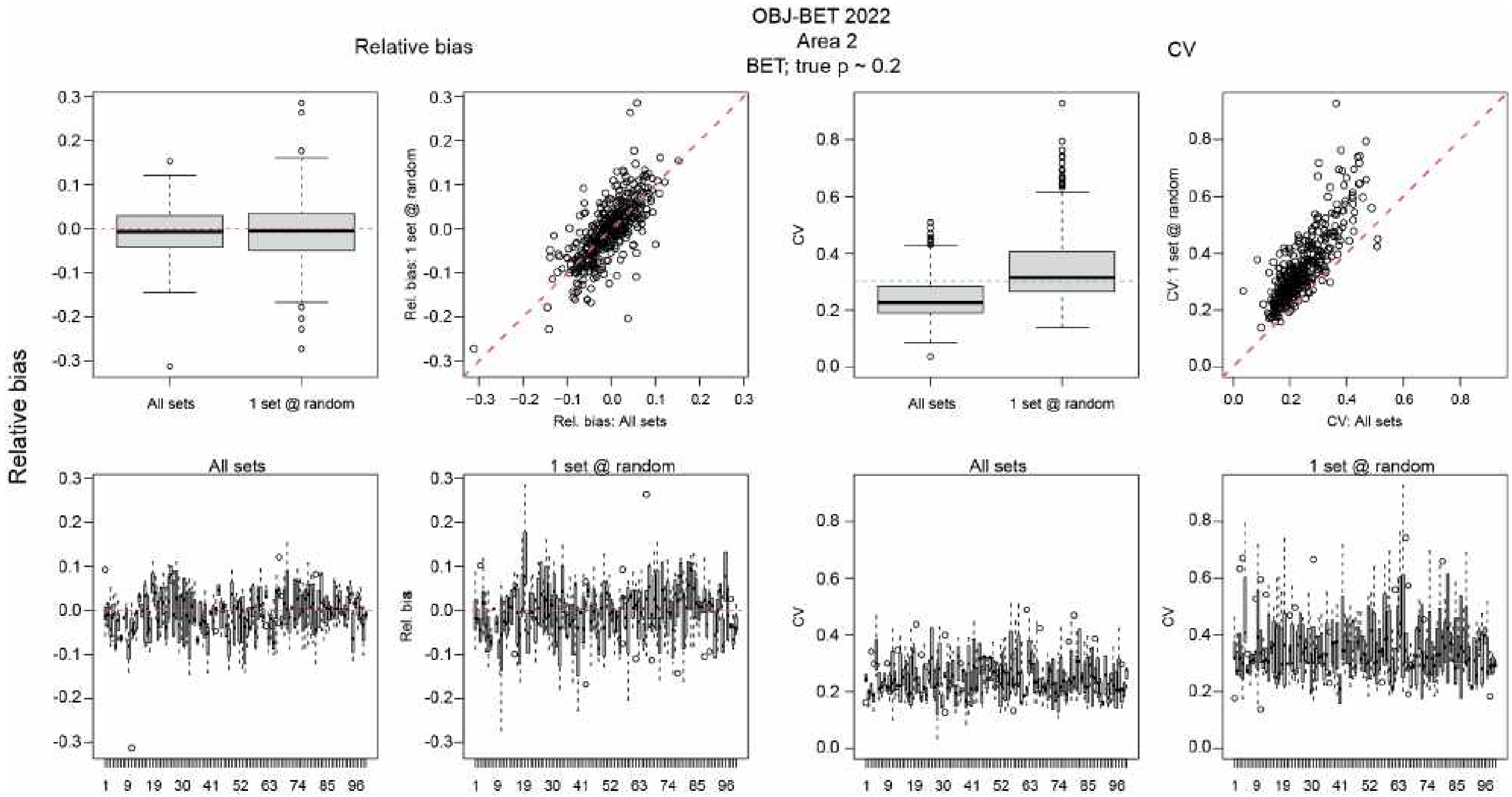

FIGURE 6. Relative bias (left 2x2 panels) and CV (right 2x2 panels) for BET for OBJ-BET Area 2 (Figure 5) for 2022, under the IPSP protocol of 3 clusters of trips, 150 trips per cluster, and 3 wells/trip. For each set of 2x2 panels: the upper left shows box-and-whisker plots of the simulation output, by within-well sampling option ('all sets'; '1 set at random'); the upper right shows a scatter plot of the simulation values for the two types of within-well sampling option, 'all sets' on the x-axis and '1 set at random' on the y-axis; the bottom row of 2 panels shows box-and-whisker plots of the simulation output by population (x-axis) for each of the 100 populations. The average, over simulated full data sets, of true species proportion for the domain, 'true p', is shown in the title of the figure.

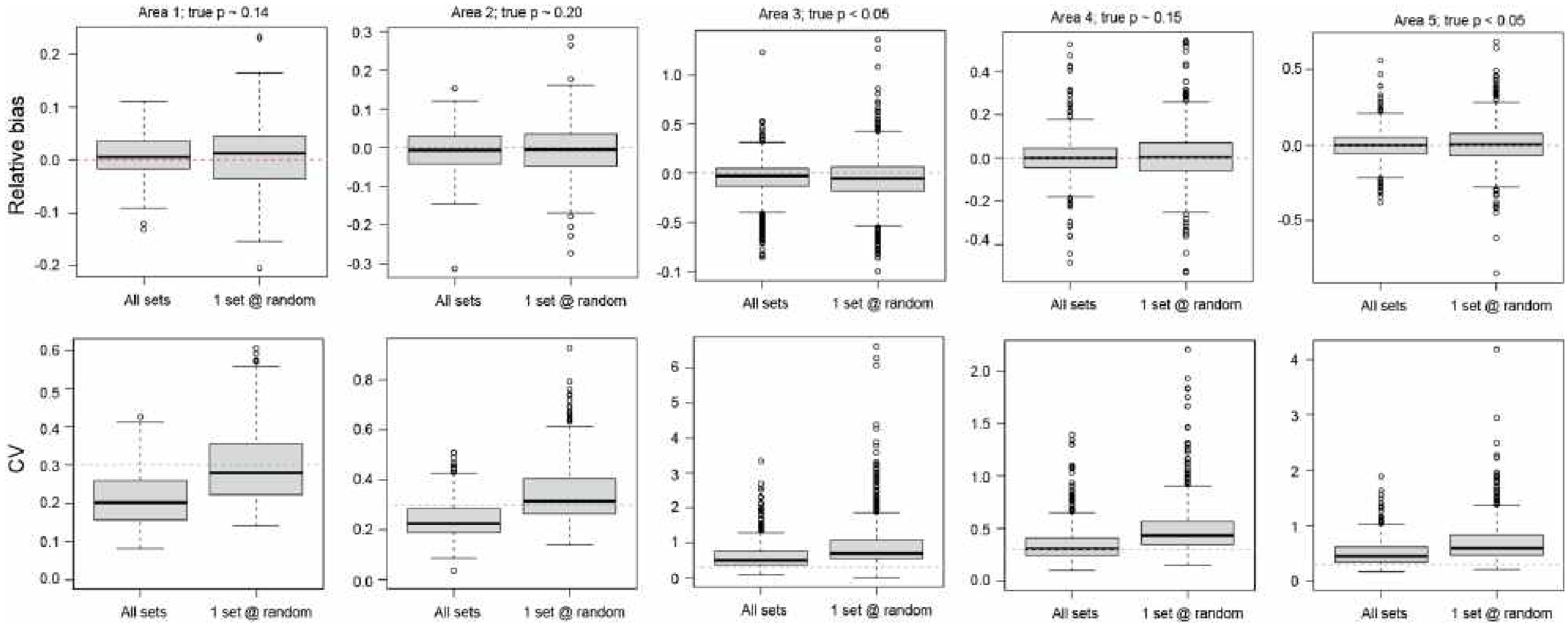

FIGURE 7. Relative bias (top row) and CV (bottom row) for the estimated fleet-level proportion of BET, for 2022 in each of the 5 OBJ-BET areas (Figure 5) for the IPSP protocol (3 clusters of trips, 150 trips/cluster, 3 wells per trip). The relative bias and CV figures are arranged by area, from area 1 on the left to area 5 on the right. The average, over simulated full data sets, of true species proportion for the domain, 'true p', is shown in the title of the figure.

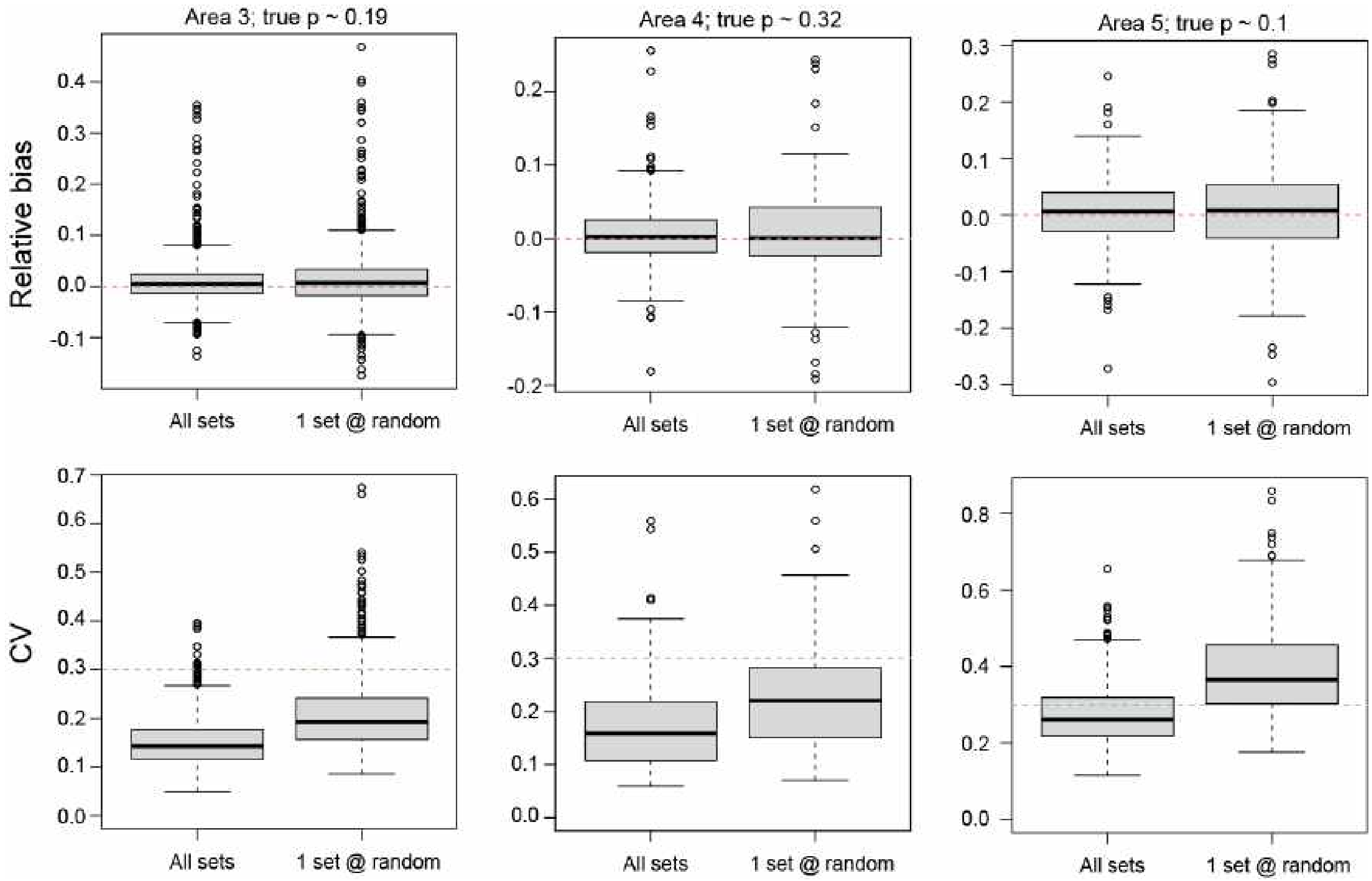

FIGURE 8. Relative bias (top row) and CV (bottom row) for the estimated fleet-level proportion of YFT in the OBJ fishery, for 2022 in OBJ-YFT areas 3 – 5 (Figure 5) under the IPSP protocol (3 clusters of trips, 150 trips/cluster, 3 wells per trip). OBJ-YFT areas 1 and 2 had fewer than 10 wells sampled per domain and estimates were not made. The relative bias and CV figures are arranged by area, from area 3 on the left to area 5 on the right. The average, over simulated full data sets, of true species proportion for the domain, 'true p', is shown in the title of the figure.

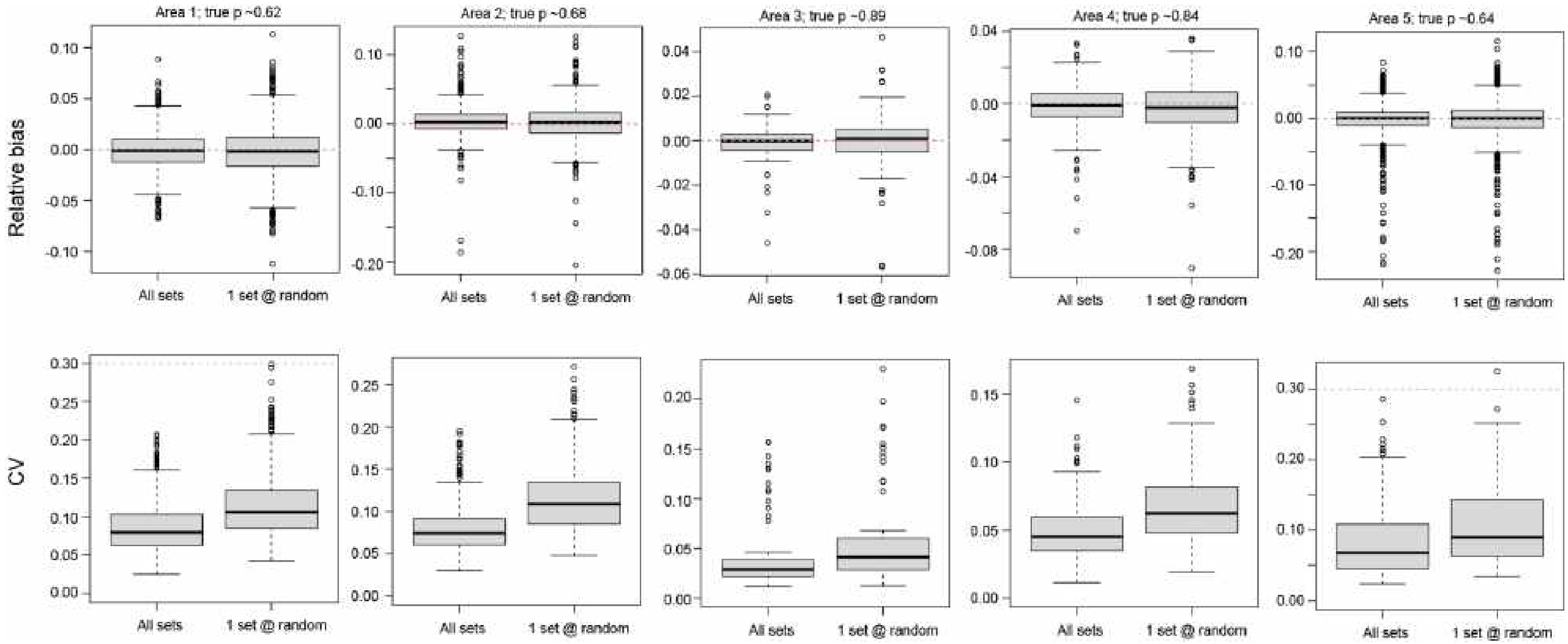

FIGURE 9. Relative bias (top row) and CV (bottom row) for the estimated fleet-level proportion of SKJ, for 2022 in each of the 5 OBJ-SKJ areas (Figure 5) for the IPSP protocol (3 clusters of trips, 150 trips/cluster, 3 wells per trip). The relative bias and CV figures are arranged by area, from area 1 on the left to area 5 on the right. The average, over simulated full data sets, of true species proportion for the domain, 'true p', is shown in the title of the figure.

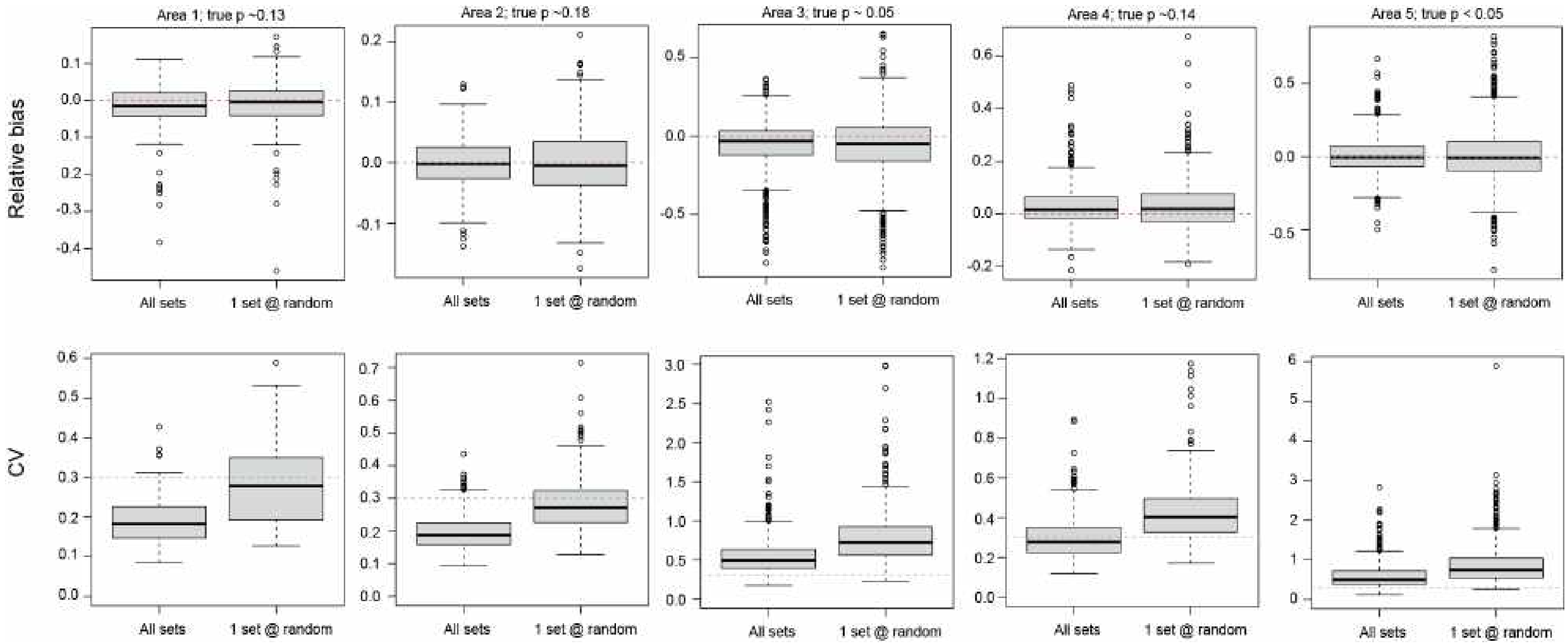

FIGURE 10. Relative bias (top row) and CV (bottom row) for the estimated fleet-level proportion of BET, for 2023 in each of the 5 OBJ-BET areas (Figure 5) for the IPSP protocol (3 clusters of trips, 150 trips/cluster, 3 wells per trip). The relative bias and CV figures are arranged by area, from area 1 on the left to area 5 on the right. The average, over simulated full data sets, of true species proportion for the domain, 'true p', is shown in the title of the figure.

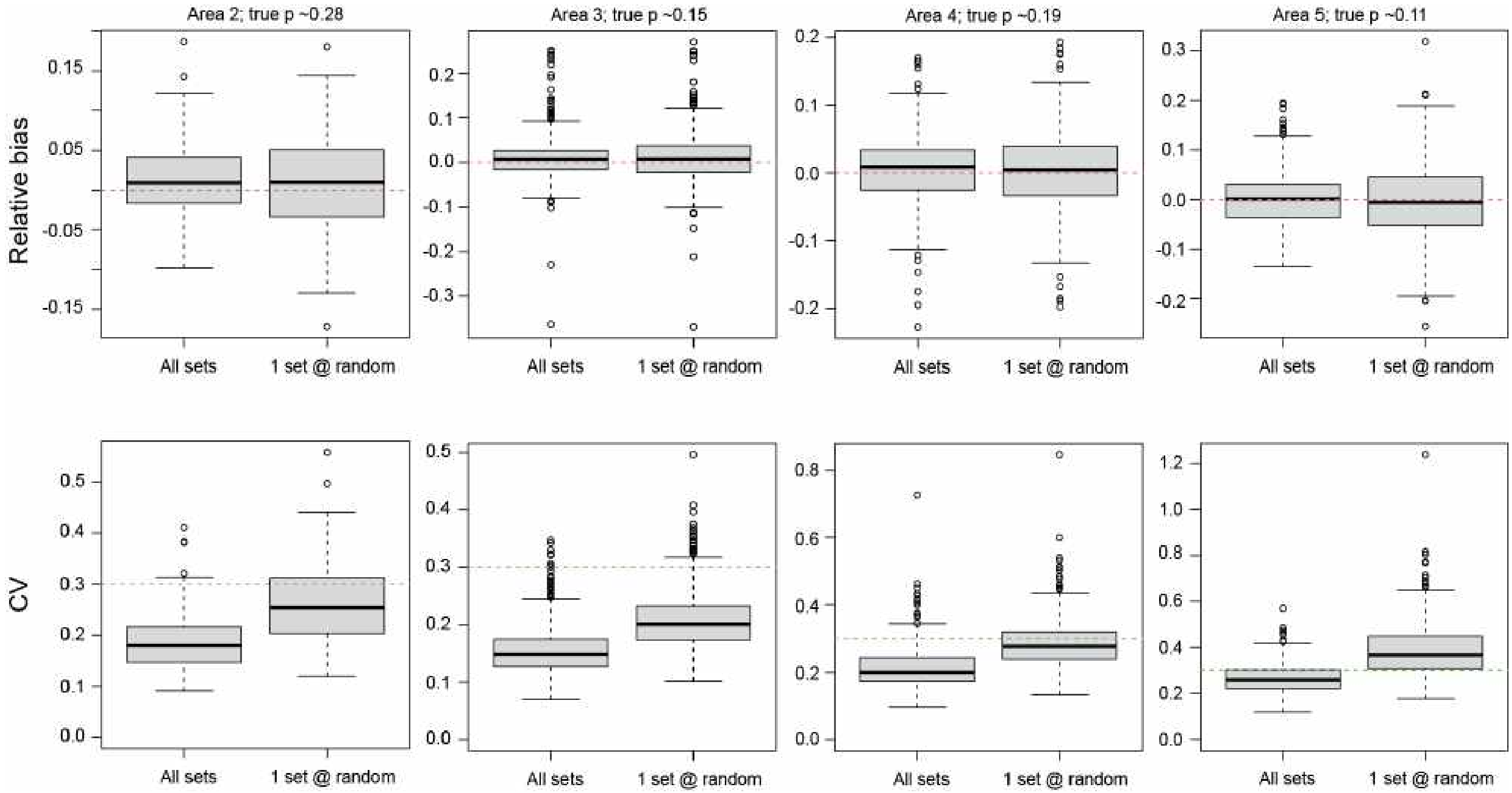

FIGURE 11. Relative bias (top row) and CV (bottom row) for the estimated fleet-level proportion of YFT in the OBJ fishery, for 2023 in OBJ-YFT areas 2 – 5 (Figure 5) under the IPSP protocol (3 clusters of trips, 150 trips/cluster, 3 wells per trip). OBJ-YFT area 1 had fewer than 10 wells sampled per domain and estimates were not made. The relative bias and CV figures are arranged by area, from area 2 on the left to area 5 on the right. The average, over simulated full data sets, of true species proportion for the domain, 'true p', is shown in the title of the figure.

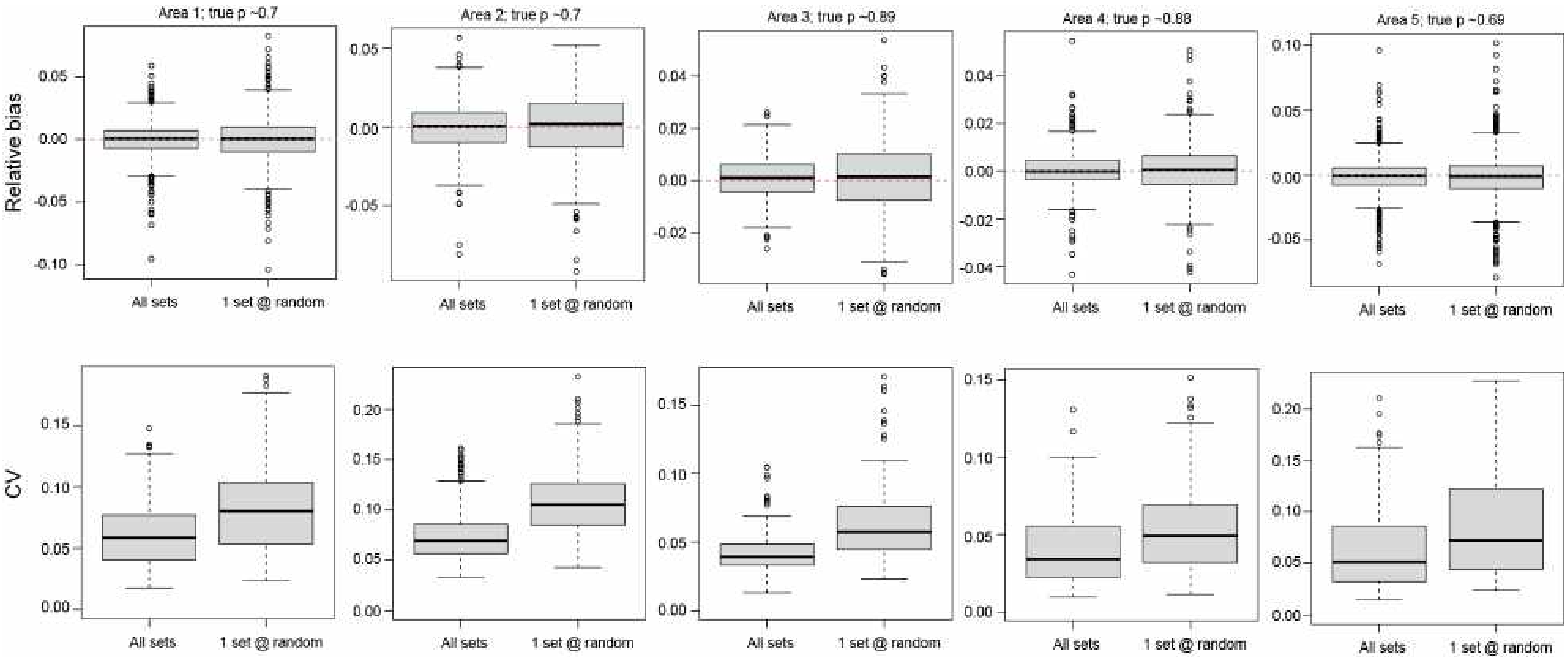

FIGURE 12. Relative bias (top row) and CV (bottom row) for the estimated fleet-level proportion of SKJ, for 2023 in each of the 5 OBJ-SKJ areas (Figure 5) for the IPSP protocol (3 clusters of trips, 150 trips/cluster, 3 wells per trip). The relative bias and CV figures are arranged by area, from area 1 on the left to area 5 on the right. The average, simulated full data sets, of true species proportion for the domain, 'true p', is shown in the title of the figure.

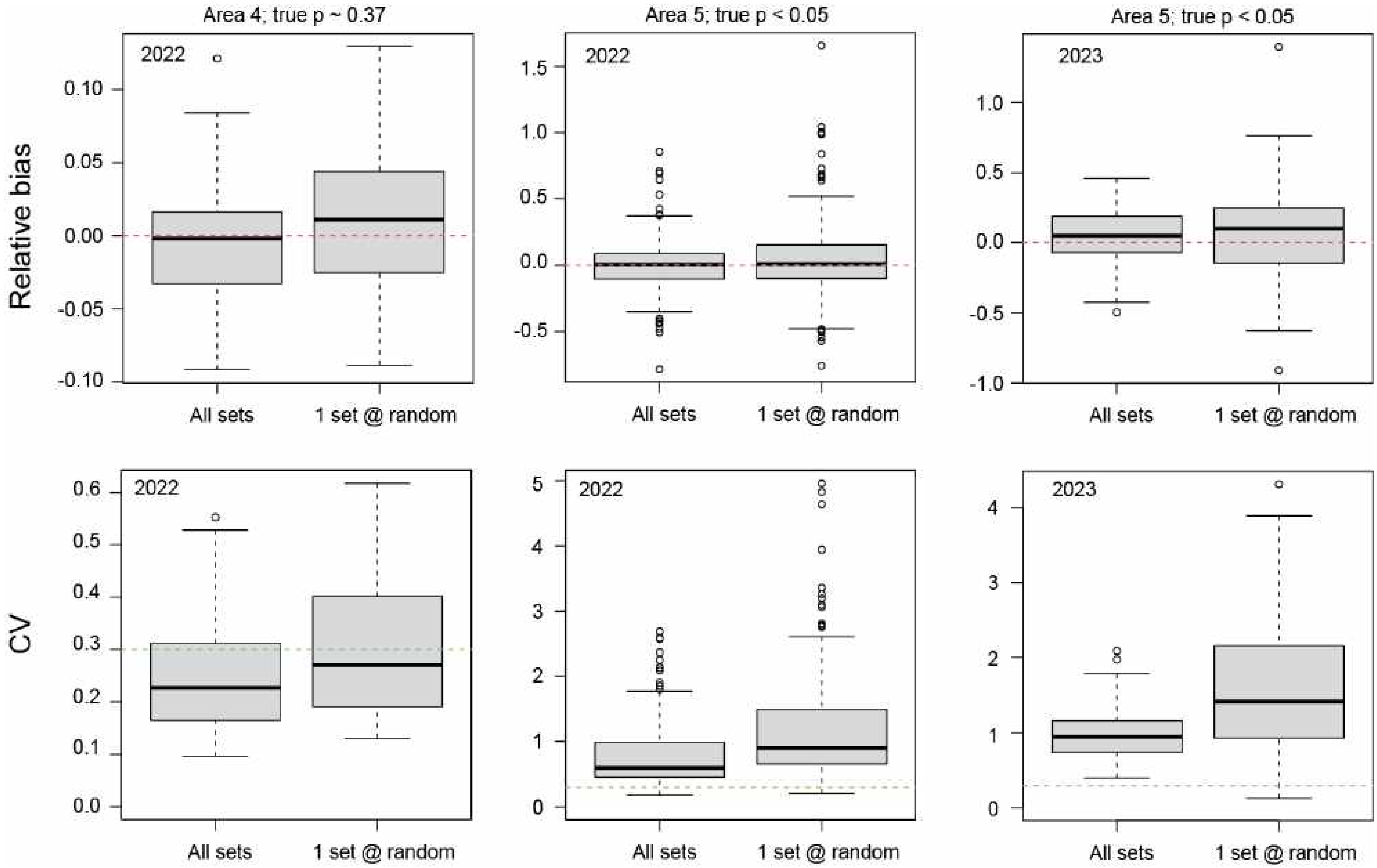

FIGURE 13. Relative bias (top row) and CV (bottom row) for the estimated fleet-level proportion of YFT in the NOA fishery, for 2022 and 2023 in those NOA-YFT areas (Figure 5) with at least 10 sampled wells per domain, under the IPSP protocol (3 clusters of trips, 150 trips/cluster, 3 wells per trip). The average, over simulated full data sets, of true species proportion for the domain, 'true p', is shown in the title of the figure.

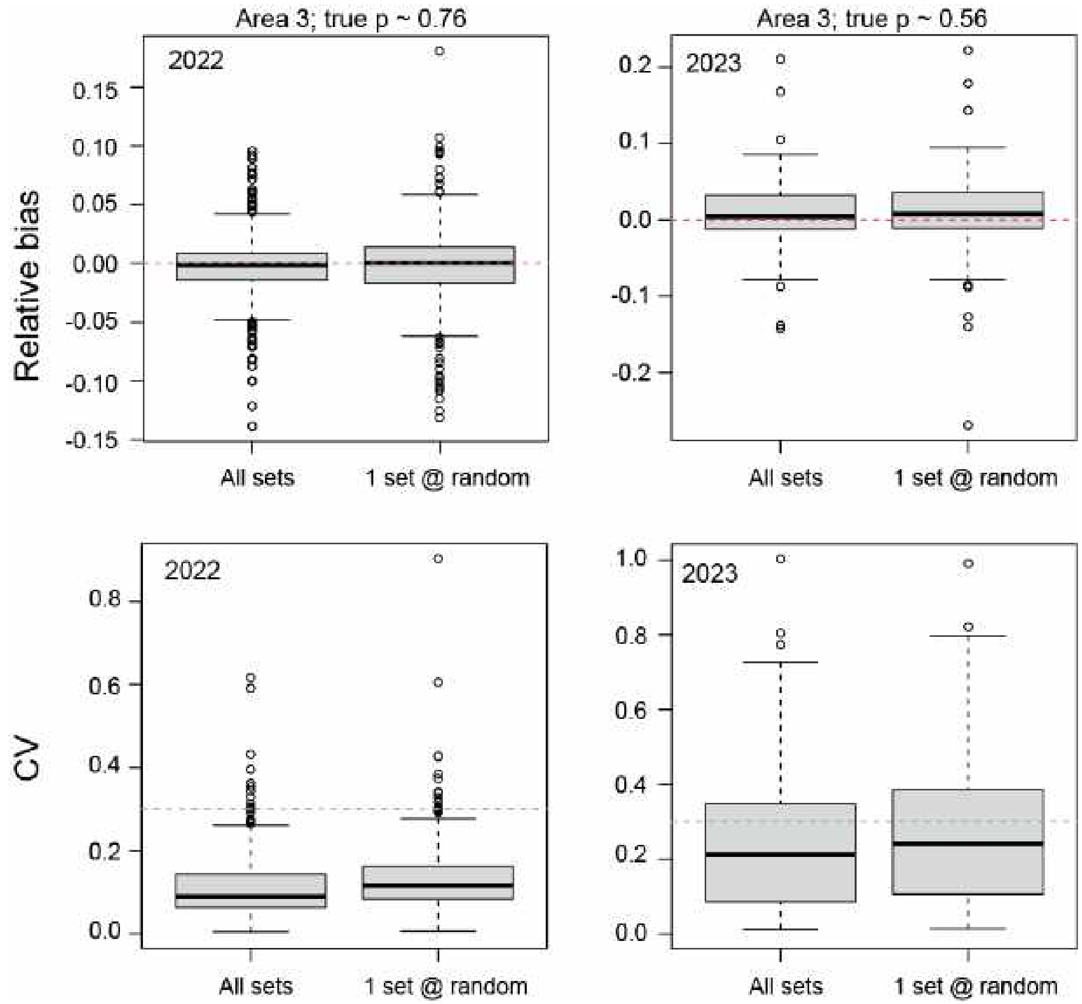

FIGURE 14. Relative bias (top row) and CV (bottom row) for the estimated fleet-level proportion of SKJ in the NOA fishery, for 2022 and 2023 in OBJ-SKJ Area 3 (Figure 5) with at least 10 sampled wells per domain, under the IPSP protocol (3 clusters of trips, 150 trips/cluster, 3 wells per trip). The average, over simulated full data sets, of true species proportion for the domain, 'true p', is shown in the title of the figure.

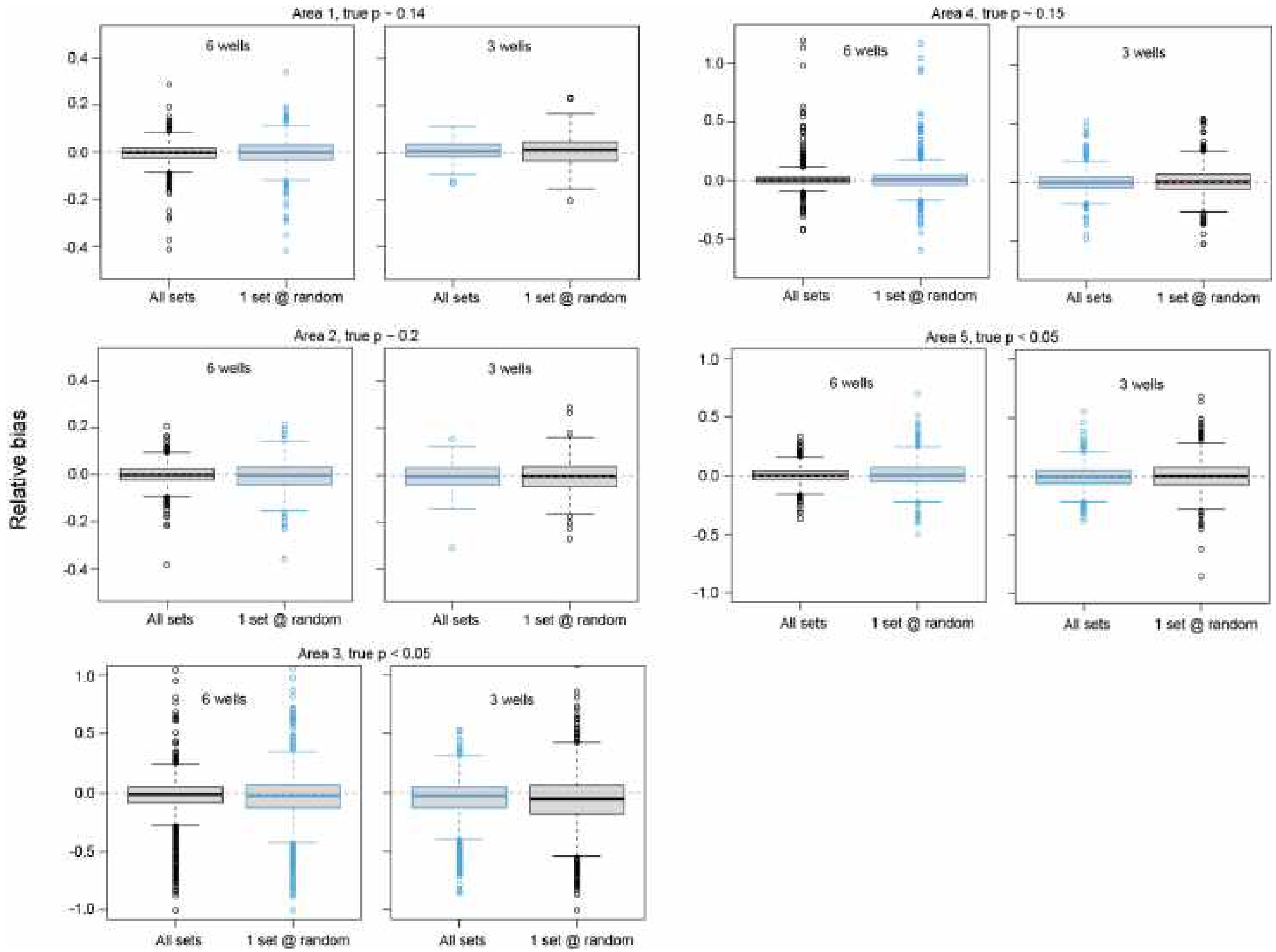

FIGURE 15. Relative bias comparison for the estimated proportion of BET in 2022 for each of the 5 OBJ-BET areas (Figure 5): 6 wells per trip versus 3 wells per trip, both with 3 clusters of 150 trips. The relative bias figures are arranged by area, from area 1 - top left row to area 5 - middle right row. The average, over simulated full data sets, of true species proportion for the domain, 'true p', is shown in the title of the figure.

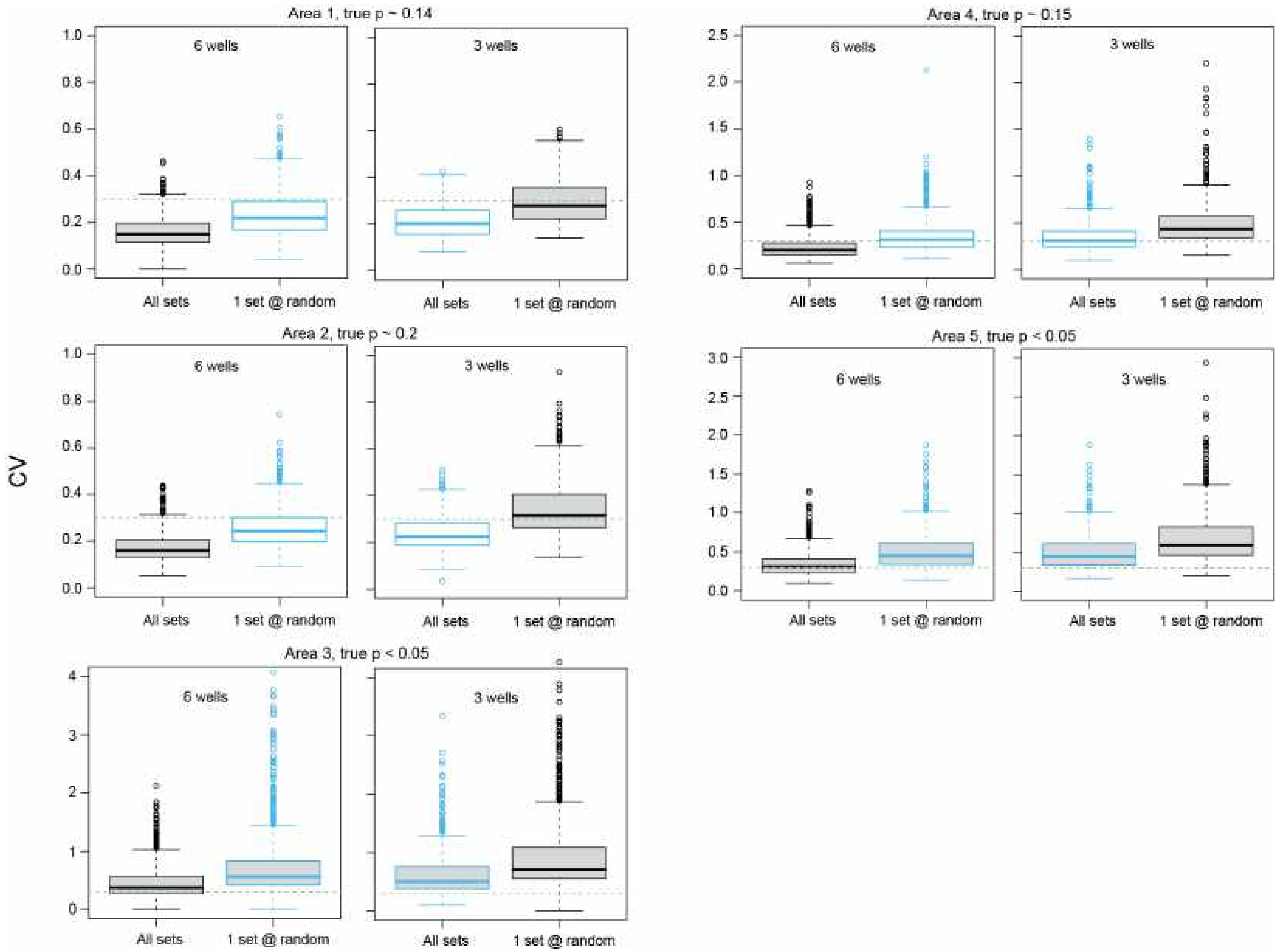

FIGURE 16. CV comparison for the estimated proportion of BET in 2022 for each of the 5 OBJ-BET areas (Figure 5): 6 wells per trip versus 3 wells per trip, both with 3 clusters of 150 trips. The CV figures are arranged by area, from area 1 - top left row to area 5 - middle right row. The average, over simulated full data sets, of true species proportion for the domain, 'true p', is shown in the title of the figure.

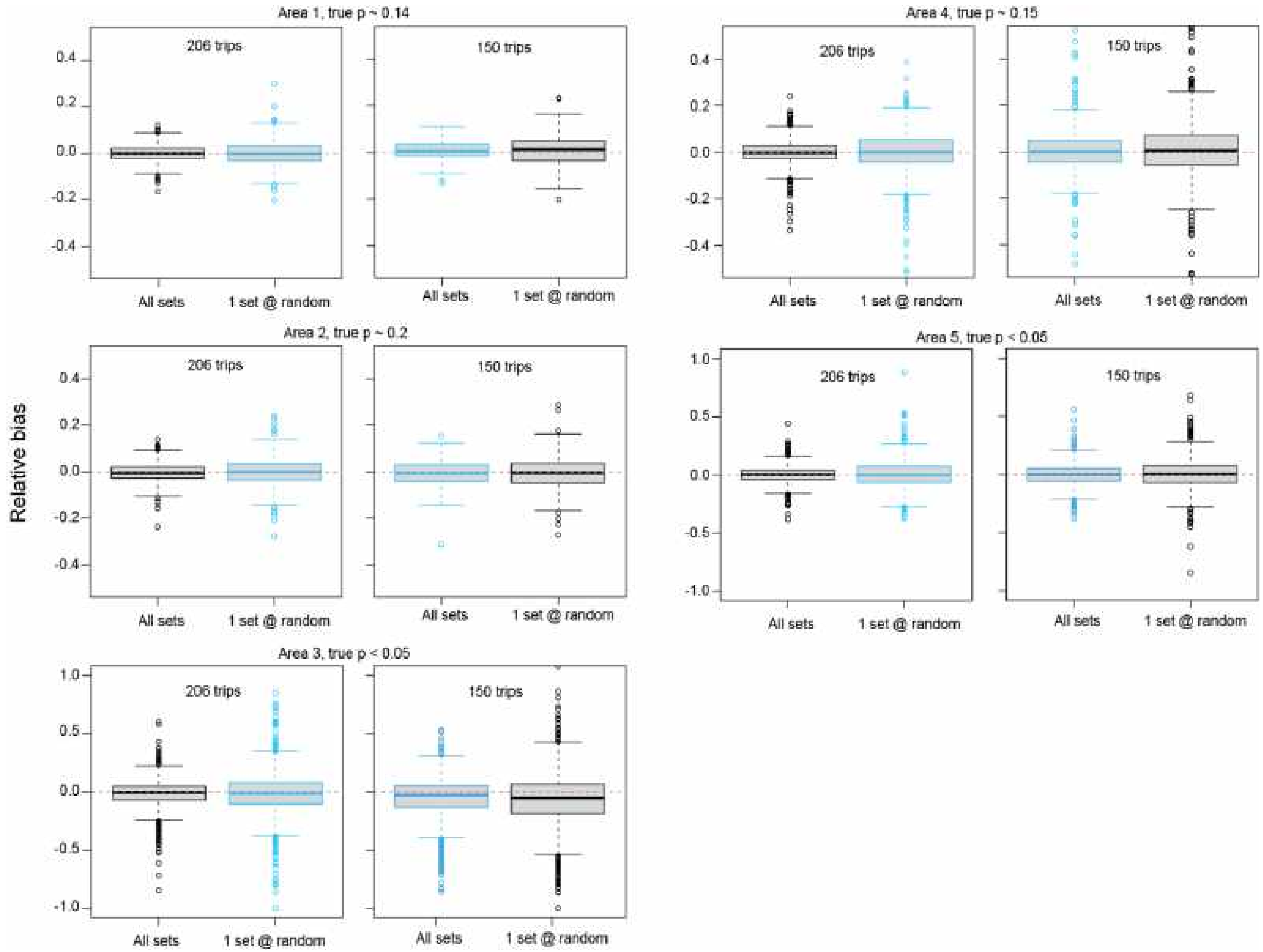

FIGURE 17. Relative bias comparison for the estimated proportion of BET in 2022 for each of the 5 OBJ-BET areas (Figure 5): 206 trips per cluster versus 150 trips per cluster, both for 3 clusters of trips and 3 wells per trip. The relative bias figures are arranged by area, from area 1 - top left row to area 5 - middle right row. The average, over simulated full data sets, of true species proportion for the domain, 'true p', is shown in the title of the figure.

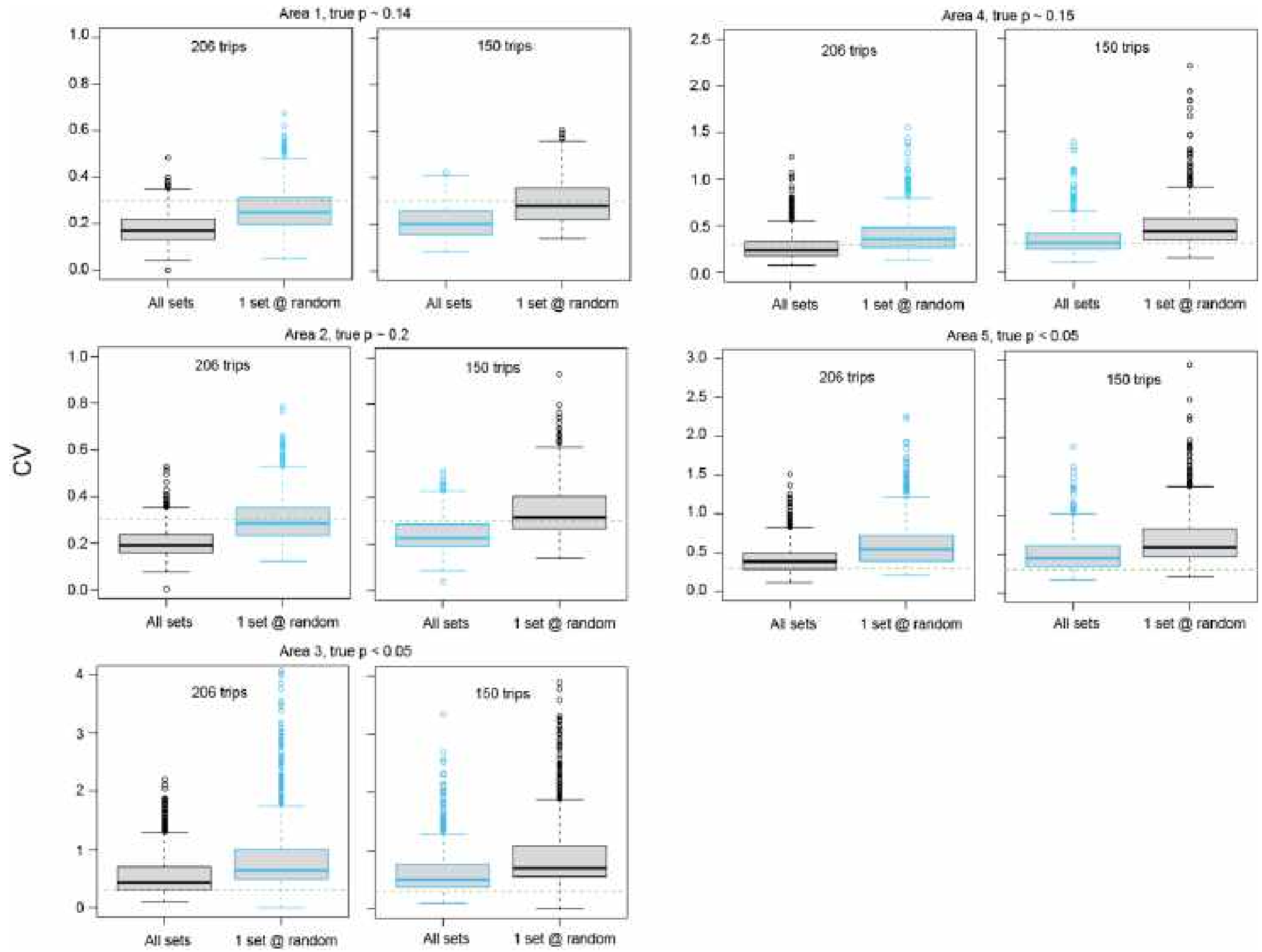

FIGURE 18. CV comparison for the estimated proportion of BET in 2022 for each of the 5 OBJ-BET areas (Figure 5): 206 trips per cluster versus 150 trips per cluster, both for 3 clusters of trips and 3 wells per trip. The CV figures are arranged by area, from area 1 - top left row to area 5 - middle right row. The average, over simulated full data sets, of true species proportion for the domain, ‘true p’, is shown in the title of the figure.

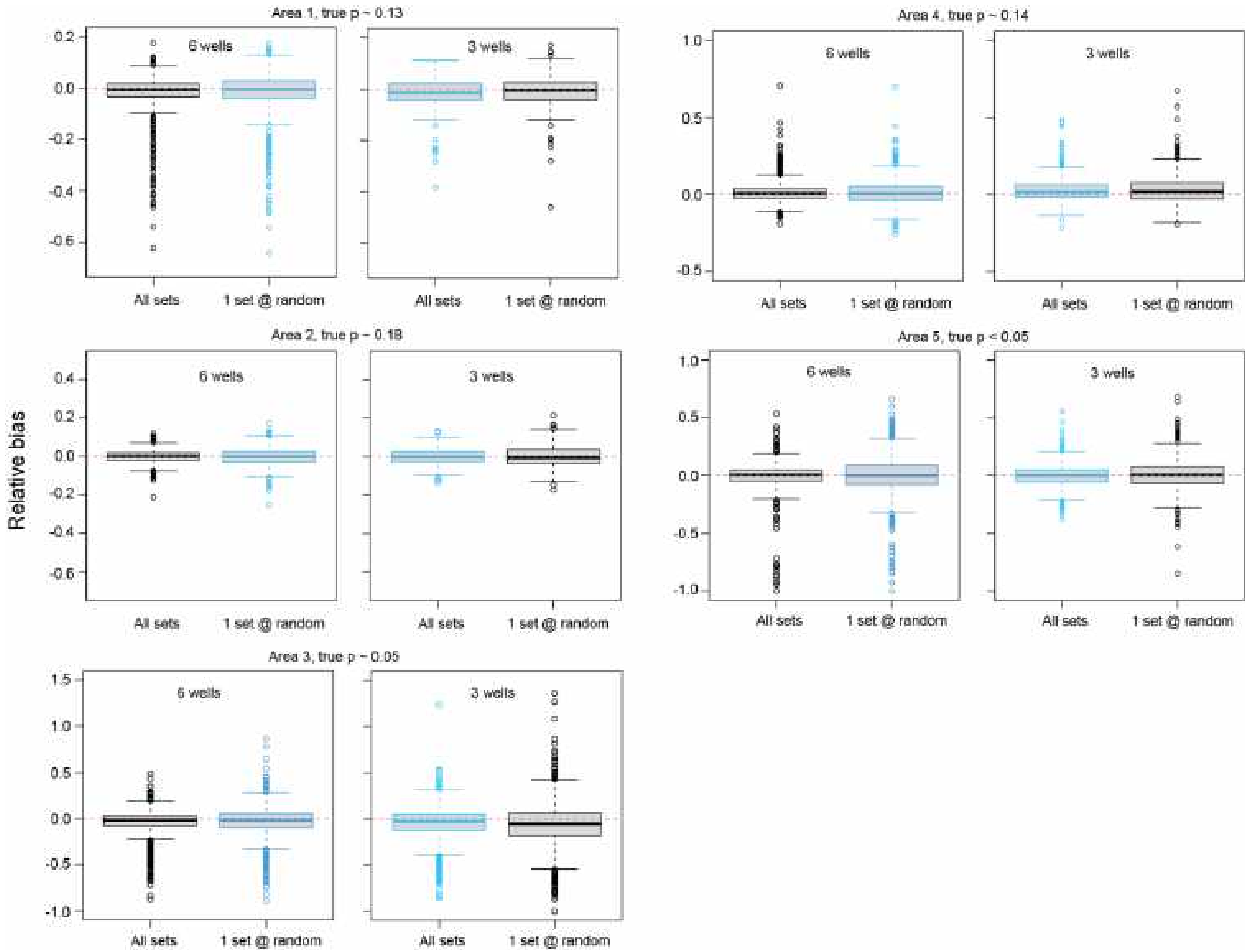

FIGURE 19. Relative bias comparison for the estimated proportion of BET in 2023 for each of the 5 OBJ-BET areas (Figure 5): 6 wells per trip versus 3 wells per trip, both with 3 clusters of 150 trips. The relative bias figures are arranged by area, from area 1 - top left row to area 5 - middle right row. The average, over simulated full data sets, of true species proportion for the domain, 'true p', is shown in the title of the figure.

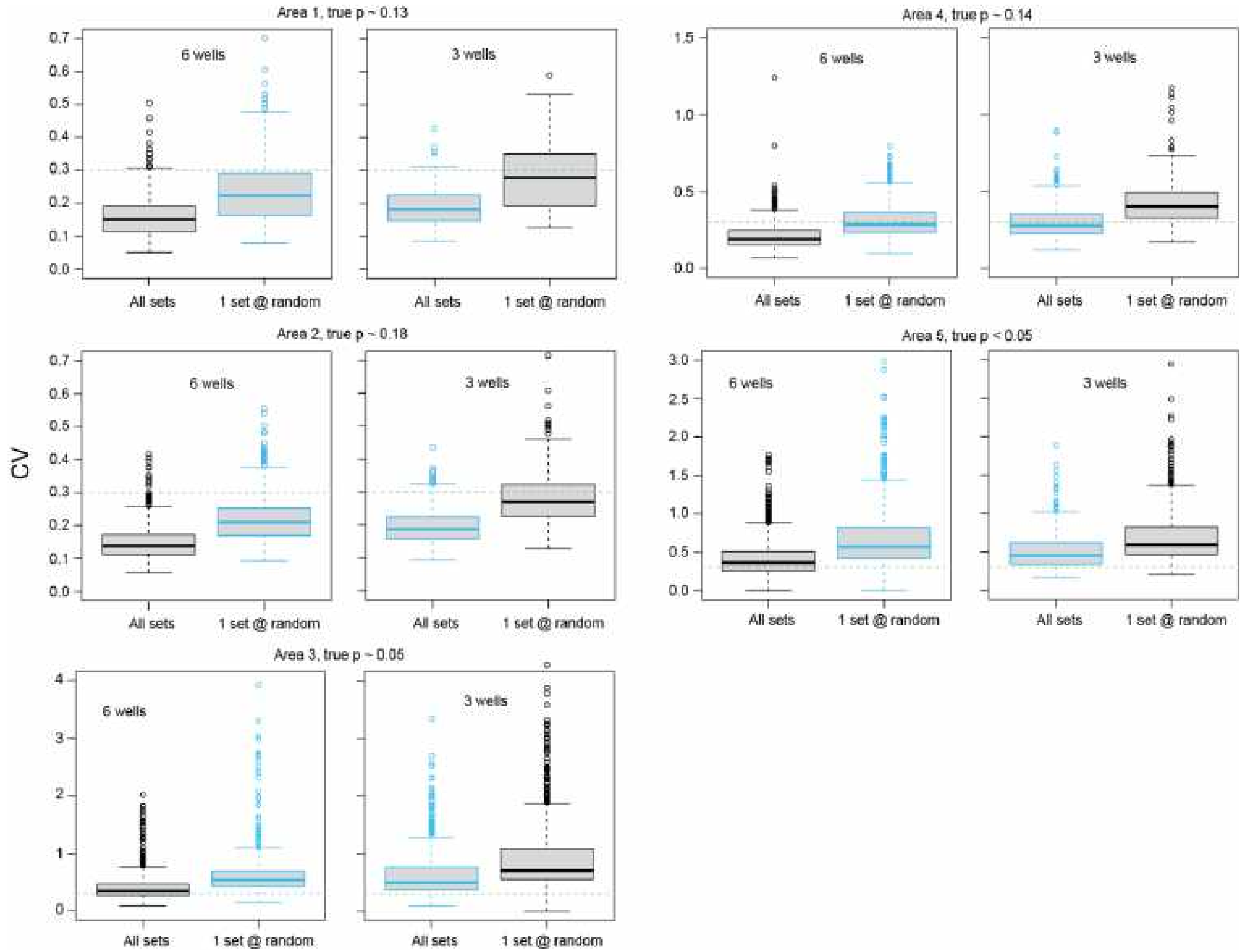

FIGURE 20. CV comparison for the estimated proportion of BET in 2023 for each of the 5 OBJ-BET areas (Figure 5): 6 wells per trip versus 3 wells per trip, both with 3 clusters of 150 trips. The CV figures are arranged by area, from area 1 - top left row to area 5 - middle right row. The average, over simulated full data sets, of true species proportion for the domain, 'true p', is shown in the title of the figure.

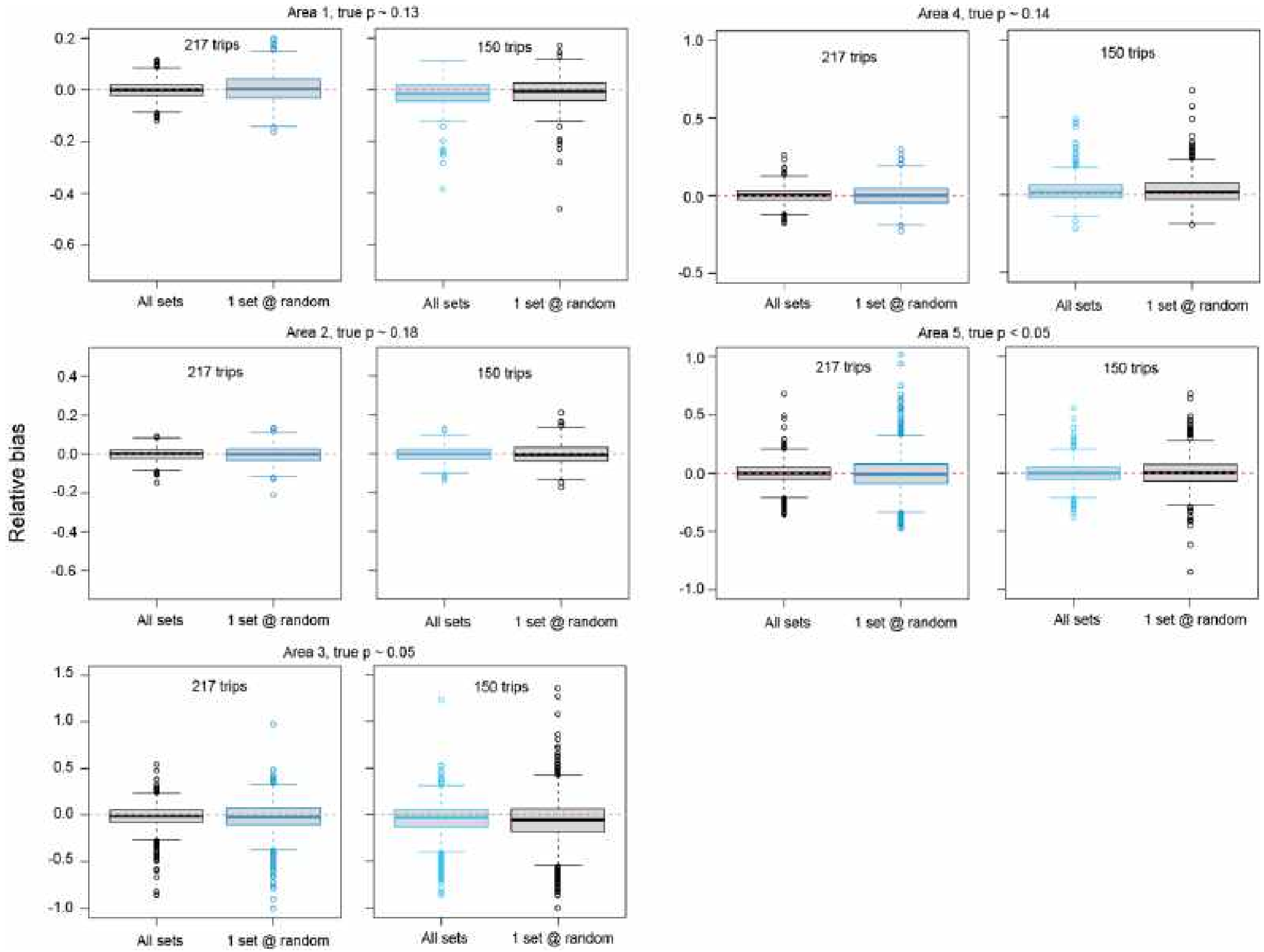

FIGURE 21. Relative bias comparison for the estimated proportion of BET in 2023 for each of the 5 OBJ-BET areas (Figure 5): 217 trips per cluster versus 150 trips per cluster, both for 3 clusters of trips and 3 wells per trip. The relative bias figures are arranged by area, from area 1 - top left row to area 5 - middle right row. The average, over simulated full data sets, of true species proportion for the domain, 'true p', is shown in the title of the figure.

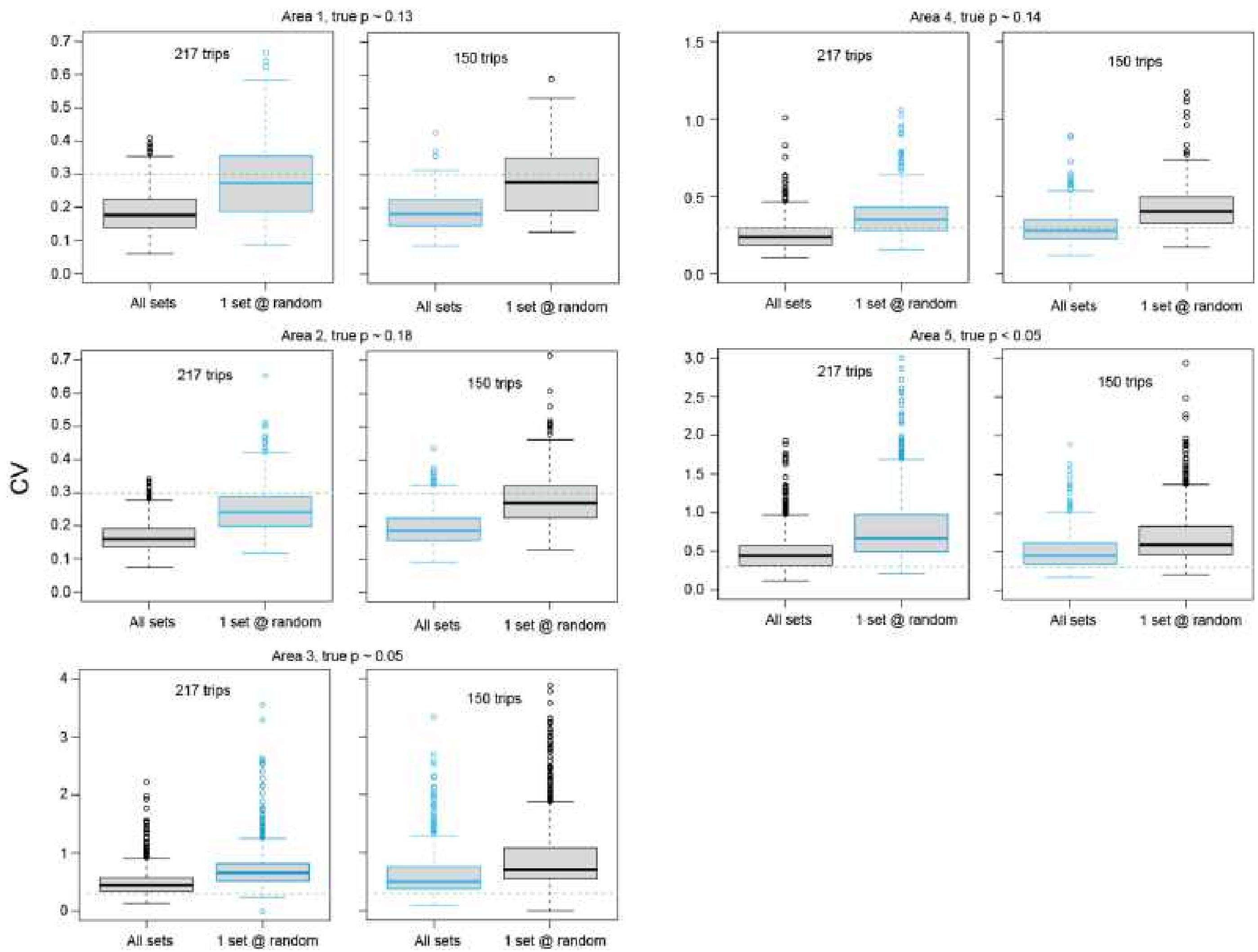

FIGURE 22. CV comparison for the estimated proportion of BET in 2023 for each of the 5 OBJ-BET areas (Figure 5): 217 trips per cluster versus 150 trips per cluster, both for 3 clusters of trips and 3 wells per trip. The CV figures are arranged by area, from area 1 - top left row to area 5 - middle right row. The average, over simulated full data sets, of true species proportion for the domain, 'true p', is shown in the title of the figure.

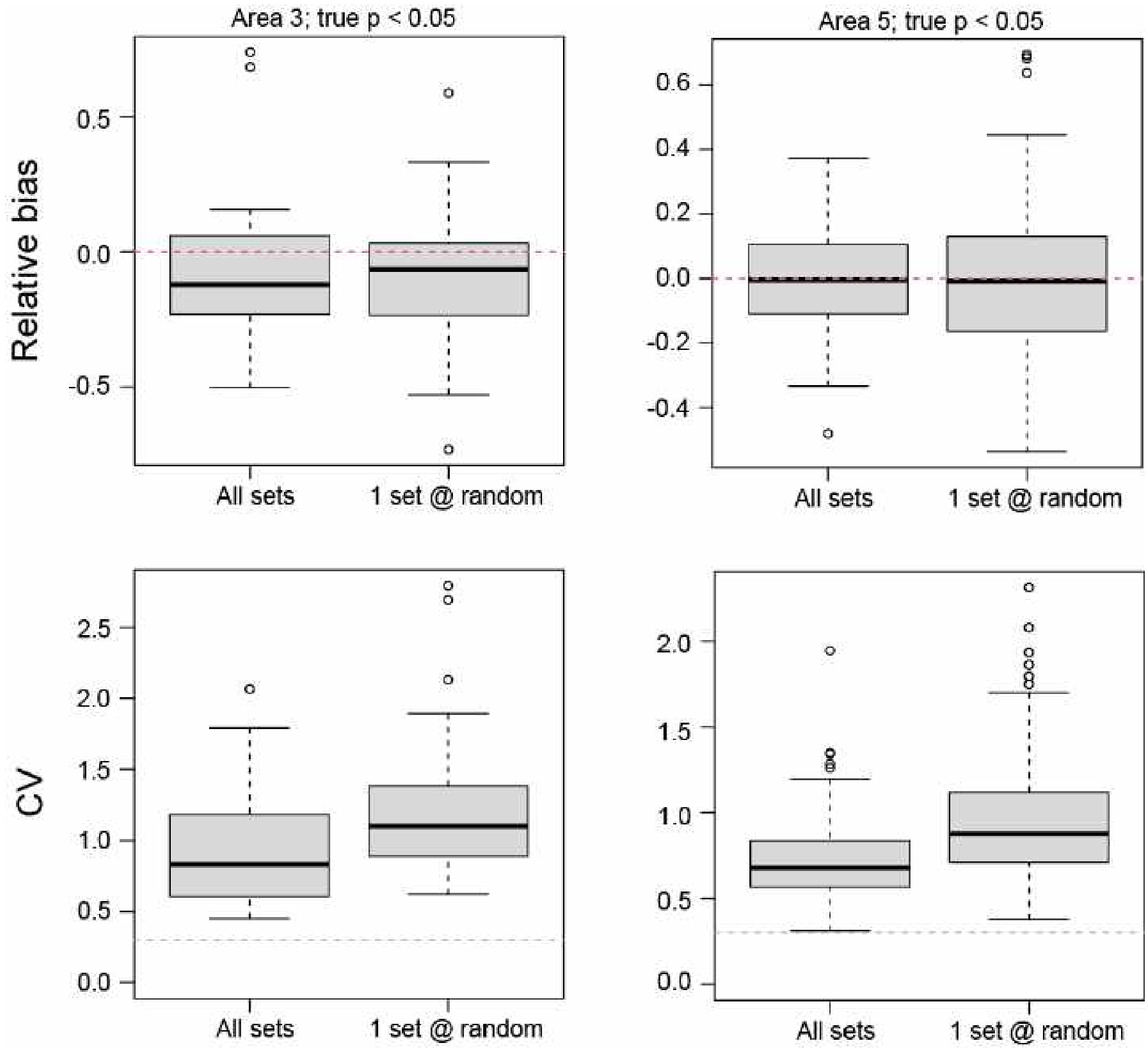

FIGURE 23. Relative bias (top row) and CV (bottom row) for the estimated fleet-level proportion of BET, for 2022 in OBJ-BET areas 3 and 5 (Figure 5), the only two areas with at least 10 wells sampled per domain, for the reduced protocol of 2 clusters of 100 trips each and 2 wells per trip. Results were similar for 2023 and hence are not shown. The average, over simulated full data sets, of true species proportion for the domain, 'true p', is shown in the title of the figure.

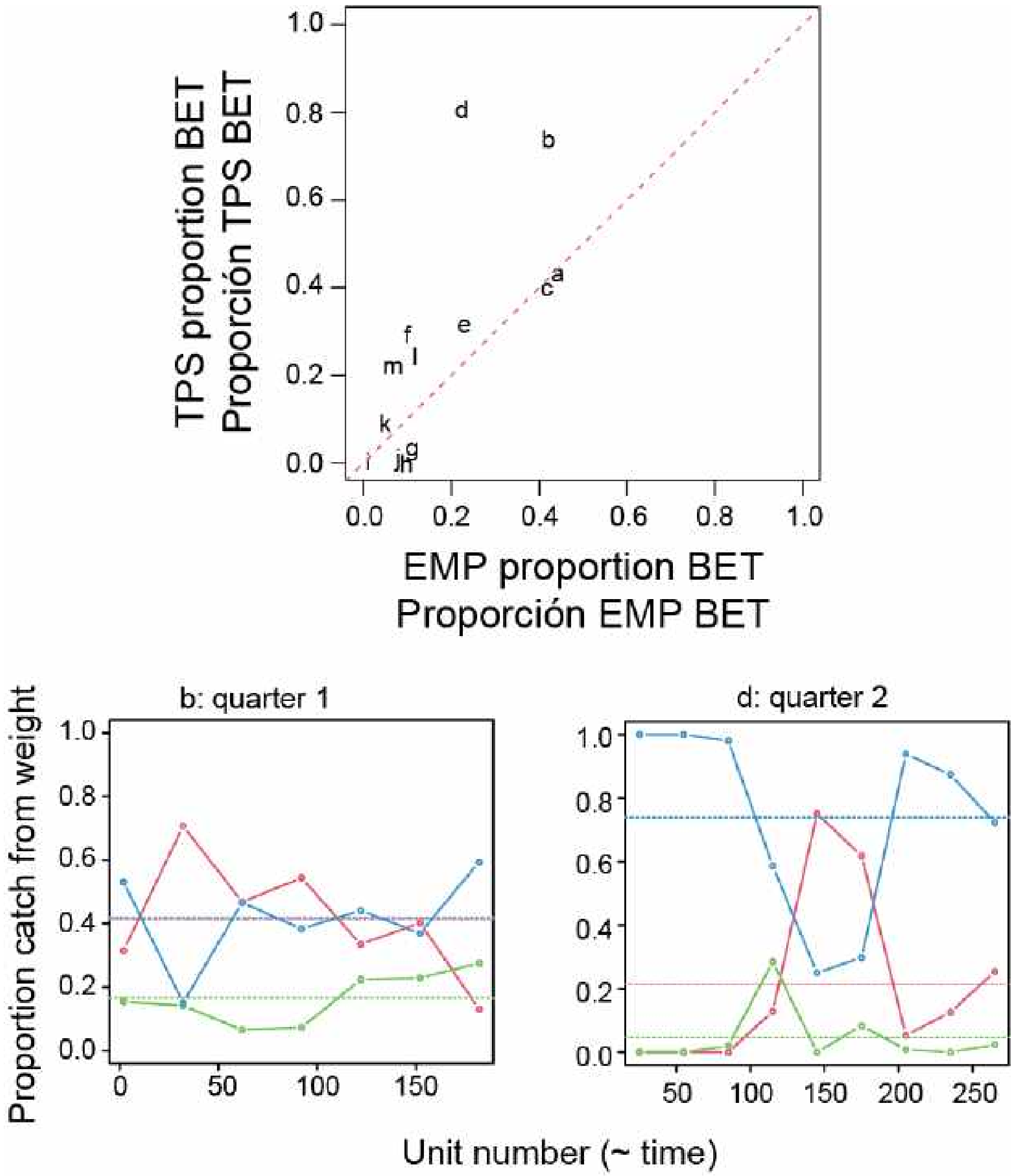

FIGURE 24. Top: estimated proportion of BET in the well, for wells sampled by both EMP and the TPS, 2023. Each letter is a well; there are 13 wells shown in the figure. The red dash line is the 1-to-1 line. Bottom: the EMP data for wells 'b' and 'd' (top graph), which are the wells furthest from the 1-to-1 line. Each open circle is the estimated species proportion for a single unit (container); red: BET; green: YFT; blue: SKJ. The EMP sampled the entire unloading of each of the 13 wells. The TPS sample from well 'b' was taken from the first quarter of the well, and the TPS sample for well 'd' was taken from the second quarter.

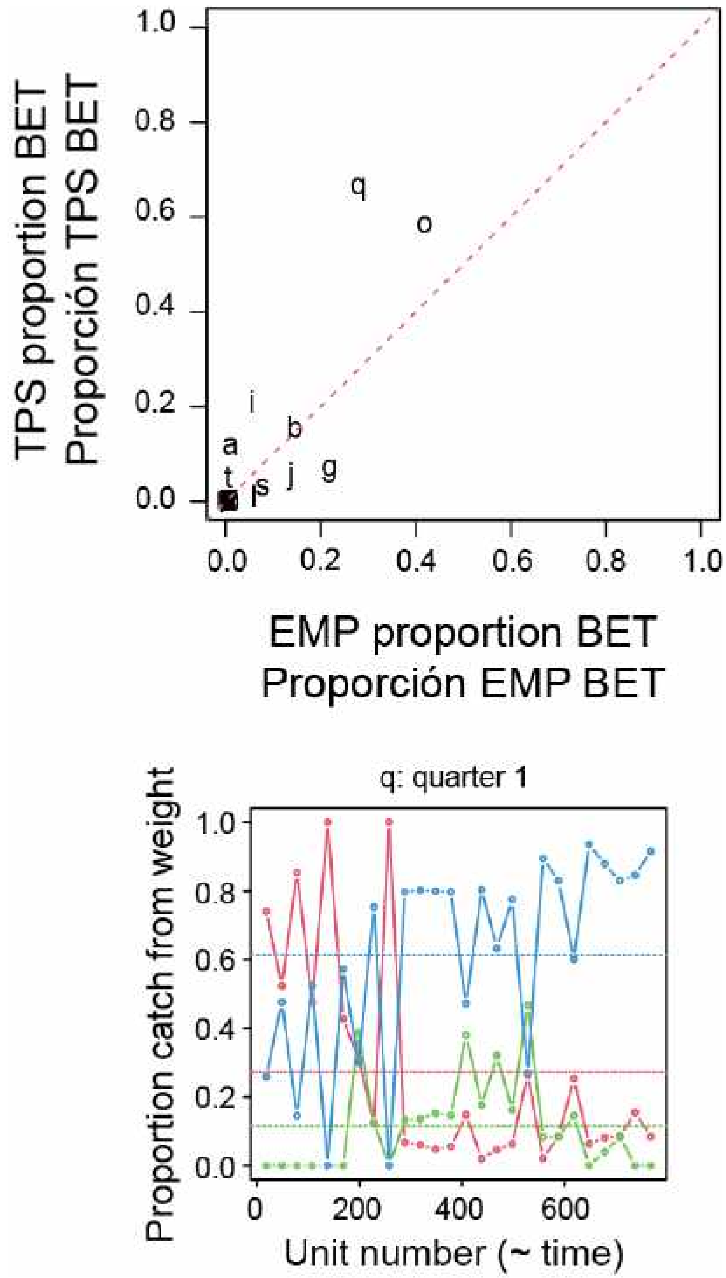

FIGURE 25. Top: estimated proportion of BET in the well, for wells sampled by both EMP and the TPS, 2024. Data for the TPS are preliminary. Each letter is a well; there are 20 wells in the figure. The red dash line is the 1-to-1 line. Bottom: the EMP data for well 'q' (top graph), which is the well furthest from the 1-to-1 line. Each open circle is the estimated species proportion for a single unit (container); red: BET; green: YFT; blue: SKJ. The EMP sampled the entire unloading of each of the 20 wells. The TPS sample from well 'q' was taken from the first quarter of the well.

APPENDIX

Determination of which wells and trip are sampleable depends on several factors: the sampling protocol; the port of unloading of the trip; and, how the catch is unloaded from the vessel. The TPS sampling protocol (appendix in Suter 2010) stipulates that sampled wells must have catch from the same purse-seine set type, market measurement area and month. Wells with catch from different set types/sampling areas/months are not to be sampled. In general, for a trip to be sampled, it must unload in a port where IATTC has a field office. However, on occasion, IATTC field office staff have travelled to other ports, sometime located in other countries, to sample. Finally, some vessels may unload their catch in a way that it is difficult to sample the contents of a single well (e.g. vessels that have 'cold room' through which catch of other wells passes during unloading; vessels that unload entirely with cargo nets, with catch passed through wells using chutes). Catch of such wells is difficult to sample for logistical reasons, without delaying the unloading process. Thus, under the TPS protocol, a well is sampleable if all the catch in the well came from the same set type, market measurement area and month, the catch from the well was unloaded in a way that made it accessible to field office staff without interfering with the unloading process, and the well was unloaded in a port where it could be sampled by field office staff. A trip is sampleable, under the TPS sampling protocol, if it has at least one sampleable well.

The logistical constraints that affect sampleability under the TPS sampling protocol, i.e., availability of samplers at the port of unloading and accessibility of catch from individual wells, would also affect sampleability under the proposed IPSP sampling protocol. However, the IPSP protocol is more flexible than the TPS protocol in that a well with catch from the year of interest - instead of the same month, from the EPO – instead of one of the 13 market measurement areas, and from a single set type, would be sampleable if it was to be unloaded in a port with samplers and in a manner that made the catch of individual wells accessible. Thus, more wells would be considered sampleable under the IPSP sampling protocol than under the TPS sampling protocol.

To compute the exact coverage of a particular sampling protocol, the set composition of each well of every trip must be known. Such information is recorded on the RDL data forms by all observers for Class-6 vessel trips. However, not all RDL forms are keypunched. Therefore, RDL data for the entire Class-6 vessel fleet must be reconstructed from the permanent observer database and from the Tuna Tracking database. Because the set number is not available in the Tuna Tracking database, the two data sources have to be linked using the trip number and set locations and dates. And, as a result, there is a small percentage of wells that cannot be linked to the set information in the permanent observer database. Thus, the coverage calculations presented in this document are approximate.

Finally, the calculation of the number of sampleable wells and trips, under the TPS sampling protocol, that are presented in this document (Tables 1 - 2) include wells and trips from any port where sampling occurred during the year. Thus, ports that were rarely sampled will be included in the calculations. Limiting the ports to only those routinely sampled will change the estimate of the approximate coverage, likely increasing coverage of sampleable trips and wells somewhat, but decreasing sampling coverage of all trips and wells. Additionally, only for the vessels sampled by the EMP is it well know how individual vessels unload catch, and thus, whether or not a vessel unloads in a manner that would make their catch unsampleable. Data on this operational aspect was collected as part of the EMP pilot study for vessels that were to be the focus of the EMP sampling (SAC-14 INF-I). Thus, the calculation of sampleable, for both the IPSP and the TPS sampling protocols, presented in this document does not take this into account because of a lack of information for all Class-6 vessels. Were such information available, including it into the calculations might affect the estimates of sampleable wells and trips somewhat.