**SCIENTIFIC COMMITTEE**
**TWENTY-SECOND REGULAR SESSION**
11-19 August 2026
Apia, Samoa (Hybrid)

# COMMISSION SCIENTIFIC DATA AND DATA PRODUCTS: DEFINITIONS, APPLICATIONS AND NUANCES

**WCPFC-SC22-2026-ST-IP04**
**11 July 2026**

**Submitted by**

Tiffany Vidal[1] and Peter Williams[2]

[1] Oceanic Fisheries Programme of the Pacific Community (SPC)
[2] Tobesus Consulting

## Executive summary

Data management to support the work of the Commission evolves over time. Data requirements change, coverage of key data provisions change, and assumptions/ processing to prepare data to support the work of the Commission adapt as a result. This document aims to provide clarity around the terminology and definitions associated with the scientific data to be provided to the Commission (SciData) and related data products, to detail the key applications of these assets, and to articulate specific data processing procedures or assumptions applied to prepare these data for the Commission, for transparency. This document is intended to serve as a guide for members to enhance their understanding of the various data sources used for statistical reporting and modeling efforts and to better understand the potential nuances within or among them. The SSP will aim to update this document, as necessary, if CCMs find it useful, to ensure members have a clear up-to-date understanding of the data used as part of the Commission process and to support effective participation.

## 1 Background

Effective participation and decision making in regional fishery bodies requires clear communication. It is commonplace for many participants in these fora to be regular attendees year after year, and although the content and language may have been foreign at the outset, many now comfortably speak in acronyms and jargon befitting the situation. This is a natural evolution and often expedites discussions, but from time-to-time it may be necessary to benchmark the language to ensure all participants have a clear and common understanding. This paper aims to provide such a benchmark around the key scientific data assets and products available to and used by the Western and Central Pacific Fisheries Commission (WCPFC) for its scientific, technical, and compliance-related work.

## 2 Scientific Data Reporting

The Commission document describing the *Scientific Data to be Provided to the Commission*, commonly referred to as 'SciData', is a key instrument to guide members in their collection, preparation, and submission of scientific data. Within the SciData, there are five key data submissions:

- Annual catch estimates (ACE);
- Number of active vessels;
- Aggregated catch and effort (often referred to as 'agg' data);
- Operational catch and effort (referred to as logbook data); and
- Size composition data.

These data have been deemed by members to be essential for the scientific work of the Commission. Each of the above data submissions will be described briefly below.

## 2.1 Annual catch estimates

### 2.1.1 Definition

The annual catch estimates (ACE) are the highest-level summation of catch. ACEs are provided at the level of the year, flag, fleet, fishing gear, species, and broad ocean area. The main provision is for the WCPFC Statistical Area, but there are several additional broad areas for which ACEs are required, for a subset of species (e.g., the North Pacific, South Pacific, IATTC-WCPFC overlap area, etc.).

Annual catches serve as the official catch estimate for a CCM, as additional information beyond the logbook data (typically visual estimates at-sea) may be available to refine the final catch estimates, including unloading receipts and VMS information (to identify trips with missing logbook data). Annual catches are used for a few key purposes: i) reporting of overall catch trends through time; ii) 'raising' the logbook data to the level of total removals; iii) determining relative coverage rates of other data sources (e.g., logbook); and iv) to evaluate compliance with conservation and management measures.

### 2.1.2 Key applications

The ACE data are used as the basis for determining coverage of the other sources of catch data submitted to the Commission and are used to raise spatially and temporally aggregated data which are used in the scientific work of the Commission. ACE data are made available to the public each year through the publication of the WCPFC Tuna Fishery Yearbook (WCPFC, 2025).

### 2.1.3 Data processing, issues, and potential improvements

The submission of ACE data varies by CCM, as for several years now, Pacific Community (SPC) member countries have used the TUFMAN2 platform to manage their fisheries data. TUFMAN2 has a dedicated module that summarises logbook data, assesses logbook coverage using VMS information, and ultimately provides a best estimate of the total annual catch by gear and species, for the respective ocean areas. For non SPC-member countries, these data have been submitted via email by the SciData submission deadline of 30 April each year. In 2026, WCPFC members began submitting their SciData through a web-portal using standardized data submissions templates (see Vidal et al., 2025).

The contemporary approach for processing the ACE data through the web-portal is for submissions that pass initial validation checks to be stored in a secure data lake. From here, the SSP has developed data pipelines to ingest these data into raw database tables, at which point they are subjected to second-level validation checks. If there are any issues detected, CCMs are contacted through the data portal and data submissions may be edited and resubmitted to restart the process. Once fully validated these submissions are consolidated with the TUFMAN2 submissions into the production database.

Although these data are fairly simple and straightforward, there are some nuances associated. In some situations, the logbook data may be higher than the reported annual catch estimates. This can result from fisher overestimation of the catches, but it can also arise due to data updates (e.g., in logbook data) after the submission of the ACEs. For example, in TUFMAN2, and potentially for CCMs submitting their ACEs directly to SSP, there is an estimation procedure to derive the ACEs based on the logbook coverage. If additional data fill those coverage gaps, the estimation procedure may no longer represent the best estimate of the annual catches (e.g., changes in average catch-per-unit-effort (CPUE) used in the raising procedures). In some cases, catch estimates from previous years have not been revisited, leading to discrepancies in the data (albeit often minor). In 2025, the SSP worked with Pacific Island member countries to update their ACEs for the previous 5 years, where discrepancies were noted. This is now an annual review activity as part of the Tuna Data Workshop. The SSP aims to conduct a similar comparison for all WCPFC members in case there may be discrepancies to address in the estimates from recent years.

There are also cases where the source of annual catch estimates is not logsheet data, but the measured (weighed) catch from landings data, which has been the case in some longline fisheries in the past.

ACE data may also be post-adjusted after submission to the SSP, in the case of purse seine catches. The Commission has developed a methodology by which the species composition from observer reports are fitted to a model which is used to adjust logbook catch estimates, as the reporting of juvenile yellowfin and bigeye is assumed to be under-reported due to the high volume nature of the fishery (i.e., Project 60; see Peatman et al., 2024 and previous papers on the topic). This adjustment, aimed at producing more accurate overall species catch estimates can introduce discrepancies in the data as a result. For example, when comparing logbook data to the ACE, there may be differences even if the logbook coverage represents 100% of the fishing activities, due to this adjustment.

The ACE data are the primary data source used to generate the ACE by EEZ. ACE by EEZ, as the name implies, is an attempt to spatially disaggregate the ACE into Exclusive Economic Zones (EEZ) and designated high seas areas (Figure 1) at an annual level. The key limitation is that ACE by EEZ attempts to provide an estimate of effort as well as catch, but effort is not reported in the ACE. Therefore, additional data sources and assumptions must be incorporated to produce effort estimates, including raising of logbook effort, application of catch-per-unit-effort (CPUE) for similar fleets, and potentially adjust effort based on vessel sizes within a fleet (e.g., to adjust small-scale vessel effort to create more comparable effort metrics across a gear type). In addition, ACE by EEZ relies on the spatial distribution of catch from logbook data. If logbook coverage is low, the resulting disaggregation may be biased. This issue will be discussed in more detail below.

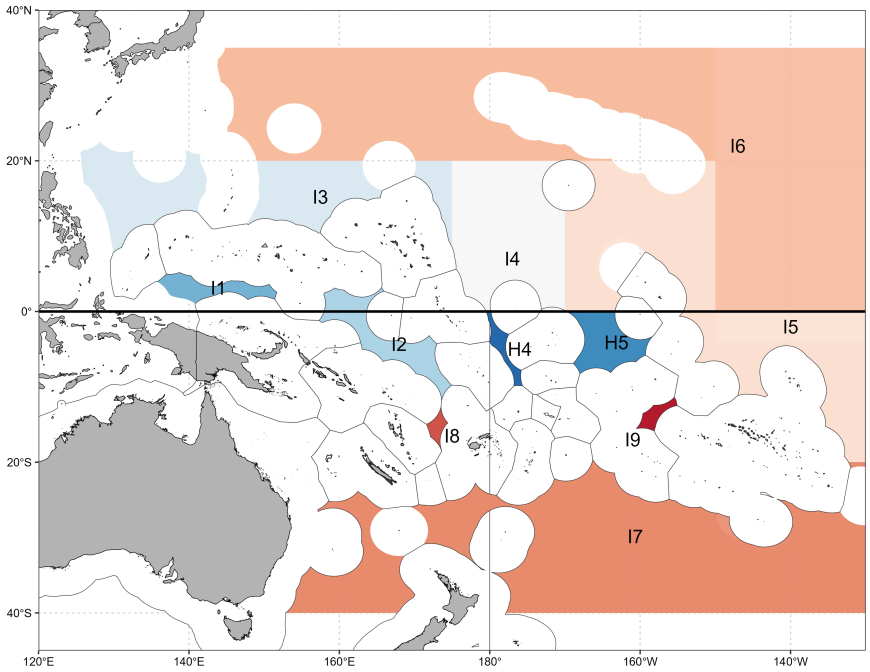


Figure 1: High seas areas of the WCPFC shown for reference[3]

Another cross-cutting challenge is the issue of chartering arrangements which are covered under the latest WCPFC CMM for Chartering (see https://cmm.wcpfc.int/measure/cmm-2024-03). For scientific data, catch and effort need to be attributed to the chartering nation, which may differ from the flag of registration (i.e., flag-state of the vessel). The proper attribution of catches to the appropriate chartering nation depends on accuracy of the chartering information provided to the Commission, and ensuring that the charting information is subsequently shared with the SSP, the chartering nation, and the flag-state to ensure consistency. If this information is not made available to all stakeholders, then catch data may be mis-attributed which can have knock-on effects for the generation of data products. Further complicating this issue is the concept of conditional chartering. Conditional chartering, as we understand it, is when a vessel is chartered under certain spatial or temporal conditions (e.g., only while in a specific EEZ is a vessel chartered, otherwise the flag reverts to the flag of registration or another chartering party). The SSP has not received chartering information with these details, but instead we are notified that for a given period of time there is a chartering arrangement between two CCMs, inclusive. We understand the importance of conditional chartering to some small island developing states, but if conditional chartering is occurring, this has broad implications for data collection and attribution that should be considered by the Commission.

Lastly, it is worth noting that the ACE for other areas, specifically broad ocean areas that include

[3]The boundaries shown on this map are for the purposes of illustration only and are not the expression of any opinion on maritime disputes or the delimitation of maritime claims.

areas under the jurisdiction of the IATTC (i.e., South Pacific, North Pacific), represent the catches for WCPFC members only. In some SC reports (e.g., Vidal and Heather, 2026), South (or North) Pacific catches may be reported based on the ACE submissions. These values may sometimes differ from the estimates of catches for these areas derived from combined WCPFC/IATTC data capturing reporting for members in both RFMOs (i.e., the aggregate catch and effort data). The SSP tries to note the data source when these catch statistics are presented, but this distinction is important.

## 2.2 Number of active vessels

### 2.2.1 Definition

The SciData specifies the fisheries (i.e., gear types) for which CCMs must submit the number of vessels that actively participated in the reporting year. The reporting of active vessels is based on vessel size category, which also varies by gear type (purse seine, longline, and pole-and-line). Reporting of vessels active in other smaller-scale gear types is not yet required by the Commission.

### 2.2.2 Data processing, issues, and potential improvements

The number of active vessels are linked to the ACE data, by year, flag, fleet, and gear.

One issue with the number of active vessels occurs when the number reported by the CCM in their SciData submission differs from (e.g., is less than) the number of vessels that have been reported in the operational logbook data. Ideally, the SSP would work with the CCM or chartering nation to reconcile/confirm the vessel numbers between logbook and their SciData submissions. This has not yet been incorporated into the standard work flows, but it could be considered as time allows.

## 2.3 Aggregate catch and effort data

### 2.3.1 Definition

The aggregate catch and effort data are often referred to as the 'BEST' estimate of spatially and temporally aggregated data (sometimes disaggregated from ACE). In an ideal world, the aggregate data would simply be an aggregation of the logbook data to the appropriate spatial (e.g., 1x1 degree or 5x5 degree spatial cells) and temporal (e.g., monthly) resolution.

WCPFC and the IATTC have a data sharing MOU whereby each year key data sets including aggregate catch and effort for longline and purse seine are exchanged between the organizations for scientific purposes, including assessment and harvest strategy work.

### 2.3.2 Key applications

BEST data are one of the key inputs into assessment models as it represents total catch and effort at a relatively fine-scale. Unlike logbook data, the BEST data are raised to the level of total annual

catch estimates. These data inform assessment models on total catches in space and time at the level of the flag-state and sub-fleet.

### 2.3.3 Data processing, issues, and potential improvements

If a CCM submits 100% (or approximately 100%) logbook data, there is no need to submit aggregate data separately. In cases where 100% logbook data are not available, aggregate data are to be submitted to provide the best estimate of raised catch and effort to represent total annual catch and effort. In certain situations where CCMs have diverse small-scale fisheries taking significant quantities of tuna (e.g., Indonesia and Philippines small-scale fisheries), it may not be possible to implement adequate data collection systems so neither 100% logbook data nor aggregate data can be provided. In these situations the SSP works with these members to disaggregate their annual catch estimates, as best they can. In this latter case, there are many simplifying assumptions involved (e.g., the same catch is taken each month); however, it offers the best estimate at that point in time.

It should be noted that the SSP has been actively engaged with Indonesia, Philippines, and Vietnam, through the WPEA-SPF project to prepare aggregate catch and effort data using all available data sources. In 2026, all three countries have provided these data[4], which are likely to provide a more representative assessment of the spatial and temporal characterization of their catch and effort for key gear types (i.e., purse seine, longline, and pole-and-line). One of the longstanding challenges with the effort from these three countries is that the small-scale nature of some of their fleets means there is an issue with effort parity (i.e., a day of fishing for a small-scale purse seiner is not the same as an industrial vessel). To address this, an adjustment factor has been developed based on vessel size and average catch per day. This is applied to scale the fishing days to values more comparable to larger fleets. These bespoke adjustments are challenging to maintain consistently and document transparently. One potential solution, perhaps not without caveats, could be to implement additional gear codes (ideally coordinated with the FAO Coordinated Working Parties (CWP) for interoperability) to capture the small-scale nature of some of the gears, which may differ from their industrial counterparts. The interpretation of a day at sea for a small-scale gear may be different from the industrial version, but it may mitigate these post-hoc adjustments that may or may not be appropriate through time.

The raising procedure used by the SSP involves comparing the key tuna catch (by weight), by gear (e.g., BET, SKJ, YFT for purse seine; ALB, BET, YFT for longline), from logbook data to the annual catch estimates (ACE). This ratio provides a raising factor that is used to raise all catch and effort at the level of aggregation. For example, if the logbook data provided (by year, flag, fleet, and gear) has been aggregated to a 1x1 degree and monthly resolution, all estimates at that resolution are raised by the same factor to ensure the total catch matches the annual catch estimate for the

[4] At the time of the preparation for this paper, aggregate data from 2 of these 3 CCMs were loaded as there were some unresolved issues with the third, but significant progress has been made nonetheless.

sum of key tunas. It should be noted that non-tuna species are not raised; the raw aggregate values are used in these data products. Improving aggregate data for non-tunas is an area identified for improvement, and one the SSP aims to focus on in the coming years.

There are obvious uncertainties associated with the raising approach. One of the key limitations is that it does not identify where and when gaps in the logbook coverage exist to estimate catches with potentially greater precision (in space and time); instead, it applies a general raising factor across the full data set (assuming that the unraised logbook data are spatially and temporally representative; which is likely a reasonable assumption as the coverage approaches 100%). Historically, catch and effort were not raised in this fashion when logbook coverage was low. More recently, logbook coverage is generally high among all fleets, so these uncertainties continue to diminish. Another uncertainty is that effort is raised in the same manner as catch. This assumes a linear relationship between the two, but that may be unrealistic. There may be operational reasons why logbook catch and ACEs differ (e.g., consistent under-reporting by fishers due to visual estimate bias). These factors have not yet been fully explored, but it would be valuable exercise.

For purse seine, there is an additional adjustment to the catches of key tuna, to adjust for the bias in logbook reporting resulting from the high-volume nature of that fishery (as discussion in Section 2.1.3). These adjustments have been well documented over the past several years as part of Project 60, which concluded in 2024 (Peatman et al., 2024). These corrections are applied based on the species composition model outputs, using models fitted to observer data. These adjustments are expected to improve the catch composition, generally by increasing the catches estimated for bigeye and yellowfin while reducing the catches for skipjack (to account for the misreporting of small yellowfin and bigeye in the catch of skipjack tuna).

It is accepted that the species composition adjustments are beneficial for improving the accuracy of purse seine catches. Even so, one of the consequences of this approach is that a disconnect between logbook and the resulting aggregated catch and effort estimates (BEST) arises as a result. The two are no longer directly comparable at the species catch level, even when logbook coverage is 100%, as the BEST data are adjusted whereas the logbook data are not. These considerations would have been discussed and weighed at the time the adjustment approach was adopted; however, it is good to be reminded of some of the consequences of those decisions.

With respect to the eastern Pacific Ocean (EPO) aggregate catch and effort data provided by the IATTC, the SSP aims to avoid any unnecessary additional processing; however, some key gaps have been identified. Specifically, in the longline catch data there are some CMMs that provide only numbers of fish and others that provide catch in weight only. This data gap is relevant for albacore due to the South Pacific-wide stock assessment and associated management procedure, but has relevance for other species as well. Given that the assessment relies on numbers of fish and yet the harvest strategy work relies on catch in weight, both are required for the work of the WCPFC. Historically, a long-term average estimate of weight per individual for albacore was applied. This process has been updated to capture temporal variability in size structure of the catch by using an

annual average to apply to the missing data fields.

## 2.4 Operational catch and effort data

### 2.4.1 Definition

Operational catch and effort data (logbook) provide fine-scale spatial and temporal information about fishing activities. These data are generally collected by fishers while at-sea and the catch in weight often relies on visual catch estimates. Operational data provide critical information to interpret trends in catch rates over time and to understand selectivity of different fleets. Because operational data are collected at the vessel and fishing event level, catch rates and fishing activities can be analyzed with respect to vessel, gear, and environmental characteristics, where this information is lost in the aggregate data sets and certainly in the ACE.

If 100% operational data could be provided to the Commission, the reporting could be greatly simplified. However, CCMs may still wish to adjust the ACEs (which could be produced from logbook data) or aggregate catch and effort data if they have more precise estimates of catches from unloadings data, for example. It should also be noted that the SSP attempts to adjust logbook catch estimates when unloading data are provided for a given trip, to ensure consistency between the logbook and the precise catch estimates obtained at the unloading event.

The current version of the SciData specifies the required information around gear and fishing activities to be reported to support the standardization of CPUE indices, among other scientific inquiries; however, the data collection frameworks can easily be adapted to include additional information deemed important by the SC to improve these analyses. In 2024, the SC adopted a suite of voluntary longline data fields aimed at improving CPUE standardization. Several CCMs have started submitting these data, which will likely prove valuable for the monitoring of relative abundance trends for key tunas.

### 2.4.2 Key applications

Operational data are primarily used in the assessments models via relative abundance indices generated by standardizing CPUE data. Pole-and-line gear has provided a relative index of abundance for the skipjack stock many years, but as that fishery contracts and wanes, attention has shifted toward trying to standardize purse seine catch rates, and more recently with a focus on free-school sets. Longline operational data have supported CPUE indices for the key tuna (albacore, bigeye, and yellowfin) and billfish stock assessments.

### 2.4.3 Data processing, issues, and potential improvements

Increasingly, operational data are being collected and submitted via electronic reporting (ER) platforms. ER data are generally subjected to more validation constraints earlier on in the data collection process, potentially mitigating errors relative to paper-based collections, particularly

when ER systems incorporate data from instruments such as GPS, for example, which ensures accurate positional information. ER data also have the added advantage of increased timeliness of data availability.

Some of the challenges associated with logbook data are not unique to the WCPFC. Logbook data relies on fisher expertise relative to the identification of catches to the species-level, consistent and precise reporting of activity dates, times, and positions, and accurate reporting even when interactions with species of special interest (SSIs) occur. Fishers are generally focused on the target and valuable catches, whereas the reporting of catch discarded at-sea may be reported with a lower degree of accuracy. Observers are generally deployed to fill some of these gaps around bycatch, discards, and to collect biological samples of the catch. Port sampling and unloadings data provide the opportunity to cross-check and update the logbook-reported catch by species at the trip level; the weighed catch by species available in unloadings data can potentially be used to adjust the visual estimates for catch (in weight) in the logbook data in some cases.

There is an opportunity to expand operational data collection, perhaps on a voluntary basis, to other fisheries. For example, in the Western Pacific - East Asia (WPEA) region, handline fishing has rapidly emerged as a preferred gear type. The countries participating in this fishery are collecting logbook data, but the Commission doesn't yet have a formal process for submitting these data nor have these data been incorporated into the assessment processes. Enhancing data collection for these fisheries could be beneficial for continued monitoring of key stocks, as other fisheries such as pole-and-line contract.

## 2.5 Size composition data

### 2.5.1 Definition

Size data (length and weight of individual fish) are generally reported through observer-collected data and port sampling activities. There are standard operating procedures for observer data collections, but even so, data collection methodologies between observer and port sampling programs may differ, and the ability to implement those standards may vary depending on fish sorting and handling by crew at-sea or during unloadings (Hamer et al., 2025).

Size data are typically (and necessarily) associated with metadata to inform on the measurement type (e.g., total length, fork length) and whether or not the animal is in a processed state (e.g., gilled and gutted) or whole. Using this information, the SSP is able to apply appropriate conversion factors to standardize the size data for aggregation and use in the assessment models. Data should not be converted by CCMs prior to submission, as the SSP must ensure consistent conversions across all CCMs providing data.

Aside from observer data, the submission of aggregated size data collected by CCMs (through other types of sampling) for the key WCPFC species is a requirement under the SciData. These size data submissions are typically stratified temporally and spatially based on the type of sampling, for

example, port sampling data usually have more coarser levels of temporal and spatial stratification than sampling at sea. Information on the size measurement increment used (e.g., 1cm or 2cm intervals) and morphometric measurement (e.g., upper jaw caudal fork, or lower jaw caudal fork) are also required.

The SSP holds comprehensive size data collected at the fishing operational level by observers through the ROP and at the trip level by port samplers from the Pacific Islands and the WPEA countries. These 'operational' level size data are consolidated into aggregate size data which are then combined with the aggregate size data from other sources submitted directly by CCMs.

#### 2.5.2 Key applications

Aggregated size data are one of the key inputs into the length-based, age-structured assessment models and are used in conjunction with the aggregated catch and effort data. Unlike the aggregate catch and effort data (i.e., BEST data sets), aggregated size data are not raised.

#### 2.5.3 Data processing, issues, and potential improvements

Hamer et al. (2025, 2026) provide detailed explanations of data processing, issues and potential improvements. Two of the overarching issues related to size composition data are the need for more information in the CCM submissions on the protocols for how the data were collected, and attempting to improve the standardisation of the variety of formats and fields in the data provided.

## 3 Scientific Data Products

The Commission produces several public domain scientific data products which provide, *inter alia*, third parties (CCMs, other agencies, the public, etc.) with access to WCPFC public domain data. The WCPFC Rules and Procedures for the Protection of, Access to, and Dissemination of Data (RAP) governs what data can be made available to the public.

### 3.1 ACE by EEZ

#### 3.1.1 Definition

One of the key data products, aside from the BEST data sets (which are a mixture between reported and estimated catch and effort estimates), is ACE by EEZ. As noted above, this data set represents a disaggregation of the ACEs into high seas and EEZs. The historical generation of this data set required complex assumptions and a deep understanding of fleet dynamics to be able to reasonably attribute catch and effort to the respective areas; this was mainly due to the paucity of available logbook data for the extent of the WCPFC area, noting that logbook data for distant-water fleets has only been submitted to the WCPFC since 2018. We won't revisit those rules, as they are complex and quite bespoke. However, we believe there is value in understanding how the

contemporary estimates are generated, and the potential uncertainties associated. Given relatively high levels of operational data, the process for generating the ACE by EEZ begins with aggregating operational catch and effort into high seas areas and EEZs using the fine-scale positions reported in those data. The total tuna catch from the aggregation of the operational data is compared to the ACE and similar to the raising applied in the BEST data sets, we raise effort for the ACE by EEZ, based on the estimated coverage level. For the catches, however, we raise the catches at the species level, instead of raising based on the coverage for the sum of the key tunas combined (as for BEST). This approach is used to ensure the final catch estimates (for key species) are consistent between the ACE and the ACE by EEZ, which is expected to produce reasonably good estimates; however, if operational data coverage is low, these estimates will become less reliable.

### 3.1.2 Data processing, issues, and potential improvements

There are a number of exceptional circumstances applied to generate the ACE by EEZ, however. The first is for Indonesia, Philippines, and Vietnam as they do not yet provide much operational data. Given these countries are primarily operating in their own EEZs, we generally attribute their catch to their respective EEZs. Philippines also operate in high seas pocket 1 (HSP1), but they submit 100% observer data for that fishery, enabling us to confidently attribute that portion of their catch to HSP1.

A small fleet of Philippine-flagged vessels have fished in the archipelagic waters of PNG under a special arrangement for many years, predating the establishment of the WCPFC. Their catches are landed in and attributed to PNG; therefore, those catches continue to be assigned to the PNG flag, even though notifications of these charters may not be reported annually to the WCPFC.

In addition, if there is unassigned catch (i.e., the difference between operational catch and the ACE) from the longline fisheries for China and Chinese Taipei, those catches are allocated to certain high seas areas, depending on the species and flag-state. This rule was applied based on an understanding of the fishing activities and the logbook reporting (for example, that the available logbook data were biased to fishing in the EEZs). In 2023, in particular, there was a large amount of catch attributed to the high seas for China due to low coverage of operational data available at the time. These assumptions are continually being refined as additional data become available and with improved reporting. In some cases it is possible to rely on the aggregated data to fill gaps between logbook and ACE (depending on how the BEST data were generated), but for longline, 5°x5° cells often span two or potentially more fishing zones, which adds complexity.

In many cases, the SSP takes a precautionary approach with unidirectional raising. For example, if longline logbook data are paired with an unloading event that suggests higher catch weights, the logbook catches are raised to match the more precise weights obtained at the unload; however, if the unloaded catch is lower, we don't down-weight the logbook data, as we don't always know whether an unloading event is a full or partial unload. When preparing the ACE by EEZ, we typically do not reduce catches, if for example, the logbook data suggests more catch than the ACE. These

discrepancies in the data matter, however, especially when cross-validating catch estimates from different sources. This is especially challenging with the ACE by EEZ for purse seine catches because both the ACE and the BEST have been adjusted for species composition. Therefore, if we sum the skipjack catch from the operational data (where coverage is high) and compare to the species composition adjusted ACE (e.g., to determine the species-specific raising factor for ACE by EEZ), it will generally be higher, whereas the bigeye and yellowfin would be lower, and thus raised. Therefore, we would generally raise the yellowfin and bigeye catches from the logbook, but not the skipjack (as the species composition adjustment applied to the ACEs generally down-weights skipjack and increases yellowfin and bigeye); however, this could result in an overestimation of total tuna catches. In 2025, this procedure was updated to address this issue. For the purse seine fishery only, if skipjack logbook data are higher than the species composition adjusted ACE, we down-weight those catches, proportionally across all zones where skipjack was caught, to ensure that the total catch in the ACE by EEZ matches the ACE.

Due to potential updates to logbook and/or ACE data, the ACE by EEZ values may change through time, as with most other data sets. The ACE by EEZ is a public domain data product and available through the WCPFC website from 2018 on. The full data set is available upon request, but given the uncertainties associated with the more historical estimates, we have refrained from publishing it on the Commission site.

## 3.2 Public domain aggregated catch and effort data

### 3.2.1 Definition

The WCPFC has agreed on the Rules and Procedures around sharing of Commission data in the public domain. These rules define the agreed criteria to classify data as shareable in the public domain or not. The SSP employs a precautionary approach when interpreting these rules, and confers with the Secretariat when additional interpretation is required. The key criterion for sharing data is that there can be no fewer than 3 vessels at the finest level of aggregation. Strata which contain 3 or more vessels may be shared in an aggregated manner. Operational data are never shared beyond their data owners without explicit consent.

Aggregate catch and effort data are prepared in various public domain products, at different levels of aggregation (e.g., monthly, quarterly, annually) and made available through the Commission website. Data associated with strata that cannot be shared due to the confidentially rules are removed from data products before sharing. These omissions often result in considerable reductions in the available data, especially in the areas along the margins of the main fishing areas (e.g., higher and lower latitudes) where there is low fishing effort. One of the key caveats with preparing public domain data is that for much of the historical data, logbook data are lacking. This means that when aggregating data we may not have precise estimates on the number of vessels operating in an area at a given time. In these situations we estimate the activity as best we can by estimating fishing days, and applying conservative estimates that stipulate there must be more than 61 days

in a strata, for example, if aggregated monthly, for that cell to be included in the public domain. This is possibly an unreasonable assumption, but in the absence of other information it is the most appropriate, in our view, to ensure the confidentiality of members' data is respected.

It may be prudent for the SC to consider the intention behind the provision of public domain data. Strong filtering of data to remove information that does not align with the Commission confidentiality standards certainly means provisioning incomplete data, which may lead to misinterpretation or inappropriate conclusions drawn from these data sets, even though every effort to document the caveats is included with the public domain data. The SC may wish to consider the pros and cons of providing heavily filtered domain data to assess whether the benefits outweigh the risks, and potentially explore alternative considerations for making more complete data available to the public, perhaps with different rules around anonymization, acknowledging that the Commission has a framework for access to non-public domain data in the RAP, and also third parties can potentially propose collaboration, which provides access to non-public domain data, through the authorised WCPFC projects.

## 3.3 Public domain aggregated bycatch data

### 3.3.1 Definition

Bycatch refers to the unintentional catch of non-target species such as sea turtles, sharks, marine mammals, or seabirds during fishing operations. Using the Bycatch Data Exchange Protocol (BDEP), WCPFC publishes aggregated bycatch statistics alongside related fishing effort data and observer information (see Clark et al., 2015 and https://www.wcpfc.int/sustainability/bycatch-mitigation/public-domain-bycatch-data). The BDEP provides a standardized way for members to share bycatch data, ensuring that the information is comparable across different fleets and time periods. In this section, you can find summary tables of bycatch (e.g. number of incidents or catch of key species) and associated effort data, based on observer reports and other data submissions.

Each year the SSP updates the public domain bycatch data sets which include both total fishing effort and observer data (to interpret coverage rates), including non-ROP observer data held by SPC. These data sets include the following:

- Total fishing and observed effort per year and fishery.
- Bycatch in Longline Fisheries (year): Observed and estimated captures/mortalities by species group.
- Bycatch in Purse Seine Fisheries (year): Observed and estimated captures/mortalities by species group.
- Bycatch in Longline Fisheries (year/area): Breakdown by fishing area.
- Bycatch in Purse Seine Fisheries (year/area): Breakdown by fishing area.

The inclusion of tables on observer data coverage is important to provide the public with an indication on the level of uncertainty with the bycatch data presented with, at best, only 5%

coverage of observer data for the longline fishery but much higher levels of coverage (and so more certainty) with observer data for the purse seine fishery.

### 3.4 Public domain aggregated size data

#### 3.4.1 Definition

Aggregate size data are included in the classification of 'Biological data' in the RAP, with a confidentiality classification of 'lowest' risk. The public domain version of aggregated size data has therefore been extracted from the aggregated size database generated for the WCPFC stock assessments (see Section 2.7 above) without any filtering and with a similar structure.

#### 3.4.2 Data processing, issues, and potential improvements

Since the public domain version of aggregated size data closely aligns to the aggregated size data used for the WCPFC scientific work, the issues and nuances described in Section 2.5.3 and in Hamer et al. (2025) are relevant.

## 4 Acknowledgments

We acknowledge the entire data management team for their continual efforts to process, review, and prepare data to support the work of the Commission. We sincerely thanks the WCPFC membership for their collaborative engagement around data and their continual support to improve these data assets including through enhancements to data systems and processes. We thank the Commission members for their funding in support of TV and PW's engagement, which has contributed to this paper.

# 5 References


Clark, S., Nicol, S., and Williams, P. (2015). Proposal for a Bycatch Data Exchange Protocol (BDEP) amongst the t-RFMOs. *WCPFC SC11-2015.EB-WP-06, Pohnpei, Federated States of Micronesia, 5-13 September*.

Hamer, P., Castillo-Jordán, J., Panizza, A., and Peatman, T. (2026). Project 127: reconciliation of size data collections. *WCPFC-SC22-2026/ST-WP-02, Apia, Samoa 11-19 August*.

Hamer, P., Schneiter, E., Vidal, T., and Williams, P. (2025). Review and reconciliation of size data collected in the WCPFC-CA for stock assessment purposes (WCPFC Project: 127). *WCPFC-SC21-2025/ST-WP-02, Nuku'alofa, Tonga, 13-21 August*.

Peatman, T., Vidal, T., Williams, P., and Nicol, S. (2024). Project 60: Progress towards achieving SC18 recommendations. *WCPFC SC20 ST–WP–02, Manila, Philippines, 14-21 August*.

Vidal, T. and Heather, F. (2026). Estimates of annual catches in the WCPFC Statistical Area. *WCPFC-SC22-2026/ST-IP-01, Apia, Samoa 11-19 August*.

Vidal, T., Pohl, B., and Loganimoce, E. (2025). Data standardization to improve efficiency of WCPFC data provisions. *WCPFC TCC21-2025-25, Pohnpei, Federated States of Micronesia, 23 September - 1 October*.

WCPFC (2025). WCPFC Tuna Fishery Yearbook - 2025. *Western and Central Pacific Fisheries Commission, Pohnpei, Federated States of Micronesia*.