

A Methodological Framework for Integrating Heterogeneous Fishery-Dependent and Independent Tuna Data in India

S. Ramachandran, Sijo P. Varghese and Sreenath K. R.

Fishery Survey of India, Department of Fisheries, Ministry of Fisheries, Animal Husbandry and Dairying, Government of India

Purpose, expectation and action intended

Discussions (for information and open discussion):

The paper is to inform meeting members that India faces a fundamental challenge in meeting the species-wise, georeferenced catch and effort reporting requirements of IOTC, FAO and other fisheries management organisations, as set out in Sections 1–3. This challenge arises because India's tuna fishery is predominantly artisanal and small-scale in nature, characterised by a vast fleet of over 238,000 craft operating from 1,537 landing centres, where fishing-incident-wise information on species caught, catch, effort, size and biological parameters is inherently difficult to collect through landing surveys alone. The paper describes how India's well-established fishery-independent data collection system, based on exploratory and experimental fishing surveys conducted by the Fishery Survey of India (FSI) across the entire Exclusive Economic Zone (EEZ) and adjoining seas, provides a critical complementary source of species-wise, georeferenced information that can be collated with landing-centre and fishing-harbour data to generate the required catch, effort, size and biological information. The paper proposes a four-step integration framework, preserve, standardise, link and reconcile, designed to combine these heterogeneous sources while maintaining their statistical identity. No fitted calibration model or combined abundance index is presented, the framework is proposed as a national methodology for discussion.

Feedback (to collect input and suggestions):

The paper is to request review and comments on the proposed framework and the specific challenge of deriving species-wise, georeferenced catch and effort information from artisanal and small-scale fisheries, as set out in Sections 2–5. The authors seek feedback from WPM17 participants on: (i) the scientific defensibility of using fishery-independent survey data as the primary source of species-wise, georeferenced information when fishery-dependent data lack fishing-incident-level resolution; (ii) the treatment of catchability and uncertainty when collating survey hooking rates with commercial landings; (iii) appropriate spatio-temporal

stratification and record-linkage rules for artisanal fleets; and (iv) the proposed approach to documenting unresolved discrepancies without forcing artificial agreement between data sources. Input is also sought on how the framework might align with ongoing MSE developments for Indian Ocean tuna stocks and with the reporting requirements of IOTC Resolutions 15/01, 15/02 and 26/06.

Abstract

India's tuna fishery is predominantly artisanal and small-scale, with over 238,000 craft operating from 1,537 landing centres. This structure presents a fundamental challenge for meeting the species-wise, georeferenced catch and effort reporting requirements of IOTC, FAO and other fisheries management organisations. Fishing-incident-wise information on species caught, their catch, effort, size and other biological parameters is inherently difficult to collect through landing surveys alone, because landings are dispersed, multi-gear, multi-species and often recorded at coarse spatial and temporal resolution. India has addressed this challenge through a well-established fishery-independent data collection system, in which research and survey vessels of the Fishery Survey of India (FSI) conduct exploratory and experimental fishing across the entire Exclusive Economic Zone (EEZ) and adjoining seas. Data collected during these voyages, including species-wise catch, effort, geo-referenced fishing positions, size composition and biological information, are collated with data from landing centres and fishing harbours to generate the species-wise, georeferenced information required for stock assessment and management. This paper explains the nature of India's artisanal and small-scale tuna fishery, describes the FSI fishery-independent survey system, and proposes a rule-based methodological framework for integrating these heterogeneous sources. The framework follows four steps, preserve, standardise, link and reconcile, designed to maintain the statistical identity of fishery-dependent and fishery-independent datasets while enabling complementary interpretation. The paper clarifies that fishery-independent survey data serve as the primary source of species-wise, georeferenced information, while fishery-dependent landing data provide coverage, magnitude and fleet-level context. No fitted calibration model, combined abundance index or completed validation is presented; the framework is proposed as a national methodology for discussion and refinement.

Keywords: Tuna fisheries, data integration, fishery-dependent data, fishery-independent surveys, artisanal fisheries, georeferenced catch and effort, India

1. Introduction

The Indian Ocean Tuna Commission (IOTC), the Food and Agriculture Organization of the United Nations (FAO) and other regional fisheries management organisations require Contracting Parties and Cooperating Non-Contracting Parties (CPCs) to submit species-wise, georeferenced information on catch and effort. IOTC Resolution 15/01 requires standardised recording of catch and effort at the fishing-vessel level, while Resolution 26/06 establishes mandatory statistical reporting requirements, and Recommendation 14/07 and Resolution 15/10 provide further context for catch and effort recording and scientific information (IOTC, 2026a, 2026b). These requirements are designed to support stock assessments, management strategy evaluation and informed decision-making.

For many CPCs with industrial or semi-industrial fleets, meeting these requirements is relatively straightforward: vessel logbooks, observer programmes and vessel monitoring systems (VMS) record fishing positions, effort and catch at the level of individual fishing operations. However, for CPCs with predominantly artisanal and small-scale fisheries, such as India, the collection of fishing-incident-wise information on species caught, their catch, effort, size and other biological parameters is inherently difficult. This paper explains the nature of this challenge in the Indian context, describes how India's fishery-independent survey system addresses it, and proposes a methodological framework for integrating heterogeneous data sources to generate the required species-wise, georeferenced information.

India's Artisanal and Small-Scale Tuna Fishery

India's marine fisheries sector is characterised by a vast, dispersed and highly diverse artisanal and small-scale fleet. The 2025 national report recorded 238,442 marine fishing craft and 1,537 landing centres (Sreenath et al., 2025). These craft operate from landing centres distributed along the mainland coast and island territories, using a wide range of gear types including gillnets, hooks and lines, ringseines, trawls and traditional craft. The tuna fishery is predominantly artisanal in nature: tropical tunas, yellowfin (*Thunnus albacares*), skipjack (*Katsuwonus pelamis*) and bigeye (*Thunnus obesus*), are caught as part of multi-species, multi-gear operations, often as incidental or supplementary catch alongside other pelagic and demersal species.

For 2024, tropical tuna catch was 60,530 t, comprising yellowfin (29,141 t), skipjack (30,943 t) and bigeye (446 t), compared with 221,665 t for tuna and tuna-like species together (Sreenath et al., 2025). These totals require clear species groupings and reporting periods. While India reported a nominal catch of 88,676.58 t of tropical tunas in 2025, the dispersed nature of the fishery means that catch is landed at hundreds of locations, by thousands of individual craft, on a daily basis.

The artisanal and small-scale character of the fishery has several important implications for data collection:

Dispersed landing sites: With 1,537 landing centres, many of which are small and remote, systematic sampling of all landings is logistically challenging. Landing surveys must rely on stratified sampling designs, with raising factors applied to estimate total catch.

Multi-species, multi-gear operations: Individual fishing trips often target multiple species using multiple gear types. Catch is not sorted by species at the point of landing in all cases, and effort cannot always be attributed to a single gear or target species.

Limited fishing-incident resolution: Fishermen do not routinely record fishing positions, effort or species-wise catch at the level of individual fishing operations. Logbooks are not universally maintained, and observer coverage is limited.

Coarse spatial and temporal resolution: Landing surveys typically record catch by landing centre and day (or trip), not by precise fishing ground. Inferred fishing grounds may be estimated from fishers' information, but these are approximate and may lack the precision required for georeferenced reporting.

Size and biological information: Collection of size composition and biological information (e.g., length, weight, sex, maturity) requires physical sampling of catches, which is labour-intensive and difficult to achieve representatively across a dispersed fleet.

These constraints mean that fishery-dependent data alone cannot fully meet the species-wise, georeferenced catch and effort reporting requirements of IOTC and FAO. This is not a failure of data collection effort, but an inherent characteristic of artisanal and small-scale fisheries.

India's Fishery-Independent Data Collection System

India has addressed this challenge through a well-established fishery-independent data collection system. The Fishery Survey of India (FSI), under the Department of Fisheries,

Ministry of Fisheries, Animal Husbandry and Dairying, operates a fleet of research and survey vessels that conduct exploratory and experimental fishing surveys across the entire Exclusive Economic Zone (EEZ) and adjoining seas. These surveys are designed to provide scientific observations independent of commercial fishing operations.

FSI surveys collect data that are fundamentally different in nature from fishery-dependent landing data:

Species-wise catch: All catch is identified to species level by trained scientific personnel. Species composition is recorded precisely, without the aggregation or misidentification that can occur in landing surveys.

Georeferenced fishing positions: The precise geographic position of each fishing operation (set) is recorded using GPS, along with depth, sea surface temperature and other environmental parameters. This provides the fine-scale spatial resolution required for georeferenced reporting.

Standardised effort: Effort is recorded in standardised units (e.g., number of hooks, number of sets, soak time), allowing calculation of catch per unit effort (CPUE) and hooking rates that are comparable across time and space.

Size and biological information: Biological sampling is conducted systematically, with length, weight, other morphometric characters, food and feeding, sex, maturity and other parameters recorded for representative samples of the catch. This provides the biological information required for stock assessment.

Consistent methodology: Survey design, gear specifications including size of hook, bait, fishing depth, and thermocline depth operational protocols are standardised across voyages, ensuring consistency and comparability of data over time.

The historical FSI inventory contains 10,703 hauls, 5,920,965 hooks, 71,761 fish and 1,641,256 kg of catch through 2021 (WPM17 proposal presentation, 2026). These figures will be checked against the original records before analysis. FSI surveys provide abundance indices for yellowfin tuna and other pelagic species (Varghese et al., 2011), while studies on longline bycatch and yellowfin tuna population structure provide additional biological information (Varghese et al., 2013; Varghese et al., 2019). Kuriakose et al. (2025) demonstrated the use of landing observations and fishers' information to identify probable fishing grounds, and the FSI survey data provide a complementary source of verified georeferenced information.

The Rationale for Data Integration

The fundamental insight of this paper is that fishery-dependent and fishery-independent data provide different but complementary types of information, and that their integration, properly conducted, can generate the species-wise, georeferenced catch and effort information required for stock assessment and management.

Fishery-dependent data (e.g., from the Fish Catch Survey and Analysis system, FCSA) provide:

- Coverage of the commercial fleet across all landing centres
- Estimates of total landed catch by species group, gear and fleet
- Information on fleet size, composition and operational patterns
- Context for interpreting the magnitude and distribution of catches

Fishery-independent data (from FSI surveys) provide:

- Species-wise catch and effort at fine spatial and temporal resolution
- Georeferenced fishing positions and environmental parameters
- Size composition and biological information
- Standardised CPUE and hooking rates that can serve as abundance indices

Neither source alone can meet all reporting requirements. Fishery-dependent data lack the spatial, species and biological resolution required; fishery-independent data lack the coverage and commercial context required to estimate total removals. Their integration, while maintaining their statistical identity, can provide a more complete picture.

However, integration must be conducted with scientific rigour. Fishery-dependent and fishery-independent data represent different observations, collected under different sampling designs, with different sources of bias and uncertainty. They should not be combined without scientific justification. As noted by Mukesh et al. (2019), the two components of Indian marine fisheries data collection, landing surveys and scientific surveys, have different coverage and sampling methods. The framework proposed in this paper is designed to preserve the statistical identity of each source while enabling complementary interpretation.

Scope and Structure of This Paper

This paper proposes a reproducible, rule-based framework for integrating heterogeneous fishery-dependent and fishery-independent tuna data in India, with the specific objective of generating species-wise, georeferenced information on catch, effort, size and other biological parameters. The framework follows four steps, preserve, standardise, link and reconcile, and is designed to be transparent, traceable and statistically defensible.

The paper is structured as follows. Section 2 explains the nature of the data integration challenge in detail, describing the characteristics of fishery-dependent and fishery-independent data and the specific reporting requirements they must meet. Section 3 presents the proposed methodological framework, describing each of the four steps and the protocols for data fusion, cross-validation and uncertainty treatment. Section 4 discusses the application of the framework, including the use of FSI survey data as the primary source of species-wise, georeferenced information and the role of landing data in providing coverage and context. Section 5 presents illustrative examples of how the framework might be applied. Section 6 concludes and identifies next steps.

No fitted calibration model, combined abundance index or completed validation is presented in this paper. The framework is proposed as a national methodology, not an algorithm specifically mandated by IOTC. Calibration will depend on adequate overlap and treatment of catchability and uncertainty.

2. The Data Integration Challenge

2.1 Characteristics of Fishery-Dependent Data in India

Fishery-dependent (FD) data in India are collected through the Fish Catch Survey and Analysis (FCSA) system, operated by ICAR-Central Marine Fisheries Research Institute (CMFRI) in collaboration with State/Union Territory fisheries departments. The FCSA system uses stratified sampling designs to estimate total landings by species group, gear and fleet.

Key characteristics of FD data include:

Observation unit: Landings at landing centres, recorded by trip or day

Spatial resolution: Landing centre location; fishing ground may be inferred from fishers' information but is not directly observed

Temporal resolution: Daily to monthly, depending on the landing centre and sampling intensity

Species resolution: Species or species groups, depending on the capacity of enumerators and the condition of landed catch

Effort: Number of craft, fishing days, gear units; not always attributable to specific fishing operations

Size and biological information: Collected through separate sampling programmes, with limited coverage

Raising factors: Applied to scale sample observations to total landings, with associated sampling errors

The FCSA system provides essential information on the magnitude and composition of landings, but it does not provide fishing-incident-wise, species-wise, georeferenced data. The spatial resolution is limited to the landing centre, and the link between landed catch and specific fishing operations is often uncertain.

2.2 Characteristics of Fishery-Independent Data in India

Fishery-independent (FI) data are collected through FSI exploratory and experimental fishing surveys. These surveys are designed to provide scientific observations of fish stocks independent of commercial fishing operations.

Key characteristics of FI data include:

Observation unit: Individual fishing operation (set), with precise georeferenced position

Spatial resolution: GPS position of each set; depth and environmental parameters recorded

Temporal resolution: Date and time of each set

Species resolution: All catch identified to species level by trained personnel

Effort: Standardised units (number of hooks, soak time, number of sets)

Size and biological information: Systematic biological sampling of representative catches

Survey design: Standardised protocols across voyages, with consistent gear, bait and operational procedures

The FSI survey data provide the fine-scale, species-wise, georeferenced information that FD data lack. However, they represent scientific observations, not commercial removals. Survey catch rates are influenced by survey design, gear selectivity, bait, fishing depth and other factors that differ from commercial operations.

2.3 The Integration Challenge

The challenge of integration arises from the fundamental differences between FD and FI data:

Different observation units: FD data record landings at landing centres; FI data record catches at sea, by fishing operation. Linking them requires matching landings to specific fishing operations, which is often not possible with current data.

Different spatial references: FD data are referenced to landing centres; FI data are referenced to fishing positions. Inferred fishing grounds from FD data are approximate and may not correspond precisely to FI survey positions.

Different effort units: FD effort is recorded in craft-days or gear-units; FI effort is recorded in hooks or sets. Converting between them requires assumptions about fishing power and catchability that may not be verifiable.

Different species resolution: FD data may aggregate species; FI data identify all species. Reconciliation requires careful matching of species codes and groups.

Different sources of bias: FD data are subject to reporting biases, recall errors, raising errors and sampling biases; FI data are subject to survey design biases, gear selectivity and catchability effects. These biases cannot be assumed to cancel out.

Different temporal coverage: FD data provide continuous coverage; FI data provide intermittent coverage, with surveys conducted at specific times and locations.

Given these differences, integration cannot proceed by simply combining datasets. Instead, a structured framework is required that preserves the statistical identity of each source, establishes verified links where possible, and documents unresolved differences.

2.4 Meeting the Reporting Requirements

The reporting requirements of IOTC and FAO demand species-wise, georeferenced information on catch and effort. In the Indian context, this information can be generated through the following approach:

Species-wise catch: FI survey data provide species-wise catch composition at fine resolution. FD data provide total landings by species group. By collating these sources, species-wise catch can be estimated for the commercial fleet, using FI data to inform species composition and FD data to inform total magnitude.

Georeferenced effort: FI survey data provide georeferenced effort at fine resolution. FD data provide effort at landing-centre resolution. By linking FD effort to inferred fishing grounds (using fishers' information and passive georeferencing), and by using FI data to characterise the spatial distribution of fishing effort, georeferenced effort can be estimated.

Size and biological information: FI survey data provide size composition and biological parameters at fine resolution. FD data provide limited biological information. By using FI data as the primary source of biological information, and FD data for context, size and biological parameters can be estimated for the commercial catch.

Catch and effort integration: The integration of FD and FI data allows the generation of species-wise, georeferenced catch and effort information that neither source alone can provide. This integration must be conducted with scientific rigour, maintaining the statistical identity of each source and documenting all assumptions and uncertainties.

The framework proposed in this paper is designed to achieve this integration in a transparent, reproducible and statistically defensible manner.

3. Proposed Methodological Framework

3.1 Overview

The framework aims to produce a source inventory, common data dictionary, verified record links, separate fishery indicators and a record of unresolved differences. It has not yet been tested with paired data. Similar landing and survey trends may support interpretation, but landed tonnage and survey hooking rate measure different things. Differences may reflect effort, sampling, gear, fishing grounds or catchability and do not necessarily mean that one source is wrong. Statistical relationships should be reported with the underlying data, sampling details and analysis methods.

The framework follows four sequential steps:

Preserve: Maintain original records unchanged, with full traceability.

Standardise: Apply common formats, codes and units without altering original information.

Link: Establish verified links between records describing the same activity, where sufficient evidence exists.

Reconcile: Examine differences between linked records, classify them, and document corrections and unresolved uncertainties.

Each steps are described in detail below.

3.2 Step 1: Preserve

The selected source files will be preserved unchanged, with processing restricted to working copies. Each extract will retain its version, date, coverage and record identifiers. Transformations and corrections will record the original value, revised value, method and reason, allowing analytical records to be traced to their sources. Later submissions will form new versions rather than replace the frozen inputs. Access to vessel-level information will follow the data custodians' conditions

Preservation ensures that the statistical identity of each source is maintained and that all subsequent processing is traceable. This is essential for transparency and reproducibility. (The original source files will be kept unchanged, and all processing will use working copies)

3.3 Step 2: Standardise A common dictionary will specify species and gear codes, catch units, ISO 8601 dates and decimal-degree coordinates while retaining original values. Landing dates and locations will remain distinct from fishing dates and measured or inferred fishing positions. Effort will retain gear-specific units; multi-gear catches will be allocated only when documented. Landing strata, raising factors and sampling errors, and survey gear, bait, depth and coverage will accompany the observations (Mukesh et al., 2019). Missing values will be separated from zeros. Format standardisation will not substitute for statistical treatment of fishing power and catchability.

Standardisation enables comparison and linkage without altering the original information. It is a prerequisite for the subsequent steps.

3.4 Step 3: Link

Verified vessel, trip and operation identifiers will link records describing the same activity. Separate commercial and research-vessel observations will instead be related by species,

fishing area, period and depth, without combining their catches. Spatial and temporal matching rules will reflect coverage and positional uncertainty. The $0.25^\circ \times 0.25^\circ$ grid of Kuriakose et al. (2025) offers a starting resolution, not a universal requirement; inferred fishing grounds will retain their distance, bearing and depth information. Unmatched records and the basis of every accepted link will remain identifiable.

Linking is the critical step that enables integration. It must be based on sufficient evidence and must not force links where evidence is lacking. Records that cannot be linked remain separate and are analysed independently.

3.5 Step 4: Reconcile

Discrepancies will be investigated after checking catch components, units and observation periods. Differences will be classified as explained, partially explained, unresolved or demonstrable errors. For example, hypothetical logbook and observer catches of 150 and 120 kg require confirmation that both describe the same operation and catch component. Measurement, identification, transcription and raising procedures will then be examined. Only verified errors will be corrected; legitimate or unresolved differences will retain both values and supporting notes.

Reconciliation ensures that differences are investigated systematically and that corrections are made only when justified. Unresolved differences are retained and reported, not forced to agree.

3.6 Data Validation

Validation will check species codes, units, duplicates, missing periods, positions and catch-effort consistency. Observer records may verify catches, while vessel tracks may corroborate activity but cannot independently measure catch or effective effort. Landing sampling uncertainty and survey coverage, missed stations and changes in catchability will be retained. The actual design of each survey series must be established (Mukesh et al., 2019). Calibration (estimate the relationship between the survey data and commercial fishery data) will be attempted only with credible commercial effort (the number of hooks used or fishing trips made) and comparable survey observations, using a model that considers the limits and possible errors in both datasets. The results will show data problems, differences that remain unexplained, and the steps used to process the data

3.7 Comparative Analysis

Separate FD estimates of catch by species, gear, fleet and period will be compared with FI sets, hooks, positive sets, catches and hooking rates. Comparisons will examine seasonal and spatial patterns within compatible strata while accounting for effort, gear and environmental conditions. A hypothetical Gulf of Mannar survey rate of 1.2 and a landing of 3 t by six boats cannot be compared directly without the rate denominator, capture dates, fishing locations and commercial effort. A hooking rate measures catch relative to survey effort, not biomass. Matching tolerances and statistical models will be documented.

3.8 Generation of Species-Wise, Georeferenced Information

The integrated dataset will be used to generate species-wise, georeferenced information on catch, effort, size and biological parameters. The approach is as follows:

Species-wise catch: FI survey data provide species composition at fine resolution. FD data provide total landings by species group. As landing records show only the total catch for a group of fish, the researchers may use the FI survey's species percentages to estimate how much of that total belongs to each species.

Georeferenced effort: FI survey data provide georeferenced effort. FD data provide effort at landing-centre resolution. Inferred fishing grounds from FD data (using fishers' information and passive georeferencing) are combined with FI spatial distributions to estimate georeferenced effort.

Size and biological information: FI survey data provide size composition and biological parameters. These are applied to FD catch estimates, within compatible strata, to estimate size and biological parameters for the commercial catch.

Uncertainty: Some results are estimates, so it may not be exact. We will report how much data were collected and explain what information is missing. Where possible, we will show how close each estimate may be to the actual value.

This approach recognises that FI data provide the primary source of species-wise, georeferenced information, while FD data provide coverage, magnitude and fleet-level context. The integration is conducted within a structured framework that maintains the statistical identity of each source and documents all assumptions and uncertainties.

4. Discussion

4.1 The Role of Fishery-Independent Data

Commercial landings and scientific survey catches provide different types of information and should be treated separately. Research-vessel catches should not be added to estimate commercial landings, and their catch per unit effort (CPUE) values should not be considered directly comparable. Passive georeferencing can help compare fishing locations, but using a finer grid does not improve the accuracy of the estimated positions (Kuriakose et al., 2025). Earlier survey studies provide useful information on abundance, bycatch and stock structure, but these differences between data sources must still be considered (Varghese et al., 2011, 2013, 2019).

In the Indian context, FSI survey data serve as the primary source of species-wise, georeferenced information. This is not because survey data are inherently superior to landing data, but because the artisanal and small-scale nature of the fishery means that landing data cannot provide the required resolution. The FSI survey system is well-established, with standardised protocols, trained personnel and comprehensive coverage of the EEZ and adjoining seas. It provides the species-wise, georeferenced, size and biological information that the reporting requirements demand.

4.2 The Role of Fishery-Dependent Data

Fishery-dependent data provide essential information on the magnitude and composition of commercial landings. They provide coverage of the fleet, information on fleet size and composition, and context for interpreting survey results. Without FD data, it would not be possible to estimate total removals or to understand the commercial context of the fishery.

FD data also provide information that FI data cannot: the actual landings of the commercial fleet, including species that may not be well-sampled by surveys, and the operational patterns of the fleet. This information is essential for stock assessment and management.

4.3 Integration Without Conflation

The framework proposed in this paper is designed to integrate FD and FI data without conflating them. The key principles are:

Maintain statistical identity: FD and FI datasets remain distinct throughout the process. They are never simply combined or added together.

Link only with evidence: Records are linked only when sufficient evidence indicates that they relate to the same activity. Unmatched records remain separate.

Document differences: Differences between linked records are investigated, classified and documented. Only verified errors are corrected; unresolved differences are retained.

Use complementary strengths: FD data provide coverage and magnitude; FI data provide species-wise, georeferenced, biological information. Each is used for what it does best.

Report uncertainty: All estimates are accompanied by measures of uncertainty. Calibration is attempted only with credible commercial effort and comparable survey observations.

4.4 Consistency with IOTC Frameworks

India has several independent sources of tuna data, including fishery-independent surveys conducted by the Fishery Survey of India (FSI), fishery-dependent landing estimates from CMFRI/FCSA, fishing vessel logbooks, observer records and vessel monitoring systems. These sources provide different but useful information on tuna fisheries. However, differences in their coverage, sampling methods, species identification, measurement practices and reporting procedures may result in variations in the data. Therefore, the integration of these datasets requires a systematic and transparent approach. IOTC Resolution 15/02 provides the relevant context for improving the reliability, completeness and consistency of fisheries data, while Resolution 15/01 is relevant to the standardised recording of catch and effort at the fishing-vessel level. The proposed rule-based data integration and validation framework follows four steps: preserve, standardise, link and reconcile.

The original records are preserved to maintain their identity and traceability. Data fields, units, species codes and reporting formats are standardised without altering the original information. Records are linked where sufficient evidence indicates that they relate to the same vessel, fishing trip or fishing operation. Finally, differences between linked records are examined, and their possible reasons, corrections and unresolved uncertainties are documented. Such differences may arise from variations in weighing practices, estimation methods, fish processing or evisceration onboard, sampling coverage, reporting errors or differences in the timing and purpose of data collection. Importantly, a difference should not automatically be treated as an error, and records should not be forced to agree without supporting evidence. Fishery-dependent and fishery-independent datasets should remain statistically distinct and be used for complementary validation wherever appropriate. For example, FSI survey indices may be compared with commercial landing trends to examine broad patterns, but differences in gear, fishing depth, effort and survey design must be considered before drawing conclusions. This

approach is consistent with the use of the best available scientific information, as reflected in the relevant IOTC scientific framework, including Recommendation 12/15. The four-step sequence is proposed as a scientifically defensible national methodology to improve data quality, transparency and traceability; it is not presented as an algorithm specifically mandated by IOTC. Its application can support more reliable interpretation of India's tuna data and provide a clear basis for future validation and scientific reporting.

4.5 Limitations and Next Steps

The framework proposed in this paper has not yet been tested with paired data. The illustrative examples are hypothetical and the displayed statistics are unverified. No fitted calibration model or combined abundance index is presented. The framework is proposed as a national methodology for discussion and refinement.

Next steps include:

Pilot application: Apply the framework to a pilot dataset covering overlapping periods of FD and FI data.

Verification: Verify historical inventory figures against original records.

Calibration: Develop and test calibration models, where adequate overlap exists, accounting for catchability and uncertainty.

Validation: Validate integrated estimates against independent sources.

Reporting: Report results, including unresolved discrepancies and uncertainties, to IOTC and other relevant bodies.

Conclusion

This paper has explained the challenge of meeting species-wise, georeferenced catch and effort reporting requirements in India's artisanal and small-scale tuna fishery, and has proposed a rule-based methodological framework for integrating heterogeneous fishery-dependent and fishery-independent data to address this challenge. The key points are:

The reporting challenge: IOTC, FAO and other organisations require species-wise, georeferenced information on catch and effort. In India's artisanal and small-scale fishery, with over 238,000 craft and 1,537 landing centres, fishing-incident-wise information on

species caught, their catch, effort, size and biological parameters is inherently difficult to collect through landing surveys alone.

The fishery-independent solution: India has a well-established fishery-independent data collection system, in which FSI research and survey vessels conduct exploratory and experimental fishing across the entire EEZ and adjoining seas. These surveys provide species-wise, georeferenced information on catch, effort, size and biological parameters at fine resolution.

The integration framework: The proposed framework follows four steps, preserve, standardise, link and reconcile, designed to maintain the statistical identity of FD and FI datasets while enabling complementary interpretation. Fishery-independent data serve as the primary source of species-wise, georeferenced information, while fishery-dependent data provide coverage, magnitude and fleet-level context.

The way forward: The framework is proposed as a national methodology for discussion and refinement. It has not yet been tested with paired data, and no fitted calibration model or combined abundance index is presented. Next steps include pilot application, verification, calibration, validation and reporting.

The authors welcome feedback from WPM17 participants on the scientific defensibility of the approach and its potential alignment with ongoing MSE developments for Indian Ocean tuna stocks.

Reference

IOTC. 2026a. Compendium of active conservation and management measures for the Indian Ocean Tuna Commission, 1 January 2026. Indian Ocean Tuna Commission, 352 pp.

IOTC. 2026b. Resolution 26/06: Mandatory statistical reporting requirements for IOTC Contracting Parties and Cooperating Non-Contracting Parties (CPCs). Conservation and management measures database. <https://iottc.org/cmm/resolution-2606-mandatory-statistical-reporting-requirements-iottc-contracting-parties-and> (accessed 14 October 2026).

IOTC Secretariat. 2026. Outcomes of the 27th and 28th Sessions of the Commission. IOTC-2026-WPNT16-04. June 2026. 4 pp.

IOTC-WPM17(MSE). 2026. Report of the 17th Session of the IOTC Working Party on Methods (Management Strategy Evaluation Task Force). Online, 23–24 March 2026. IOTC-2026-WPM17(MSE)-R[E]. 23 pp.

Kuriakose, S., Mini, K. G., Padua, S., Jayasankar, J., Varghese, E., Sreenath, K. R., and George, G. 2025. Spatial Distribution of Marine Fish Landings through a Modified Grid Framework. IOTC-2025-WPM16-24. 9 pp.

Mukesh, Varghese, S. P., Pandey, S., and Ramalingam, L. 2019. Marine fisheries data collection methods in India - an update. IOTC-2019-WPM10-21. 5 pp.

Sreenath, K. R., Varghese, S. P., Tiburtius, A., Siva, A., Das, A., Sinha, M. K., Pawar, R. U., Kizhakudan, S. J., Jayasankar, J., Koya, M., and Pandey, S. 2025. India's National Report to the Scientific Committee of the Indian Ocean Tuna Commission 2025. IOTC-2025-SC28-NR08_rev1E. 23 pp.

Varghese, S. P., Tiburtius, A., Vijayakumaran, K., Premchand, and Gulati, D. K. 2011. Interannual and geographic variations in the abundance indices of yellowfin tuna, billfishes and sharks in the Indian EEZ. IOTC-2011-WPTT13-18. Indian Ocean Tuna Commission.

Varghese, S. P., Vijayakumaran, K., and Gulati, D. K. 2013. Pelagic megafauna bycatch in the tuna longline fisheries off India. IOTC-2013-WPEB09-36. Indian Ocean Tuna Commission.

Varghese, S. P., Mukesh, Pandey, S., and Ramalingam, L. 2019. Recent studies on the population delineation of yellowfin tuna in the Indian Ocean - considerations for stock assessment. IOTC-2019-WPM10-18. Indian Ocean Tuna Commission.

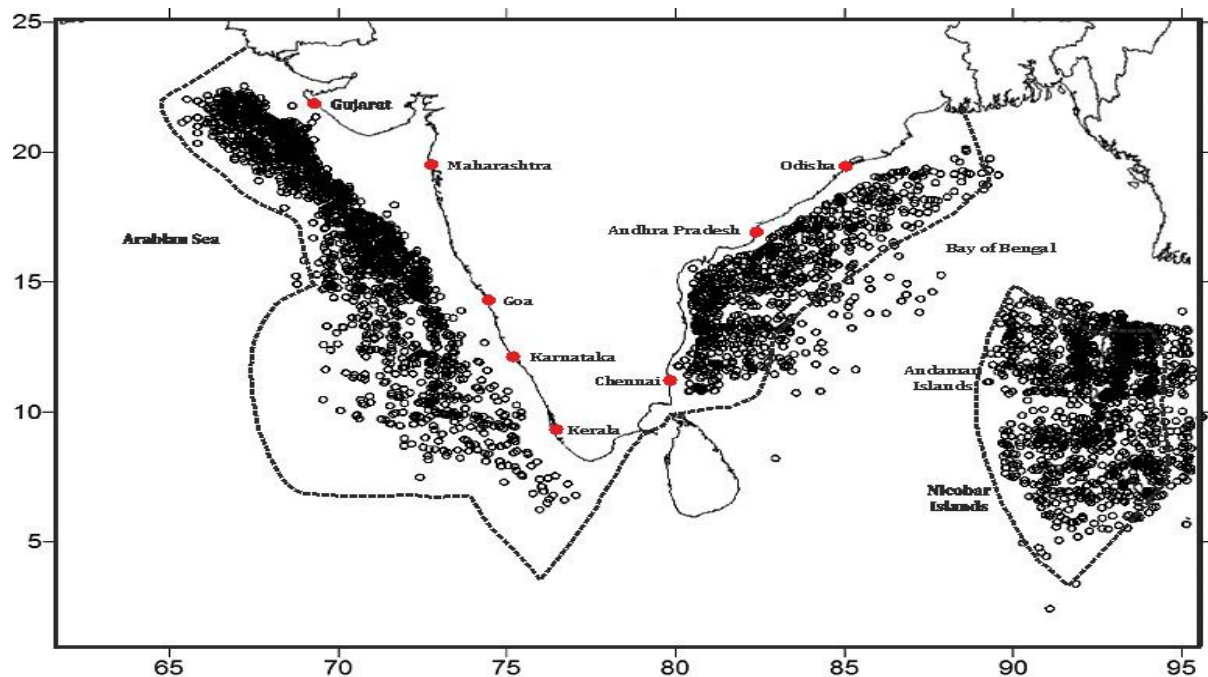


Fig. 1. Oceanic tuna longline survey sampling stations shown in the historical FSI inventory