Partnering with Regional Fisheries Management Organisations to build global data sets of tuna fisheries catches

Emmanuel Chassot1,2, Emmanuel Blondel2,, Peter Williams3, Nick Vogel4, Carlos Palma5, Colin Millar6, Tim Jones7, Sylvain Caillot4, Carlos Mayor5, Marc Taconet2, Fabio Fiorellato1,2, Bastien Grasset8, and Julien Barde8

# Affiliations

1. Indian Ocean Tuna Commission (IOTC)
2. Food and Agriculture Organisation of the United Nations (FAO)
3. Secretariat of the Pacific Community (SPC)
4. Inter-American Tropical Tuna Commission (IATTC)
5. International Commission for the Conservation of Atlantic Tunas (ICCAT)
6. Commission for the Conservation of Southern Bluefin Tuna (CCSBT)
7. Western and Central Pacific Fisheries Commission (WCPFC)
8. French national Research Institute for Sustainable Development (IRD)

Corresponding author: <Emmanuel.Blondel@fao.org>

# Abstract

We built the most complete global data set of annual catches of tunas and tuna-like species and pelagic sharks and rays for the period 1950-2021 by compiling public domain data available from the five tuna Regional Fisheries Management Organisations: the Commission for the Conservation of Southern Bluefin Tuna (CCSBT), the Inter-American Tropical Tuna Commission (IATTC), the International Commission for the Conservation of Atlantic Tunas (ICCAT), the Indian Ocean Tuna Commission (IOTC), and the Western and Central Pacific Fisheries Commission (WCPFC).

The data set covers medium- and large-sized pelagic species monitored by the five t-RFMOs and occurring in both neritic and oceanic habitats over the world oceans, including tunas, billfish, bonitos, Spanish mackerels, and wahoo, which belong to the Scombridae, Istiophoridae, and Xiphiidae families. In addition, catches reported for several pelagic sharks and rays that may be targeted or incidentally caught in tuna and tuna-like fisheries are included in the data set.

# Background and summary

*The Background & Summary should provide an overview of the study design and the data generated, including any background information needed to put this study in the context of previous work and the literature, and should reference literature as needed. The section should also briefly outline the broader goals that motivated collection of the data, as well as its potential use. We also encourage authors to include a figure that provides a schematic overview of the study design or workflow (if applicable).*

# Methods

*The Methods should include detailed text describing any steps or procedures used in producing the data, including full descriptions of the experimental design, data acquisition and any computational processing. See* [*the detailed section in our submission guidelines*](https://www.nature.com/sdata/publish/submission-guidelines#sec-5) *for advice on writing a transparent and reproducible methods section. Related methods should be grouped under corresponding subheadings where possible, and methods should be described in enough detail to allow other researchers to interpret and repeat, if required, the full study. Specific data outputs should be explicitly referenced via data citation (see Data Records and Citing Data, below). Authors should cite previous descriptions of the methods under use, but ideally the method descriptions should be complete enough for others to understand and reproduce the methods and processing steps without referring to associated publications. There is no limit to the length of the Methods section.*

# Data Records

The Data Records section should be used to explain each data record associated with this work, including the repository where this information is stored, and to provide an overview of the data files and their formats. Each external data record should be cited as described below. A data citation should also be placed in the subsection of the Methods containing the data-collection or procedure(s) used to derive the corresponding record.

Tables can be used to support the data records, specifying the data output resulting from each data-collection or analytical step and the names of the corresponding files should these form part of the archived record.

# Technical validation

The Technical Validation section should present any experiments or analyses that are needed to support the technical quality of the dataset. This section may be supported by figures and tables, as needed. *This is a required section*; authors must provide information to justify the reliability of their data.

Possible content **may include**:

* experiments that support or validate the data-collection procedure(s) (e.g. negative controls, or an analysis of standards to confirm measurement linearity)
* statistical analyses of experimental error and variation
* phenotypic or genotypic assessments of biological samples (e.g. confirming disease status, cell line identity, or the success of perturbations)
* general discussions of any procedures used to ensure reliable and unbiased data production, such as blinding and randomization, sample tracking systems, etc.
* any other information needed for assessment of technical rigour by the referees

Generally, this **should not include**: - follow-up experiments aimed at testing or supporting an interpretation of the data - statistical hypothesis testing - exploratory computational analyses

# Usage notes

*This section is optional*

The Usage Notes can contain brief instructions to assist other researchers with reuse of the data. This may include discussion of software packages that are suitable for analysing the assay data files, suggested downstream processing steps (e.g. normalization, etc.), or tips for integrating or comparing the data records with other datasets. Authors are encouraged to provide code, programs or data-processing workflows if they may help others understand or use the data. Please see our code availability policy for advice on supplying custom code alongside Data Descriptor manuscripts. For studies involving privacy or safety controls on public access to the data, this section should describe in detail these controls, including how authors can apply to access the data, what criteria will be used to determine who may access the data, and any limitations on data use.

# Code availability

For all studies using custom code in the generation or processing of datasets, a statement must be included under the subheading “Code availability”, indicating whether and how the code can be accessed, including any restrictions to access. This section should also include information on the versions of any software used, if relevant, and any specific variables or parameters used to generate, test, or process the current dataset.

# Acknowledgments

The Acknowledgements should contain text acknowledging non-author contributors. Acknowledgements should be brief, and should not include thanks to anonymous referees and editors or effusive comments. Grant or contribution numbers may be acknowledged.

# Author contributions

Each author’s contribution to the work should be described briefly, on a separate line, in the Author Contributions section.

# Competing interests

A competing interests statement is required for all papers accepted by and published in Scientific Data. If there is no conflict of interest, a statement declaring this must still be included in the manuscript.

# Figures

Figure images should be provided as separate files and should be referred to using a consistent numbering scheme through the entire Data Descriptor. We discourage the inclusion of figures in the Supplementary Information - all key figures should be included here in the main Figure section.

For initial submissions, authors may choose to supply a single PDF with embedded figures. You will later be asked for separate files closer to publication.

Authors are encouraged to consider creating a figure that outlines the experimental workflow(s) used to generate and analyse the data output(s).

# Figure legends

Figure legends begin with a brief title sentence summarizing the purpose of the figure as a whole, and continue with a short description of what is shown. Legends should ideally be no more than 350 words, and may contain literature references. The first sentence of the legend will be used as the title for the figure. It should contain no references of any kind, including to specific figure panels, bibliographic citations or references to other figures or panels.

# Tables

Tables in the manuscript should generally not be used to present primary data (i.e. measurements). Tables containing primary data should be submitted to an appropriate data repository.

Authors may provide tables within the Word document or as separate files (tab-delimited text or Excel files). Legends, where needed, should be included in the Word document. Tables may be of any size, but only tables that fit onto a single printed page will be included in the PDF version of the article (up to a maximum of three).

Due to typesetting constraints, tables that cannot be fit onto a single A4 page cannot be included in the PDF version of the article and will be hosted as Supplementary Tables. Any such tables must be labelled in the text as ‘Supplementary’ tables and numbered separately from the main table list e.g. ‘Table 1, Table 2, Supplementary Table 1’ etc. Please note bibliographic references cannot be included within Supplementary Tables and should not be listed in the reference list, which only refers to references used in the main article file. If you do wish to formally cite information used in any supplementary file, please find a means of mentioning these references on the main text. Finally, please note it may be preferable to host large tables within your repository-deposited dataset, as highlighted above.

# References