Open ecological data for tuna: The time has come!

Emmanuel Chassot[1,2], Julien Barde[3], Laurent Floch[4], Laurent Ibanez[2,5], Nathalie Bodin[1,2,*]

[1] Institut de Recherche pour le Développement, SFA, BP570, Victoria, Seychelles
[2] Seychelles Fishing Authority, PO BOX 449, Victoria, Seychelles
[3] Institut de Recherche pour le Développement, Indian Ocean Commission, Ebène, Mauritius
[4] Institut de Recherche pour le Développement, Sète, France
[5] LIENSs, UMR 7266, CNRS-Université de La Rochelle, 2 Rue Olympe de Gouges, F-17000, La Rochelle, France

[*] Corresponding author: emmanuel.chassot@ird.fr

Abstract
We describe the methodological approach developed to collect and manage a large range of ecological data for tuna sampled throughout several routine monitoring and research programs conducted in the Indian Ocean in the last decades. To cover the large habitats of occurrence of tunas, collaboration with fishermen and processing factories was instrumental in fish sampling. We propose a minimum set of standard variables to be collected to describe and keep track of the fishing environment. In industrial fisheries, information from logbooks and traceability of tunas through the storage process is key to fully represent the spatio-temporal uncertainty associated with the capture of each fish. Morphometric measurements and ecological tracers are hosted in a relational database in a way that allows for any type of measure to be added without modifying the database design and structure. Standard code lists and metadata are used to describe the data managed in the database, which facilitates their visibility and diffusion. We argue that the use of standard, generic, well described sampling and analytical protocols combined with standard code lists is key for sharing and merging data sets at the scales of interest of large pelagic species such as tunas. The comprehensive description of the methods used to collect and produce ecological data requires interactions with computer scientists to make the most of ecological data that are costly and mostly financed by public funds.

Keywords: biochemical tracers; otoliths; morphometric data; spawning

1- Introduction

Ecological data collected on fish populations are managed by the different administrative and scientific institutes, Universities, and NGOs that conduct the monitoring and research programmes. The cost of the data includes the logistics, materials, and human resources employed during the design of the studies, the acquisition of funding, the collection of the samples, their preparation and analysis with methods that can be very sophisticated in some cases (e.g. magnetic resonance spectroscopy), and the management of the data. Data acquisition is therefore a long process with culminates with the publication of scientific results in reports and peer-review articles and can promote careers and support the recognition of an individual or research group. For this reason, ecological data are difficult to access and their sharing is generally restricted to short-term collaborations within a specific project. Even when not used, data are generally not released or shared as they are expected to be used in the future. Another key limiting factor for data release is the cost in time associated with the description of the data necessary for the (potential) user and the lack of knowledge on metadata methods. Consequently, it is not only difficult to access published data in fisheries ecology but having a good knowledge of the sets of ecological data existing in a given marine region is currently intractable.

Sharing ecological data is however key to promote the broad, integrative analyses that are required for fisheries management. The development of standards for metadata and data formats and access protocols has recently enhanced interoperability functions in information systems, supporting open access to ecological data (Reichman et al. 2011). This is particularly relevant for tuna populations that cover areas of millions square kilometres and are exploited by several fishing nations. A large range of data on physiological and ecological traits of tuna are collected everyday through routine monitoring programs and research projects globally. Data include simple size measurements of body components and organs, descriptors of spawning activity, estimates of age and individual growth, and several biochemical tracers which provide information on habitat use and trophic interactions. In fisheries monitoring, morphometric data are key to derive conversion equations to standardize size data, raise size samples to catch weight, and compute sampling coverage. In longline fisheries, conversions of length into weight has been shown to be critical for stock assessment as they can affect the information provided on gear selectivity (Langley et al. 2006, Minte-Vera et al. 2016). In purse seine fisheries, the species composition of the catch is generally derived from size sampling at landing after conversion of length into weight (Pallarés & Hallier 1997, Pianet 1999). Poor estimation of length-weight relationships can bias estimates of species composition, with potential effects on management (Marsac et al. 2017). The IOTC Secretariat manages a database of length, weight, and sex data (IOTC 2005) which are however limited in spatio-temporal coverage and lack visibility. Morphometric data are also essential for research purpose, including monitoring changes in population age structure and spawning activity over time as well as modelling energy allocation (e.g. Jusup et al. 2011). In addition, biochemical tracers such as stable isotopes ratios of carbon and nitrogen combined with information on reproduction and growth have given new insight into the mechanisms of energy allocation in Indian Ocean tuna (Zudaire et al. 2014, 2015, Sardenne et al. 2016, Grande et al. 2016).

In the present paper, we describe the methodological approach developed throughout several routine monitoring and research programs conducted on tuna in the Indian Ocean, including some data from the Indian Ocean Tuna Tagging Program (IOTTP), to showcase how ecological data for tuna can become open-access. The overarching objective of the study is to promote discussions on the opportunities and challenges raised by data openness. First, we describe the main features of the fish sampling operations and how collaboration with fishermen was instrumental in retrieving information on the origin of each fish. We argue that confidentiality issues can easily be addressed and solved for ecological data through removal of individual names of vessels or companies. Second, we provide the general principles we adopted to describe the different steps of the process conducting from fish collection to final data acquisition in a database built for hosting all ecological data collected. We emphasize the need to use open, standard protocols for data collection and analysis, and describe each step of the process with metadata to promote full understanding and transparency of the data and foster collaboration through data sharing. Our approach allows for full traceability of the samples throughout their lifetime, including storage methods and location, to help in data interpretation in case of apparently aberrant data records, and support the development of a bank of samples for future studies on tuna ecology.

2- Fish sampling

More than 44,000 tunas were sampled throughout several projects conducted in the Indian Ocean during 1987-2015 from five species: albacore tuna (*Thunnus alalunga*; ALB), bigeye tuna (*Thunnus obesus*; BET), frigate tuna (*Auxis thazard*; FRI), skipjack tuna (*Katsuwonus pelamis*; SKJ), and yellowfin tuna (*Thunnus albacares*; YFT) (**Table 1**). Species are defined following the standard ASFIS classification (Garibaldi and Busilacchi 2002). Sampling operations took place either at sea by observers onboard fishing vessels, or at land during unloading of the vessels at landing sites or at processing plants in Seychelles, Mauritius, South Africa, and La Réunion. The sampling location, status of the fish (i.e. fresh or frozen) and conditions of storage (mode and duration) of the samples were recorded as it may affect the morphometric measurements and the nature of the analyses of ecological tracers that can be conducted thereafter (see **Section 3**).

Table 1. Annual number of tunas sampled by species

Year	ALB	BET	FRI	SKJ	YFT
1987	0	23	0	0	615
1988	0	240	0	0	707
1989	0	68	0	0	481
1990	0	625	0	0	899
1991	0	5	0	0	32
2003	0	0	0	2	316
2004	0	0	0	41	1199
2005	0	0	0	0	2413
2006	0	45	0	5	4542
2007	0	185	0	72	2299
2008	0	69	0	4	2986
2009	0	6	0	782	4286
2010	1	7	0	582	4611
2011	0	33	0	0	2895
2012	47	16	0	0	3653
2013	367	208	4	164	1851
2014	1437	430	0	983	1075
2015	337	558	0	857	1121

Information on the fishing environment was retrieved thanks to the collaboration of fishermen and staff at processing factories (**Table 2)**. Access to skipper logbooks and well plans that describe the location of the fish storage in purse seiners, as well as information on the traceability of the fish processed at tuna factories enabled linking tunas sampled with fishing operations. Uncertainty arose when tunas were sampled from purse seiners brine-freezing wells that can contain fish from multiple sets. The uncertainty was larger for tunas sampled at the IOT Ltd. cannery of Victoria where the tuna is tracked back to the fishing trip. In such cases, the range of fishing dates was recorded and information on spatial uncertainty was computed as the surface area (km²) inside the convex hull of the positions of the potential fishing sets. Also, only the range of dates and longitude/latitude was available for some tunas collected from deep-freezing longliners landing in Port-Louis, Mauritius. In this case, spatial uncertainty was computed as the surface area (km²) of the polygon linking the extreme fishing positions collected from fishermen. For each tuna, the fishing gear of the commercial and recreational fisheries that captured the fish was recorded as it provides qualitative information on fishing depth and size/age selectivity based on the FAO standard classification of fishing gears (FAO 2016). In addition, the type of school association (i.e. free-swimming school or school associated with a floating object) was recorded as it affects habitat use and trophic interactions (e.g. Jaquemet et al. 2011). No standard classification for the type of school association exists among tuna RFMOs yet but some initiatives have recently started to propose a standard classification (Gaertner et al. 2016).

Table 2. List of variables used to describe the fishing environment of the tunas sampled

Variable	Description
ocean_code	Acronym of the ocean where the fish was caught
gear_code	The gear that caught the fish
vessel_name	Name of the fishing vessel that caught the fish
vessel_code	Code of the fishing vessel that caught the fish
landing_date	Date of arrival of the fishing vessel in the port
landing_site	Place of arrival of the fishing vessel in the port
fishing_date	Date the fish was caught
fishing_date_min	Minimum date of fishing of the fish when the accurate date could not be collected
fishing_date_max	Maximum date of fishing of the fish when the accurate data could not be collected
activity_number	Sequential number of activity in the day
fishing_time	Time of fishing operation
latitude_deg_dec	Latitude of the location of capture
latitude_deg_dec_min	Minimum latitude of the location of capture when the accurate position could not be collected
latitude_deg_dec_max	Maximum latitude of the location of capture when the accurate position could not be collected
longitude_deg_dec	Longitude of the location where the fish was caught
longitude_deg_dec_min	Minimum longitude of the location of capture when the accurate position could not be collected
longitude_deg_dec_max	Maximum longitude of the location of capture when the accurate position could not be collected
school_type	Aggregated fishing mode in the case of purse seine fishing: nature of the association of the fish, i.e. fob-associated school or free-swimming school
well_number	The number of the freezing-well in the fishing vessel
well_position	The position of the tank in the fishing vessel (portside or starboard)
vessel_storage_mode	The mode of conservation of fish in the vessel (brine, freeze, chill)

Different morphometric measurements for the whole fish and organs were collected for each tuna (**Table 2**). The information is stored in the database in a table with a column 'measure_name' and a column 'measure_value' for each fish identifier, which provides a very generic structure where any type of new measurement can be added to the database. Each measurement type is described in a table of metadata. Information on the tool used for measurement and associated precision is described. In addition, the sex (M: Male; F: Female; Ind: Indeterminate) and macroscopic maturity stage for females were derived from visual exam (**Table A2**). Although the macroscopic maturity stage does not accurately describe the spawning oocyte development stage in tunas (Albaret 1977), it provides qualitative information on the spawning activity of the fish (i.e. developing vs. mature) and it was found useful to identify gonads to be further analysed with histological methods.

Table 2. Description of morphometric measurements collected on tunas

Notation	Definition	Description	Precision	Tool
L_F	Fork length	Projected straight distance from the tip of the upper jaw (snout) to the shortest caudal ray (fork)	0.1 cm	Calliper
L_{CF}	Curved fork length	Projected curved-body distance from the tip of the upper jaw (snout) to the shortest caudal ray (fork)	0.1 cm	Tape measure
L_D	Predorsal length	Projected straight distance from the tip of the snout to the anterior base of the first dorsal fin	0.1 cm	Calliper
L_{TG1}	First thorax girth	Projected curved-thorax length just behind the first pectoral fins, i.e. where the fish height is greatest	0.1 cm	Tape measure
L_{TG2}	Second thorax girth	Projected curved-thorax length in front of the second dorsal fin and the anal fin	0.1 cm	Tape measure
W_T	Total weight	Weight of the whole fish	0.1 g	Scale
W_G	Gonad weight	Weight of the two gonads	0.1 g	Scale
W_L	Liver weight	Weight of the liver	0.1 g	Scale
W_{FS}	Full stomach weight	Weight of the full stomach (stomach tissue and contents)	0.1 g	Scale
W_{ES}	Empty stomach weight	Weight of the stomach after removing the contents	0.1 g	Scale
W_V	Rest of viscera weight	Weight of all viscera (i.e., heart, oesophagus, pylorus, intestine, pyloric caeca and anus) without liver and stomach	0.1 g	Scale

3- Sample analysis

Hard and soft tissues from different organs of the tunas were sampled according to the sampling protocols in use within each project (**Fig. 1**).

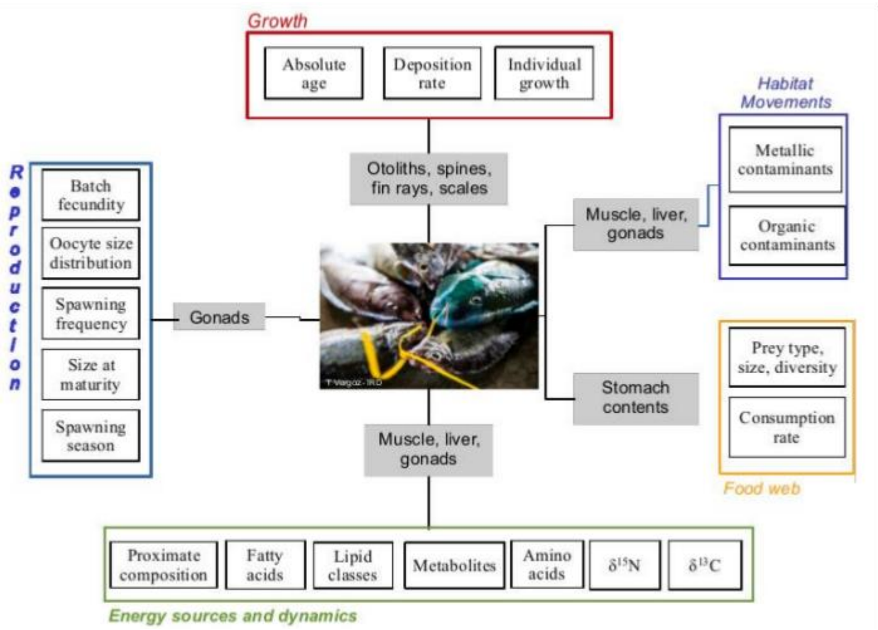

Fig. 1: Scheme illustrating the different tracers collected from tunas tissues and their domain of application

The conditions of sampling (i.e. fish state, tools, storage conditions, etc.) determine the subsequent analyses that can be carried out. For instance, microchemistry of otoliths or estimation of contaminant concentration require specific tools and methods of sample handling that will not affect the chemical composition. For this reason, it is essential to (i) develop generic sampling protocols that will facilitate further studies not planned at the time of the project, (ii) comprehensively describe and make available the protocols used for data collection (e.g. https://www.protocols.io/), and use lab procedures that are internationally recognized (e.g. ISO/IEC 17025). For instance, delipidation of tissue samples has been shown to be essential for analysis of stable isotopes ratios in fish tissue (Bodin et al. 2006) and should be systematically conducted to promote comparative analyses between populations and oceans. Also, conditions of storage (freezer, etc.) and duration can have major repercussions on analysis results such as fatty acid profiles. Storage of fish in brine-freezing wells for periods corresponding to typical purse seine fishing trips has however been shown not to affect the lipid and fatty acid composition, carbon and nitrogen stable isotope ratios, and organochlorine contaminant levels in tuna muscle (Bodin et al. 2014).

Similarly to morphometric measurements, the database is conceived to host any type of ecological tracer and biological descriptor. Generic information (i.e. type, description, lab, operator, mode, date, mass, mass unit) and values of analysis are stored in two linked tables. Additional tables enable to add specific features for each type of analysis. Information on all fields is described with metadata tables.

4. Data management

[TO DO]

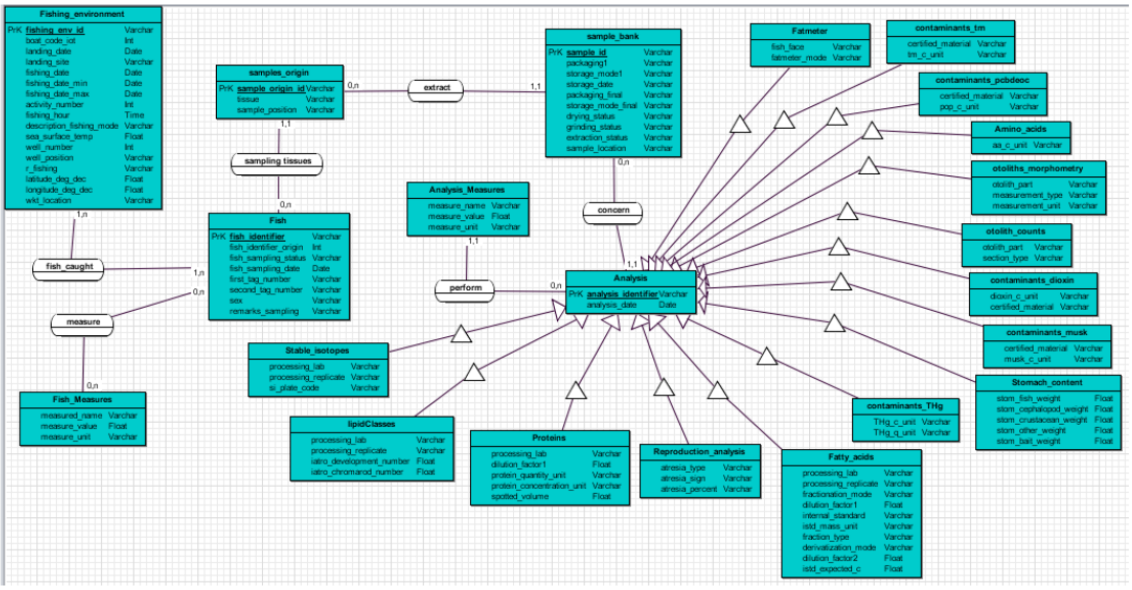

Fig. 2. *Simplified structure of the database*

5. Free ecological data, now[1]!

[TO DO]

1 https://www.ted.com/talks/tim_berners_lee_the_year_open_data_went_worldwide

Acknowledgements. We thank the many fishermen, observers, technicians, managers and staff at processing factories, and students who assisted with tuna collection and sample processing in the laboratory of SFA (Victoria, Seychelles) and the analytical platforms of the LIENSs (La Rochelle, France), MetaToul-LIPIDOMIQUE (Toulouse, France) and EME/MARBEC-Ecotrace (Sète, France). We are particularly grateful to Fany Sardenne, Iker Zudaire, Maitane Grande, Zahirah Dhurmeea, Hilario Murua, Hervé Guillou, Catherine Munschy, Paco Bustamante, Wendy West and Jérôme Bourjea for close collaboration over the last five years. This work benefited from funding from the French National Research Agency (ANR-11-JSV7-007-01), the Thaï Union Europe Group, France Filière Pêche (FFP), and the Fisheries European Programme (FEP 759/DMSOI/2013).

References

Albaret J-J (1977) La reproduction de l'albacore (*Thunnus albacares*) dans le Golfe de Guinée. Cah ORSTOM Sér Océan 15:389–419

Bodin N, Le Loch F, Hily C, Abarnou A (2006) Influence of lipid extraction on carbon and nitrogen stable isotopes in crustacean tissues. Assessment of polychlorobiphenyl bioaccumulation in the spider crab food web using stable isotopes. In: Belfast, Ireland

Bodin N, Lucas V, Dewals P, Adeline M, Esparon J, Chassot E (2014) Effect of brine immersion freezing on the determination of ecological tracers in fish. Eur Food Res Technol 238:1057–1062

Grande M, Murua H, Zudaire I, Arsenault-Pernet EJ, Pernet F, Bodin N (2016) Energy allocation strategy of skipjack tuna *Katsuwonus pelamis* during their reproductive cycle. J Fish Biol 89:2434–2448

IOTC (2005) Biological data on tuna and tuna-like species gathered at the IOTC Secretariat: Status report. In: IOTC Proceedings.p 22

Jaquemet S, Potier M, Ménard F (2011) Do drifting and anchored Fish Aggregating Devices (FADs) similarly influence tuna feeding habits? A case study from the western Indian Ocean. Fish Res 107:283–290

Jusup M, Klanjscek T, Matsuda H, Kooijman SALM (2011) A full lifecycle bioenergetic model for bluefin tuna. PLoS ONE 6:e21903

Langley A, Okamoto H, Williams P, Miyabe N, Bigelow K (2006) A summary of the data available for the estimation of conversion factors (processed to whole fish weights) for yellowfin and bigeye tuna. In: Second regular session of the Scientific Committee of the Western-Central Pacific Fisheries Commission. WCPFC, p 21

Minte-Vera CV, Aires-da-Silva A, Satoh K, Maunder MN (2016) Changes in longline size-frequency data and their effects on the stock assessment models for yellowfin and bigeye tunas. In: Seventh meeting of the Scientific Advisory Committee of the Inter-American Tropical Tuna Commission, 09-13 May 2016, La Jolla, California, USA. IATTC, La Jolla, California (USA), p 22

Pallarés P, Hallier J-P (1997) Analyse du schéma d'échantillonnage multispécifique des thonidés tropicaux. IEO/ORSTOM

Pianet R (1999) Evolution du système de collecte et de traitement des données de la pêche thonière des senneurs européens et assimilés de 1981 à 1998. In: IOTC Proceedings.p 74–96

Reichman OJ, Jones MB, Schildhauer MP (2011) Challenges and Opportunities of Open Data in Ecology. Science 331:703–705

Sardenne F, Bodin N, Chassot E, Amiel A, Fouché E, Degroote M, Hollanda S, Pethybridge H, Lebreton B, Guillou G, Ménard F (2016) Trophic niches of sympatric tropical tuna in the Western Indian Ocean inferred by stable isotopes and neutral fatty acids. Prog Oceanogr

Zudaire I, Murua H, Grande M, Goni N, Potier M, Ménard F, Chassot E, Bodin N (2015) Variations in the diet and stable isotope ratios during the ovarian development of female yellowfin tuna (*Thunnus albacares*) in the Western Indian Ocean. Mar Biol 162:2363–2377

Zudaire I, Murua H, Grande M, Pernet F, Bodin N (2014) Accumulation and mobilization of lipids in relation to reproduction of yellowfin tuna (*Thunnus albacares*) in the Western Indian Ocean. Fish Res 160:50–59

Appendix tables

Table A1. Macroscopic maturity scale for male and female tunas (source: ICCAT)

Code	Label	Description - Males	Description – Females
1	Immature	Testes extremely thin, flattened and ribbon-like, but sex determinable by gross examination	Gonads elongated, slender, but sex determinable by gross examination
2	Developing	Enlarged testes, triangular in cross section, no milt in central canal	Gonads enlarged but individual ova not visible to the naked eye
3	Late developing	Milt flows freely if testes pinched or pressed	Gonads enlarged, individual ova visible to the naked eye
4	Spawning	Testes flabby, bloodshot, surface dull red, little or no milt in central canal	Ovary greatly enlarged, ova translucent, easily dislodged from follicles or loose in lumen of ovary
5	Spent, spawned, regressing, regenerating	Testes flabby, bloodshot, surface dull red, little or no milt in central canal	Includes recently spawned and postspawning fish, mature ova remnants in various stages of resorption, and mature ova remnants about 1.0mm in diameter
999	Unknown	Not determined	Not determined