Indian Ocean Tuna Commission
Commission des Thons de l'Océan Indien
iotc ctoi

Food and Agriculture Organization of the United Nations

ELECTRONIC TOOLS IN SUPPORT OF THE IOTC ROS DATA COLLECTION AND REPORTING WORKFLOW

PREPARED BY: IOTC SECRETARIAT[1], 23RD NOVEMBER 2018

BACKGROUND

The ROS electronic tool suite is a set of data models, software components and applications developed by the IOTC Secretariat – with additional support from NOAA, WWF and SIOTI – as part of the Regional Observer Scheme pilot project (Resolution 16/04 *On the implementation of a pilot project in view of promoting the Regional Observer Scheme of IOTC*).

The purpose of the ROS tools is to facilitate the process of collecting, reporting and managing ROS data at national and regional level, by providing a fully integrated and flexible solution to all IOTC CPCs that currently lack a comprehensive data management workflow for their scientific observer data (to be collected according to the requirements set forth by Resolution 11/04 *On a regional observer scheme*).

Furthermore, by adopting selected sub-components belonging to the ROS tool suite, all CPCs already collecting and managing scientific observer data in a streamlined manner – according to the ROS guidelines – could further improve the efficacy and timeliness of the submission of these same data to the IOTC Secretariat, and therefore contribute to reduce the *time-to-market* of the information collated and eventually disseminated by the ROS Regional Database to the scientific community.

Additionally, the availability of tools and standardized data exchange formats built on top of the ROS standard specifications[2] will help reduce the proliferation of custom, incompatible data reporting standards that have represented one of the major challenges to the effective incorporation of scientific information collected by national observer programs in the region.

ROS DATA COLLECTION AND MANAGEMENT WORKFLOW

According to the current ROS requirements, the information to be collected by scientific observers is categorized in three distinct sets:

- Fields marked as **mandatory for collection**, that must be recorded by observers as part of their data collection routine, and that will be submitted *only* to the corresponding national focal point;

[1] F. Fiorellato (fabio.fiorellato@fao.org); J. Geehan (james.geehan@fao.org); S. Martin (sarah.martin@fao.org); L. Pierre (lucia.pierre@fao.org)
[2] http://iotc.org/sites/default/files/documents/science/IOTC-2015-ROS_11_04_Observer_Manual_v1.2.pdf

- Fields marked as **mandatory for reporting**, that must be recorded by observers and submitted to the corresponding national focal point and to the IOTC Secretariat;
- Fields marked as **optional** or **recommended**, that can *optionally* be recorded by the observer and will be submitted (when available) to the national focal point only.

Table 1 summarizes the data exchange requirements depending on the category to which the ROS data fields belong:

Data field category	To be collected by the observer	To be submitted to:	
		National focal point	IOTC Secretariat
Mandatory for **collection**	**Always**	**Yes**	**No**
Mandatory for **reporting**	**Always**	**Yes**	**Yes**
Recommended	*Optionally*	*When collected*	**No**

Table 1. Summary of the collection and reporting requirements for all ROS data fields according to their category

The *expert review workshop on standards for the IOTC Regional Observer Scheme* recently held in Victoria, Mahé (September 24–28, 2018) reviewed the rationale and practicalities of the current data collection (and data reporting) fields, with the outcomes of this workshop described in a separate paper[3].

Regardless of the actual operational details adopted by CPCs for the implementation of their scientific observer data collection and management processes **under the IOTC ROS**, the general workflow does usually include four distinct *actors*, exchanging information across the domain boundaries as described in Figure 1.

- The **scientific observer**: an accredited individual responsible for the independent collection of the scientific information during a trip;
- The **national focal point**: generally, the national fisheries or research institution responsible for the collation of information collected by scientific observers operating under the flag state;
- The **IOTC Secretariat**: receives official data submissions of scientific observer data from CPCs by each flag state (i.e., a sub-set of the information collated by the respective national focal points), for their incorporation within the ROS regional database;
- The **Scientific community**: accesses the observer information reported to the IOTC Secretariat and disseminated according to the rules and policies dictated by Resolution 12/02 (*Data confidentiality policy and procedures*).

[3] IOTC-2018-WPDCS14-35:http://www.iotc.org/sites/default/files/documents/2018/11/IOTC-2018-WPDCS14-35_-_ROS_standards.pdf

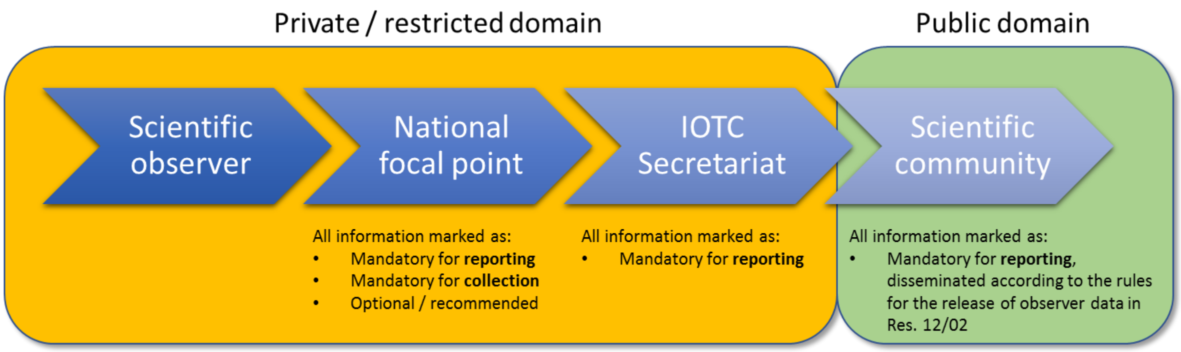

Figure 1. General ROS data collection and dissemination workflow across the private / public domains

How the different steps of the scientific observer data collection workflow are implemented in practice can vary by CPC.

The ROS suite of tools addresses the need to standardize the data collection, manage and report scientific observer data for all those CPCs that do not yet have well-established data collection mechanisms already in place: each ROS tool is designed to support a specific phase in the management workflow, from the data collection to the data dissemination.

ROS E-COLLECTION AND MANAGEMENT INTERFACE

The ROS e-Collection and management interface (ROS CI) is a tool designed to support scientific observers in organizing the information collected during a trip, validate its adherence to the ROS requirements and eventually submit the results – including the data themselves – to the national focal points for their further analysis and collation. As such, the ROS CI implements the **first** step of the ROS data collection workflow (see Figure 1).

The ROS CI is a multi-platform tool that does not require Internet connectivity to work out of the box, and its technical specifications are as follows:

- **Current version:** 0.9.9
- **Operating system:** Windows / Mac OS X / Linux (in preparation)
- **Minimum RAM**: 4GB
- **Minimum free space required on the HD**: 1GB
- **Recommended screen resolution**: 1280x1024 pixels or above
- **Pre-requisites**: Java 8 (either JDK or JRE) installed on the same machine
- **Language**: English / French
- **Download link**: https://goo.gl/3iJ3po (Windows) / https://goo.gl/kq9ha4 (Mac OSX)

When network connectivity is available (typically, before the observer is deployed onboard) the ROS CI can communicate with the IOTC ROS Regional Database and provide additional features such as:

- Update to a newer version (when available);
- Synchronize the reference code tables;
- Retrieve observer information from the list of IOTC accredited observers;
- Retrieve vessel information from the IOTC RAV (Record of Authorized Vessels).

The ROS CI supports the management of the ROS data requirements currently expressed by the ROS observer manual[4], and provides data entry facilities for **all** data fields (i.e., *mandatory for collection*, *mandatory for reporting* and *optional* / *recommended*) such as auto-completion, dropdown menus, value spinners, time pickers, geolocation selection through interactive maps, field validators and error reporting procedures, to ensure that the collected information is entered in the system in the most user-friendly and error-free way as possible.

A *trip report* feature is also available, showing the major events recorded by the observer as these unroll during the trip.

Access to the ROS CI is currently restricted to users in the list of IOTC accredited observers: to enroll / remove an observer from this list, CPCs should submit official request to the IOTC Secretariat including references to the required documentation (basic training and medical certificates, etc.)

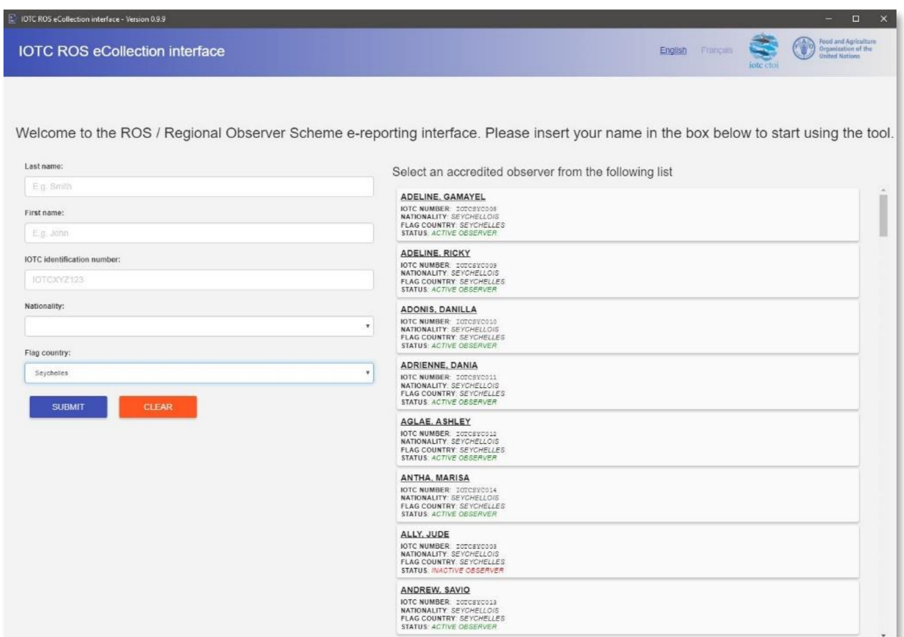

Figure 2. The ROS CI entry page, showing the list of currently accredited observers for a single CPC

[4] http://iotc.org/sites/default/files/documents/science/IOTC-2015-ROS_11_04_Observer_Manual_v1.2.pdf

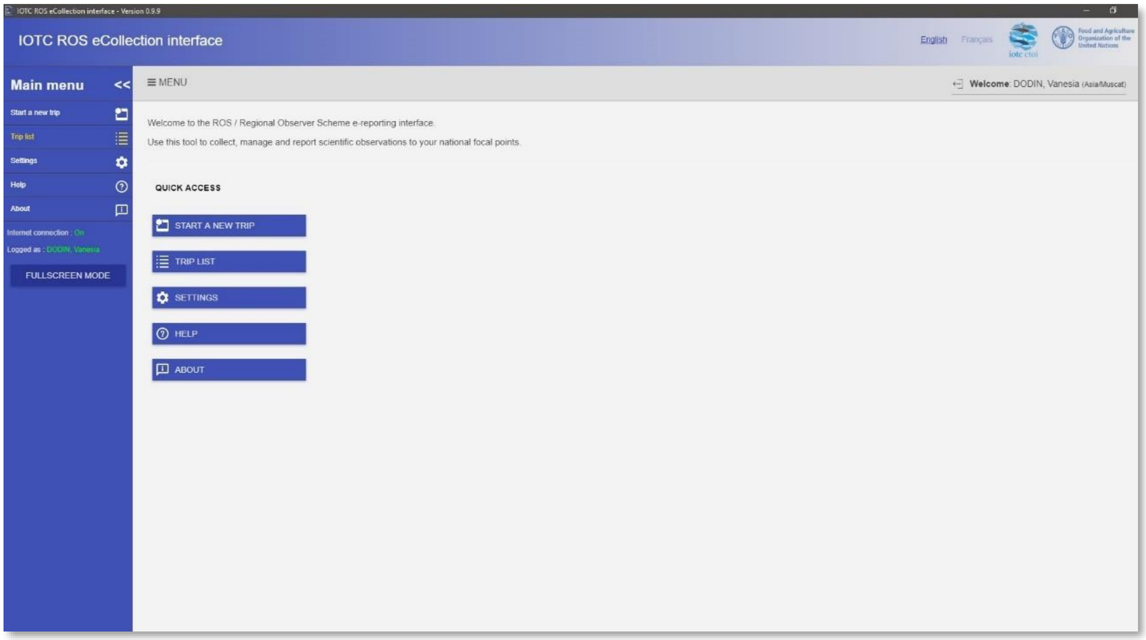

Figure 3. The ROS CI dashboard, showing the possible actions available to the user

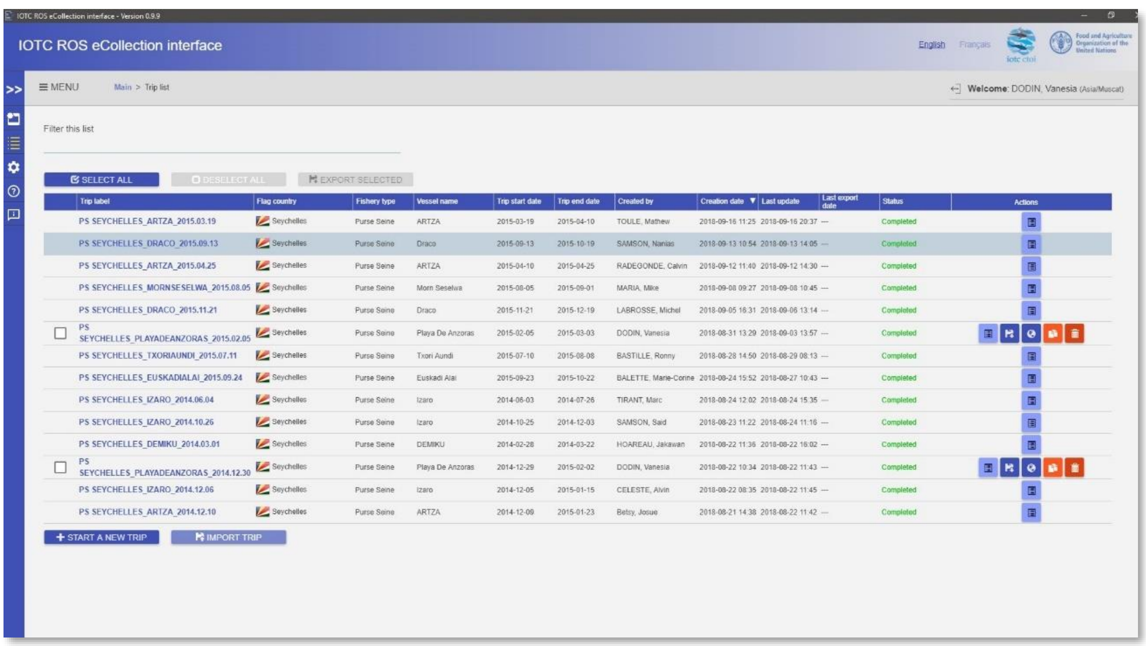

Figure 4. The list of scientific observer trip data available to the currently logged in observer

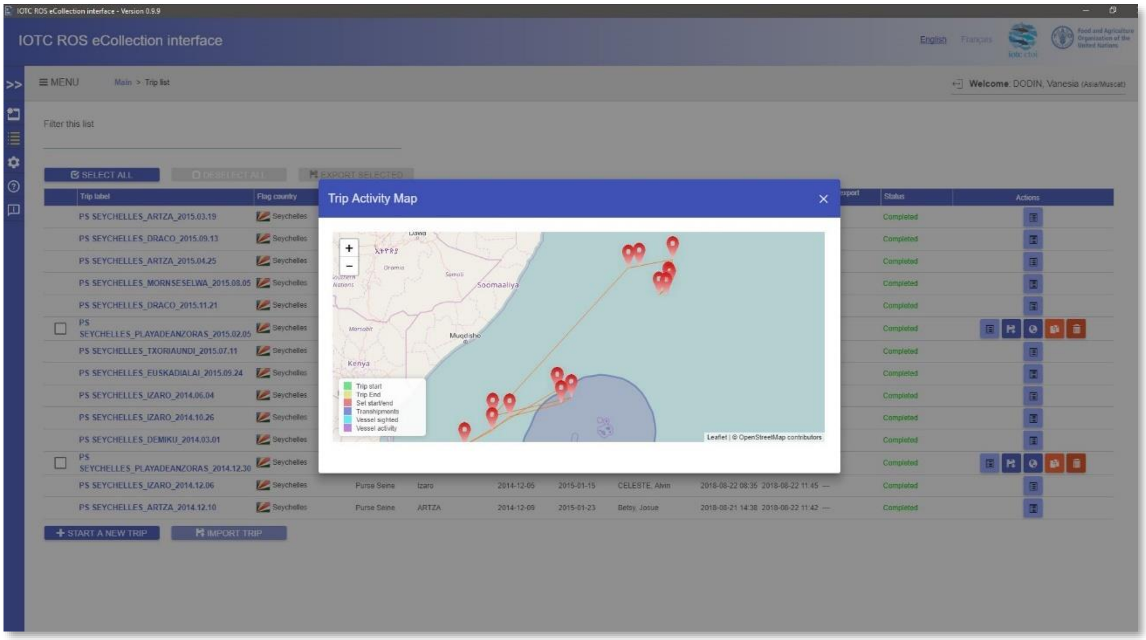

Figure 5. The trip activity map for a sample trip

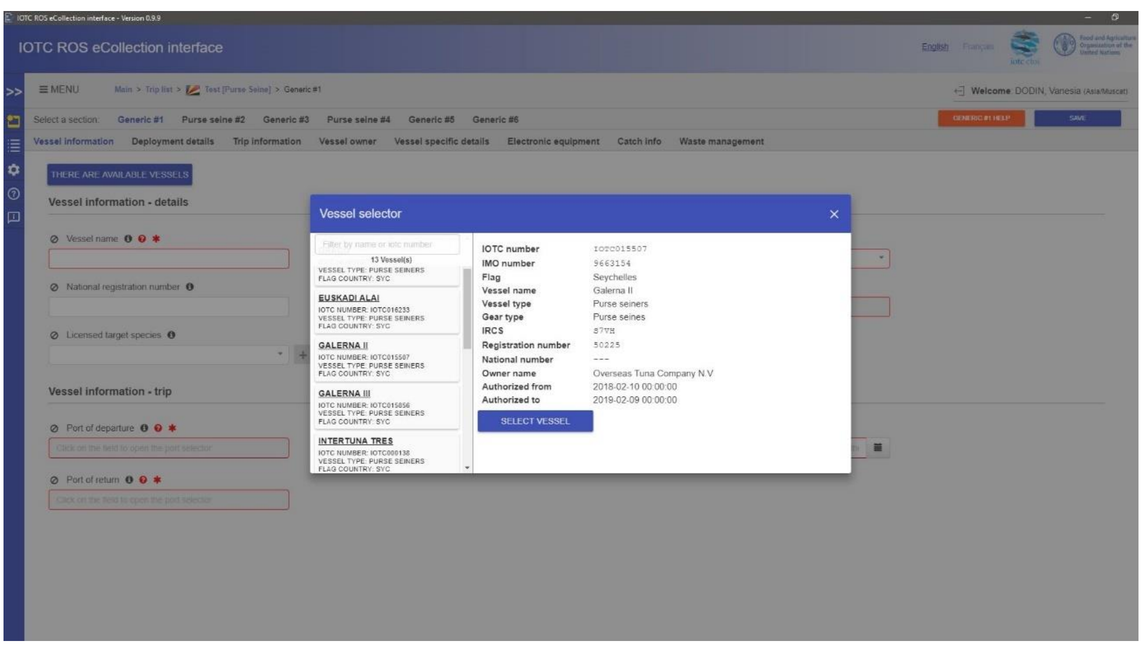

Figure 6. An example of vessel selection using information provided by the RAV

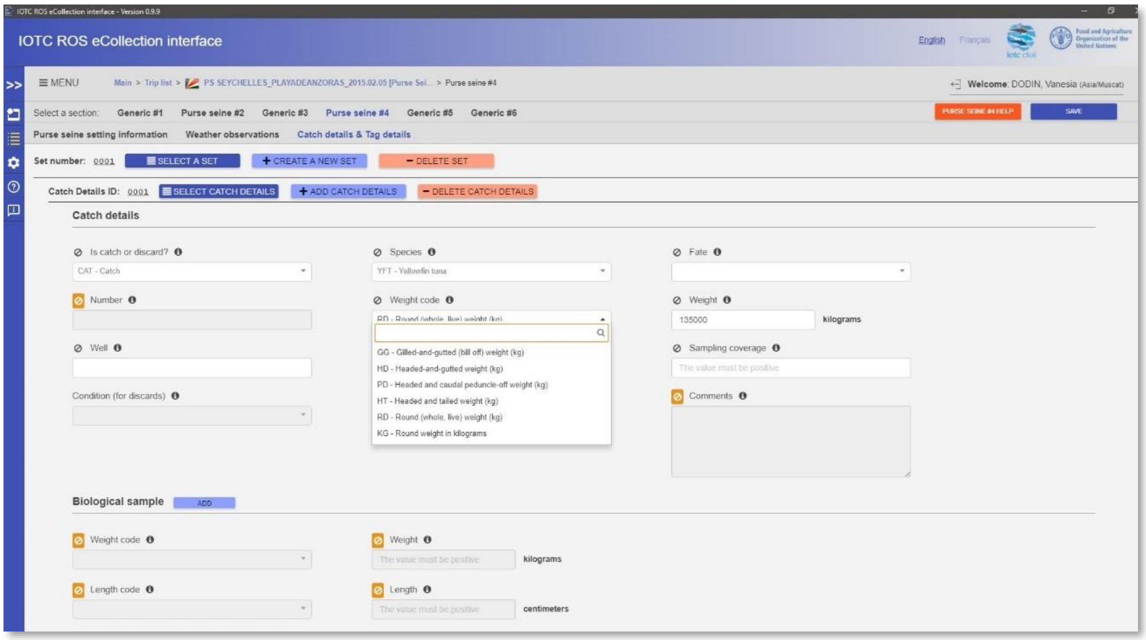

Figure 7. An example of the data entry capabilities for a PS trip (catch by species for a given fishing set)

ROS NATIONAL DATABASE

The ROS National Database (ROS NDB) is the tool designed to complement the ROS CI and allow national institutions to receive, validate and collate all scientific observer data collected by national observers through the ROS CI and submitted at the end of a trip.

National focal points can use the ROS NDB to analyze the content of the scientific observer data submitted by their observers by extracting several pre-packaged metrics from the reported information, and eventually submit the data marked as *mandatory for reporting* to the IOTC ROS Regional Database. As such, the ROS NDB implements the **second** step of the ROS data collection workflow (see Figure 1).

The ROS NDB is a multi-platform tool that requires Internet connectivity to authenticate national focal point users and submit information to the IOTC ROS Regional Database. Its technical specifications are as follows:

- **Current version**: 0.9.2
- **Operating system:** Windows / Mac OS X / Linux (in preparation)
- **Minimum RAM**: 4GB
- **Minimum free space required on the HD**: 128MB
- **Recommended screen resolution**: 1280x1024 pixels or above
- **Pre-requisites**: Java 8 (either JDK or JRE) installed on the same machine
- **Language**: English / French

- **Download link**: https://goo.gl/Yysvs5 (Windows)

National focal points successfully logging in on their local instance of the ROS National Database can perform the following actions:

- Examine the overall status of all currently imported trip data (including the trip completion status and whether data for the trip have already been submitted to the IOTC ROS Regional database);
- Import new observer trip data as *.ros* binary files produced through the ROS CI tool;
- Submit all information marked as *mandatory for reporting* for a completed trip to the IOTC ROS Regional database;
- Export the entire current content of the ROS National Database as an Access file for off-line analysis;
- Extract and analyze summary metrics from the database using one of the pre-defined extraction queries;
- Edit an existing data extraction query;
- Create or import new data extraction queries;
- Export summary metrics as CSV files;
- Issue a trip deletion request to the ROS Regional database (if the trip was already submitted).

Access to the ROS NDB is restricted to national focal point whose credentials are granted by the IOTC Secretariat on a case-by-case basis: as already discussed for the list of IOTC accredited observers, a formal procedure for the provision of national focal point credentials should be considered as part of future ROS standards reviews, or independently defined by the Secretariat.

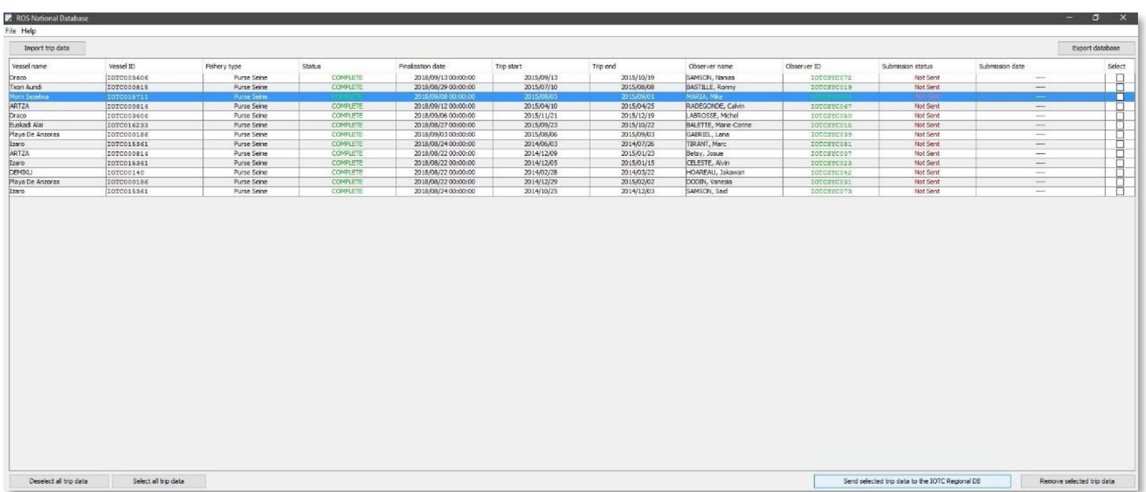

Figure 8. The ROS NDB main interface, showing the status of all imported trips received from observers managed by the currently logged-in national focal point

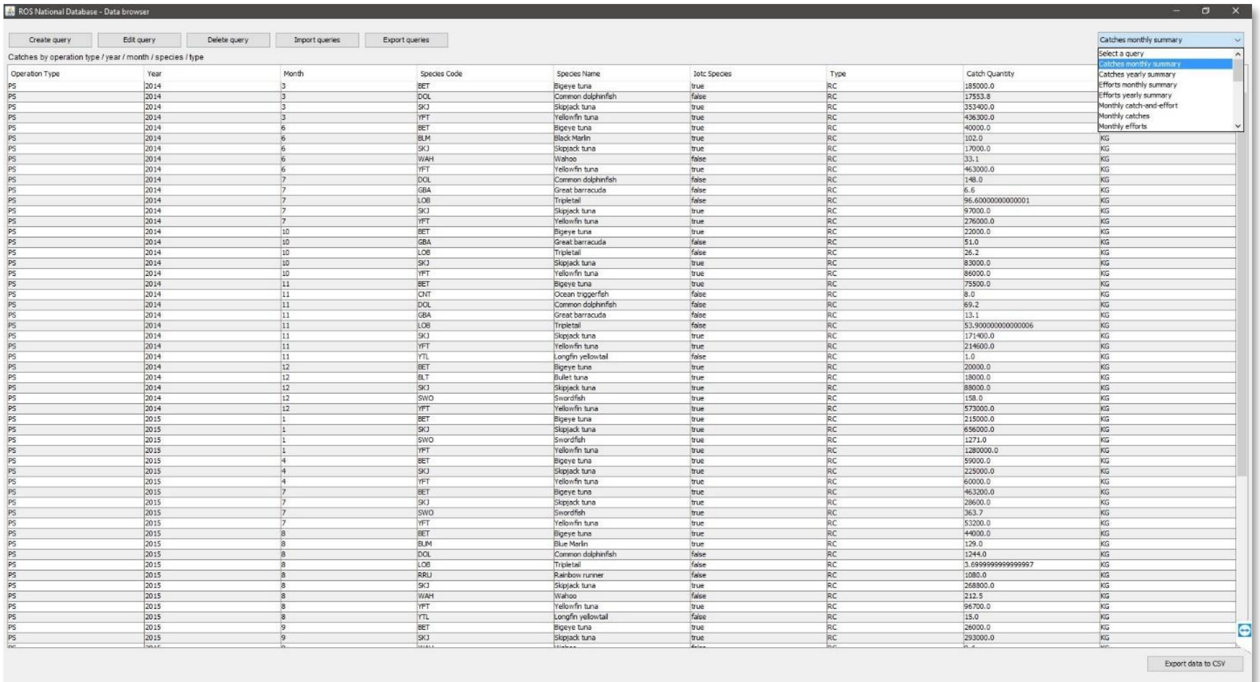

Figure 9. The results of extracting one of the pre-packaged metrics ("*Catches monthly summary*") from the currently imported scientific observer data in the local ROS NDB instance

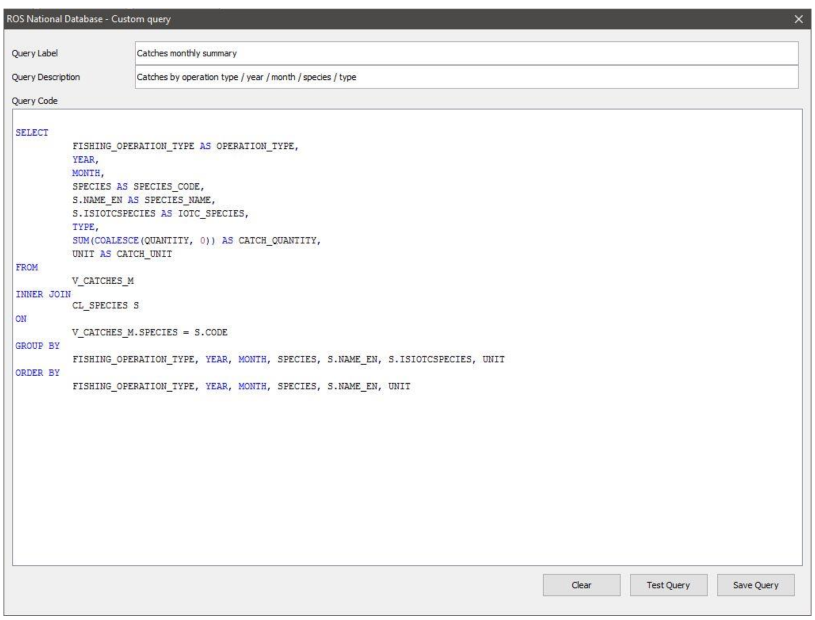

Figure 10. Editing a data extraction query for the calculation of specific summary metrics

ROS REGIONAL DATABASE

The ROS Regional Database (ROS RDB) differs from the ROS e-collection interface (ROS CI) and the ROS National Database (ROS NDB) as it exists a single, centralized entity hosted by the IOTC Secretariat. It is integrated with the other tools of the ROS suite through HTTP(s) based remote data exchange mechanisms and will eventually be made available to the public to disseminate data in aggregated form according to the criteria specified by Resolution 12/02.

It currently consists of a set of different remote interfaces implementing:

- the ROS API (*Application Programming Interface*) to enable reference data synchronization, user authentication and ROS data exchange from / to the ROS CI and the ROS NDB, and
- a set of **prototype** data visualization widgets to filter and aggregate the information currently stored within the ROS Regional Database according to the ROS requirements (only data originally marked as *mandatory for reporting*) and to the confidentiality policies set forth by Resolution 12/02[5].

Currently, the IOTC ROS Regional Database has been populated with historical data received by the Secretariat in previous years, covering purse seine and longline fleets in the 2005-2017 period.

The process of importing these historical observer trip reports has required major efforts in dealing with the various, often non-standardized formats reported by the various CPCs (see Table 2). With the adoption of the ROS tools by CPCs lacking data entry and management tools, alongside the integration of other CPCs with pre-existing observer data management workflows, we expect the content of the ROS Regional Database to be updated seamlessly and as soon as the information is made available by the original data providers.

As of November 2018, the status of the ROS data available to the IOTC Secretariat and incorporated within the ROS Regional Database is as follows:

Flag	Gear	Total trips		Total sets	Format of remaining submissions
		Reported	ROS RDB	ROS RDB	
AUS	LL	51	*N/A*	*N/A*	IOTC preliminary format (.pdf & .doc) and non-standard .xls format (including EMS data)
CHN	LL	14	*N/A*	*N/A*	IOTC preliminary format (.doc) and non-standard .xls format
TWN, CHN	LL	102	*N/A*	*N/A*	IOTC preliminary format (.pdf)
EU, PRT	LL	7	*N/A*	*N/A*	IOTC preliminary format (.pdf) and IOTC ad interim format (.xls)
EU, ESP	PS	42	16	545	IOTC preliminary format (.pdf, handwritten)
EU, ESP	LL	5	*N/A*	*N/A*	IOTC preliminary format (.doc), non-standard report format (.doc), ST09

[5] "*Observer data grouped by 1° longitude by 1° latitude for surface fisheries and by 5° longitude by 5° latitude for longline, stratified by month and by fishing nation are considered to be in the public domain, provided that the activities /catch of no individual vessel can be identified within a time/area stratum*" – IOTC Resolution 12/02, paragraph 2.c

EU, ITA	PS	10	*N/A*	*N/A*	Non-standard report format (.pdf)
EU, FRA	LL	492	492	2838	Non-standard report format (ST09)
EU, FRA	PS	275	266	5732	Non-standard report format (ST09)
EU, GBR	LL	2	*N/A*	*N/A*	IOTC ad interim format (.xls)
FRAT	PS	23	9	203	IOTC preliminary format (.pdf)
IDN	LL	11	*N/A*	*N/A*	IOTC ad interim format (.xls)
IDN	PS	1	*N/A*	*N/A*	IOTC ad interim format (.xls)
JPN	LL	70	51	2681	Non-standard (.xls)
KEN	LL	1	*N/A*	*N/A*	Non-standard (.xls)
KOR	LL	13	*N/A*	*N/A*	IOTC preliminary format (.doc)
KOR	PS	16	6	169	IOTC preliminary format (.doc)
MDG	LL	32	*N/A*	*N/A*	SWIOFP handwritten forms and IOTC preliminary format (.doc)
MDG	PS	6	*N/A*	*N/A*	Letters (.pdf)
MDV	PL	1	*N/A*	*N/A*	IOTC ad interim format (.xls)
MUS	PS	17	17	184	IOTC preliminary format (.doc) & non-standard .xls
MOZ	LL	11	*N/A*	*N/A*	Non-standard Excel
SYC	PS	103	52	1221	IOTC preliminary format (.pdf)
ZAF	LL	66	*N/A*	*N/A*	IOTC preliminary format (.doc & .pdf)
LKA	LL	9	*N/A*	*N/A*	IOTC ad interim format (.xls)
LKA	PS	1	*N/A*	*N/A*	IOTC forms
TZA	LL	1	*N/A*	*N/A*	IOTC preliminary format (.doc)
	TOTAL	**1374**	**909 (66%)**	**13573**	

Table 2. Current status of the ROS data available to the IOTC Secretariat: received vs. processed information

Of all the scientific observer trips imported in the ROS Regional Database, 91 were entered with support from the *Sustainable Indian Ocean Tuna Initiative* (SIOTI) that funded the work of a consultant to manually enter purse seine data from Seychelles, Mauritius, Korea and France (OT) previously reported to the Secretariat as Word/.pdf files.

The remaining observer trips in the ROS Regional Database were converted from other standard formats, such as ICCAT ST09 for French purse seiners and longliners as well as Spanish purse seiners, or the custom template adopted by the Japanese government for its longline fleet.

Scientific observer data from Maldives pole-and-line fisheries is in the process of being reviewed by the Secretariat and eventually incorporated within the ROS Regional Database in the coming months.

Dissemination of ROS Regional Database data

Most of the information currently stored in the ROS RDB relates to observed catches, discards and other basic operational data; while very little additional information (such as the effect and the impact of bycatch mitigation measures) has been captured by the historical reports or from the other standard formats used to submit information to the IOTC Secretariat.

Therefore, the design of the prototype interface to disseminate the public domain information currently within the ROS RDB reflects the limited extent of the information set available.

The situation is expected to improve as the ROS tools (and the "*connectors*" enabling data exchange with other existing data collection systems) help to ensure that observers collect and report the complete set of information required by the ROS data specifications.

Currently, the dissemination interfaces for the ROS RDB public data are available at the following URLs and are password protected with credentials to be provided during the WPDCS14:

- Observed catches and discards: http://ros-browser.iotc.org/catches/
- Observed efforts: http://ros-browser.iotc.org/efforts/
- Observed effort coverages: http://ros-browser.iotc.org/coverage/

We recommend using a recent version of Google Chrome™ or Mozilla Firefox™ to access the dissemination interfaces, as some of the front-end libraries they use may not work correctly with other browsers.

All three interfaces are plotting the corresponding information sets on regular grids within the Indian Ocean boundaries. According to the confidentiality policies specified by Resolution 12/02, longline data is aggregated to the level of 5x5 degrees grids, whereas all surface fisheries data (exclusively purse seine data, as of today) is aggregated to the level of 1x1 degrees grids.

Information on vessel flag was deliberately excluded from these prototype interfaces, as the data incorporated within the ROS DB is still in a preliminary phase – as highlighted by issues such as grids referring to fishing operations in mainland Africa, discrepancies between recorded fates and corresponding product codes, unusually high levels of effort coverage – that require further analysis before a proper public release of the interfaces could be announced.

Observed catches and discards

The first dissemination interface allows users to filter and display observed retained catches and discards (in numbers) according to different criteria.

Data can be filtered by gear type, species group, species, fate, condition at release, product type and timeframe.

Once filtered, the information is displayed onscreen using different layers that can be customized for data related to retained and discarded individuals that can be exported as a .csv file using the dedicated control on the bottom-left of the main control widget.

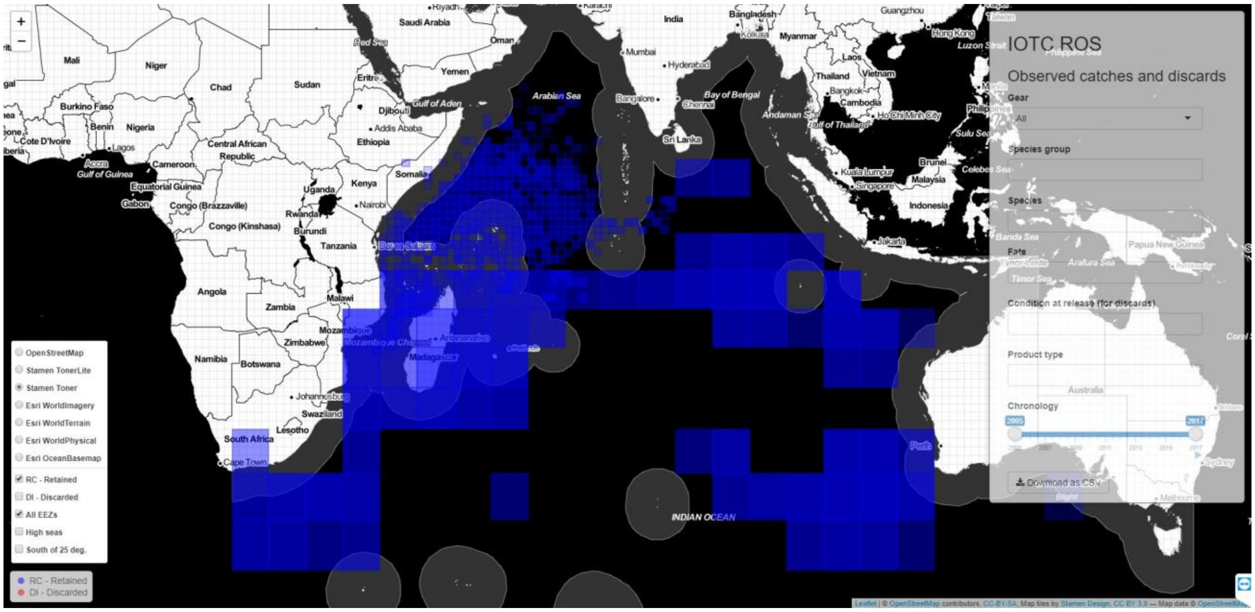

Figure 11a. *Retained* catches observed for all gears and species (2005-2017)

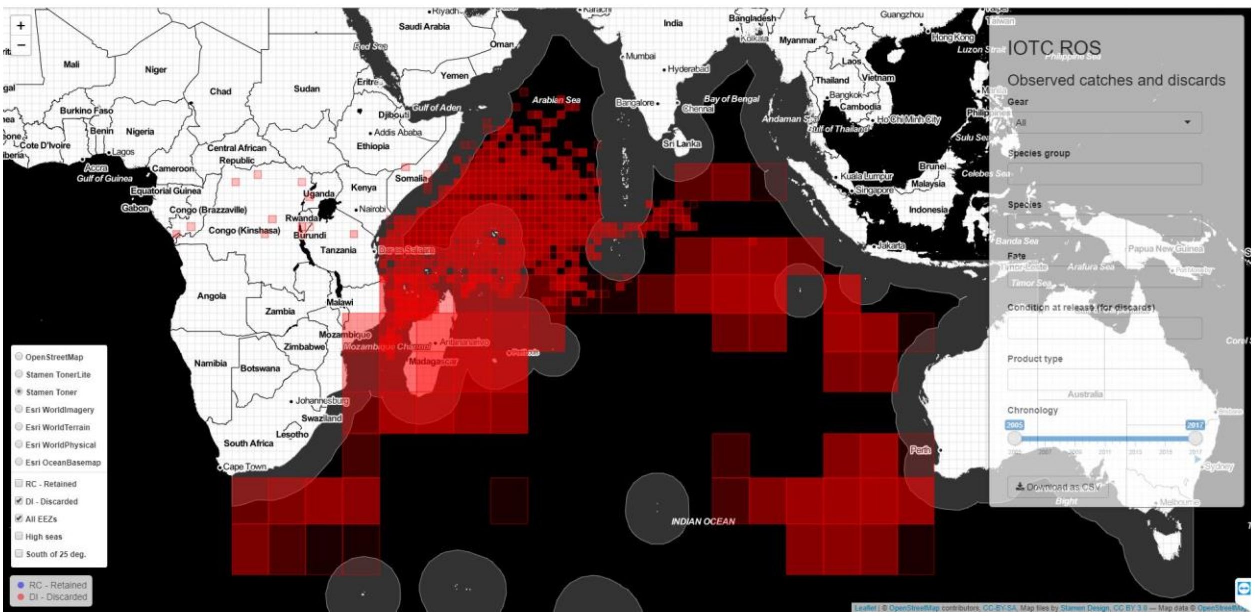

Figure 11b. *Discarded* catches observed for all gears and species (2005-2017)

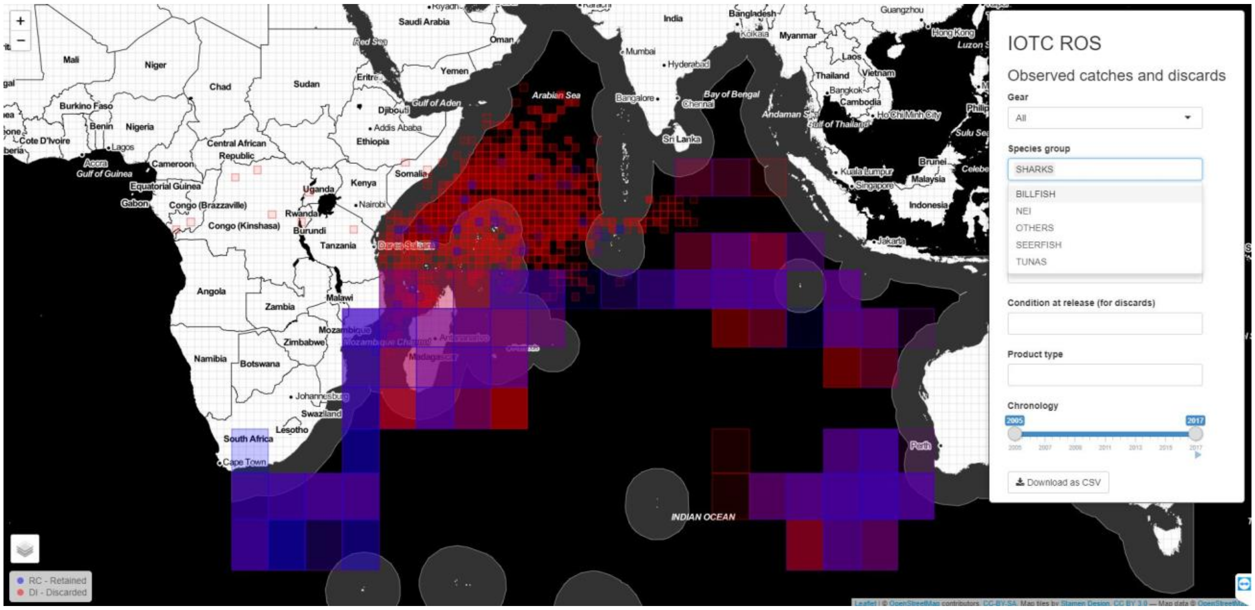

Figure 11c. *Retained* and *discarded* catches observed for all gears and all *sharks and rays* species (2005-2017)

Observed efforts

The second dissemination interface allows users to filter and display observed fishing effort according to different criteria.

Data can be similarly be filtered by gear, effort unit type and timeframe and once filtered, the information is displayed onscreen and can be exported as a .csv file using the dedicated control on the bottom-left of the main control widget.

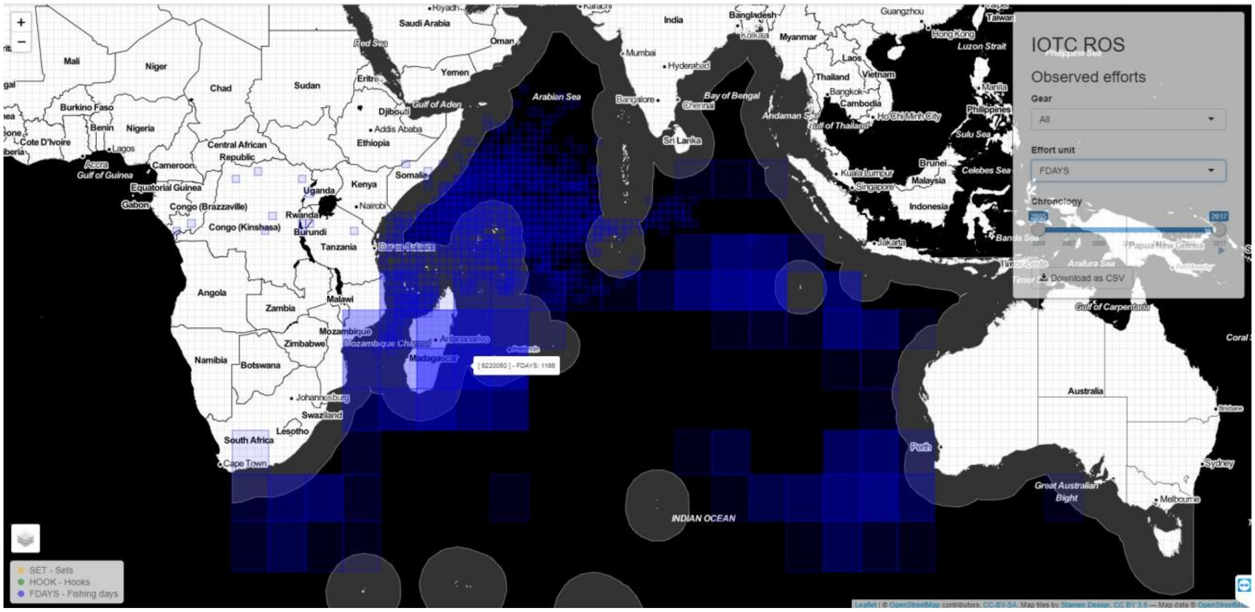

Figure 12a. Observed efforts (*fishing days*) for all gears (2005-2017)

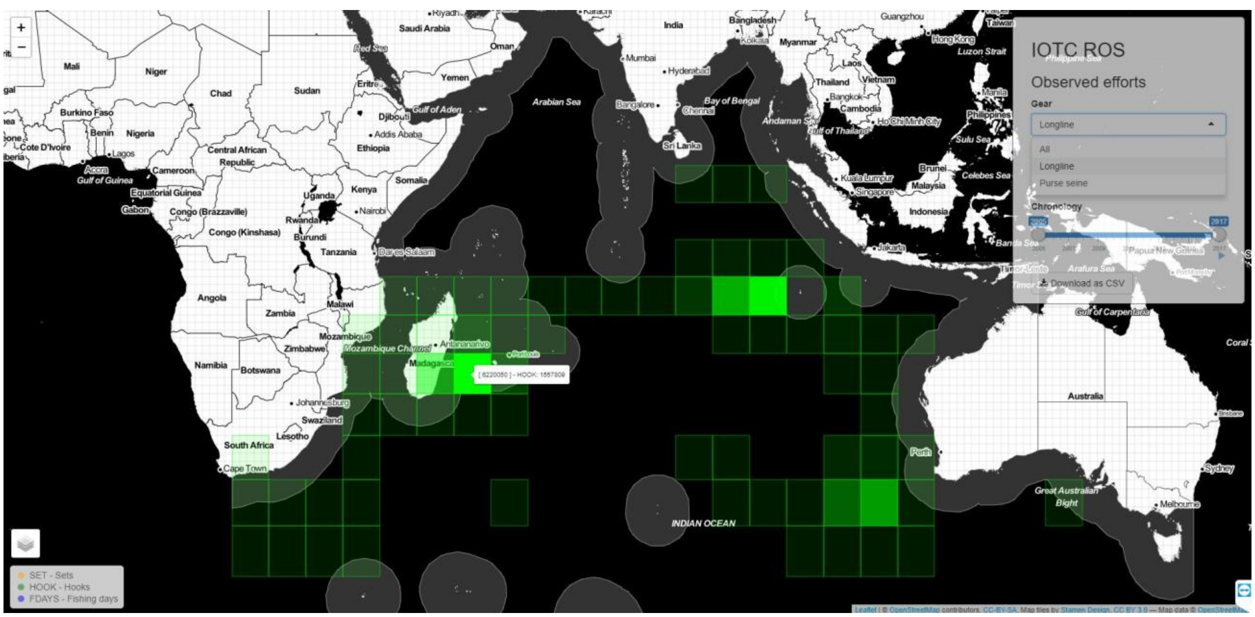

Figure 12a. Observed efforts (*hooks*) for longline gears (2005-2017)

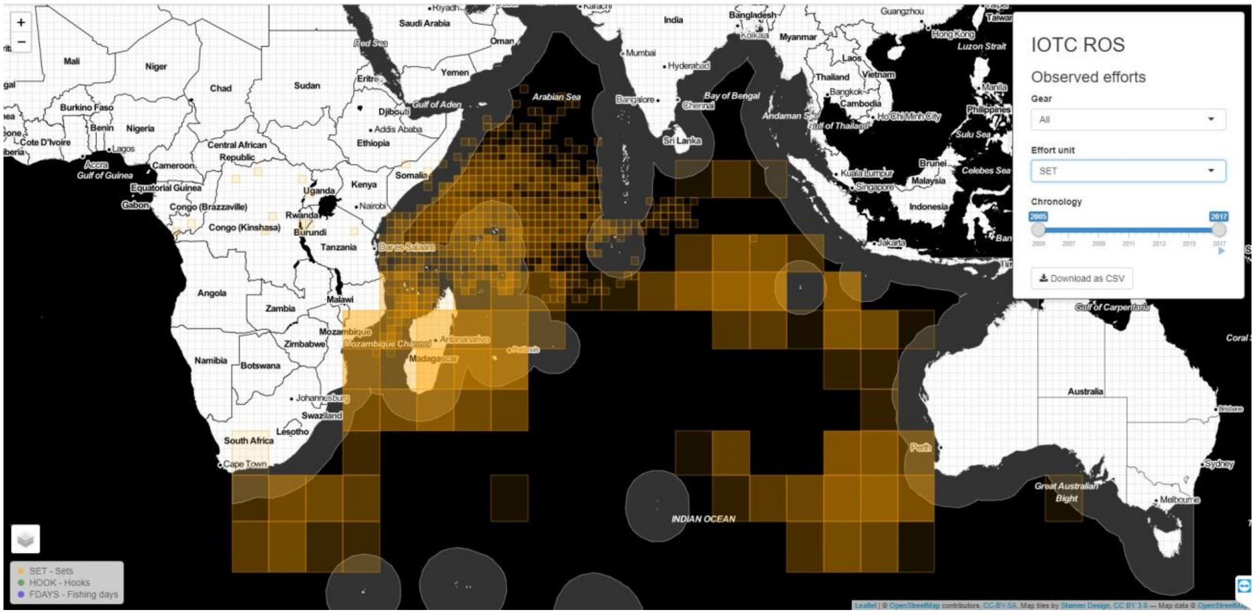

Figure 12c. Observed efforts (*sets*) for all gears (2005-2017)

Observed coverages

The third and last dissemination interface allows filtering and displaying observed fishing effort coverage by regular grid, according to different criteria.

Data can be filtered by gear, effort unit type and timeframe and once filtered, the information is displayed onscreen and can be exported as a .csv file using the dedicated control on the bottom-left of the main control widget.

The reference effort against which the coverage level is calculated is derived from the officially reported catch-and-effort data sets provided by CPCs to the IOTC Secretariat as part of Resolution 15/02 mandatory statistical data submissions.

It was noted that sometimes the observed data is not backed up by the corresponding information in the IOTC database by fishing gear, fleet (not shown by the dissemination interface), year and grid, and this might be due to three possible reasons:

- the original catch-and-effort submissions do not reach 100% coverage, and some reported strata – for which scientific observer data is available – are currently missing from the official data sets;
- the original catch-and-effort submissions are using a different, non-convertible effort unit compared to the scientific observer data available for the same strata (e.g. PS efforts are reported as *sets* only by a limited number of CPCs);
- the scientific observer data are inaccurate, as they record fishing events for gears, grids and years for which no official information is available.

Conversely, there are strata for which the calculated observer coverage is higher than 100% of the fishing effort reported for the strata by Resolution 15/02 catch-and-effort data.

Again, this might be due to differences in the reporting coverage at the level of individual strata for which scientific observer data is available, but for which there may be only be *partial* coverage of fishing sets recorded by the catch-and-effort, thus bringing the observed coverage level to values above 100% for the specific stratum.

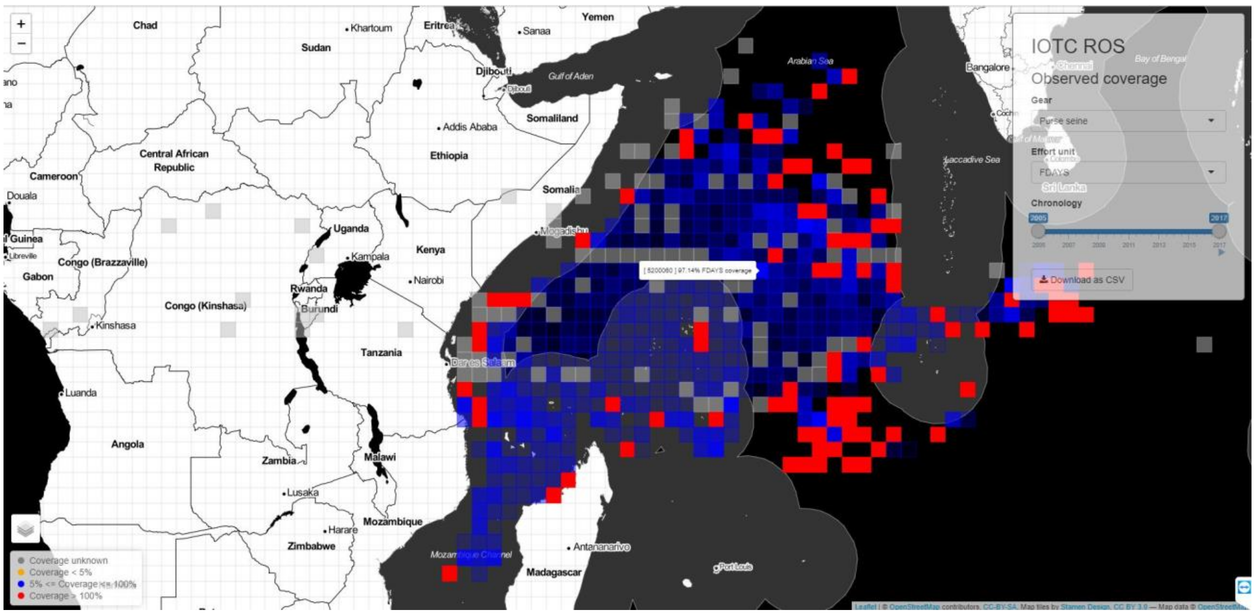

Figure 13a. Observed coverage (*fishing days*) for purse seines gears (2005-2017)

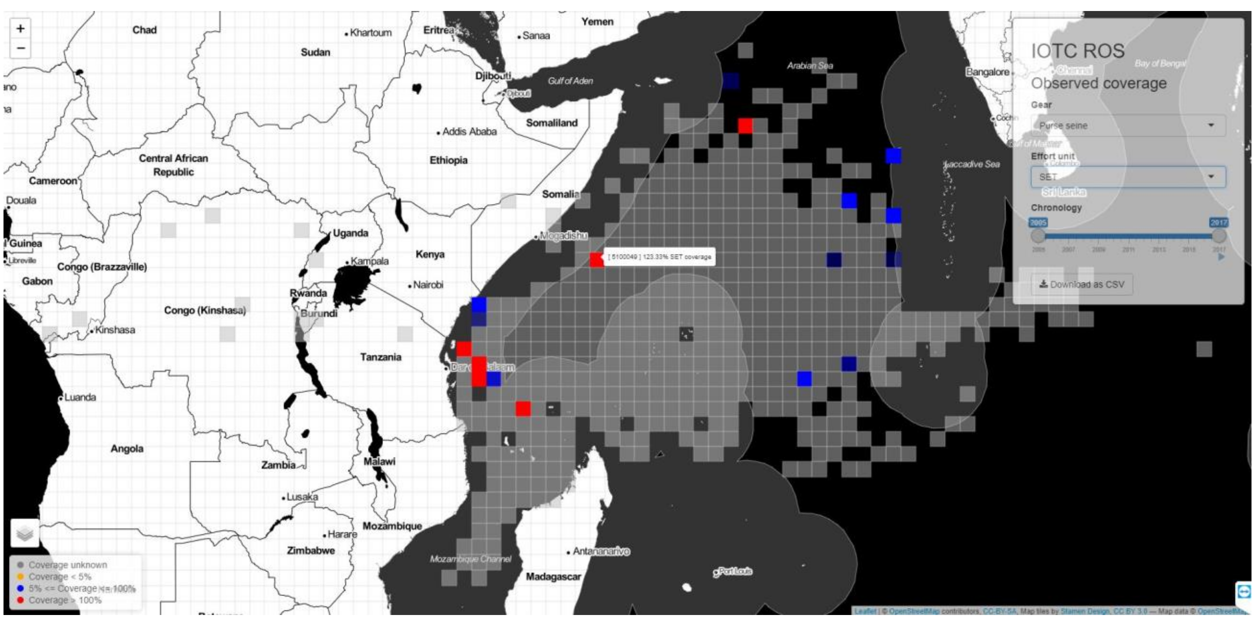

Figure 13b. Observed coverage (*sets*) for purse seines gears (2005-2017)

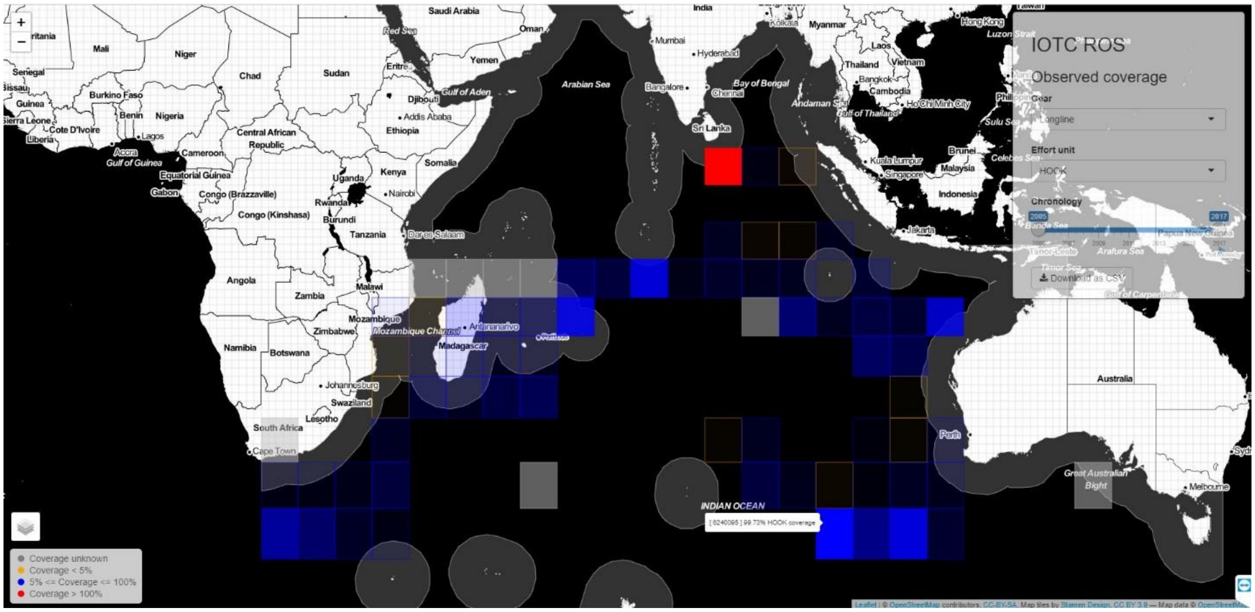

Figure 13c. Observed coverage (*hooks*) for longline gears (2005-2017)

The ROS tools and the general ROS data management workflow

Figure 1 described the *abstract* steps involved in the ROS data management workflow, from the collection of scientific observation to the dissemination of (aggregated) information in the public domain derived from more fine-grained data.

The same figure can be enriched by showing the tools of the ROS suite that implement the different workflow steps, as in figure 14:

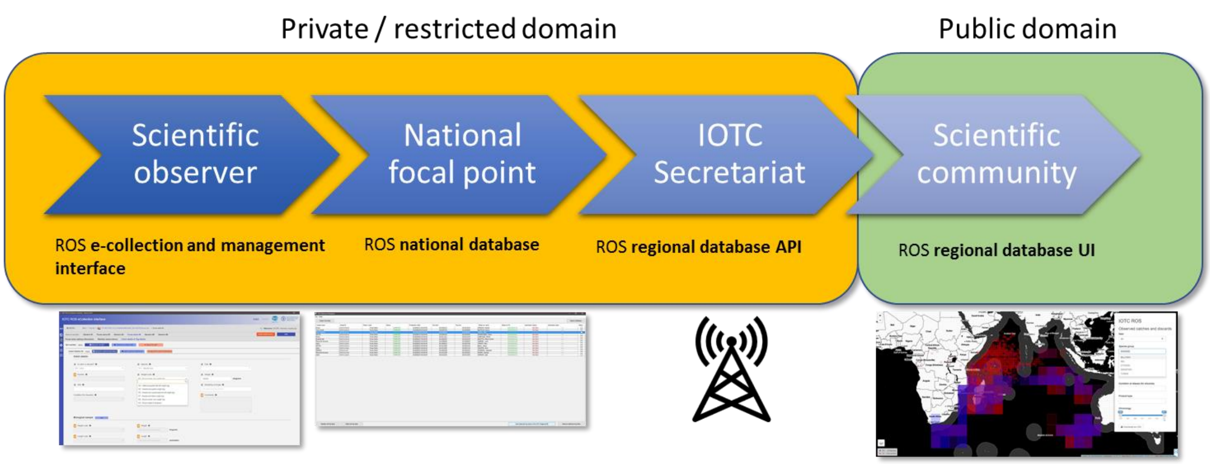

Figure 14. The ROS tools and their relationships with the steps of the ROS data management workflow

One of the key components involved in the implementation of the ROS tools is the *ROS data model*; that is the formal specification of the structure of the information reflecting the ROS requirements.

This can be materialized in two distinct ways:

- As the structure (tables, relationships) of a relational database that can be used to effectively *store* and *retrieve* ROS data;
- As the *grammar* (XSD) defining the structure of the XML documents used to exchange information across the boundaries of the ROS tools.

The two materialized forms of the ROS model are mutually interchangeable and contain / express the same level of information.

The ROS model is the *glue* that links together the ROS tools and enables the exchange, storage and retrieval of the information across components and boundaries (Figure 15).

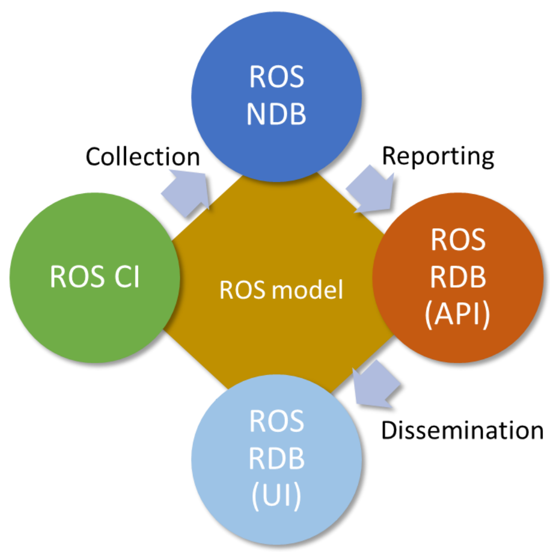

Figure 15. The role of the ROS model in the data flow across ROS components and boundaries.

Another advantage of having a formal model for the ROS data specification is that it can be used as the *target container* for the conversion of third-party data models into a format ready for ingestion within the ROS RDB (Figure 16).

This is the case of the ST09 converter, developed as part of the ROS tools not meant for public dissemination, and is also the key idea behind the future integration scenarios linking the ROS Regional Database and other data collection and management frameworks adopted in the region (e.g. ObServe, SWIOFP database etc.).

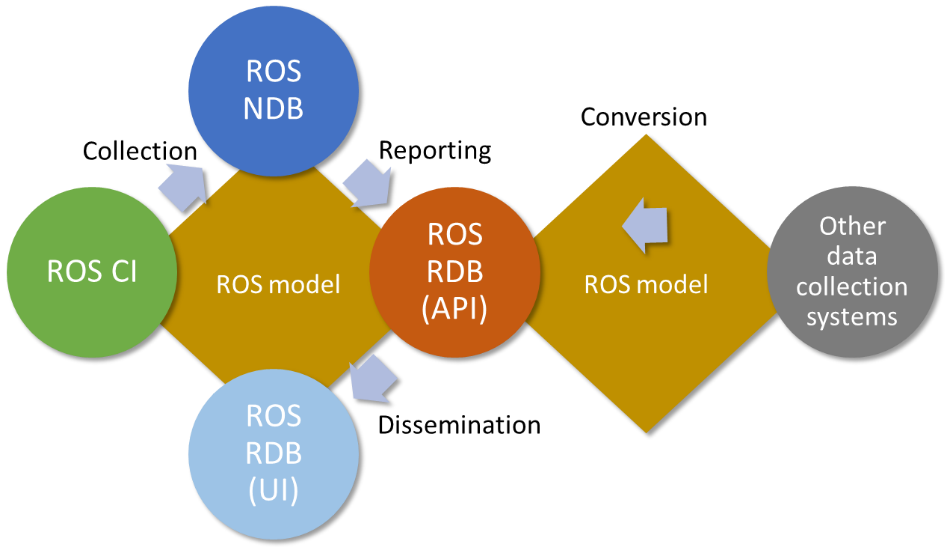

Figure 16. The role of the ROS model in the data flow across ROS and external components and boundaries

FURTHER DEVELOPMENTS

The current ROS model is based on the ROS data collection and reporting requirements expressed by the latest version of the ROS manual (v1.2, October 2015).

The *Expert review workshop on standards for the IOTC Regional Observer Scheme* held in Victoria, Mahé from 24-28 September 2018, led to the identification of several data fields requiring updates (in terms of their definition and / or status) as well as the identification of existing data fields to be removed and new data fields to be incorporated within the revised ROS specifications.

Further consultations are also expected during the WPDCS14, in order to acknowledge and bring forward the requests for updates to the SC for its final adoption: eventually, the ROS data model and the accompanying ROS tools will be updated to reflect changes in the model and in the nature of the data fields.

In addition, more formal mechanisms for the registration of accredited observers and national focal points need to be established and adopted to efficiently update the list of credentials and observer personal information within the system. Lastly, the implementation of comprehensive data exchange mechanisms between ObServe and the ROS Model should be finalized following the formal adoption of the revised ROS specifications.

ACKNOWLEDGEMENTS

The IOTC Secretariat would like to acknowledge the contribution from:

- NOAA and the EU (DG-MARE) for the funding in support of the design and development of the ROS tools;
- SIOTI for the funding in support of the incorporation of historical observer trip data for purse seine fleets operating in the Indian Ocean (both from coastal countries and distant water fishing nations);
- WWF and Mr. Enrico Anello (FAO official consultant) for the development of the tools in collaboration with the Data Section of the IOTC Secretariat;
- Ms. Danielle Jupiter (SIOTI consultant) for the actual data entry of several PS trip reports available as PDF / Word documents;
- Mr. Sihle Victor Ngcongo and Mr. Chris Heinecken (CAPFISH ZAF), Ms. Teresa Athayde (independent consultant), Mr. James Clark (MRAG UK), Mr. Kuruppuge Suraj Chandrakuamara (LKA) and Mr. Bram Setyadji (IDN) for the valuable contributions and feedback provided during the testing phase of the ROS tools.