



Improving the core IOTC data management processes

PREPARED BY: IOTC SECRETARIAT¹, 25th August 2016

Summary

The current state of the art related to the internal IOTC core data management processes is described, depicting benefits and shortcomings as they emerged after more than one decade of adoption. Reasons for a radical change in the process implementation are listed, together with the improvements that the envisaged changes will bring to the internal data flow – as part of the Secretariat's daily operations – and outside its boundaries (targeting mostly scientists, data analysts, policy makers, country-level focal points as well as national and regional management bodies).

The proposed changes aim at rationalizing the entire data management chain, all the way up from the data ingestion to the data dissemination steps, at the same time enabling data consumers to have a simpler and more effective way to get access to the data while still enforcing the confidentiality policies currently adopted by the Commission.

The most ambitious goal of this exercise is to increase the overall value of the data, transforming raw information into a valuable asset from the very first stages of the process, at the same time reducing the time-to-market prior to the final dissemination of regular information updates.

A description of the core and ancillary tools that the new data management processes will make available is given, detailing the impact that these tools will have on the Secretariat's staff daily operations as well as on the broader community that relies on the disseminated information. The strong interactions between the new data management processes and the revised data collection forms are also highlighted, stressing out the need for the revised forms to be adopted to the largest extent possible.

An example of the extended functionalities that the new integrated data management system will provide is also shown, demonstrating the added value that these functionalities could bring if made publicly available through the IOTC website.

Also, preliminary results indicate that the time required to produce the expected data sets for assessment of data-poor species is greatly reduced with respect to the current processes. For more complex species (Tropical or Temperate tunas), we expect the same gain in terms of efficiency once all the involved processes are successfully implemented and tested.

As one of the key tasks required by the successful implementation of the new data management processes is the harmonization of the current reference data sets (including fleet codes, gear codes and species codes) we will also propose *ad-interim* procedures that we expect will be adopted by the end-users to ensure a smooth transition between the two systems.

¹ Fabio Fiorellato (<u>fabio.fiorellato@iotc.org</u>), James Geehan (<u>james.geehan@iotc.org</u>) and Lucia Pierre (<u>lucia.pierre@iotc.org</u>)

Ca	nt n	nto
CO	nie	nis

SUMMARY	1
INTRODUCTION	3
CURRENT STATE OF THE ART	3
Major changes	5
Data aliasing	
FEATURES OVERVIEW	6
Dashboard	6
User management	8
Codelists and other reference data management	9
Datasets management	
Nominal catches	
Catch and effort	
Size-frequency	
Fishing crafts	27
Discards	27
Country indicators	27
Fish prices	27
Other tools	
FUTURE DEVELOPMENTS	28
Feedbacks / improvements	
APPENDIX	30
A1. Programmatically accessing IOTC data services via remote APIs	
A2. The Nominal Catch disaggregation process	
Example of disaggregation results	
Using the Nominal Catch disaggregation to reconstruct catch time series	

Introduction

The current data management processes in place at the Secretariat are the results of more than one decade of evolution and refinements, following frequent changes in recommendations, data management procedures as well as data submission and requirement policies.

The principle driving the implementation of the original processes was to enable users (the Data and Science section of the Secretariat) into timely and effectively respond to data updates, validate and finalize the received information and produce the expected outcomes, for dissemination and scientific purposes, in a simple and repeatable way.

Although great efforts were put into ensuring that the overall processes should behave and be managed as an *integrated system*, this approach was not fully pursued since the beginning (mostly for contingent reasons) thus resulting in a collation of tools and data storage mechanisms that – though fit for the purpose – turned out to be less resilient to changes and updates than expected.

As a result, the processes grew in complexity over time, requiring a well-established experience in order to be successfully mastered. Furthermore, due to the necessity of favouring quick response times over safe and controlled access to the data, some of the process components were designed as *stateful* systems, preventing – *de facto* – concurrent operations on different, non-overlapping subsets of the data, and therefore implicitly increasing response times to any change or update.

The need of disseminating the curated information through the IOTC website (e.g. the *Online Query Panel* and the Working Parties-specific datasets) added further complexity to a system that was beginning to strive to reach its goals, at the same time introducing unneeded complexity that impacted on its maintenance and extension.

Acknowledging the limit (and benefits) of the current systems, and considering the needs for a faster and more streamlined production of the datasets required for stock assessment purposes, the Secretariat took the chance to revise and start redesigning the entire data management chain.

This redesign process – very demanding in terms of efforts – is still ongoing, although it has recently reached a very advanced stage. The results achieved so far are encouraging, and the new workflow has proven to be a concrete step forward towards the achievement of the goals and objectives that the system was originally designed for.

The changes introduced with this process do also have very relevant (and beneficial) side effects: the new data management systems have indeed been designed - since their inception - in order to be remotely accessible: this means that all entitled users could log-in to the data management console from any location, as long as they have network connectivity. In principle, although with some limitations mostly due to screen sizes, the data management processes could also be accessed through a *smartphone* or a *tablet*, increasing the flexibility of a data management system that is crucial for the Secretariat. At the same time, the inherent *remote* nature of the new system might enable data consumers to link their statistical and analytical processes directly to the new data dissemination services, thus gaining direct access (with the due limitations) to the data currently available at the Secretariat.

Current state of the art

Data are reported to the Secretariat (from each CPCs) at different times of the year, usually at the end of June and at the end of December, or are collected from other sources (published data, national institutions etc.) depending on their availability.

Data is usually provided in multiple formats, most of the times as custom Excel files that need to be transformed into the target format (still an Excel file, as per the expected data submission formats) prior to be validated and ingested within the system.

The IOTC Data Section receives the updates, performs the due conversion and validates the contents, following up with the data submitter if any request for clarification is needed or if alternative sources provide contrasting information.

Once data are ready for ingestion, they are incorporated within the system by means of *ad-hoc* forms and procedures that do also perform sanity checks and multiple validation tasks. Only when all checks are successfully completed, the reported data are stored within the IOTC database.

Beside enriching the content of the IOTC database, finalized data can serve multiple additional purposes:

- It can be made available through the Online Query Panel (confidential records are not published if they don't come in aggregated form);
- It can be used to produce the different datasets required by stock assessment scientists prior to each Working Party (here, as well, confidential records are published only if they come in aggregated form);
- It can be used to respond to specific requests coming from scientists, CPCs or other national / regional institutions.

The production of the stock assessment datasets is the most complex of the tasks, as it requires multiple processing steps, including:

- the disaggregation of nominal catch records that come aggregated by species or gear;
- the reallocation of reported catches and efforts on standard areas and timeframes;
- the cleanup and conversion of size-frequency records to the default length or weight units (by species);
- the raising of reported catches to nominal catches (including size-frequency samples to scale down the results);
- the production of catch at size / catch at age datasets.

From a technical point of view, all the above processes are implemented as a collection of different *ad-hoc* solutions including multiple Relational Databases (Microsoft SQL Server and MySQL Server), Access databases with embedded forms, macros and VBA scripts, Excel templates for the production of summaries and the dissemination of dataset-specific records, R scripts for the production of charts, geo-spatial plots and reports.

All the components of this technology stack are linked together by means of shared data sources and files, that due to their nature require that users are connected to the IOTC Local Area Network, as the involved databases and files cannot be accessed from the outside.

In terms of the number of single components involved, we can currently identify:

- 3 different SQL Server databases on the same server instance (one containing the raw data plus the reference codelists and configuration tables, two containing the outcomes of different processes mostly related to the generation of raised catches and size-frequencies)
- 1 MySQL Server database for the storage of data to be disseminated through the Online Query Panel;
- 3 Access databases for the ingestion / management / disaggregation / dissemination of Nominal Catch data;
- 2 Access databases for the ingestion / management of Catch-and-Effort data and the dissemination of reallocated data;
- 2 Access databases for the ingestion / management of Size-Frequency data and the production and dissemination of converted length and weight data;
- 4 Access databases for the production of raised catches;
- 1 Access database for the production of Catch-at-Age data;
- 1 Access database for the production of each of the stock assessment files (excluding stocks being assessed with data-poor methodologies);
- 9 different Excel templates whose content is feeding from a subset of the above databases, used to produce the final Excel files disseminated for each working party.

The IOTC database is responsible for the storage and management of all codelists (including higher level aggregations of species, fleets and gears) as well as for the storage of the raw data.

Submitted data (mostly coming from CPCs) are managed through *ad-hoc* Access databases that serve the main purpose of providing a graphical interface for the ingestion and validation of the provided data. Once these are finalized, they become part of the datasets stored by the main IOTC database: all information related to the data submitter (including contact details for the CPC focal point) is lost, and at this stage it's not possible to get back to the original data if not by storing them aside (as archived files) for later reference.

Keeping track of historical records for all of the main datasets is only possible by either archiving the status of the different databases involved in the process (mostly the Access database files above) or by backing up the content of the main relational databases.

Given the complexity of the tasks to be accomplished and the long lifespan of the current processes (since their initial inception) the proliferation of components and sub-systems depicted above should not come as a surprise and is perfectly legitimate (yet open to improvements).

The current efforts into redesigning the data management systems while still maintaining the same level of functionality of the legacy one, have currently materialized in the following components:

- 3 different SQL Server databases on the same server instance (one containing the raw data plus the user credentials and roles, the reference codelists, the configuration tables and the original reported data files, one containing the outcome of the different processes including nominal catch disaggregation, catch-and-effort / size-frequencies distribution and catch-at-size and catch-at-age datasets, and one containing the history log of all the operation performed on the reported data);
- One web application exposing the REST² remote services that serve as backbone for any operation available through the system;
- One web application exposing the User Interface to interact with the system.

Furthermore, beside the need of having a SQL Server database license, there is no other vendor lock-in and all the components are implemented with fully open source technologies (namely Java 8, Spring, Jersey, Angular JS, Bootstrap) all available as production-level components.

As anticipated, the major benefits implied by this approach are:

- Limited number of components (reduces maintenance and evolution costs);
- Data processes are stateless (multiple users can perform the same process at the same time without interfering with each other and producing repeatable, deterministic results);
- The system is inherently accessible from anywhere (does not require being physically within the IOTC Local Area Network);
- Fully-fledged access control at record-level (users do generally have limited capabilities: they can perform only the operations they're entitled to);
- Inclusion of geospatial features as natively provided by the RDBMS;
- Data extraction could be made available to external users as well (depending on user roles and capabilities);
- Data consumers could use the REST services to get access to the data they need (see above). This means, in principle, that consumers should incorporate REST calls e.g. in their R scripts to get live data from the system instead of having to download the datasets once they're made publicly available;
- Extended features could be plugged-in within the IOTC website to complement what is currently available through the Online Query Panel;
- Dissemination of curated data sets through the IOTC website is greatly simplified.

Major changes

A preliminary step for the implementation of the new data management processes consisted in a complete redesign of the existing main database.

The data structure was almost completely *normalized* in order to reduce redundancy to a minimum. The definition and content of many core codelists was revised to be more consistent with the actual data: in particular, some gear and species aggregations have changed not only in terms of adopted codes (aggregates do now have a symbolic code starting with 'AGxx') but also in terms of the entries that are part of such aggregations.

Furthermore, the gear definition does now incorporate the concept of 'school type' within the gear itself (*Purse Seines* are now split between *Purse Seines – Free School* and *Purse Seines – Log School*).

Area codes (fishing grounds) have been hugely revised as well: in addition, thanks to the native geospatial extension provided by the database engine, it is now possible to incorporate the proper geographical boundaries for each area, including most of the irregular ones. As a consequence, all operations on the data that require the availability of proper fishing ground geometries to be performed, can now take advantage of the geospatial features of the new database.

The notion of '*alternative effort*' has also been introduced: in the legacy database, multiple effort records with different effort units could refer to the same strata, and a prioritization of effort codes was necessary in order to tell which of the different effort units for the same strata should be considered as *primary*.

² REST stands for *REpresentational State Transfer*, which is a coordinated set of components, connectors, and data elements within a distributed hypermedia system – See: https://en.wikipedia.org/wiki/Representational_state_transfer

This was particularly evident when producing the reallocated catch-and-effort dataset, as efforts not referring to the main effort unit (by gear) were simply discarded. Now, with the addition of a complementary effort unit, we increased the number of effort records that could be managed by the database (as long as they refer either to the primary or to the alternative effort for a given gear) and this introduced changes in effort values stored within the disseminated datasets, enhancing their accuracy.

Data aliasing

To enable a proper transition from the *old* (legacy) to the *new* coding systems, IOTC will update and disseminate the transposition mappings between current and new codelists. These transpositions are not always 1:1 and under specific circumstances they might introduce *aliasing* that will prevent exact backward conversions between datasets.

As an example, the old 'Species' codelist was listing these three distinct codes:

- BAR Barracudas
- **BVV** Yellowmouth barracuda
- YRS European Baracuda

for different types of barracudas that now, under the new 'Species' codelist, are all equivalent to the aggregate:

• AG32 – Barracudas

This means that it will not always possible to match records when comparing a dataset produced by the current process with the same dataset produced by the new system.

The same also happens with a few gear codes, namely:

- HOOK Hook and line
- HATR Hand line and Troll line

that are now, under the new 'Gear' codelist, all equivalent to the aggregate:

• AG08 – Handline and Trolling

Other relevant changes related to the 'Gears' codelist include renaming the Baitboat-related gears to their Pole-and-line equivalent, namely:

Legacy database	New database
BB – Baitboat	PL – Pole and line
BBM - Baitboat mechanized	PLME - Pole and line (mechanized boats)
BBN - Baitboat non-mechanized	PLNM - Pole and line (non-mechanized boats)
BBPS - Baitboat and purse seine	AG01 - Baitboat and purse seine

Features overview

Here we will present a brief list of the features as currently available within the new integrated IOTC statistical data management system.

The data management console is accessible as a remote web application and as such it just requires a web browser (we recommend using recent, W3C-compliant browsers such as Google Chrome and Mozilla Firefox) and a working, DSL-grade Internet connection. A screen resolution of at least 1600x900 pixels is also recommended for a better browsing experience.

Dashboard

The entry point of the system is the Tasks dashboard, that lists the available tasks grouped by category.

It is complemented by a chart showing the size over time (in terms of number of records available per year) for each dataset, and can display the data as different types of charts filtered by dataset and time interval.

The bottom-left panel of the Tasks dashboard displays the overall details for each dataset, reporting - for each of them - the total number of records currently stored in the database, the number of records marked as *final* or *confidential* and the number of records with remarks set, plus the date of last update of each dataset.

Users can have a quick glance at the current status of each dataset and directly access its content by clicking on the dataset label within the 'Current dataset' section.

IOTC Statistics Database 🗧 🖉 Quick links 🗸	Welcome, Mr. Administratori
Tasks dashboard Select any of the available tasks by category	
✓ Available tasks (based on current logged user capabilities)	♥ Records per dataset by year
Administration - Perform common administration tasks	\odot Include records from 1950 to 2015 as $\Delta \Sigma$ LN BAR STK
Reference data - Manage codelists and other reference data	C - Nominal Catch DI - Discards FC - Fishing Crafts EF - Catch and Effor CA - Catch and Effor SA - Size and Exercise FD - Size and Exercise PD - Eich Pricese - P
U Data sets - Manage all current data sets	PR - Fish Prices - P CI - Country Indicat
Mominal catches	195381
Discards	180000
✓ Fishing crafts	160000 NC - Nominal Catch 1013
Catch and Effort	FC - Fishing Crafts 85
Size and Frequency	120000 Er - Catch and Effort - Catches 30456
¥ Fish prices	SA - Size and Frequency - Samples 46/1 FR - Size and Frequency - Frequencies 109377
Country indicators	PD - Fish Prices - Products 98 PR - Fish Prices - Prices 98
Tools and utilities - Miscellaneous, general-purpose tools	CI - Country Indicators
Image: Interpretation - Manage data dissemination to public / restricted audience	10000
✓ Current datasets ✓ Refresh	1950 1960 1970 1980 1990 2000 2010 2016 Year
Dataset Records Confidential Final With remarks Last update	1900 1970 1980 1980 2000 2010 2015
Q DI 409 31 409 0 2016/03/17 09:38:48	
	© 2016-08-17 17:06 (+0400) 25m : 365

Figure 1 – The Tasks dashboard

The overall list of currently available tasks is the following:

• Administration

- 0 Users list / filter / create / update / enable / disable users, assign API keys
- o User logins list / filter / delete recorded user logins
- User actions list / filter / delete recorded user actions
- User roles list / create / update / delete user roles
- User capabilities list / create / update / delete user capabilities
- User sessions list / terminate user sessions
- Reference data
 - o Codelists list available codelists, create / update / delete codelists entries
 - Other reference data list other available reference data (aggregations, mappings etc.), create / update / delete other reference data entries
- Data sets
 - Nominal catch manage nominal catch records / produce nominal catch summaries and reports / upload nominal catch data / export nominal catch records / disaggregate nominal catches
 - **Discards** manage discard records / upload discards data
 - $\circ \quad \textbf{Fishing crafts} \text{manage fishing crafts records / upload fishing crafts data}$

- Catch and Effort manage catch and effort records / upload catch and effort data / export catch and effort records summary / reallocate catch and effort records / display reported catch effort CPUE data on an interactive map / produce catch effort and CPUE data reports
- Size / Frequency manage size and frequency records / upload size and frequency data / convert and distribute size and frequency records / display reported sample distributions on an interactive map / produce size / frequency summary plots
- Fish prices manage fish prices records / upload fish prices data
- o Country Indicators manage country indicator records / upload fish prices data
- Tools and utilities
 - **Geospatial tools** browse fishing ground geometries and filter fishing grounds by their current Indian Ocean area size

The list of available tasks will differ, based on currently logged user grants (see Figures 2a and 2b).

 Available tasks (based on current logged user capabilities) 	 Available tasks (based on current logged user capabilities)
Administration - Perform common administration tasks	Administration - Perform common administration tasks
L Users	Reference data - Manage codelists and other reference data
-D User logins	U Data sets - Manage all current data sets
* User actions	Nominal catches
★ User roles	Discards
★ User capabilities	✓ Fishing crafts
► User sessions	Catch and Effort
Reference data - Manage codelists and other reference data	✓ Size and Frequency
I Codelists	¥ Fish prices
S Other reference data	Country indicators
III Data sets - Manage all current data sets	Tools and utilities - Miscellaneous, general-purpose tools
Nominal catches	Issemination - Manage data dissemination to public / restricted audience
f Discards	
Catch and Effort	
Size and Frequency	
¥ Fish prices	
Ocuntry indicators	
Cools and utilities - Miscellaneous, general-purpose tools	<
Dissemination - Manage data dissemination to public / restricted audience	

Figure 2a – Available tasks for an Administrative user

Figure 2b – Available tasks for another, limited user

User management

This section is accessible only to users with proper administrative grants. It provides all the required features to create new users, roles and capabilities as well as update user details and their grants / capabilities. Additionally, it gives administrative users the option to disable / enable an existing user as well as assigning / revoke API keys for non-interactive access to a subset of the remote data services provided by the application.

OTC Statis	stics Database	🖉 Quick links 🗸			Welcome, Mr. Administratori	C+ Logou
‡ A	dministration -					
L Curr	ent users + Add	user 5 10 15 20 50 All 📿	Refresh Reset Showing 12 filtered ent	ries over a total of 12		
	ID	Name	E-Mail	API key	Roles	
1		Alejandro Anganuzzi (Mr.)	alejandro.anganuzzi@fao.org	NOT SET	* ACCESS_ALL_RECORDS • * DATASET_RO • * NC_DISAGGREGATOR • * REF_RO •	
1	ADMIN	Administrator (Mr.)	stats.admin@iotc.org	NOT SET		
		David Wilson (Mr.)	david.wilson@iotc.org	NOT SET	* ACCESS_ALL_RECORDS - * REF_DATA_MANAGER -	
1	E_ROS	e-ROS	e-ros@iotc.org	a25bd6a2035d582dcc68f7bca9ccdbe5	bes ★REF_RO -	
10	E EANELLO	Enrico Anello (Mr.)	enrico.anello@fao.org	NOT SET	★ REF_RO -	
1	FFIORELLATO	Fabio Fiorellato (Mr.)	fabio.fiorellato@iotc.org	NOT SET	* ACCESS_ALL_RECORDS • * CE_MANAGER • * * REF_RO • * SF_MANAGER •	
1	GUEST	Guest User	guest@iotc.org	7b5147f7546c73437a066bb9c0d086f0	5f0 ★ DATASET_RO + ★ REF_RO +	
1		lago Mosqueira (Mr.)	iago.mosqueira@jrc.ec.europa.eu	819e0e9042283d364effd3190767bdc5	CE_MANAGER → ★ FC_RO → ★ REF_RO →	
1	🖸 💄 JGEEHAN	James Geehan (Mr.)	james.geehan@iotc.org	NOT SET	★ CE_CREATE C_UPLOADER → ★ REF_DATA_MANAGER →	
1		Lucia Pierre (Ms.)	lucia.pierre@iotc.org	NOT SET	★ CE_DELETE ATA_UPLOADER ▼ ★ NC_MANAGER ▼ ★ REF_DATA_MANAGER ▼	
1	MHERRERA	Miguel Herrera (Mr.)	miguel.herrera@opagac.org	NOT SET	* CE_FINALIZE_RAISING	
II 0	SMARTIN	Sarah Martin (Ms.)	sarah.martin@iotc.org	NOT SET	EGATOR → ★ REF_RO →	
					★ CE_READ	
					★ CE_REALLOCATE	
					★ CE_UPDATE	

Figure 3 – The current users' list

The tracing of user logins (*who* accessed the system, at *what time* and from *which location*) and user actions (*who* performed an action on the system, at *what time* and from *which location* – see Figure 4) does further increase the level of security and ensures that any non-legitimate access to the application could be easily identified, analyzed and accounted.

TC Statisti	cs Database		🔗 Quick links 🗕												,	Welcome, N	Ir. Adm	ninistratori	0 - C+Logout
🌣 Adr	ninistration -	-																	
✤ Filter	user actio	ons 🧧	i 10 15 20 50	All 💭 Refr	resh 🗙 Rese	t 🏦 De	lete all Showing	115 fi	Itered actions	s ove	er a total of (115)								
1 2	User ID	ţ١.	Token	1t	Date Date	1t	IP address	ţţ.	Method	11	Target	11	URL	ţ١	Time 1	Status	1t	Message	ţ١
1	ADMIN		96881A005F9E186A4D37	9BB77E32F5CD	2016-08-22	11:25:24	<i>§</i> 192.168.9	8.127	GET		http://statistics.iotc.org/rest/servi	lces/	admin/users/logins		0.008 s	290		ок	
1	1 ADMIN		96881A005F9E186A4D37	98877E32F5CD	2016-08-22	11:24:58	Ø 192.168.9	8.127	GET		http://statistics.iotc.org/rest/servi	lces/	admin/users/logins		2.579 s	200		ок	
1	ADMIN		96881A005F9E186A4D37	9BB77E32F5CD	2016-08-22	11:24:43	<i>8</i> 192.168.9	8.127	GET		http://statistics.iotc.org/rest/servi	ices/	reference/data/retrieve/S	izeDa	7.485 s	200		ОК	
1	ADMIN		96881A005F9E186A4D37	98877E32F5CD	2016-08-22	11:24:43	\$ 192.168.9	8.127	GET		http://statistics.iotc.org/rest/servi	ices/	reference/data/meta/descri	ibe/S	0.034 s	200		ок	
1	ADMIN		96881A005F9E186A4D37	98877E32F5CD	2016-08-22	11:24:41	Ø 192.168.9	8.127	GET		http://statistics.iotc.org/rest/servi	ices/	reference/data/meta/list/a	911	2.200 s	200		ОК	
1	[NOT SE	T]	[NOT SET	1	2016-08-22	11:24:41	Ø 192.168.9	8.127	POST		http://statistics.iotc.org/rest/servi	lces/	users/login		0.059 s	200		ОК	
1	ADMIN		DD4036E167568727B917	11B3C637364F	2016-08-22	10:46:58	Ø 192.168.9	8.127	GET		http://statistics.iotc.org/rest/servi	lces/	reference/data/retrieve/Si	izeDa	30.476 s	290		ОК	
1	ADMIN		DD4036E167568727B917	11B3C637364F	2016-08-22	10:46:56	Ø 192.168.9	8.127	GET		http://statistics.iotc.org/rest/servi	lces/	reference/data/meta/descri	ibe/S	2.457 s	290		ок	
1	ADMIN		DD4836E167568727B917	1183C637364F	2016-08-22	10:46:54	\$ 192.168.9	8.127	GET		http://statistics.iotc.org/rest/servi	ices/	reference/data/meta/list/a	811	1.684 s	290		ОК	
İ	ADMIN		DD4036E167568727B917	1183C637364F	2016-08-22	10:46:36	\$ 192.168.9	8.127	GET		http://statistics.iotc.org/rest/servi	ices/	reference/data/meta/list/a	911	12.024 s	200		ОК	
İ	ADMIN		DD4036E1675687278917	1183C637364F	2016-08-22	10:46:26	\$ 192.168.9	8.127	GET		http://statistics.iotc.org/rest/servi	ices/	data/stats/inTime		0.935 s	200		OK	
1	ADMIN		DD4036E1675687278917	1183C637364F	2016-08-22	10:46:25	Ø 192.168.9	8.127	GET		http://statistics.iotc.org/rest/servi	lces/	data/stats/overall		1.289 s	290		ок	
1	[NOT SE	T]	[NOT SET	1	2016-08-22	10:46:02	Ø 192.168.9	8.127	POST		http://statistics.iotc.org/rest/servi	lces/	users/login		22.453 s	200		ОК	
1	[NOT SE	т <u>ј</u>	[NOT SET	1	2016-08-21	09:17:17	Ø 41.86.3	9.58	POST		http://statistics.iotc.org/rest/servi	lces/	data/ce/catch/reallocation	n/byA	0.003 s	488		Bad Request	
1	[NOT SE	T]	[NOT SET	1	2016-08-21	09:17:16	8 41.86.3	9.58	POST		http://statistics.iotc.org/rest/servi	ices/	data/ce/catch/reallocation	n/byA	0.019 s	488		Bad Request	
									(1 2	3	4 5 8								

Figure 4 – Listing recorded user actions

Codelists and other reference data management

Users with access to this section can perform changes to existing codelists or other types of reference data (including mappings, aggregations and configuration tables). These operations have an impact across the entire system, as changes to either codelists or reference data will indirectly introduce changes to the data dissemination and management processes (e.g. changing a higher level species classification will include / exclude the species from appearing within the disseminated data).

Codelists and *reference data* are different types of entities: the first are usually referenced by all datasets, whereas the second are mostly used to relate codelists among themselves and provide mappings (for backward compatibilities with the legacy codes) and configuration parameters to the data management processes.

Examples of codelists are:

- The SPECIES codelist (lists species by codes, scientific name, English name, French name etc.);
- The FLEETS codelist (lists fleets by codes, country code, reporting country code etc.);
- The FISHERIES codelist (lists fisheries by codes, fishery type etc.)

whereas examples of other reference data are:

- The SPECIES_AGGREGATIONS reference data (listing species by code that are part of a given aggregation);
- The FISHERIES_AGGREGATIONS reference data (listing fisheries by code that are part of a given aggregation);
- The SPECIES_MAPPING reference data (listing the mapping between legacy species codes and new species codes);
- The SIZE_DATA_LIMITS reference data (listing the minimum / maximum valid measurements by type for any relevant species)

IOTC Statistics Da	atabase 🔳 🖉 Quick lint	KS 🕶	Welcome, Mr. Administrator!	C+Logout
Referenc	e Data 🗸 Manage o	odelist and other reference data		
🔳 Current c	odelists 5 10 15	20 50 All C Refresh Reset Showing 64 filtered entries over	a total of 64	
Type 👫	Name 👫	Description	Last updated Last updated Last updated	11
FLAT	III Activity	Q + The Activity codelist	2016/07/08 11:25:24	21
STRUCTURED	🗮 Bait	Q + The Bait codelist		0
FLAT	🗮 BaitType	Q + The Bait Type codelist	2016/07/08 11:27:34	6
FLAT	🗮 BoatType	Q + The Boat Type codelist	2015/01/01 00:00:00	32
FLAT	i≣ Certainty	C + The Certainties codelist	2015/01/01 00:00:00	3
FLAT	I≣ Country	Q + The Country codelist	2015/01/01 00:00:00	83
FLAT	I≣ CoverageRate	Q + The Coverage Rate codelist	2015/01/01 00:00:00	26
FLAT	E Currency	C + The Currency codelist	2015/01/01 00:00:00	4
FLAT	🗮 Data Set	C + The Data Set codelist	2015/01/01 00:00:00	8
STRUCTURED	🧮 Data Source	C + The Data Source codelist	2015/10/21 13:48:31	28
FLAT	E DiscardReason	C + The Discard Reason codelist	2015/01/01 00:00:00	3
FLAT	E DistributionRange	Q + The Distribution Range codelist	2015/09/06 00:30:27	3
STRUCTURED	I Equipment	Q + The Equipment codelist	2016/07/08 11:40:28	6
FLAT	EquipmentType	Q + The Equipment Type codelist	2016/07/08 11:38:30	1
STRUCTURED	EstimationProcedure	Q + The Estimation Procedure codelist	2016/03/17 17:26:15	63

« < 1 2 3 4 5 » »

Figure 5 – Listing the currently available codelists

OTC Statistic	Statistics Database 🗧 🖉 Outchinks - Welcome, Mr. Administratori 🗛 💽 - Cologan											
II Refe	Reference Data - Manage codelist and other reference data											
I∎ Species ▼	Especies ▼ + Add entry 5 10 15 20 50 AII 🗶 2 Retesh ¥ Reset Showing 211 filtered entries over a total of 211 in 1620 mSec.											
	ID	Code	Species group	Species official	Species scientific	Is IOTC species	Scientific name	Name EN	P.			
			11		11	G O 11		tt tt				
📋 🗈 🗹	1	YFT	🔗 TUNAS - Tunas	🔗 YFT - Yellowfin tuna	🔗 YFT - Yellowfin tuna	C	Thunnus albacares	Yellowfin tuna	Albacore			
📋 🗈 🧭	2	BET	🔗 TUNAS - Tunas	🔗 BET - Bigeye tuna	🔗 BET - Bigeye tuna	C	Thunnus obesus	Bigeye tuna	Patudo; Thon obès			
📋 🗈 🗹	3	SKJ	🔗 TUNAS - Tunas	🔗 SKJ - Skipjack tuna	🔗 SKJ - Skipjack tuna	C	Katsuwonus pelamis	Skipjack tuna	Listao			
🗊 🗈 🗹	4	ALB	🔗 TUNAS - Tunas	🔗 ALB - Albacore	🔗 ALB - Albacore	C	Thunnus alalunga	Albacore	Germon			
🛍 🗈 🗹	5	SBF	🔗 TUNAS - Tunas	🔗 SBF - Southern bluefin tuna	🔗 SBF - Southern bluefin tuna	C	Thunnus maccoyii	Southern bluefin tuna	Thon rouge du Sua			
1	6	SWD	🔗 BILLFISH - Billfishes	🔗 SWO - Swordfish	🔗 SWO - Swordfish	C	Xiphias gladius	Swordfish	Espadon			
1	7	BLM	🔗 BILLFISH - Billfishes	🔗 MAR - Marlins nei	🔗 BLM - Black Marlin	C	Makaira indica	Black Marlin	Makaire noir			
1	8	BUM	🔗 BILLFISH - Billfishes	🔗 MAR - Marlins nei	🔗 BUM - Blue Marlin	C	Makaira nigricans	Blue Marlin	Makaire bleu			
1	9	MLS	🔗 BILLFISH - Billfishes	🔗 MAR - Marlins nei		C	Tetrapturus audax	Striped marlin	Marlin rayê			
📋 🗈 🗹	10	SFA	🔗 BILLFISH - Billfishes	🔗 BIL - Billfish nei	🔗 SFA - Indo-Pacific sailfish	C	Istiophorus platypterus	Indo-Pacific sailfish	Voilier indo-pacifiq			
📋 🗈 🗹	11	LOT	🔗 TUNAS - Tunas	🔗 LOT - Longtail tuna	🔗 LOT - Longtail tuna	C	Thunnus tonggol	Longtall tuna	Thon mignon			
ii 🗈 🗹	12	KAW	🔗 TUNAS - Tunas	🔗 KAW - Kawakawa	🔗 KAW - Kawakawa	C	Euthynnus affinis	Kawakawa	Thonine orientale			
ii 🗈 🗹	13	FRI	🔗 TUNAS - Tunas	🔗 FRZ - Frigate and bullet tunas	🔗 FRI - Frigate tuna	C	Auxis thazard	Frigate tuna	Auxide			
📋 🗈 🗹	14	BLT	🔗 TUNAS - Tunas		🔗 BLT - Bullet tuna	C	Auxis rochei	Bullet tuna	Bonitou			
📋 🗈 🗹	15	COM		🔗 COM - Narrow-barred Spanish mac	🔗 COM - Narrow-barred Spanish mac.	C	Scomberomorus commerson	Narrow-barred Spanish mackerel	Thazard rayé indo-			
€									Þ			
					2 3 4 5 6 7 17 5	ъ						

Figure 6 – Displaying the content of a specific codelist (SPECIES)

The available tasks for codelist and other reference data management include changing (updating / deleting) existing entries as well as creating new ones.

During the creation or editing of an entry, the system will support users into ensuring that all mandatory fields are properly set, and will provide '*search as you type*' facilities to identify related reference data or codelists (see the 'Species official' field in Figure 7) that the entry being edited might depend from.

IOTC Statistics Database 🛛 🖉 Quick links 🗸		Welcome, Mr. Administrator!	
Reference Data - Manage codelist and other reference data	ta		
Add Species : new entry RAdd Rack			
💻 Data			
♥ Code:	String(4) REQUIRED	A Please review the following issues before proceeding:	
FOO		Value for field <u>Species official</u> is required Value for field <u>Species scientific</u> is required Value for field <u>Bick assessment</u> is required	
✤ Species group:	SpeciesGroup OPTIONAL		
Select an entry in the SpeciesGroup codelist	•		
Lor Longtail tuna YFT Yellowfin tuna	SpeciesOfficial REQUIRED		
LO			
✓ Species scientific:	SpeciesScientific REQUIRED		
Select an entry in the SpeciesScientific codelist	•		
✓ Is IOTC species:	Boolean REQUIRED		
E Yes D No			
> Scientific name:	String(256) OPTIONAL		
> Name EN:	String(512) OPTIONAL		
L 1.8.6-20160822:0652-IOTCS09 (C) 2016 & Indian Ocean Tuna Commission - • • Recommended settings		◎ 2016-08-22 11:41 (+0400) 29m : 56s	∃ • I• ♥•

Figure 7 – Editing a codelist entry (SPECIES)

Datasets management

The *dataset management tasks* are common to all datasets currently managed by the system and include operations and functionalities as:

- *Filtering* a dataset content
- o Creating / Editing / Deleting a dataset entry
- Uploading new data (by using one of the forms currently available for download through the IOTC website)

Additional operations could also be available to entitled users, depending on the dataset type and on current user grants.

The new interfaces used to filter the content of a given dataset (see Figure 8) and edit any of its record (see Figure 9) are consistent across all datasets: in particular, when editing a dataset entry, the system provides support to ensure that all mandatory fields are properly set, and will display searchable *dropdown lists* to simplify the selection or update of any referenced codelist (as required by the specific dataset being edited).

TC Statistics I	Database	🖹 🔗 Quick links 🗸									Welcome, Mr. Administrator!	41	- C+Logo
🛗 Nomina	al Catch 🗸	Manage this dataset											
Filter ava	ailable da	ata or + Add a new record											
dditional data:	Q Show Q I	Hide Results per page: 5 1 Quarter	0 15 20 50 × Re	Showing	15 filtered re	Ecords over a total of	54312	n 209 mSec. Fi	ishing ground		Species		Sc
		- • 11	Select a codelist entry	↓ ↑	Longline	,	1t	Select a codelis	st entry	ţ١.	Select a codelist entry	.↓†	
0 💼 🕜	1950	Entire year (Jan 1st - Dec 3	🔗 IND - India		AG05	Gillnet / Iongline			dian Ocean - We	ster	🔗 FRI - Frigate tuna		Auxis thaza
8 🖻 🗹	1950	Entire year (Jan 1st - Dec 3	🔗 MDV - Maldives		AG10	Longline / handlin	ie		dian Ocean - We	ster	🔗 FRI - Frigate tuna		Auxis thaza
0 🛍 🗹	1950	Entire year (Jan 1st - Dec 3	🔗 IND - India		AG11	Longline and trol	ing	(idian Ocean - We	ster	🔗 KAW - Kawakawa		Euthynnus
0 🗊 🗹	1950	Entire year (Jan 1st - Dec 3	🔗 IND - India		LL	Drifting longline (over 1800) hooks)	idian Ocean - We	ster	🔗 SKJ - Skipjack tuna		Katsuwonu
0 💼 🗹	1950	Entire year (Jan 1st - Dec 3	🔗 MDV - Maldives		LLCO	Small longline		<u> </u>	dian Ocean - We	ster	🔗 SKJ - Skipjack tuna		Katsuwonu
0 🗊 🗹	1950	Entire year (Jan 1st - Dec 3	🔗 IND - India		LLEX	Drifting longline (explorato	ry) .	dian Ocean - We	ster	🔗 YFT - Yellowfin tuna		Thunnus al
0 🗊 🗹	1950	Entire year (Jan 1st - Dec 3	🔗 MDV - Maldives		🔗 PL - Po	le and line		🔗 IRWESIO - :	Indian Ocean - We	ster	🔗 YFT - Yellowfin tuna		Thunnus a
0 🗊 🗹	1950	Entire year (Jan 1st - Dec 3	🔗 IND - India		🔗 BS - Be	ch seine		🔗 IRWESIO - :	Indian Ocean - We	ster	🔗 BIP - Striped bonito		Sarda oriel
0 🗊 🗹	1950	Entire year (Jan 1st - Dec 3	🔗 IND - India		🔗 BS - Be	ach seine		🔗 IRWESIO - :	Indian Ocean - We	ster	🔗 BLT - Bullet tuna		Auxis roche
8 🗊 🗹	1950	Entire year (Jan 1st - Dec 3	🔗 IND - India		🔗 BS - Be	ach seine		🔗 IRWESIO - 3	Indian Ocean - We	ster	🔗 FRI - Frigate tuna		Auxis thaza
0 🗊 🗹	1950	Entire year (Jan 1st - Dec 3	🔗 IND - India		🔗 BS - Be	ach seine		🔗 IRWESIO - 3	Indian Ocean - We	ster	🔗 KAW - Kawakawa		Euthynnus
0 前 🗹	1950	Entire year (Jan 1st - Dec 3	🔗 IND - India		🔗 BS - Be	ach seine		🔗 IRWESIO - 3	Indian Ocean - We	ster	🔗 SKJ - Skipjack tuna		Katsuwonu
0 🛍 🗹	1950	Entire year (Jan 1st - Dec 3	🔗 IND - India		🔗 BS - Be	ach seine		🔗 IRWESIO - :	Indian Ocean - We	ster	🔗 YFT - Yellowfin tuna		Thunnus al
8 🗊 🗹	1950	Entire year (Jan 1st - Dec 3	🔗 YEM - Yemen		🔗 AG03 - (illnet, handline ar	id tr	🔗 IRWESIO - 3	Indian Ocean - We	ster	🔗 COM - Narrow-barred Spanis	n mac	Scomberon
0 🛍 🗹	1950	Entire year (Jan 1st - Dec 3	🔗 YEM - Yemen		🔗 AG03 - (Gillnet, handline ar	d tr	🔗 IRWESIO - 3	Indian Ocean - We	ster	🔗 FRI - Frigate tuna		Auxis thaza
4													

Figure 8 – Filtering the content of a dataset (Nominal Catch)

		IOTC-2016-WPB14-09_Rev1
IOTC Statistics Database 🛛 🖉 Quick links -		Welcome, Mr. Administratort C-Logout
Manage this dataset		
Add a new NC record RAdd X Back		
E Data		
❤ Year:	Integer REQUIRED	Please review the following issues before proceeding:
1975		Value for field Fishery code is required Value for field Species code is required
♥ Quarter:	Set REQUIRED	Value for field Catch quantity is required Value for field Catch quantity unit code is required Value for field Source code is required
0 - Entire year (Jan 1st - Dec 31st)	-	Value for field Quality code is required Value for field Data source code is required Value for field Estimation procedure code is required
✓ Fleet code:	& Fleet REQUIRED	Value for field Coverage rate code is required
ITA - Italy	•	
PSFS Free-school luna purse seine	S Fishery REQUIRED	
PSFS		
❤ Area code:	S FishingGround REQUIRED	
IREASIO - Indian Ocean - Eastern Area	•	
> Target species code:	Species OPTIONAL	
Y Species code:	Species REQUIRED	
L 1.0.0-20160822:0652-IOTC509 (C) 2016 & Indian Ocean Tuna Commission - O Recommended settings	*	⊙ 2016-08-22 11:48 (+0400) 29m : 575

Figure 9 – Manually editing a dataset entry (Nominal Catch)

The bulk upload of new dataset records to the system is a process that is basically dataset independent. The only changes required, by dataset, are related to the adoption of a different data upload template among those currently available for download under the IOTC website³ and by the type of checks applied to verify the integrity and completeness of the data.

As of today, it is not mandatory for CPCs to provide data to the Secretariat using these dataset-dependent, standard forms. Nevertheless, we strongly *encourage* the adoption of this standard for all the required mandatory data reporting tasks: adopting these forms will have the beneficial effect of ensuring that the sanity checks and error fixing procedures available within the system could be effectively used to perform a first quality assessment of the provided information as soon as data is received by the Secretariat and reduce – consequently – the time needed to successfully import new data in the system.

The preliminary step of the upload process consists in displaying the metadata available within the uploaded file (submitter name / organization / contact details etc.) and present a preview of the uploaded data, including the first ten rows of the form content (see Figure 10).

³ http://www.iotc.org/data/requested-statistics-and-submission-forms

▲ 1 - C+Logout

Welcome, Mr. Administrator!

Manage this dataset

IOTC Statistics Database 🛛 🛢 🖉 Quick links 🗸

Upload NC data

elect an .	xism file Form_	1RC.xlsm							
etadata	(as extracted fr	om the uploaded file):			Additional (mandatory) detail	:			
Submitt Submitt Submitt Submitt Submitt	tername: terName tere-mail: ter@mail.me terphone: SET]	Organization name: Submitter Organization Organization e-mail: organization@mail.me Organization phone:	Reporting country ARE Flag country: ARE	: Year: 1950 Catch units: MT Comment: Foo comments	U Liaison Officer U - Unknown Unkn Confidentiality: [PUBLIC]				
der and	first 10 content	112233	IC worksheet.		Public data Confident * To enable worksheet content v	ial 🔗 Validate works alidation, please select a	heet content*	quality level for th	e file
der and Header	first 10 content	112233 rows as available in the f	NC worksheet:] DATA SOURCES [5] DATA PROCES	Public data Public data Public data Confiden To enable worksheet content v	alidation, please select a	heet content [*]	quality level for th	e file [11]
der and Header Row #1	first 10 content	112233 rows as available in the f	IC worksheet: 3] TYPE OF DAT [4 alse] DATA SOURCES [5] DATA PROCES	Public data O Confiden 'To enable worksheet content v [6] TARGET SPEC[7] COVERAGE	ial 🛛 Validale works alidation, please select a [[8] SPECIES ALB	heet content* source and a data [9] BET	quality level for th	e file [11]
der and Header Row #1 Row #2	first 10 content	112233 rows as available in the f [2] AREA [f f f	NC worksheet: 3] TYPE OF DAT [4 alse] DATA SOURCES [5] DATA PROCES	Public data O Confiden To enable worksheet content v [6] TARGET SPEC[7] COVERAGE	alidation, please select a alidation, please select a compared by the select a ALB Thunnus alalunga	heet content* source and a data (9) BET Thunnus obesus	quality level for th	e file [11]
der and Header Row #1 Row #2 Row #3	first 10 content	112233 rows as available in the f (2) AREA (f	NC worksheet: 3] TYPE OF DAT [4 alse IN RCOI] DATA SOURCES [5] DATA PROCES RCLG	Public data O Confiden 'To enable worksheet content v [6] TARGET SPEC [7] COVERAG ALB N0	alidation, please select a alidation, please select a ALB Thunnus alalunga 1223	heet content* source and a data [9] BET Thunnus obesus 456	quality level for th	e file
der and Header Row #1 Row #2 Row #3 Row #4	first 10 content	112233 rows as available in the P <u>[2] AREA [</u> F F F F F	NC worksheet: 3] TYPE OF DAT [4 alse TIN RCOIL TIN RCOIL] DATA SOURCES [5] DATA PROCES RCLG	Public data O Confiden 'To enable worksheet content v 6) TARGET SPEC [7] COVERAGE ALB N0	alidation, please select a alidation, please select a (8) specifies AL8 Thunnus elelunge 1223 	heet content* source and a data [9] BET Thunnus obesus 456 	quality level for th	e file
der and Header Row #1 Row #2 Row #3 Row #4 Row #5	first 10 content	112233 rows as available in the P	IC worksheet: 3) TYPE OF DAT [4 alse TM RCOI] DATA SOURCES [5] DATA PROCES	Public data Confiden 'To enable worksheet content v [6] TARGET SPEC[7] COVERAGE ALB N0	al Validate works alidation, please select a alidation, please select a alidation, please select a AL8 Thunnus elalunga 1223 	et content* source and a data	[10]	e file [11]
der and Header Row #1 Row #2 Row #3 Row #4 Row #5 Row #6	first 10 content 1 FISHERY BS	112233 rows as available in the P [2] AREA [IC worksheet: 3] TYPE OF DAT [4 alse IN RCOI] DATA SOURCES [5] DATA PROCES RCLG 	Public data Confiden 'To enable worksheet content v [6] TARGET SPEC [7] COVERAG ALB N0	alidation, please select a alidation, please select a ALB Thunnus alalunga 1223 	eet content source and a data [9] BET Thunnus obesus 456 	(10)	e file
der and Header Row #1 Row #2 Row #3 Row #4 Row #5 Row #6 Row #7	first 10 content	112233 rows as available in the P (2) AREA (F F	J TYPE OF DAT 4 alse] DATA SOURCES [5] DATA PROCES	Public data O Public data Confiden 'To enable worksheet content v To enable worksheet content v Confiden 'To enable worksheet content v 'To enable wo	alidation, please select a alidation, please select a compared and alignment ALB Thunus alalunga 1223 	(9) BET Thunnus obesus 456 	quality level for th [10]	e file
der and Header Row #1 Row #2 Row #3 Row #4 Row #5 Row #6 Row #7 Row #8	first 10 content	112233 rows as available in the P	IC worksheet: 3) TYPE OF DAT [4 alse TM RCO] DATA SOURCES [5] DATA PROCES	Public data Confiden 'To enable worksheet content v [6] TARGET SPEC [7] COVERAGE ALB N0 ALB N0	al Validate works alidation, please select a alidation, please select a AL8 Thunnus slalunga 1223 	heet content* source and a data g 9) BET Thunnus obesus 456	[10]	e file

Figure 10 – Displaying / editing metadata for an uploaded dataset (Nominal Catches)

Provided metadata are stored and linked to each corresponding record within the dataset (once finalized): this ensures that it will always be possible to trace back the origin of each and every record stored in the database and possibly contact the data provider in order to get further clarifications in case discrepancies are noted at a later date.

Once the required information is provided by the data-entry user, data are ready to be formally validated by applying preliminary checks that verify:

- that all mandatory fields are provided;
- that there is no gap in the records;
- that all codelist references are correctly set.

If any of the uploaded records is not triggering one of these rules, the upload process halts and presents the user with the list of identified issues, some of which might require updating the content of the involved codelists (e.g. to add a new species or a new fishing ground that is not yet part of the reference data) whereas other types of issues might require an update to the data stored in the file sent for upload.

As an example, the system may identify that within an uploaded form there are five fields that were supposed to be mandatory and that have not been provided: it's up to the data-entry user to decide whether these can be properly filled (e.g. by looking at past data sent by the same CPC) or they require getting back to the original data submitter and issue a request for clarification.

Once everything has been fixed and all validation rules are successfully triggered, the system will check whether the uploaded data are conflicting with existing records of the same dataset. This in order to prevent overwriting data that is already consolidated and furthermore to assess the trends of the data being updated: in some circumstances it is perfectly legitimate for a data submitter to provide updates to past records. For this reason, the system allows users to decide – on a colliding record by record basis – whether updates should be committed or not.

Nominal catches

Users entitled to perform additional tasks on the Nominal Catch dataset can access a growing set of features mostly meant to support the identification of possible inconsistencies within the data (prior or following the bulk upload of new nominal catch records) and produce *reports showing trends and composition of current records* within the dataset.



Figure 11. Nominal Catch reports by fleet and species

These type of reports can display catch amounts (for a given fleet) during a specific user-selectable timeframe (the last 50 years, by default). Cell colors vary with the range of catches recorded for each species / year, and the superimposed symbols show any relevant discrepancy (in terms of catch quantities) in data reported between two consecutive years as well as repeated patterns and trends.

Figure 11, as an example, shows that catches of Striped marlin for United Arab Emirates reported for 1999 have exactly the same value as for 1998 (~ 4.17 MT).

Similar reports by fleet can also be limited to specific gears or fishing zones (eastern or western Indian Ocean) and in all cases they can be exported as PNG images (for inclusion in other reports) or as raw CSV data for later analysis.

Users with the proper grants can also display *summary reports for currently available Nominal Catch records*: this type of summary reports provide the same type of information as currently available in the Excel files disseminated – as part of the Nominal Catch dataset – prior to each working party.

With this specific feature, users can select a time range, a subset of the available fleets, gears, gear types (industrial / semi industrial / artisanal), catch areas and species, and have the corresponding reported Nominal Catches displayed as a line chart (see Figure 12), a cumulative bar chart (see Figure 13), a stream chart (see Figure 14) or a table (see Figure 15).



Figure 12 – Nominal Catch summary report as a line-chart

ourmany grapha by ye	or (MT) [10	50 2015 l au	lasta / All araga / A	Il fabor i turor	(All faboring)	Small funce IIN	RAD STK TRI	Logond: R		load -			
a CReset	Cumulative cate	thes (5 species)	Nominal	catches by yea	ar (MT) [1950 20	15] - All fleets / All a	areas / All fishery ty	bes / All fisheries	s / 5 species / Fi	nal / Public rec	ords		
20/40'	Species	MT		O Grouped	Stacked		YFT - Yellowfin t	una 🔍 BET - Bio	ieve tuna 😐 SK.	- Skipiack tuna	ALB - Albaco	e SWO - Sword	fish
ange.	Ø YFT	11,988,019.381	1,268.47				-		-	1			
: 1950	🔗 BET	4,307,860.083	1,200.00										
2015	& SKJ	12,527,713.025	1,100.00										
	🖉 ALB	1,470,101.436	1.000.00										
by:	🖉 SHO	862,569.601									- II.		
0+ fleets			900.00										
			800.00										
0+ areas			Ê 700.00									2005	aak tuma . 666 0.
0+ fishery types			2 600 00									SKJ - SKipja	ick tuna 355.01
0. feborios			Cato										
0+ lisitenes			500.00										
ALB × SKJ × SWO ×			400.00					_					
YFT × BET ×			300.00										
non long species			200.00										
d types:			100.00										
★ FINAL ★ PRELIMINARY			0.00										
DNFIDENTIAL OPUBLIC			19	50	1929	1900	1973 198	su 1: Year	981	1994	2001	2008	2015
es classification:								This summ	ary graph has bee	n generated with	n data that were c	urrent as of 2016/08	/22 @ 12:00:31

Figure 13 – Nominal Catch summary report as a stacked bar-chart





OTC Stati	stics Data	base 🔳	& Quick	links 🗸														1	Velcome, Mr.	Administrato	r! 🔺 🚺	← Logou
	ominal Ca	itch -	Manage	e this data	set																	
NC	summai	ry grapl	ns by y	ear (MT	T) [1950	0 2015	5] - All fle	ets / All area	s / All fisher	y types / All t	isheries - (Graph type:	LIN BAR	STK TBL	Legend:	6 D	⊕ Download +					
D Nor	ninal catche	s by year (I	MT)[1950.	. 2015] - Al	I fleets / All a	reas / All fish	ery types / All	fisheries / Al	I IOTC specie	es / Final / Pu	blic records											
Species	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971
YFT	4,295.635	4,740.217	8,392.958	11,518.546	26,968.650	50,688.656	66,465.540	40,170.955	31,042.115	30,941.359	44,372.236	42,796.537	56,068.316	35,383.468	36,026.495	37,360.840	56,917.866	46,521.281	88,446.658	65,222.822	42,548.765	43,482.657
BET	21.179	45.545	321.454	1,695.053	6,899.783	9,795.495	12,904.021	12,211.609	11,883.257	10,105.300	16,353.512	15,244.564	18,882.793	13,801.964	18,299.622	19,531.084	24,081.863	24,924.387	37,514.263	29,479.111	25,678.979	21,124.880
SKJ	10,832.701	13,681.413	13,466.282	14,326.636	15,076.730	15,342.144	15,991.635	17,653.896	16,782.343	16,905.969	16,699.964	16,574.748	18,920.021	21,328.795	22,312.626	28,457.246	33,689.049	36,182.015	35,935.684	38,478.008	44,314.102	42,554.335
ALB	8.146	18.372	80.447	1,113.650	2,846.619	3,358.029	5,628.661	5,340.850	7,298.538	11,659.703	12,146.064	16,647.720	19,013.251	14,177.509	19,378.542	13,033.597	15,465.122	21,918.969	18,349.396	20,300.379	13,424.630	12,458.334
SBF			647.000	4,040.000	2,594.032	3,235.175	15,287.090	15,067.538	9,549.000	50,455.999	72,696.388	66,152.101	35,717.199	51,781.894	40,001.354	38,483.272	39,606.649	60,097.071	53,498.985	49,100.474	30,657.000	29,065.817
SWO	43.377	40.801	43.727	64.996	211.285	269.492	1 470 306	411.652	614.830	Striped marli	/11.296	917.583	1,0/5.358	973.698	1,180.445	1,282.233	1,360.364	1,807.511	1,/18.104	1,993.440	2,4/3.105	2,063.830
DUM	40.330	5 631	395 120	1 268 213	3 008 331	3 597 679	4 949 137	3 775 75	2,480.923 MT	r'		3 431 227	3 200 100	2 180 384	3 491 578	3 628 152	3 507 004	4 057 115	3 541 008	3 235 570	3 637 893	2 926 539
MIS	1,122	6,466	85.028	274.134	816.054	859,174	1,908,735	1,887,924	1.886.914	2,480,923	2,277,807	2,671,448	1,987,701	1,795,924	2,001,147	3,401,806	4,201,818	4,511,060	3,284,041	4.012.593	3,336,657	1,923,240
SFA	336.257	316.974	359.399	427.717	576.696	803.538	1,008.519	786.501	696.809	1,013.891	1,305.164	1,257.477	1,179.815	1,053.625	1,047.370	1,048.429	1,226.255	1,346.373	1,388.831	1,119.219	1,025.752	1,206.349
LOT	2,808.836	2,784.997	3,054.176	3,321.261	3,563.853	3,599.238	3,285.899	4,664.114	3,709.211	4,482.114	4,504.035	4,418.528	5,300.618	6,095.977	7,159.419	7,736.699	9,078.341	9,387.662	9,424.680	8,836.924	7,430.138	6,455.482
KAN	5,556.339	3,235.223	3,263.453	3,220.602	4,472.911	5,358.876	5,845.078	5,380.002	5,056.449	5,254.202	6,959.705	8,668.094	5,977.266	8,250.503	10,138.395	8,760.243	8,805.809	9,859.179	10,475.048	10,433.541	9,551.606	10,993.111
FRI	2,415.756	2,717.080	2,814.080	3,373.699	3,749.928	3,845.109	4,029.136	3,008.371	2,913.138	3,029.363	3,387.211	4,081.394	3,778.024	4,279.483	4,344.899	5,541.296	7,511.411	7,690.409	7,956.810	7,969.553	6,358.673	6,500.942
BLT	98.759	113.558	120.117	127.008	165.352	176.394	197.857	205.361	186.380	199.131	246.838	267.759	227.773	297.730	311.280	332.783	458.225	472.085	497.314	486.161	426.140	414.753
СОМ	9,088.024	9,727.024	9,607.154	9,486.723	10,954.920	10,059.832	14,190.783	13,640.428	12,452.569	12,876.027	13,262.072	15,324.741	16,869.287	17,499.572	19,666.288	19,518.099	23,254.242	25,227.404	26,330.496	24,943.130	23,370.195	25,287.040
GUT	6,744.128	5,431.152	4,871.185	3,082.727	3,460.752	4,368.259	6,034.766	4,636.105	3,823.719	3,843.538	4,970.966	6,026.016	6,414.382	6,281.647	7,414.874	7,230.051	7,779.566	7,802.669	9,678.384	9,080.879	9,132.198	10,739.858
A614					10.636	32.722	62.769	77.697	100.000	300.000	200.000	100.000	200.000	300.000	468.000	491.000	274.000	88.000	262.000	46.000	154.000	
AG04																		100.000	400.000	400.000	1,439.000	1,387.000
AG03																						
AG08	100.000	100.000	100.000	200.000	100.000		100.000	100.000	100.000	200.000			176.856	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
AG05														17.000	77.000	56.000						
1.0.0-201 host:9080/#/i	60822:0652-T ic/filter/1959/*/	DTC589 (C).2	16 🧟 Indian O	rean Tuna Com	mission . A Der	nmmended settin	ne										1,177.699	9 2016-08-22	11:57 (+0400) 29m : 51s	Т. Т	· •

 $Figure \ 15-Nominal \ Catch \ summary \ report \ in \ tabular \ format$

The selected data subset can then be exported as a CSV file and charts can be downloaded as images for the inclusion in documents and reports.

The last Nominal Catch-specific feature currently available is the *Nominal Catch Disaggregation*: this is the process of breaking down Nominal Catch records that refer to either a species or a gear aggregate into their component (by species or gear) and is the preliminary step for the preparation of the basic datasets required for the stock assessment of all working-parties' species.

It works by applying substitution rules – to all records involved – that look for other records by proxy fleet / gear / area of operation / time interval from which proportions by species or gear can be derived.

So far, within the legacy system, the disaggregation process is implemented as a specific Access database that links to the existing Nominal Catch dataset and applies the disaggregation procedures to produce a fully disaggregated dataset (i.e. whose records do all refer to single, non-aggregated gears or species).

A first improvement over the existing Nominal Catch Disaggregation process is that users are now able to select just a subset of the Nominal Catch records to be disaggregated, and apply the same disaggregation procedures to non-IOTC species as well.

Additionally, disaggregation results can now be stored within the database (as 'Disaggregation runs') and re-loaded at any time, to provide further insights into historical results.

Once the user has selected a subset of the Nominal Catch to disaggregate (or possibly the entire dataset for IOTC-species, as it is required most of the times), the system will process the records and – based on the current disaggregation procedures – provide as output the disaggregated results.

IOTC Statistics Database 🗧 🖉 Quick links -	Welcome, Mr. Administrator!	A 1 - C+ Logout
AG14 - Bullish nel		
▼ Disaggregate NC data for: 1+ years 1+ fisheries 1+ areas		
O Please note that you might find, among available choices for fishery and species, also references to non-aggregated entries. This is perfect normal, and is due to at least one record existing for that fishery / species that is in combination with an aggregated entries.	the other type.	
Additional filters: limit to 🕑 IOTC 🔹 NON-IOTC species 👁 PUBLIC 🗢 CONFIDENTIAL records with a 🗙 FINAL 🖈 PRELIMINARY status 💐 Process 🗳 Retresh 🚍 Load		
Disaggregation completed in 2762 mSec. producing 333 disaggregated records out of the original 33 aggregated records.		
Disaggregation procedures triggered by number of original records.		
Procedure #3 Was triggered by 77 records Procedure #3 Was triggered by B records Procedure #3 Was triggered by 9 records		
> Disaggregation results Q. Show Pinalize Q. Hude filters Fload		
Browse results by procedure # 1 2 3 4 5 6 7 8 9 and severity: 1 W E 5 10 15 20 50 At X + c 1 2 3 4 5 .		
👔 🚺 🕅 🖓 👔 🕐 10: 601152 / 1954 / TuN / LL / IREASIO / AG14 / 9-088815 MT 🛛 Maxedward Tu 🔇 disaggregated records deriving from 🔇 anginal Kominal Catch records		
DI: / 1954 / TWN / LL / IREASIO / BLM / 1.717430 MT		
10: / 1954 / TWH / LL / IREASIO / BUM / 3.417868 MT		
D: / 1954 / TWH / LL / IREASIO / MLS / 2.248353 MT		
10: / 1954 / THN / LL / IREASIO / SNO / 1.705164 MT		
Hide V 🚺 original NC records that yield the previous disaggregation results (as per criteria triggered by procedure #1)		
10:001154 / 1954 / TWN / LL / IREASIO / BLM / 13:424966 MT Region: EASIO Operation: IN Quality: Poor quality Source: Lisison Officer		
DI: 001156 / 1954 / TAN / LL / IREASIO / BUN / 26.717106 NT Begion: EASIO Operation: IN Quality: Poor quality Source: Lisison Officer		
D10: 001158 / 1954 / TAN / LL / IREASIO / MLS / 17.575134 MT Region: EASIO Operation: IN Quality: Poor quality Source: Liaison Officer		
10: 001165 / 1954 / TAN / LL / IREASIO / SNO / I3.329000 NT Region: EASIO Operation: DN Quality: Poor quality Source: Lisison Officer		
Criginal > ID: 001030 / 1954 / Tulk / LL / IRHESTO / AG14 / 1.547032 MT Proceedings I d disaggregated records denving from d original Nominal Calch records		
E 1.0.0-20156222-0552-1017559 (7) 2016 // Induit Ocean Tuina Commission / O Cocommension /	-08-22 12:05 (+0400) 29m : 16s	π . Σ . Φ .

Figure 16 – Disaggregation of 'AG14 - Billfish nei' Nominal Catch records

Under some circumstances, it might happen that a record cannot be automatically disaggregated by using different records from proxy fleets / gears / areas / time intervals: in this case, the user is supposed to provide his / her estimates about how to assign catches by single gear / species.

The reason why still some records might need to be disaggregated *manually* is that all of the current procedures are looking at different time intervals and proxy fleets / gears to identify records to be used for the disaggregation. Under some circumstances – and with the current procedure configurations – not all records can identify alternate records to be used for the purpose (data is particularly poor or sparse for the involved gears / species combination).

Figure 16 shows an example of the preliminary report produced by the disaggregation process. In particular, we see that of the original 94 'AG14 - Billfish nei' Nominal Catch records to be disaggregated the system was able to produce 351 distinct (non-aggregated) new records by applying 3 out of the 8 currently available disaggregation procedures.

OTC Statistics Database 🔳 🖉 Quick links -	Welcome, Mr. Administratori	A 1 - C+Logo
Manage this dataset		
Tisaggregate NC data for: 1+ years 1+ fisheries 1+ areas A614 K		
• Please note that you might find, among available choices for fishery and species, also references to non-aggregated entries. This is perfectly normal, and is due to at least one record existing for that fishery / species that is in combination with an aggregated entries.	ation of the other type.	
> Disaggregation results Q show O Finalize Q show filters		
🕈 Browse results by procedure # 1 2 3 4 5 6 7 8 9 and severity: 1 W E 5 10 15 20 50 All 🔀 a c 1 2 5 6		
💶 189 🕐 Walkal 2 IDI: 948197 I 2412 I UMI I LLFK I INNESIU I Ak14 I 18-65255 MI 🕨 KKKKKK H 🔮 design general berving mom 🅑 organi kommal calch records		*
18: 01gmai > ID: 849981 / 2813 / MDV / HL / IRWESIO / AG14 / 85.117580 MT Processor 2 1 disappropated records deriving from 1 original Nominal Calch records		
🕫 Disanger 🕶 ID: 856127 / 2813 / MDV / LLCO / IRWESID / AG14 / 288.348330 MT Processor 2 3 disappropated records deriving from 10 original Nominal Catch records		
TD: / 2013 / MOV / LLCO / IRWESIO / BLM / 18.251297 MT		
10: / 2013 / MOV / LLCO / IRNESIO / BUM / 38.616615 MT		
DI: / 2013 / MOV / LLCO / IRNESIO / ML5 / 6.726511 MT		
10: / 2013 / MOV / LLCO / IRWESIO / SFA / 9.339110 MT		
ID: / 2013 / MDV / LLCO / IRNESIO / SNO / 135.406800 MT		
Hide 🗸 (10) original NC records that yield the previous disaggregation results (as per oriteria tiggered by procedure #2)		
D: 051760 / 2014 / NDV / LLCO / IRNESIO / BLM / 28.565000 NT Region: MuLDI Operation: AR Quality: Fair quality Source: Liaison Officer		<u>^</u>
1D: 053393 / 2015 / NDV / LLCO / IRNESIO / BLM / 13.972000 NT Region: MALDI Operation: AA Quality: Good quality Source: Liaison Officer		
DI: 051761 / 2014 / NDV / LLCO / IRNESIO / BLM / 80.482000 NT Region: MuLDI Operation: AR Quality: Fair quality Source: Liaison Officer		
D: 653395 / 2015 / MDV / LLCO / IRNESIO / BUM / 9.519000 MT Region: MALDI Operation: AR Quality: Good quality Source: Liaison Officer		
DD: 051766 / 2014 / NDV / LLCO / IRWESTO / NLS / 13.589000 NT Region: MALDI Operation: AR Quality: Fair quality Source: Liaison Officer		
DD: 853402 / 2015 / MDV / LLCO / IRWESIO / MLS / 2.088000 MT Region: MALDI Operation: AR Quality: Good quality Source: Llaison Officer		-
Original > ID: 849577 / 2013 / MDV / PL / IRNESIO / AG14 / 80.747000 NT Procedure #2 1 disaggregated records deriving from 1 original Nominal Catch records		

Figure 17 – Example of Nominal Catch disaggregation results

Figure 17 shows a subset of the disaggregation outputs for results related to the above mentioned example (disaggregation of 'AG14 - Billfish nei' catches only). In particular, we can see the five disaggregated records that will be used to replace the single aggregated record referring to 2013 / MDV / LLCO / IRWESIO (Western Indian Ocean) / AG14 (Billfish nei).

This record – that was originally referring to an aggregation of species - has been broken down to **five** non-aggregated records by using **ten** proxy Nominal Catch records identified by **procedure #5** (*Same fleet, same type of operation, same region, same IOTC area, up to 5 years before or after the fishing year*).

Thanks to these ten proxy records, the system can split the original catches (~ 208.34 MT) into five different, disaggregated catch records referring to the same strata and to one among **Black marlin**, **Blue marlin**, **Striped marlin**, **Indo-pacific sailfish** and **Swordfish** as species. The original catches are proportionally assigned based on how catches for the same species contribute to the total within the subset of proxy records identified by the triggered disaggregation procedure.

In case a record should be manually disaggregated, users will be presented with the original record and the required controls to add specific disaggregation values for any of the gear or species involved.

The disaggregation results can be finalized (that is, stored within the system as a *disaggregation run*) only when all the records requiring manual disaggregation have been correctly updated.

Disaggregation results can also be downloaded (either immediately, after they've been produced, or by recalling the content of a previous disaggregation run) as a CSV file that has the same format of the expected Nominal Catch disaggregated dataset produced prior to each working party.

For a formal definition of the disaggregation process, please refer to Appendix A2.

IOTC Statistics Da	atabas	se 🔳 🔗 Quick li	nks 🗸													Welco	me, Mr. Administrator!	40	C+ Logout
Nominal	Catch	Manage	this dataset																
▼ Disaggreg	gate	NC data for:	1+ years	1	+ fleets		1+ fisheries	1+ areas				1+ 5	pecies						
• Please note that you other type.	might fi	ind, among available cho	ices for fishery and	d species, al	so reference:	s to non-aggregate	d entries. This is perfec	ly normal, and	d is due	to at l	east on	e record	l existin	g for that	fishery	/ speci	es that is in combination w	ith an aggre	gation of the
Additional filters	: limit t		DN-IOTC specie	es 💿 PU	BLIC 🐠	CONFIDENTIAL	records with a \star	FINAL 🖈	r PRELI	IMIN/	ARY	status	≭Pr	ocess	SF	Refrest	Load		
Current filters ide	entify ta to disa	5875 Nominal Catch aggregate is limited to rec	records that can ords referring to a fi	be disaggro ishery and / o	egated [*] vs. (r species agg	41619 total reco	ords matching provide	d criteria											
> Available disa	aggre	gation runs:																	
	ID	Date	Processed by	Year(s)	Fleet(s)	Fishery(es)	Area(s)	Species	I.F	PF	NC	NCd	NCnd	NCdp	NCtd	NCf	Comment		
1002	6	2016-08-09 15:16:45	ADMIN					AG22			54312	2172	52140	8936	7866	61076	BAZ		
100	5	2016-08-04 18:09:20	ADMIN					AG22		0	54311	2171	52140	8930	7866	61070	Foo		
i 0 0 🖻	4	2016-08-03 16:40:19	1 ADMIN					AG22			54312	2172	52140	8936	7866	61076	Shark disaggregation t	est v2	
1 0 0 🗁	3	2016-08-03 16:33:11	1 ADMIN					AG22	- 6	. (54286	2172	52114	8936	7866	61050	Testing Sharks (AG22)	disaggregat	ion

Figure 18 – Listing currently available disaggregation runs

Catch and effort

Beside the common management processes already described for all the datasets, entitled users are enabled to perform additional sets of operations on the Catch and Effort records. These operations are mostly meant to support the data assistant into identifying possible inconsistencies within the data (prior or following the bulk upload of new Catch and Effort records) as well as provide the ground for the production of the datasets required by each working party.

The first type of additional operation that can be performed on Catch and Effort records is their **reallocation in space** and / or time.

Users can select a time frame, a set of fleets (both optional) and one fishery type (*Longlines*, *Purse seines and bait boats*, *Other coastal gears*) and then decide the level of temporal (by month vs. no reallocation) and spatial (original fishing grounds vs. $1^{\circ}x1^{\circ}$ vs. $5^{\circ}x5^{\circ}$ grids) reallocation.

If spatial reallocation is applied (that is, users select to reallocate records to either $1^{\circ}x1^{\circ}$ or $5^{\circ}x5^{\circ}$ grids) the system will take advantage of the geospatial features available within the RDBMS and proportionally allocate catches and efforts from the original grid to any of the overlapping regular grids by resolution.

The outcome of the reallocation process can be either stored (for later reference) or downloaded as a CSV file in the same format currently expected for dissemination prior to each Working Party.

IOTC Stat	istics Data	abase		🔗 Quid	k links 👻													Welcon	me, Mr. Administrat	orl 🗛 2	- C+Logo
* 0	atch and I	Effort 🗸	N	lanag	e this d	ataset															
🔻 Rea	allocate	CE da	ata f	or:	Select a	a preset	• Select	a preset	тот	- Other gears (coasta	l) v										
					0+ years		0+ fleets														
Reallo	ocation compl	eted in 10	6751 m ♥ Fina	Sec pro	Oducing	17462) reallo	ocated records ou Show filters	of the (80553) origination	ginal rec	cords.											
🕈 Brow	se results	by 1	+ years			1+ flee	rts	1+ fisheries		1+ species		©c									
5 10	15 20	50 100	All	×	e c	1 2 3	4 5 6 7	1747 > >						-	YFT	_	BE	r	- SK	3	
Flee	t Fishe	ery Year	From	To #	Months	Area	Area (regular)	Effort	Unit	Effort (2nd.)	Unit (2nd.)	Quali	ty Source		МТ	NO	т	NO	MT	NO	
MDV	TL	1970	1	1	1	6100070	6100070	0.000	TR					1	5.121				64.059		-
MDV	TL	1970	1	1	1	6200070	6200070	262.000	TR						1.983				0.743		
MDV	TL	1970	1	1	1	6105070	6105070	648.035	TR						7.835				53.358		
MDV	TL	1970	2	2	1	6100070	6100070	0.000	TR			-			7.569				33.002		-
MDV	TL	1970	2	2	1	6200070	6200070	60.000	TR			21	/IDV / TL / 1970 / 3 - 3 5.000 TR 1 YFT - Yel	3 / 6200070 / llowfin tuna					2.984		
MDV	TL	1970	2	2	1	6105070	6105070	410.123	TR			(Т	hunnus albacares) :						14.665		-
MDV	TL	1970	3	3	1	6100070	6100070	0.000	TR			1.	2660007817 MT						84.534		-
MDV	TL	1970	3	3	1	6200070	6200070	215.000	TR						1.266				0.739		
MDV	TL	1970	3	З	1	6105070	6105070	422.069	TR					1	1.938				36.280		-
MDV	TL	1970	4	4	1	6100070	6100070	0.000	TR					1	5.858				32.062		-

Figure 19 – Spatial / temporal reallocation of catches and efforts for Other gears (coastal)

Another additional feature available for the Catch and Effort dataset is the production of geospatial plots - including animation of data over time - for reported catches / efforts / CPUE and any given data subset.

Users can filter the data by year, month, fleet, fishery and produce heatmaps at different level of resolution and with different tiles as background.

Also, data can be plotted either by available number of records per grid or by overall quantities (catches, efforts or CPUE). Furthermore, it is possible to limit the geospatial plot to any custom area by simply providing the vertex coordinates in a WKT-like format and superimpose (and intersect) layers for common IO areas, including high seas and all EEZs.



Figure 20 – Plotting YFT catches for 1985 - 2015 plotted on $1^{\circ}x1^{\circ}$ grids

Figure 20 shows an example of the geographical distribution of YFT catches in 1985-2015, with grids having a resolution of $1^{\circ}x1^{\circ}$ degrees. In this specific figure, reported catches (limited to final and public records) are plotted using a logarithmic scale. The IO EEZ layer is superimposed to the plot.



Figure 21 – Plotting YFT catches for 1985 – 2015 plotted on 1°x1° grids and intersected with EEZ areas

Figure 21 outputs are based on exactly the same filtering criteria as Figure 20, with catches intersected (and retained, in the final dataset) with the IO EEZ areas.

When setting a time-frame for the display of reported catches, the system will allow users to calculate and present two types of animations: on a month-by-month basis (for the entire timeframe) or on a year-by-year basis.

This feature – which is also available for reported efforts and CPUE - is extremely useful to identify seasonality and common patterns within the reported data as is the possibility of limiting the displayed data to a given month over the entire timeframe (see Figures 22 - 23a - 23b - 23c).







Figure 23a. YFT reported catches for ESP (MT, logarithmic scale over 1°x1° degree grids) during January (all years)







Figure 23c. YFT reported catches for ESP (MT, logarithmic scale over 1°x1° degree grids) during October

These plots can be produced with close-to-realtime response times: this allows using the feature to actually 'animate' the reported catch / effort / CPUE distributions and show how they evolve during different period of the year (or across years).

Data can be downloaded as images (to be embedded within reports and other documents) or as CSV files for further analysis.

Size-frequency

Beside the common management processes already described for all the datasets, entitled users can perform additional sets of operations on the Size and Frequency records. These operations are mostly meant to support the data assistant into identifying possible inconsistencies within the data (prior or following the bulk upload of new Size and Frequency records) as well as provide the ground for the production of the datasets required by each working party.

The first type of additional operation that can be performed on Size and Frequency records is the **conversion of non-standard length or weight units to standard ones**, followed by the **conversion of lengths to weights** (based on the available conversion equations) and the **redistribution of samples across size bins**. Users can select a time frame, a set of fleets, a set of gears and one or more species currently grouped by Working Party species (Billfishes / Neritic Tunas / Temperate Tunas / Tropical Tunas).

The outcome of the redistribution process can be either stored (for later reference) or downloaded as a CSV file in the same format currently expected for dissemination prior to each Working Party.

Users can identify and browse strata for which - due to the reported length units for a given species - there is no length - weight conversion equation available and apply a different color shade (*heatmap*) to the number of fishes per size bin, in order to show where the highest concentration of samples for each specific strata lies.

отс з	statistic	s Data	base	8 0	Quick link	s -																				We	lcome, N	Ir. Admini	stratori	^ 6	- Gu	ogout
▼ [Distrik	oute S	SF dat	ta for:	Last 2007 2010 2013 2016	10 years × 2008 × 2011 × 2014 ×	(10) × 2009 × 2012 × 2015	• : × 0+ ×	Select a pr	reset	¥	Select 0+ fisher	a preset		 N BL1 KA1 	T - Neritio T × Con W × LO	c tunas (6 M × FRI T ×) × Gut	×													
Di:	stributior stributi	on resu	ed in 627 Ilts 💽	4 mSec Finalize	producing) 642 di ownload •	istribute	d records	s out of the	20203	original	records.																				
⊕ Br	owse r	esults I	oy 1+	years		/ 1	1+ fleet	s 4 5	6 7	1+	fisheries			1+ specie	s		⊗L	⊗ v	V Hea	Itmap												
009	C010	C011	C012	C013	C014	C015	CØ16	C017	C018	C019	C020	C021	C022	CØ23	C024	C025	[ESP 62000	/ PSLS / 50] FRI	2014 (4 - - Frigate t	4)/ tuna (Auxis		C031	C032	C033	C034	C035	C036	C037	C038	C039	C040	C
									86 43		86 43	52			66		thazar cm FL	d): 1473 category	fish(es) i /	n the [37, 3	38)	257	197	311	113 56	132	109	23			44	
									189		189	54					206	540	1,473	1,185	871	1,039	257	688	249		242				97	
											13	42	39	25	42		59	176	217	374	502	1,151	552	566	217	178		59				
												42	40	26	42		61	179	222	381	511	1,173	562	577	221	182		61				-
											8	27	62	69	27		98	256	308	459	543	1,077	648	639	294	150	64	99	12	48	16	
																			28	56	56	31	118		143	32	21					
																			55	112	112	62	236	11	285	63	41					
																																-
4																																- F
1.0.6	-2016052	5:1145-I	OTC509																						O 2016	-05-30 16	5:09 (+04	00) 29m	: 26s	₹.	E. O	



In terms of reported Size and Frequency data, users can also **display the raw and reported samples size distribution** (**by species and measurement type**), eventually filtering the results by fleet, fishery and timeframe.

Produced results can be displayed as charts of three types (stacked bar chart, stream chart and expanded chart) and exported either as PNG images (for inclusion within other reports and summaries) or as CSV files.

Displayed data is reported in terms of raw number of samples per length class and *estimated* samples number per length class, based on the information provided by the original data submitter.



Figure 25a – Swordfish size distribution (FL – Fork length) as a stacked chart



Figure 25b – Swordfish size distribution (FL – Fork length) as a stream chart



Figure 25c – Swordfish size distribution (FL – Fork length) as an expanded chart

Another set of additional features available for the Size and Frequency dataset is the **production of geospatial plots of sampled numbers / raw samples** for any given subset of the data.

Similar to what already described for the Catch-and-Effort dataset, users can filter the data by year, month, fleet, fishery and species, and eventually produce static or animated *heatmaps* at different level of resolution.

In this case, data can be plotted either by number of raw samples or by reported sample numbers and it is also possible to limit the geospatial plot to any custom area by simply providing the vertex coordinates in a WKT-like format or intersect the results with well-known IO areas.

Figure 26a shows an example of the geographical distribution of Skipjack (SKJ) **reported raw samples** for any available year, with grids having a resolution of $1^{\circ}x1^{\circ}$ degrees. In this specific figure, reported raw samples (limited to final and public records) are plotted using a logarithmic scale.

Figure 26b shows an example of the geographical distribution of Skipjack (SKJ) **reported sampled numbers** for any available year, with grids having a resolution of $1^{\circ}x1^{\circ}$ degrees. In this specific figure, reported sampled numbers (limited to final and public records) are plotted using a logarithmic scale.



Figure 26a. Reported raw samples distribution for SKJ (logarithmic scale over 1°x1° degree grids)



Figure 26b. Reported sampled numbers distribution for SKJ (logarithmic scale over 1°x1° degree grids)

Fishing crafts

No specific operations are available for the 'Fishing crafts' dataset.

Discards

No specific operations are available for the 'Discards' dataset.

Country indicators

No specific operations are available for the 'Country indicators' dataset.

Fish prices

No specific operations are available for the 'Fish prices' dataset.

Other tools

The new IOTC data management system provides also a few other tools for the analysis and management of the geospatial information currently stored within the database.

These provide the following features:

- Filter and display the geospatial details for all areas (of any type) and download their definition as WKT text
- Analyze and display the geospatial details for areas (of any type) by the fraction of Indian Ocean covered

The first type of functionality (see Figure 27) is useful to assess the placement and shape of the geospatial areas used by the system and export their WKT definitions (when needed). Geospatial areas can be displayed either in their raw form (as regular grids, when is the case) or in their *intersected* form, taking into account the current geospatial definition of the entire Indian Ocean.

tausuus i	Database	o quick links +									vveicome, i	Ar. Administra	tori	
🔇 GIS	tools -	Manage fishing ground geospatia	al data											
rowse	fishing grou	ind geographies												
lable fishing	g grounds by type (SRID10x20[51]:	C	Selected fishin	ng grounds overview with	ESRI WorldImagery	• tiles					\$ 578	px.	Downle
ode	GRID10x20 v	Name	<u>^</u>	+		and the		12		n Star I	Allers 8			
100040 -	GRID10x20	00-10 Latitude North; 040-060 Longitude East	C 0	-	SHITT -		2130040	Were	· · ·					
00060 -	GRID10×20	00-10 Latitude North; 060-080 Longitude East	60			2120020	2120040		E AL	and the				
00080 -	GRID10×20	00-10 Latitude North; 080-100 Longitude East	60			21100	20	2120060	2120080					
00100 -	GRID10x20	00-10 Latitude North; 100-120 Longitude East	60		Contract Contractor	- Aller	2110040	2110060	2110080					
00020 -	GRID10x20	00-10 Latitude South; 020-040 Longitude East	0	- Con	and the second second									
200040 -	GRID10×20	00-10 Latitude South; 040-060 Longitude East	0			- 1 - V	2100040	2100060	2100080 21	00100				
100950 -	GRTD18x28	00-10 I atitude South: 060-080 I onoitude Fast			ALC: N		2200040	2200060	2200080	. WW Ker	a dette	5		
urrently sek	ected fishing grour	ids:			The second	22000	020		-	2200100	2200120			
Code	Type	Name	*		(Hear)	221002	2210040	2210060	2210080	2210100	2210120			
00840 -	GRID10×20	00-10 Latitude North; 040-060 Longitude East						2220060	2220080	2220100				
100060 -	GRID10x20	00-10 Latitude North; 060-080 Longitude East				2220020	2220040	2220000	2220000	2220100	10.00			
99989 🗸	GRID10x20	00-10 Latitude North; 080-100 Longitude East	1				2230040	2230060	2230080	2230100	Dale			
88188 -	GRID10×20	00-10 Latitude North; 100-120 Longitude East	1			2230020				2233100	2230120	2230140		
10020 -	GRID10×20	10-20 Latitude North; 020-040 Longitude East	1				2240040	2240050			2240420	W I		and the second
10840 -	GRID10×20	10-20 Latitude North; 040-060 Longitude East	1			2240020	2240040	2240060	2240080	2240100	2240120	2240140		
10000	GRTD18v28	10-20 / atitude North: 060-080 / onnitude East					Leafet 1	Tiles © Esri — Source:	Esri, i-cubed, USDA, U	SGS, AEX, GeoEye, Ge	Imapping, Aerogrid,	IGN, IGP, UPR-E	GP, and the (IS User Co
				Display INT RAW	geometries and SHOW	abels with area CODES	TYPES and NA	MES	This map has	been generated with	data that were c	urrent as of 201	6/08/22 @	13:39:54

Figure 27 – Displaying 10x20 grids and their intersection with the Indian Ocean

The second type of functionality (see Figure 28) is particularly useful to support the internal data-curation process, as it allows identifying those areas which – although part of the available grids – do not actually intersect with the current geospatial definition of the entire Indian Ocean.

M					
	anage fishing ground ge	eospatial data			
ng ground	d geographies by	ocean area	9		
in ocean area o	of less than 1 km ² [56]	Currently selected fishing grounds with an ocean area less than 1 km ² and ESRI Worldimagery • tiles	t 578 рх	c. ODownload
types v	Name	Area km ²		·	
GRID05×05	15-20 Latitude North; 075	0.000			
id					
£					
5					
rds			4 5122057 51 5119056		
Ishing grounds:		1	514035 EE 515045 S00037 P		
Туре	Name	Area	200035		
SRID05x05	15-20 Latitude North; 075	0.000 📋 💽	52160		
JRID01x01	01-02 Latitude North; 099	e.eee 🛅 💽	5220045 5220045		
GRID01×01	01-02 Latitude North; 102	0.000 📋 💽	5223034 F372031		
SRID01×01	02-03 Latitude North; 042	0.000 📋 💽			
SRID01×01	02-03 Latitude North; 043	0.000	5233020 24	51 5236144 149	
SRID01×01	03-04 Latitude North; 042	0.000			
GRTD01v01	03-04 I atitude North: 044-	e eee 📼 👩	Leafel Ties ® Esri - Source: Esri, Houbed, USDA, USDA, USDA, GopEve, Getma	5242146	the GIS User Community
	n ocean area of types v of another of the second se	Type Name rdocan area of less than 1 km² [types • Name eRD05x85 15-20 Latitude North; 075 di shing grounds: 15-20 Latitude North; 075 type Name shing grounds: 15-20 Latitude North; 075 statistic grounds: 15-20 Latitude North; 075 statistic grounds: 15-20 Latitude North; 075 statistic grounds: 01-02 Latitude North; 075 statistic grounds: 01-02 Latitude North; 075 statistic grounds: 01-02 Latitude North; 075 statistic grounds: 01-02 Latitude North; 075 statistic grounds: 01-02 Latitude North; 075 statistic ground group (22-03 Latitude North; 075) 02-03 Latitude North; 072 statistic group (22-03 Latitude North; 072) 03-04 Latitude North; 042 statistic group (22-03 Latitude North; 043) 03-04 Latitude North; 042 statistic group (22-03 Latitude North; 042) 03-04 Latitude North; 042	ng ground geographies by ocean area size n ocean area of less than 1 km ² [56] Kries Name Area (RID05x05 15-20 Lastude North, 075 0.000 C 0 this grounds: Type Name Area Sisting grounds: Type Name Original Control (1994) Sisting area of the state of	ng ground geographies by ocean area size	



Both functionalities allow identifying all records (by dataset, including catch and effort and size-frequency) that do refer to any of the displayed areas.

Future developments

One of the core concepts behind the adoption of the new data management system is that some of the features it provides to internal users could also be made available to a broader audience.

We envisage that the reported catch and effort reallocation, the reported size and frequency redistribution as well as the summary and geospatial plots that could be built on top of the currently available datasets can be extremely useful features for the scientific community.

The Secretariat will ensure (and this is already enforced by the adoption of the built-in access control mechanism in place within all layers of the new system) that only data that can be safely disseminated is made publicly available, whereas external users with special grants could get access to more fine-grained datasets and functionalities.

Also, the possibility of sharing data and processes in a formal and automated way (by means of the REST services APIs – see the Appendix A1 for some basic examples) will open up endless integration possibilities for scientists worldwide.

Feedbacks / improvements

Some of the processes implemented so far should definitely benefit from feedback coming from the community.

For the time being, the **Nominal Catch Disaggregation process** is using proxy fleets / gears to identify non-aggregated records that could break down aggregated ones.

The current disaggregation procedures can reference any number of records from proxy fleet / gears and the outputs – when available – are used to proportionally assign catches for aggregated records.

Currently, the system does not set any lower threshold in terms of the minimum number of identified records that the process needs in order to consider a disaggregation output as '*reasonably accurate*': this threshold might depend on the species / fisheries involved and – at the same time – on the specific disaggregation procedure that is triggered.

When implemented, such a threshold might provide additional information to scientists in term of the uncertainty related to the produced disaggregated Nominal Catch records.

This is only an example of the feedback that the community could provide to the Secretariat. We expect – as soon as the new system and its remote APIs are made available to a broader audience – that the suggestions and requests coming from the scientific community will greatly contribute to increase the usefulness and efficacy of the data management processes as a whole. This will have, at the same time, the positive consequence of enabling users to have a more indepth and formal understanding of the entire IOTC data management chain.

Appendix

A1. Programmatically accessing IOTC data services via remote APIs

APIs (*Application Programming Interfaces*) specify the definition - in terms of business processes performed, data exchange protocol and access point - required to invoke services on a remote system.

In this case, the new IOTC data management system is inherently built with the facility of exposing all of its business processes through REST APIs.

For this to be effective, data-consumers (i.e. the users that want to incorporate *live* data coming from the IOTC data management system) need to be assigned an API-key that will uniquely identify their business requests and limit the features / functionalities they can have access to.

Two simple examples of the new IOTC data management processes integration within R follows: in this case, two remote services (one for the export of aggregated nominal catch data by species and another for the production of reallocated catch-and-effort data) are invoked and the returned data is used – within R itself – to plot sample graphs and charts.

The returned data is - at any time - the *live* content of the IOTC database: this means that a scientist that wants to produce charts and graphs from the available IOTC dataset does not need to wait until these datasets are publicly disseminated through the IOTC website.

As long as he / she has an API key that provides valid access to the dataset of interest, he / she could use this API key (in a similar way to what displayed in the R examples) to effectively access the data through the IOTC data management system.



Listing A1.1 – R-script to retrieve and plot the aggregated Nominal Catch data for the five major IOTC species (Gear: PS / LL – Years: 1975 .. 2015)



Figure A1.1 – Execution results for the R-Script in Listing A1.1



Listing A1.2 – R-script to retrieve and plot the reallocated Catch data (in MT) for BET / SKJ / SWO (Gear: PS FS / PS LS / LL – Years: 2010 .. 2015)



Figure A1.2 – Execution results for the R-Script in Listing A1.2

A2. The Nominal Catch disaggregation process

The *Nominal Catch disaggregation* is the process of breaking down all records – from the Nominal Catch dataset – that do refer to either a gear or a species aggregate (or both) in order to produce a dataset that contains only disaggregated records (i.e. referring only to single species and gears).

Its outputs are crucial for the achievement of valid stock assessments results because, being fully disaggregated, they help scientists in providing more accurate estimates and advices.

It is a non-linear yet repeatable process that can also be used to reconstruct Nominal Catch time series for all those combination of species and gears whose information is *sparse* in time (especially for past decades, when data was collected from hard copies and other 'indirect' sources).

The key concept behind the process is that catch quantities from every aggregated record can be proportionally assigned to a different combination of species or gears (in the list of entries belonging to the species / gears aggregates being processed).

In order to identify these proportions, the process applies a sequence of multiple disaggregation procedures that can identify relevant *proxy records* from within the original Nominal Catch dataset.

Once these proxy records are identified, the proportions of catches by species or gears are used to assign the original catch quantity to a combination of (potentially) multiple, *synthetic* records that are the output of the disaggregation process itself.

The procedures used to identify the proxy records do filter the original dataset by *fleet*, *type of operation* (Artisanal / Semi-Industrial / Industrial), *region*, *area* and *timeframe*: they rely on a specific configuration table that assigns – to each and every current combination of fleet / gear / area for which at least one Nominal Catch record exists – a *region* of most-likely operation.

Currently, the process adopts eight different disaggregation procedures that are complemented by a ninth procedure (*manual disaggregation*) triggered when no proxy record can be identified by any of the others.

The high-level data flow is as follows:



Figure A2.1 – The Nominal Catch Disaggregation process

Its pseudo-code implementation would be similar to the following:

```
var DisaggregationProcedures := { P1, P2, ..., P8 };
var NC_Original := { NC1, NC2, ..., NCn };
var NC_NonAggregated := { };
var NC_Aggregated := { };
var NC_Proxies := { };
var NC_Disaggregated := { };
for each NC in NC_Original:
    if !NC.isAggregated
        NC_Disaggregated.add(NC);
        NC_NonAggregated.add(NC);
        end if;
```

Listing A2.2 – Pseudo-code for the Nominal Catch Disaggregation process

Where:

end for;

- P_1 , ..., P_8 are the eight currently available Disaggregation Procedures;
- *NC*¹, ..., *NC*ⁿ is the input Nominal Catch dataset;
- procedure.apply (<NC record>, <non aggregated NC record>) returns the proxy records (according to the current disaggregation procedure) for the aggregated record <NC record> being processed, as these are identified within the full set of <non aggregated NC record>;
- *record*.**breakdown** (<*NC proxies*>) breaks down the original, aggregated *record* into multiple disaggregated records, whose catch quantities (and species / gears) are proportionally assigned based on the identified <*NC proxies*>;
- *record*.manualBreakdown prompts users for their own, manual breakdown of the original aggregated *record*, as none of the disaggregation procedure was able to identify any valid proxy record for it;

The process by itself is quite simple and straightforward: its ability in identifying proper *proxy records* for the proportional breakdown of the original catches lies in the definition of the disaggregation procedures and in the *fleet / type of operation / region / area* mappings.

Procedure #	Fleet	Type of operation	Region	Area	Years
1	Same	Same	Same	Same	Same
2	Same	Same	Same	Same	+ / - 5 years
3	Any	Same	Same	Same	Same
4	Same	Same	Same	Same	+ / - 10 years
5	Same	Same	Any	Same	Same
6	Any	Same	Any	Same	Same
7	Any	Same	Any	Same	Any
8	Any	Same	Any	Any	Any

The currently available Disaggregation Procedures are defined in the following table:

 Table A2.3 – Disaggregation procedures definition

When moving from procedure #1 to procedure #8, the constraints for the identification of proxy records become more relaxed.

The first procedure, basically, looks for proxy records for the same fleet / type of operation / region / area and fishing year that do refer to any of the species / gear belonging to the aggregates involved.

Conversely, the last procedure will look for proxy records from any fleet, any region, any area and any fishing year, as long as the type of operations are the same, and that at the same time do refer to any of the species / gear belonging to the aggregates involved.

Country	Rep. country	Gear	Area	Region	Type of operation
			•••		
ESP	ESP	LLEX	IREASIO	EASIO	IND
ESP	ESP	LLEX	IRWESIO	WESIO	IND
ESP	ESP	PS	IREASIO	EASIO	IND
ESP	ESP	ELL	IREASIO	SWEIO	IND
ESP	ESP	LL	IREASIO	SWEIO	IND
ESP	ESP	ELL	IRWESIO	SWEIO	IND
ESP	ESP	LL	IRWESIO	SWEIO	IND
ESP	ESP	BB	IRWESIO	WESIO	IND
ESP	ESP	PS	IRWESIO	WESIO	IND
ESP	ESP	SUPP	IRWESIO	WESIO	IND
FRA	FRA	HAND	IRWESIO	MOZCH	ART
FRA	FRA	TROL	IRWESIO	MOZCH	ART
FRA	FRA	ELL	IRWESIO	SWEIO	IND
FRA	FRA	PS	IREASIO	EASIO	IND
FRA	FRA	PS	IRWESIO	WESIO	IND
FRA	REU	LLCO	IRWESIO	SWEIO	ART
FRA	REU	HAND	IRWESIO	SWEIO	ART
FRA	REU	HATR	IRWESIO	SWEIO	ART
FRA	REU	TROL	IRWESIO	SWEIO	ART
FRA	REU	ELL	IRWESIO	SWEIO	IND
FRAT	FRA	PS	IREASIO	EASIO	IND
FRAT	FRA	HAND	IRWESIO	MOZCH	ART
FRAT	FRA	HATR	IRWESIO	MOZCH	ART
FRAT	FRA	TROL	IRWESIO	MOZCH	ART
FRAT	FRA	ELL	IRWESIO	SWEIO	IND
FRAT	FRA	PS	IRWESIO	WESIO	IND

Intermediate procedures have a behaviour that is a mixture of these two.

 Table A2.4 – A sample of the fleet / type of operation / region / area mappings

The *fleet / type of operation / region / area* mappings (see Table A2.4) is used to identify – whenever needed – the region of most likely operation for a combination of fleet (country / reporting country), gear, area and type of operation.

As an example, if an aggregated record for FRA / HAND / IRWESIO / ART needs to be broken down into its disaggregated components, the disaggregation procedures might look for proxy records as available for any other fleet that is likely to operate in the Mozambique Channel (MOZCH) thus restricting the possible records based on the knowledge of the fishery.

Example of disaggregation results

To give a better understanding of the overall process, let's assume we run the disaggregation for all records referring to the *AG14 – Billfish nei* species aggregate.

So far, the IOTC database contains 94 Nominal Catch records referring to that species aggregate. When the process is completed, it produces 351 *disaggregated* records out of the 94 *original* ones.

If we look at the disaggregation results for the following record:

1986 / SUN / LL / IRWESIO / AG14 / 6 MT (Nominal Catch)

We see that the process was able to apply procedure #2 (same fleet, same type of operation, same region, same area and +/-5 years of difference with respect to the fishing year) to identify the following proxy records:

1983 SUN \mathbf{LL} IRWESIO BUM 3.000000 MT 1 1 1 Region: WESIO Operation: IN I. Ouality: Poor quality Source: Liaison Officer / SUN / LL / IRWESIO / BUM 1 3.000000 MT 1984 Region: WESIO Operation: IN Quality: Poor quality Source: Liaison Officer / SWO 1987 / SUN / LL / IRWESIO / 37.000000 MT Region: WESIO Operation: IN Quality: Poor quality | Source: Liaison Officer

Which refer to **BUM** (Blue Marlin, the first two) and **SWO** (Swordfish, the last one), with total catches for BUM accounting for 6.0 MT and total catches for SWO accounting for 37.0 MT (over a total of 43.0 MT).

Therefore, the original 6.0 MT for the aggregated record are proportionally assigned to **BLM** for a total of 0.837209 MT (6 MT * (6 MT / 43 MT)) and to **SWO** for a total of 5.162791 MT (6 MT * (37 MT / 43 MT)).

The sum of the catches assigned to each disaggregated record **is always equal** to the original catches for the aggregated record, as this process is not changing the overall Nominal Catches quantities but just reallocating these to the disaggregated records.

Using the Nominal Catch disaggregation to reconstruct catch time series

Nominal catches for AG22 - Sharks nei can be broken down to reconstruct and improve the catch time series for BSH - Blue shark and FAL - Silky shark (among others).



Figure A2.5 - Raw nominal catches for AG22, BSH and FAL



Figure A2.5 shows that catches for the AG22 aggregate (*Sharks nei*) are quite consistent since the '50s, whereas catches for **BSH** (*Blue shark*) and **FAL** (*Silky shark*) are of a certain relevance only from the '90s onwards.

Figure A2.6 – Raw nominal catches for AG22, BSH and FAL



Figure A2.7 – AG22 and BSH raw vs. BSH raw + disaggregated catches

Scientists seem to agree that most of the AG22 catches from the '50s onwards are indeed including mostly catches for BSH and FAL under the aggregate. For this reason, it would be interesting to apply the disaggregation procedure to aggregated AG22 catches and evaluate the produced results.

As can be seen in Figure A2.6 and A2.7, most of the **AG22** catches, especially in the '60s and '70s, are assigned to **FAL** (Silky sharks) with **BSH** getting far minor contributions over the entire time series.

These results are explained by the fact that in the IOTC database, most of the AG22 – Sharks NEI catches are recorded under *Artisanal* or *Semi-Industrial* gears. Between FAL - Silky Sharks and BSH - Blue Sharks, is the first that is mostly caught (as can be seen from the raw Nominal Catch data) by means of *Semi-Industrial* gears, whereas the latter is mostly caught with *Artisanal* and *Industrial* gears.

Therefore, based on the way in which the Nominal Catch Disaggregation process works (and based on the currently available disaggregation procedures), the reconstructed time series are perfectly in line with the available data.

It can be easily verified how the process does provide quite a substantially different outcome than either applying a constant factor to convert from AG22 to FAL / BSH catches or by applying a moving average that reconstructs catches using the proportions available for both species in the closest n years.