README FILE (T2SZ/T2CS: Task 2 size/weight frequencies)

ICCAT Secretariat

[Version 2024.01]

Summary

This document briefly describes the two types of Task 2 size/weight frequencies, compiled by ICCAT and publicly available on the ICCAT webpage (www.iccat.int/en/accesingdb.html):

1. **T2SZ** - Task 2 size/weight frequencies sampled (also known as actual size samples [“*de facto*” measured fish]) contains the species-based biological samples measured/weighted by the ICCAT CPCs and reported to ICCAT.
2. **T2CS** – Task 2 catch-at-size, contains the species-based estimations of the size composition of the catches, extrapolated by the ICCAT CPCs and reported to ICCAT. Only 6 species (ALB - albacore; BFT - bluefin tuna, BET - bigeye tuna, YFT - yellowfin tuna, SKJ skipjack, SWO - swordfish) have this reporting requirement. And, for each species, only the ICCAT flag CPCs with relevant catches of those species on their fisheries within a given year, should do those estimations and report them together with the corresponding methodology to the SCRS. From 2024 onwards, the T2CS annual provision became optional (to be explicitly requested by each SCRS species group, whenever required).

For each species-stock, the flag CPC list requiring both T2SZ and T2CS datasets, can be found on the SCRS species-stock based catalogues, updated every year and published on the ICCAT website (www.iccat.int/en/accesingdb.html) under section “Catalogues”.

For both dataset types T2SZ and T2CS, it is given an overall view of the content, formats and file structures.

CAVEAT: the CSV (comma separated values) files, one for each of the major species and groups of species (small tunas, others), replaces temporarily the standalone MS-ACCESS database published in the past (due to MS-ACCESS limitations). These files can be read with any text editor (e.g.: Notepad++), spreadsheet (e.g.: MS-EXCEL, LibreOffice) or other specialised software (such as R-CRAN).

The Secretariat is working to replace these CSV files (and the former MS-ACCESS databases) in the mid-term, with SQLite standalone relational database versions (more powerful and efficient).

Content

Both T2SZ and T2CS dataset types contains information for the period 1950 to 2022 reported by the ICCAT CPCs until 15 January 2024. Some historical datasets collected by ICCAT (port sampling, special programmes, etc.) in the past are also included to complement the officially reported data. The field “FileTypeCode” categorises the origin of those data sources. **Table 1** presents the number of datasets per “FileTypeCode” and dataset type (T2SZ/T2CS).

Table 1. Number of datasets in T2SZ and T2CS by data source origin (FileTypeCode) together with the overall ratios (%). The column Published indicates if they were included in the CSV files (YES) or not (NO).

		Dataset number		Ratio (%)		Published	Remarks
FileTypeCode	FileType	T2SZ	T2CS	T2SZ	T2CS	(Yes/no)	
OF-REP	Normal (Off. Rep)	14476	4730	91	91	YES	
WG-EST	WG estimations (CAS or CEF)	92	250	1	5	YES	
CONF	Confidential (not published)	57	0	0	0	NO	
REFF	Reference File (contains WG-RAW)	12	0	0	0	NO	Old Reference files (from SG-TROP)
OF-SEC	Secretariat estimates based on Official data	61	24	0	0	YES	Raw t2sz (used in Tropical tunas CAS estimations)
IATTC	External IATTC	19	0	0	0	YES	
P-SAMP	Port sampling (eq=11)	223	0	1	0	YES	
GBYP	Recovered under GBYP Program	108	0	1	0	YES	
ICC-SP	ICCAT Sampling Programme (old flgs: JPN.ICCAT/etc)	206	0	1	0	YES	
DR-ICC	Data recovery (payed by ICCAT, e.g. SWO-M)	42	0	0	0	YES	
WG-T2F	WG (samples obtained from farmed BFT at-wild)	69	0	0	0	NO	Only for the use of the SG-BFT
StCamR	StereoCamera measurements (cage transfers - RAW)	72	0	0	0	YES	
szAvgW	t2sz obtained from mean weights (set based)	61	0	0	0	NO	Old SG-BFT estimations of T2SZ (mean weight based)
StCamS	StereoCamera measurements (cage transfers - ST06)	30	0	0	0	YES	
StCaR2	StereoCamera measurements (RAW_v7 - temp)	83	0	1	0	YES	
MXFIS0	Mixed (FIS) unknow port (?)	17	164	0	3	YES	ongoing work to discriminate flags
MXFIS1	Mixed (FIS) PtNoire (Congo FRA?)	51	0	0	0	YES	ongoing work to discriminate flags
MXFIS2	Mixed (FIS) Abidjan (CIV?)	49	0	0	0	YES	ongoing work to discriminate flags
MXFIS3	Mixed (FIS) Dakar (SEN?)	76	3	0	0	YES	ongoing work to discriminate flags
K+Punk	Mixed(KR+PA)[no port]	11	0	0	0	YES	ongoing work to discriminate flags
K+Pabi	Mixed(KR+PA)[abidjan]	6	0	0	0	YES	ongoing work to discriminate flags
K+Plpa	Mixed(KR+PA)[Las Palmas]	35	0	0	0	YES	ongoing work to discriminate flags
K+Pten	Mixed(KR+PA)[Tenerife]	39	0	0	0	YES	ongoing work to discriminate flags
K+Psma	Mixed(KR+PA)[St Maart]	25	0	0	0	YES	ongoing work to discriminate flags
K+Pcum	Mixed(KR+PA)[Cumana]	2	0	0	0	YES	ongoing work to discriminate flags
K+Pcto	Mixed(KR+PA)[Cape Town]	1	0	0	0	YES	ongoing work to discriminate flags
NO-PUB	Not to be published (waiting SCRS approval)	0	24	0	0	YES	
TOTAL		15923	5195	100.0	100.0		

More than 90% in both T2SZ and T2CS dataset types are CPC based size/weight data reported to ICCAT.

The complementary detailed catalogue on this page (https://www.iccat.int/en/accesingdb.html) contains important information and the corresponding metadata of each dataset for both major tuna & tuna like species and small tunas.

Formats & structure

The CSV formats published were based on the standard structures used by the ICCAT scientific Committee (SCRS) in the recent years. Each record identifies the number of fish within a size/weight class bin. All the remainder fields (fishing flag, gear, time strata, geographical strata, etc.) characterise each record. This format (vertical orientation) is usually more flexible than the horizontal one (class bins rotated as columns).

The structure of the CSV files, specifically the fields (name, type, constrains) and their short descriptions are presented in **Table 2**.

For additional questions please contact the ICCAT Secretariat (STATS_info@iccat.int).

Table 2 Structure of the CSV files for both T2SZ and T2CS dataset types

Field	Type	Description	Constrains	Example
InProcID	integer	Data set unique identifier (internal) [equal to field "DSet" in the detailed catalogue file]	n/a	38398
SpeciesCode	string	Species code	ICCAT codes (= FAO ISSCFG codes)	SSP
YearC	integer	Year (calendar)	{1950, 2022}	2019
Decade	integer	Decade (natural decades: 2010-19, 2000-09, 1990-99, etc.)	{1950, 2020}	2010
FlagName	string	Flag name	ICCAT codes (Flags table)	Chinese Taipei
FleetCode	string	Fleet code	ICCAT codes (Fleets table)	TAI
GearGrpCode	string	Gear group code	ICCAT codes (gear group table)	LL
GearCode	string	Gear code	ICCAT codes (gear table)	LL
LorD	string	Catch type (landed, discarded)	{(L, D), (landings, discards)}	L
FreqTypeCode	string	Frequency type code	ICCAT codes (frequency types table)	LJFL
SzInterval	integer	Size/weight class interval	{[1, 2, 5, 10]cm, [1, 2, 5]kg}	1
FreqsGroup	string	Frequency group (cm, kg) & class boundary limit	{(ll, cp, ul, un), (lower, central point, upper, unknown)}	cm (ll)
SizeInfoCode	string	Size/weight information code (dataset type)	{(siz, cas), (T2SZ, T2CS)}	siz
FileTypeCode	string	File type code	ICCAT codes (see Table 1)	OF-REP
StrataID	integer	Stratum unique identifier (across all T2SZ and T2CS datasets)	n/a	1612883
TimePeriodCatch	integer	Time period of the catch (from which the sample was taken)	{month [1, …, 12], quarter [13, 14, 15, 16], year[17]}	12
TimeStrata	string	Time period stratification type	{(mm, qq, yy), (monthly, quarterly, yearly)}	mm
SchoolTypeCode	string	School type code (fishing modes of tropical purse fisheries)	{(FAD, FSC, n/a), (on FADs, on free schools, not applicable)}	n/a
Stock	string	Species based stocks (or management unit)	ICCAT codes (see maps) [e.g.: ALB+ATS => ALB-S]	ATS
SampAreaCode	string	Sampling area codes (by species: ALB, BFT, BET, YFT, SKJ; rest of species: BIL areas)	ICCAT codes (see maps)	BIL96
GeoStrata	string	Geographical stratification type (rectangles, squares, sampling areas)	ICCAT codes (Square Types table) [see maps]	5x5
QuadID	integer	Quadrant identifier (clockwise)	ICCAT codes (see maps)	3
Lat	float	Latitude (always positive and dependent on fields QuadID)	{0, 70}	15
Lon	float	Longitude (always positive and dependent on QuadID)	{0, 99}	25
FishSampWGT	integer	Weight (kg) of the fish sampled/caught (T2SZ/T2CS) [caution: highly incomplete/unreliable INFO]	n/a	0
FishSampNUM	integer	Number of fish sampled/caught (T2SZ/T2CS) [caution: highly incomplete/unreliable INFO]	n/a	0
ClassFrq	integer	Size (cm)/weight (kg) class bins (some outliers under revision)	Watch global outliers outside ranges [10, 600]cm and [1, 800]kg	160
SexCode	string	Sex code (ICCAT codes)	{(M, F, I, U), (male, female, immature, unknow)}	U
Nr	float	Number of fish		1