Kramer DL and Chapman MR (1999) Implications of fish home range size and relocation for marine reserve function. *Environmental Biology of Fishes* 55: 65-79.
Longhurst A and Pauly D (1987) *Ecology of Tropical Oceans*. Academic Press, San Diego, 407 p.
MacCall A (1990) *Dynamic Geography of Marine Fish Populations*. University of Washington Press, Seattle.
NOAA (2004) ETOPO2 - 2-Minute Gridded Global Relief Data http://www.ngdc.noaa.gov/mgg/fliers/01mgg04.html.
Pauly D (1998) Tropical fishes: patterns and propensities. p. 1-17. *In:* T.E. Langford, J. Langford and J.E. Thorpe (eds.) Tropical Fish Biology. *Journal of Fish Biology* 53 (Supplement A).
Pauly D (2010) *Gasping Fish and Panting Squids: Oxygen, Temperature and the Growth of Water-Breathing Animals*. Excellence in Ecology (22), International Ecology Institute, Oldendorf/Luhe, Germany, xxviii + 216 p.
Pauly D and Chua Thia-Eng (1988) The overfishing of marine resources: socioeconomic background in Southeast Asia. *AMBIO: a Journal of the Human Environment* 17(3): 200-206.
Phillips SJ, Anderson RP and Schapire RE (2006) Maximum entropy modelling of species geographic distributions. *Ecological Modelling* 190: 231–259.
Pikitch EK, Santora C, Babcock EA, Bakun A, Bonfil R, Conover DO, Dayton P, Doukakis P, Fluharty DL, Heneman B, Houde ED, Link J, Livingston PA, Mangel M, McAllister MK, Pope J and Sainsbury KJ (2004) Ecosystem-based Fishery Management. *Science* 305: 346-347.
Sherman K, Aquarone MC and Adams S (2007) *Global applications of the Large Marine Ecosystem concept*. NOAA Technical Memorandum NMFS NE, 208, 72 p.
Smith CL (1997) *National Audubon Society field guide to tropical marine fishes of the Caribbean, the Gulf of Mexico, Florida, the Bahamas, and Bermuda*. Alfred A. Knopf, Inc., New York. 720 p.
UNEP-WCMC (2010) Global distribution of coral reefs, compiled from multiple sources. UNEP World Conservation Monitoring Centre. http://data.unep-wcmc.org/datasets/1
Zeller D and Pauly D (2001) Visualisation of standardized life history patterns. *Fish and Fisheries* 2(4): 344-355.

Section 4

The *Sea Around Us* databases and their spatial dimensions[8]

Vicky W.Y. Lam, Ar'ash Tavakolie, Daniel Pauly and Dirk Zeller

The individual catch reconstructions for all countries and territories (by EEZ) are all available at www.seaaroundus.org. The underlying taxonomically disaggregated time series of catch data they contain, covering all years since 1950, 4 fishing sectors (industrial, artisanal, subsistence and recreational), 2 catch types (landed versus discarded catch) and 2 types of reporting status (reported versus unreported) for the Exclusive Economic Zones (EEZs) of all maritime countries and territories of the world, or parts thereof, are part of an extensive dedicated database, which interacts with the other databases of the *Sea Around Us* to generate the spatially allocated fisheries catches for the 180,000 half degree latitude and longitude cells covering the world ocean. These data represent the core product of the *Sea Around Us*.

Catch database

The catch reconstruction database comprises all of the catch reconstruction data by year, fishing country, taxon name, catch amount, fishing sector, catch type, reporting status, input data source and spatial location of catch such as Exclusive Economic Zone (EEZ), FAO area or other area

[8] Adapted from: Lam VWY, Tavakolie A, Pauly D and Zeller D 2016. The *Sea Around Us* catch database and its spatial expression, *In:* Pauly D and Zeller D (eds.) *Global Atlas of Marine Fisheries: Ecosystem Impacts and Analysis*. Island Press, Washington, D.C.

designation (if applicable). The database is further sub-divided into three different data ‘layers’, which include a layer with the catch taken by a fishing country in its own EEZ (called ‘Layer 1’), the catch by each fishing country in other EEZs and/or the high seas (‘Layer 2’), and the catch of all tuna and large pelagic species caught by each fishing country’s industrial fleet (‘Layer 3’). The basic structure of Layers 1 and 2 are identical, while Layer 3 differs slightly in structure due to the nature of the large pelagic input data sets (see Section #2).

Data verification process

The process of data integration into the catch reconstruction database includes a data verification process, which is the first integration step of the original reconstruction dataset and associated reconstruction report. After completing the data verification process for each country dataset, each record is allocated to one of the layers based on the taxon, sector, and the area where the taxon was caught, and is formatted to fit the structure of the final database (see Figure 1 for overview). For example, the total *reported landings* presented in the reconstruction dataset of each country/territory (which represent the catches landed *and* deemed reported by national authorities from within the own EEZ of that country/territory) are compared with the reported data as present by FAO on behalf of the respective country/territory for each year. Any ‘surplus’ of FAO data are then considered to have been caught outside the EEZ of the given country/territory, and thus are treated as part of ‘Layer 2’ data. Thus, ‘Layer 2’ data are a derived data product. When any issue with the reconstructed catch data are identified, the issue is raised with the *Sea Around Us* catch reconstruction team and the original authors of the reconstruction for further checking and refining of the input data. Additional data verification steps include harmonization of scientific taxon names in the reconstruction data with the official, globally recognized and standardized taxon names via the global taxonomic authorities of FishBase (www.fishbase.org) and SeaLifeBase (www.sealifebase.org). Fishing country names and EEZ names are also checked and standardized against the *Sea Around Us* spatial databases. The fishing country and EEZ names allow us to link the catch data to the foreign fishing access database, which contains the information on which fishing country can access the EEZ of another country (see ‘*Foreign fishing access database*’ section below).

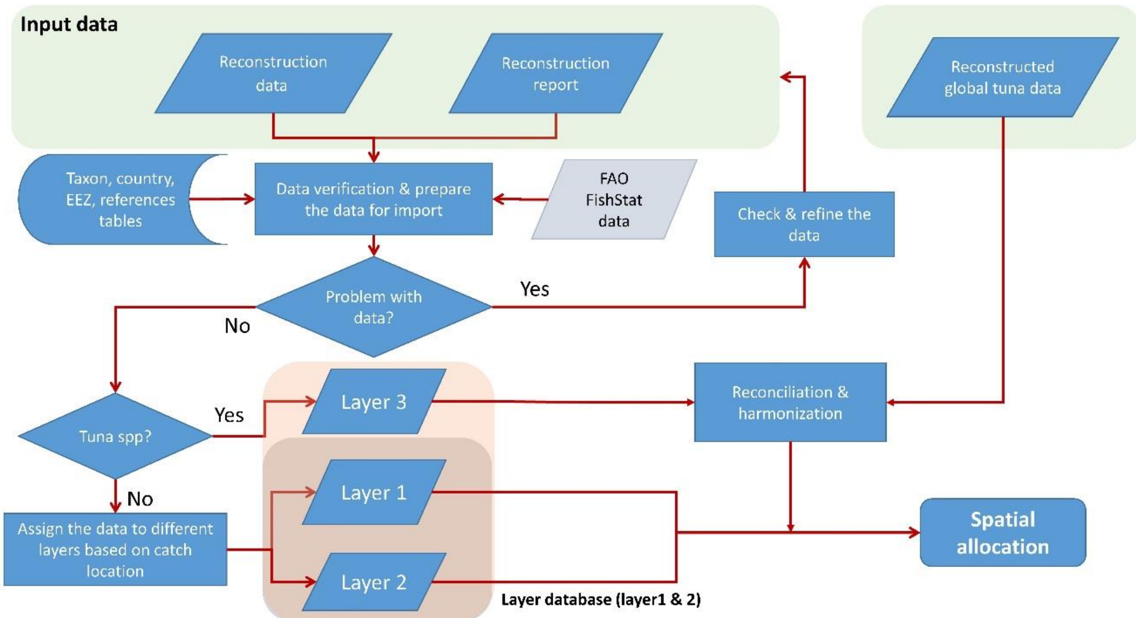

Figure 1. Data verification process for catch reconstruction data of the *Sea Around Us*.

Foreign fishing access database

The foreign fishing access database, initially derived from the fishing agreement database of FAO (1999), contains observed foreign fishing records, and fishing agreements and treaties that were signed by fishing countries and the host countries in whose EEZs the foreign fleets were allowed to fish. In addition, the database also includes the start and end year of agreements and/or observed access. The type of access is also specified as 'assumed unilateral', 'assumed reciprocal', 'unilateral' or 'reciprocal'. Also, the type of agreement is recorded in the database and the agreement can be classified into bilateral agreements such as partnership, multilateral agreements such as international conventions or agreements with regional fisheries organizations, private, licensing or exploratory agreements. Additional information contained in this database relates to the type of taxa likely targeted by foreign fleets (e.g., tuna vs. demersal taxa).

This database is used in conjunction with the catch reconstruction database and the taxon distribution database (see Section #3) in the spatial allocation process that assigns catches to the global *Sea Around Us* ½ x ½ degree latitude and longitude cell system.

The *Sea Around Us* ½ x ½ degree cell system

The *Sea Around Us* uses a spatial database where the world is divided into a global coverage of 30 arc minute cells (½ x ½ degree). A world cell structure was implemented to conform to the Land-Ocean Interactions in the Coastal Zone (LOICZ) system (www.loicz.org/). Thus, the world is partitioned into 30 x 30 arc minute cells with a top left bounding box corner coordinate at 90°N and 180°W. This results in 180,000 such ½ x ½ degree cells covering the world's oceans, which form the spatial foundation for all *Sea Around Us* data.

Spatial allocation procedure

The spatial allocation procedure - although it relies on the same global *Sea Around Us* grid of ½ x ½ degree cells that was used previously - is different from the approach used in the early phase of the *Sea Around Us* (until 2006) and described in Watson *et al.* (2004). In the earlier allocations, catches pertaining to large reporting areas (mainly FAO statistical areas) were allocated directly to the half-degree cells, subject only to constraints provided by initially derived taxon distributions for the various taxa (Close *et al.* 2006), and an earlier and more limited version of the fishing access database granting foreign fleets differential access to the EEZs of various countries (Watson *et al.* 2004). Thereafter, the catch by a given fishing country in a given EEZ was obtained by summing the catch that had been allocated to the cells making up the EEZ of that country. This process made the spatial allocation overly sensitive to the precise shape and cell-probabilities of the taxon distribution maps, and the precision of very problematic EEZ access rules for different countries. It regularly resulted in sudden and unrealistic shifts of allocated domestic catches into and out of given EEZs purely due to the lifting or imposing of EEZ access constraints. Attempts to improve the allocation procedure with more internal rules made it unwieldy, fragile and extremely time consuming, and thus the *Sea Around Us* abandoned this approach in the mid-2000s.

The more structured allocation procedure that was devised as a replacement, and is described here (Figure 2), relies on catch data that are spatially pre-assigned (through in-depth catch reconstructions, see Section #1) to the EEZ or EEZ-equivalent waters (for years pre-dating the declaration of individual EEZs) of a given maritime country or territory, and, in the case of small-scale fisheries (i.e., the artisanal, subsistence and recreational sectors), to the Inshore Fishing Areas (Chuenpagdee *et al.* 2006), and in the case of industrially caught large pelagics, to large 'tuna cells' (Section #2). This radically reduces the number of access rules and constraints that the allocation procedure must consider, reduces the chances of fish catches showing up in the EEZs of the wrong country, and dramatically reduces the processing times of the allocation procedure.

Previously, we also used the spatial allocation process to simultaneously disaggregate (i.e., taxonomically improve) uninformative taxonomic groups such as 'miscellaneous marine fishes' (FAO term: 'marine fishes nei') by relying on taxonomic information in neighboring ½ degree cells. This further added to the complexity of the allocation procedure and increased the difficulty of tracing actual country/taxon/catch entities through the process. This step was also discontinued in the new allocation approach. Instead, our 'new and improved' allocation procedure disaggregates the input catch data as part of the country-by-country catch reconstruction process (Section #1), with the associated more transparent taxonomic changes documented in the associated technical report for each reconstruction. Within the catch reconstruction database, we keep track of the quality of the taxonomic disaggregation, such that indicators sensitive to the quality of the disaggregation are not computed from inappropriate data (see '*Catch composition*' in Section #1).

These pre-allocation data processing modifications allow focusing on the truly spatial elements of the allocation, which are handled through a series of conceptual algorithmic steps. The general algorithm of spatial allocation of catches is harmonized for Layers 1, 2, and 3 (Table 1), which allows for a better software flow, while maintaining the conceptual differences in data layers.

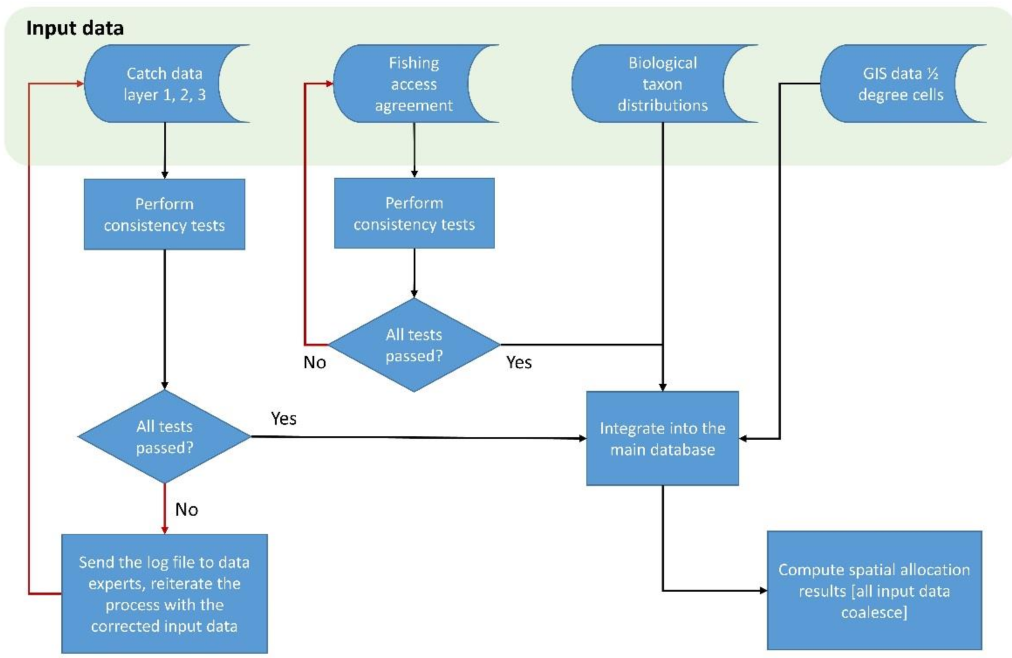

Figure 2. Spatial allocation procedure for catch reconstruction data of the *Sea Around Us*, resulting in the final ½ x ½ degree allocated cell data.

The spatial allocation of the catch is the process of computing the catch that can be allocated to each ½ degree cell based on the overlap of three main components: 1) the catch data, 2) the fishing access observations/agreements, and 3) the biological taxon distributions (Figure 5.2). The relationship/overlap amongst these components is facilitated by a series of Geographic Information System (GIS) layers, which essentially bind them together.

How each data layer is conceptually unique and how it is handled

In Layer 1, the data come spatially organized by each fishing entity's EEZ(s). The allocation algorithm allocates the small-scale catches (i.e., artisanal, subsistence, and recreational) only to the cells associated with the Inshore Fishing Area (IFA, Chuenpagdee *et al.* 2006) of that fishing entity's EEZ, while industrial catches can be allocated anywhere within that fishing entity's EEZ(s), as long as they remain compatible with the biological taxon distributions. Fishing access agreements are not applicable to this data layer, as a fishing entity (i.e., country) is always allowed to fish in its own EEZ waters. To represent the historical expansion of industrial fishing since the 1950s in each country's waters, from more easily accessible areas closer to shore to the full extent of each country's EEZ, we use the depth adjustment function for domestic industrial catches described in Watson and Morato (2013). This function takes into account that, as domestic industrial catches increase over time, an increasing fraction are being taken progressively further offshore (but within EEZ boundaries).

Table 1. Parameters of the three catch data input layers as used in the spatial allocation to ½ x ½ degree cells of the *Sea Around Us*.

	Layer 1	**Layer 2**	**Layer 3**
Taxa included	All except industrial large pelagics	All except large pelagics	Large pelagics (n =140+)
Spatial scope	Country's own EEZ	Other EEZs and high seas	Global tuna cells
Sectors	Industrial, Artisanal, Subsistence, Recreational	Industrial	Industrial
Distributions	Biological	Biological	Biological
Fishing access	Automatically granted	Used	Used
Granularity of data	EEZ, IFA[1]	EEZ, high seas, ICES, CCAMLAR, NAFO, FAO and other areas	Six types of tuna cells: 1x1, 5x5, 5x10, 10x10, 10x20, 20x20 degrees lat. long.

[1] Inshore Fishing Area (IFA), defined as the area up to 50 km from shore or 200 m depth, whichever comes first (Chuenpagdee *et al.* 2006). Note that IFAs are defined only along inhabited coastlines.

In Layer 2, the spatial granularity of the catch data can be by EEZ, high seas, or any other form of regional reporting areas, i.e., ICES, CCAMLR, NAFO, or FAO statistical areas. However, in all cases it excludes the fishing entity (fishing country's) own EEZ waters (which are treated in Layer 1). In Layer 2, the fishing access observations/agreements are used to compute the areas which allow fishing for a particular fishing entity, year, and taxon. Note that we tread EEZ areas prior to each country's EEZ declaration year as 'open access', meaning no restrictions are applied to other countries being allowed access to these waters. Once this area is computed, it is superimposed on the biological taxon distributions to derive the final Layer 2 catch allocation.

In Layer 3, which only covers industrial large pelagics and their associated bycatch and discards, the input catch data are spatially organized by larger tuna cells which range from 1 x 1 to 20 x 20 degrees (Table 1, see also Section #2). Similar to the region specific areas in Layer 2, these larger cells are intersected with all the EEZ boundaries to create a GIS layer which is suitable for use in the algorithm. Thereafter, the fishing access observations/agreements and taxon distributions are applied to calculate the final Layer 3 catch allocation.
The spatial allocation algorithm consists of 4 main steps:

1. Validating and importing the fishing access observations/agreements database;
2. Validating and importing the catch reconstruction database;
3. Importing the biological taxon distributions; and
4. Computing the catch that can be allocated to each ½ degree cell for each data layer in an iterative process (allowing for verifications and corrections to any of the input parameters).

1. Validating and importing the fishing access observations/agreements database

The fishing access observations/agreements are first verified using several consistency and 'matching' tests (Figure 2) and, upon passing, they are imported into the main allocation database. This fishing access information is subsequently used in two different processes: (a) the verification process of the catch data (Layers 1, 2, and 3); and (b) the computing of the areas where a given fishing entity (i.e., country) is allow to fish for a specific year and taxon.

2. Validating and importing the catch reconstruction database

The validating and importing of the catch data is a more complex process than the validating and importing process for the fishing access database. This process involves over 20 different pre-allocation data tests (Figure 2). These tests are designed to make sure that the data are coherent from the standpoint of database logic, and do not contain any accidental errors. These tests range from simple tests like "*is the TaxonKey valid?*" to more complex tests like "*validate if the given fishing entity has the required fishing access observations/agreements to fish in the given marine area*". Every single row of catch data is examined via these tests, and if it passes all tests the data row in question is added to the main allocation database. If it fails *any* of the tests it is returned to the relevant *Sea Around Us* data experts for review, often involving the original authors of the catch reconstruction (Figure 2). This iterative process is repeated until all the data rows pass all the pre-allocation tests.

The process of importing the catch reconstruction database includes an important sub-module for harmonizing the marine areas. This module is crucial, as the catch data come in a variety of different spatial reporting areas that are not globally homogenous in GIS definitions (e.g., the EEZ of Albania is one entity, while the EEZ of India, Brazil or the US are subdivided into sub-national areas; the north-east Atlantic uses ICES statistical areas, etc.). To harmonize these marine areas and make them accessible to the core allocation process, any given ½ degree cell is split into its constituent countries EEZs and high seas components. Then, the fishing access observations/agreements are applied to this layer to determine which of these 'shards' of ½ degree cells are allowing access to a given fishing entity. Once this is determined, these collections of 'shards' are assigned to the given row of catch data, the result is a harmonized view of all the different marine areas. Presently, we have assigned over 12,000 marine areas into their constituent 'shards' of ½ degree cells, ranging from EEZs and LMEs, to ICES, CCAMLR, NAFO, and FAO statistical areas. The procedure allows future marine areas to be readily assigned.

3. Importing biological taxon distributions

Importing the biological taxon distributions is a fairly straightforward process. The over 2,000 individual taxon distributions (see Section #3) are generated as individual text files (csv format) containing for each ½ x ½ degree cell the specific taxon's probability of occurrence. These individual taxon distribution files are compiled into a database table for further use.

4. Computing/allocating the catch to ½ degree cells

Once Steps 1-3 are completed, we perform the computations which yield the final spatial ½ x ½ degree allocation results. The catch of a given data row, *TotalCatch*, of taxon *T* is distributed amongst eligible ½ degree cells, *Cells 1...n*, using the following weighted average formula:

$$Cell_{i_{AllocatedCatch}} = TotalCatch \times \frac{Cell_{i_{SurfaceArea}} \times Cell_{i_{RelativeAbundance\ of\ Taxon\ T}}}{\sum_1^n Cell_{i_{SurfaceArea}} \times Cell_{i_{RelativeAbundance\ of\ Taxon\ T}}}$$

Throughout the allocation process, data parameters besides *year* and *taxon*, such as *sector*, *catch type*, *reporting status* etc. are preserved and carried over into the final allocated database.

The final results of the intense and detailed database preparation and spatial allocation are time series of catches by ½ degree cells that are ecologically reliable (i.e., taxa are caught where they occur, and in relation to their relative abundance) and politically likely (e.g., by fishing country and within EEZ waters to which they have access to).

Summarizing allocated data by spatial search regions

While some input data contain spatial designations, such as EEZs or FAO areas (Section #1) or large tuna cells (Section #2), no such spatial pre-designations exist for other spatial search regions we offer, such as LMEs, RFMOs, High Seas etc. in any input data. Thus, data presented for these search areas, or any other custom spatial area, are the result of combining the data from subsets of ½ x ½ degree cells (Section #4) covering the area in question.

References

Chuenpagdee R, Liguori L, Palomares MLD and Pauly D (2006) *Bottom-up, global estimates of small-scale marine fisheries catches*. Fisheries Centre Research Reports 14(8), University of British Columbia, Vancouver. 112 p.

Close C, Cheung WWL, Hodgson S, Lam V, Watson R and Pauly D (2006) Distribution ranges of commercial fishes and invertebrates. pp. 27-37 *In;* Palomares MLD, Stergiou KI and Pauly D (eds.), *Fishes in databases and ecosystems*. Fisheries Centre Research Reports 14(4), University of British Columbia, Vancouver.

FAO (1999) *FAO Fisheries Agreement Register (FARISIS)*. Committee on Fisheries, 23rd Session, 15-19 February 1999. COFI/99/INF.9E. 9 p.

Watson R, Kitchingman A, Gelchu A and Pauly D (2004) Mapping global fisheries: sharpening our focus. *Fish and Fisheries* 5: 168-177.

Watson R and Morato T (2013) Fishing down the deep: Accounting for within-species changes in depth of fishing. *Fisheries Research* 140: 63-65.