

THE SEA AROUND US CATCH DATABASE AND ITS SPATIAL EXPRESSION¹

Vicky W. Y. Lam, Ar'ash Tavakolie, Maria L. D. Palomares, Daniel Pauly, and Dirk Zeller
Sea Around Us, University of British Columbia,
Vancouver, BC, Canada

Although the fisheries catch reconstructions whose rationale and methods are described in chapters 1, 2, and 3 are all available online for checking (see www.searoundus.org), the taxonomically disaggregated time series of catch data they contain, covering 61 years (1950–2010), four fishing sectors (industrial, artisanal, subsistence, and recreational), two catch types (landed vs. discarded catch), and two types of reporting status (reported vs. unreported) for the Exclusive Economic Zones (EEZs) of all maritime countries and territories of the world or parts thereof ($n = 273$), are too large to be presented as flat tables in scientific or other articles, however detailed.

Therefore, the catch data generated by the reconstruction project of the *Sea Around Us* are stored in a dedicated catch reconstruction database, which interacts with the other databases of the *Sea Around Us* to generate various products, foremost among them spatially allocated fisheries catches to the 180,000 half-degree latitude and longitude cells covering the world's oceans. These products are freely available through our website and are likely to be widely used.

Because global catch maps and related products that are meaningful in terms of ecology as well as policy are one of the major outputs of the *Sea Around Us*, and because

the spatial allocation process is closely tied to the catch reconstruction database, this database and the spatial allocation process are described together in this chapter.

CATCH RECONSTRUCTION DATABASE

The catch reconstruction database comprises all the catch reconstruction data for 1950 to 2010 by fishing country, taxon name, year of catch, catch amount, fishing sector, catch type, reporting status, input data source, and spatial location of catch, such as EEZ, FAO area, or other area designation (if applicable). The database is further subdivided 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 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 structures of layers 1 and 2 are identical, and layer 3 differs slightly in structure because of the nature of the large pelagic input datasets (see chapter 3).

The process of data integration into the catch reconstruction database includes a data verification process, which is the first integration step undertaken after the original reconstruction dataset and associated reconstruction report are received after review by senior *Sea Around Us* staff. After the data verification process is completed for each country dataset, each record is allocated to one of the layers based on the taxon, sector, and area where the taxon was caught.

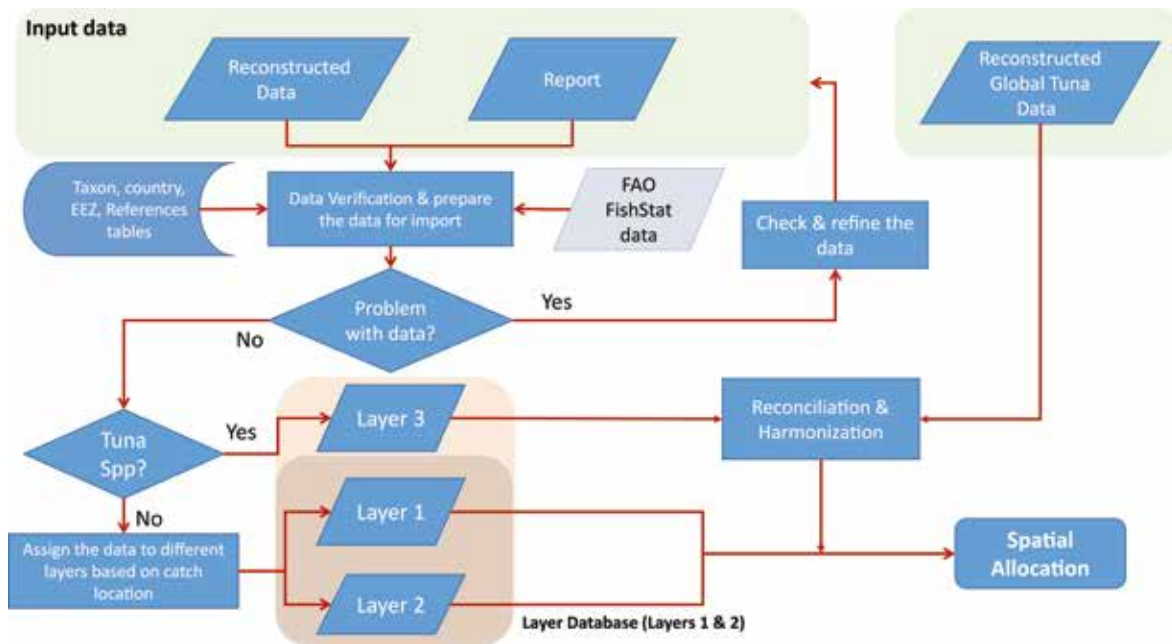


Figure 5.1. Data verification process for catch reconstruction data of the *Sea Around Us*. Details for the country- or territory-specific “Reconstructed Data” and “Report” are provided in chapters 1 and 2, and details for “Reconstructed Global Tuna Data” are described in chapter 3.

Data Verification Process

After initial, detailed review of each country or territory reconstruction dataset and associated technical report by senior *Sea Around Us* staff, the reconstruction dataset for each EEZ is further verified for accuracy and is formatted to fit the structure of the final database (see figure 5.1 for overview). For example, the total reported landings presented in the reconstruction dataset of each country or territory (which represent the catches landed and deemed reported to national authorities from within the EEZ of that country or territory) are compared with the reported data as present by FAO on behalf of the respective country or territory for each year. Any surplus FAO data are then considered to have been caught outside the EEZ of the given country or territory and thus are treated as part of layer 2 data. When an issue with the reconstructed catch data is 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 (figure 5.1). Additional data verification steps include harmonization of scientific tax-

on names in the reconstruction data with the official, recognized and standardized taxon names via the global taxonomic authorities of FishBase (www.fishbase.org) and SeaLifeBase (www.sealifebase.org; see chapter 4). 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 information on which fishing country can access the EEZ of another country (see “Foreign Fishing Access Database” section on p. 62).

Based on the location where each taxon was deemed to have been caught, each catch record is then assigned to a different layer (see “Structure of the Database” below). This includes a cross-checking process between the various reconstruction input datasets. For example, if an information source for Country A reported the landings of another country (B) in the EEZ of Country A, this catch of Country B is checked against the data in layer 2 of Country B, as provided through Country B reconstruction data. Emphasis is placed on avoiding double counting of catches.

Structure of the Database

As outlined above, the catch reconstruction database contains the catch data assigned to one of three layers.

Layer 1

This layer retains all the catches taken by a country within that country's own EEZ. It contains industrial, artisanal, subsistence, and recreational sector catches, subdivided by catch type (retained and landed vs. discarded catch) and reporting status (reported vs. unreported). However, this layer excludes all industrial catches of large pelagic fishes by a given country (see table 5.1 for the list of reported taxa excluded here), which are moved to layer 3 for later harmonization with the reconstructed global tuna data presented in chapter 3 (see also figure 5.1).

Layer 2

This layer contains data derived either directly from the reconstruction datasets and reconstruction technical reports (i.e., catches listed as being taken outside the country's own EEZ) or indirectly by subtracting the reconstructed catch identified as reported landings in a country's own EEZ (i.e., reported landings in layer 1) from the data reported by FAO on behalf of that country in the relevant (i.e., the "home" FAO area of a given fishing country) FAO area (excluding the taxa listed in table 5.1). Also, layer 2 includes catches by a given fishing country in all nonhome FAO areas (i.e., we refer to these catches as being taken by the distant-water fleets of the country in question). This layer includes only industrial fishing sector catches, as we define all fleets or gears that can operate outside of a given country's own EEZ waters as industrial (i.e., large-scale) in nature. The few documented cases where locally so-called artisanal fleets fish in neighboring EEZs, such as long-distance Senegalese pirogues operating in neighboring Mauritania or Guinea-Bissau (Belhabib et al. 2014), are internally reassigned to the industrial sector.

Table 5.1. Tuna and large pelagic taxa ($n = 29$) initially moved from country reconstruction datasets to layer 3 and later harmonized with industrially caught tuna and other large pelagic fishes (see also chapter 3).

| Common name | Taxon name |
|--------------------------|-----------------------------------|
| Albacore | <i>Thunnus alalunga</i> |
| Atlantic bluefin tuna | <i>Thunnus thynnus</i> |
| Atlantic bonito | <i>Sarda sarda</i> |
| Atlantic sailfish | <i>Istiophorus albicans</i> |
| Atlantic white marlin | <i>Tetrapturus albidus</i> |
| Bigeye tuna | <i>Thunnus obesus</i> |
| Billfishes | Istiophoridae |
| Black marlin | <i>Makaira indica</i> |
| Black skipjack | <i>Euthynnus lineatus</i> |
| Blackfin tuna | <i>Thunnus atlanticus</i> |
| Blue marlin | <i>Makaira nigricans</i> |
| Bullet tuna | <i>Auxis rochei</i> |
| Indo-Pacific blue marlin | <i>Makaira mazara</i> |
| Indo-Pacific sailfish | <i>Istiophorus platypterus</i> |
| Kawakawa | <i>Euthynnus affinis</i> |
| Little tunny | <i>Euthynnus alletteratus</i> |
| Longbill spearfish | <i>Tetrapturus pfluegeri</i> |
| Longtail tuna | <i>Thunnus tonggol</i> |
| Mediterranean spearfish | <i>Tetrapturus belone</i> |
| Pacific bluefin tuna | <i>Thunnus orientalis</i> |
| Shortbill spearfish | <i>Tetrapturus angustirostris</i> |
| Skipjack tuna | <i>Katsuwonus pelamis</i> |
| Slender tuna | <i>Allothunnus fallai</i> |
| Southern bluefin tuna | <i>Thunnus maccoyii</i> |
| Striped marlin | <i>Kajikia audax</i> |
| Swordfish | <i>Xiphias gladius</i> |
| Tuna | <i>Thunnus</i> spp. |
| Wahoo | <i>Acanthocybium solandri</i> |
| Yellowfin tuna | <i>Thunnus albacares</i> |

Layer 3

This layer initially included twenty-nine specific large pelagic taxa (table 5.1), whose reconstructed industrial catch data were moved to this layer for harmonization with the independently and globally reconstructed large pelagic dataset (chapter 3). The global tuna dataset combined taxonomically more diverse large-pelagic catch datasets and added bycatch and discards associated with the global industrial tuna and large pelagic fisheries.

Thus, the final large pelagic dataset (harmonized layer 3, figure 5.1) contains about 140 taxa and their associated catches.

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 publicly accessible 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 years of agreements 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 partnerships, 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 to be targeted by foreign fleets (e.g., tuna vs. demersal taxa), as well as any available data on fees paid or quotas included in the agreements.

Note that some agreements are not made public by either fishing country or host country. Thus, it is not readily possible to differentiate between foreign legal or illegal catches.

This database is used in conjunction with the catch reconstruction database and the taxon distribution database (see chapter 4) in the spatial allocation process that assigns catches to the global *Sea Around Us* half-degree latitude and longitude cell system.

SPATIAL ALLOCATION PROCEDURE

The spatial allocation procedure, although it relies on the same global *Sea Around Us* grid of half-degree cells that was used previously, is different from the approach used in the early phase of the *Sea Around Us* (until 2006) as described in Watson et al. (2004). In the

earlier allocations, catches pertaining to large reporting areas (mainly FAO areas; see figure 2.1 in chapter 2) 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 (Watson et al. 2004). 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 often resulted in sudden and unrealistic shifts of allocated catches into and out of given EEZs purely through the lifting or imposing of EEZ access constraints. Attempts to improve the initial 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 5.2), relies on catch data that are spatially preassigned (through in-depth catch reconstructions; see chapter 2) to the EEZ or EEZ-equivalent waters (for years predating 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 their Inshore Fishing Areas (IFAs; Chuenpagdee et al. 2006). This radically reduces the number of access rules and constraints that the allocation procedure must consider, avoids fish catches showing up in the EEZs of the wrong country, and dramatically reduces the processing times of the allocation procedure.

Watson et al. (2004) also used the spatial allocation process to simultaneously disaggregate (i.e., taxonomically improve) uninfor-

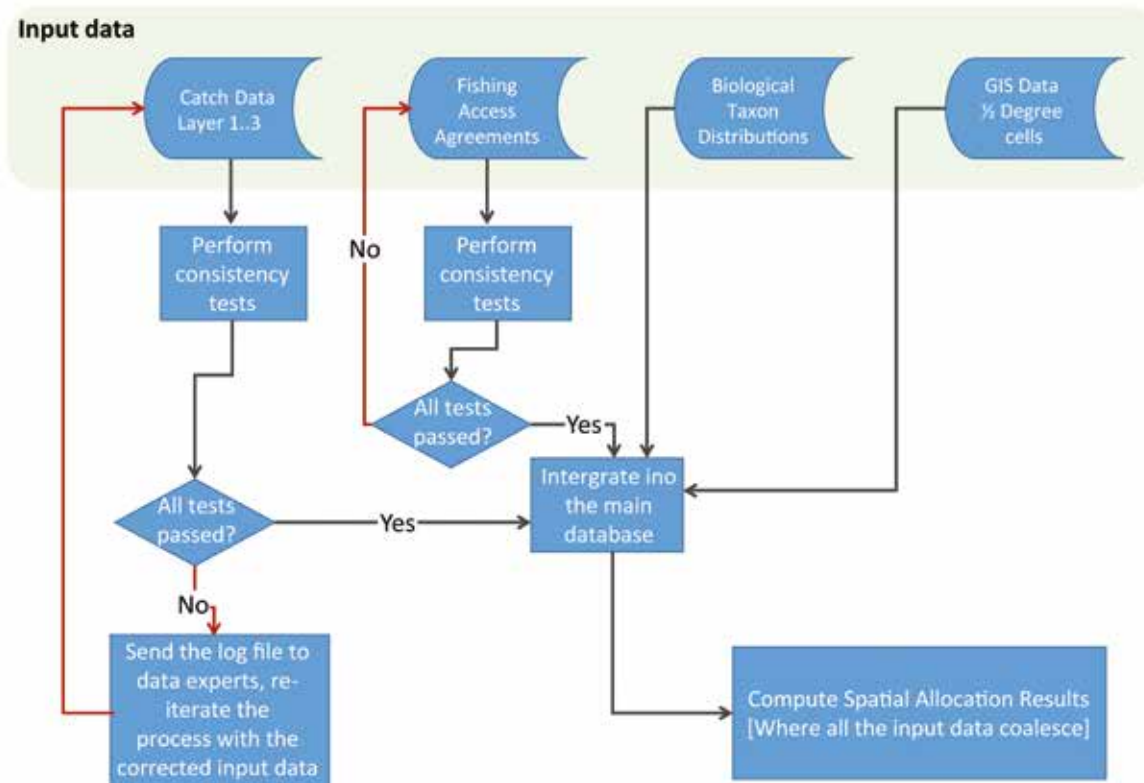


Figure 5.2. Spatial allocation procedure for catch reconstruction data of the *Sea Around Us*, resulting in the final half-degree allocated cell data. Details for the biological taxon distributions are provided in chapter 4.

mative taxa such as “miscellaneous marine fishes” (FAO term: “marine fishes nei”) by relying on taxonomic information in neighboring half-degree cells. This further added to the complexity of the allocation procedure and increased the difficulty of tracing actual country, taxon, and catch entities through the process. This step was also discontinued in the new allocation approach. Instead, our new allocation procedure disaggregates the input catch data as part of the country-by-country catch reconstruction process (chapter 2), which therefore transparently documents the taxonomic changes 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 chapter 2).

These preallocation 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 5.2), which allows a better software flow while maintaining the conceptual differences in data layers. This starts with an overview of the new allocation process (figure 5.2), followed by a description of the unique features of each data layer and how it is handled, and ends with an overview of the general algorithm of the spatial allocation.

The spatial allocation of the catch is the process of computing the catch that can be allocated to each half-degree cell based on the overlap of three main components: the catch data, the fishing access observations and agreements, and the biological taxon distributions (figure 5.2; table 5.2). The overlap of these components is facilitated by a series of geographic information system (GIS) layers, which essentially bind them together.

Table 5.2. Parameters of the 3 spatial catch data input layers as used in the catch reconstruction database and in the spatial allocation to half-degree cells of the *Sea Around Us*.

| Data layer | 1 | 2 | 3 |
|---------------------|--|--|---|
| Taxa included | All except industrial large pelagics | All except industrial 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 ^a | EEZ, high seas, ICES, CCAMLR, NAFO, FAO, and other areas | Six types of tuna cells: 1×1 , 5×5 , 5×10 , 10×10 , 10×20 , 20×20 degrees |

^aInshore 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 occur only along inhabited coastlines.

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 EEZs. The allocation algorithm allocates the small-scale catches (i.e., artisanal, subsistence, and recreational) only to the cells associated with the IFA (Chuenpagdee et al. 2006) of that fishing entity's EEZs, while industrial catches can be allocated anywhere within that fishing entity's EEZs, as long as they remain compatible with the biological taxon distributions. Fishing access agreements are not applicable to this data layer, because 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).

In layer 2, the spatial granularity of the input 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's (fishing country's) own EEZ waters (which

are treated in layer 1; table 5.2). In layer 2, the fishing access observations and agreements are used to delineate the areas that allow fishing for a particular fishing entity, year, and taxon. 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 covers only industrial large pelagic fishes and their associated bycatch and discards, the input catch data are spatially organized by larger tuna cells, which range from 1×1 to 20×20 degrees (table 5.2; see also chapter 3). Similar to the region-specific areas in layer 2, these larger cells are intersected with all the EEZ boundaries to create a GIS layer that is suitable for use in the algorithm. Thereafter, the fishing access observations and agreements and taxon distributions are applied to calculate the final layer 3 catch allocation.

An Overview of the Algorithm

The spatial allocation algorithm consists of four main steps:

- 1 Validating and importing the fishing access observations and agreements database
- 2 Validating and importing the catch reconstruction database
- 3 Importing the biological taxon distributions
- 4 Computing the catch that can be allocated to each half-degree cell for each catch data layer in an iterative process

(allowing verifications and corrections to any of the input parameters)

1. Validating and Importing the Fishing Access Observations and Agreements Database

The fishing access observations and agreements are first verified using several consistency and matching tests (figure 5.2), and upon passing they are imported into the main allocation database. This fishing access information is subsequently used in two different processes: the verification process of the catch data (layers 1, 2, and 3) and the computing of the areas where a given fishing entity (i.e., country) is allowed 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 about twenty-two different preallocation data tests (figure 5.2), designed to make sure that the data are coherent from the standpoint of database logic. These tests range from simple tests such as “is the TaxonKey valid?” to more complex tests such as “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 5.2). This process is repeated until all the data rows pass all the preallocation tests.

The process of importing the catch reconstruction database includes an important submodule for harmonizing the marine areas. This module is crucial, because the catch data come in a variety of different spatial reporting

areas that are not globally homogeneous in GIS definitions (e.g., the EEZ of Albania is one entity, whereas the EEZs of India, Brazil, and the United States are subdivided into states or provinces; the northeast Atlantic uses ICES statistical areas). To harmonize these marine areas and make them accessible to the core allocation process, any given half-degree cell is split into its constituent countries’ EEZs and high seas components. Then, the fishing access observations and agreements are applied to this layer to determine which of these “shards” of half-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. At present we have assigned more than 12,000 marine areas into their constituent “shards” of half-degree cells; these marine areas range 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 more than 2,500 individual distributions (see chapter 4) are generated as individual text files (comma-separated values [csv] format) containing for each half-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 and Allocating the Catch to Half-Degree Cells

Once steps 1, 2, and 3 are completed, we perform the computations that yield the final spatial half-degree allocation results. The catch of a given data row, TotalCatch, of taxon T is distributed among eligible half-degree cells, Cell 1 . . . n, using the following weighted average formula:

$$\text{Cell}_{i \text{ Allocated Catch}} = \text{Total Catch} \times \frac{\text{Cell}_{i \text{ Surface Area}} \times \text{Cell}_{i \text{ Relative Abundance of Taxon T}}}{\sum_1^n \text{Cell}_{i \text{ Surface Area}} \times \text{Cell}_{i \text{ Relative Abundance of Taxon T}}}$$

Throughout the allocation process, catch reconstruction parameters in addition to year and taxon, such as fishing sector, catch type, data layer, and reporting status, are preserved and carried over into the final half-degree allocated database.

Final Output

The final results of the intense and detailed database preparation and spatial allocation are time series of catches by half-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), as documented in the summaries, p. 185 onward.

CONCLUSIONS

The previous global catch data spatialization of the *Sea Around Us* was akin to hacking a first path through the jungle of official catch data: One could go from place to place (i.e., EEZs) but with the dark suspicion that behind the data foliage, thousands of tonnes of misallocated data monsters lay hidden, ready to

pounce. And pounce they did (i.e., colleagues complained of gross errors), which is why the new allocation system was developed.

This new system is more akin to a French garden, partitioned by clear lines (the EEZ boundaries), defining spaces into which known quantities (the reconstructed catches) are planted, with only the occasional poodle crossing the lines (i.e., errors will be small, easy to isolate and to correct). We invite and look forward to colleagues' assistance in tending this garden.

ACKNOWLEDGMENTS

We thank Reg Watson, whose pioneering work on spatializing the FAO catch database for the years 1950 to 2006 cleared the way for the reconceptualization presented here. We also thank Danielle Knip for initiating the reconstructed catch database. This is a contribution of the *Sea Around Us*, a research activity at the University of British Columbia initiated and funded by the Pew Charitable Trusts from 1999 to 2014 and currently funded mainly by the Paul G. Allen Family Foundation.

REFERENCES

- Belhabib, D., V. Koutob, A. Sall, V. W. Y. Lam, and D. Pauly. 2014. Fisheries catch misreporting and its implications: the case of Senegal. *Fisheries Research* 151: 1–11.
- Chuenpagdee, R., L. Liguori, M. L. D. Palomares, and D. Pauly. 2006. Bottom-up, global estimates of small-scale marine fisheries catches. *Fisheries Centre Research Reports* 14(8), University of British Columbia, Vancouver.
- Close, C., W. W. L. Cheung, S. Hodgson, V. Lam, R. Watson, and D. Pauly. 2006. Distribution ranges of commercial fishes and invertebrates. Pp. 27–37 in M. L. D. Palomares, K. I. Stergiou, and D. Pauly (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, February 15–19, 1999. COFI/99/INF.9E.
- Watson, R., A. Kitchingman, A. Gelchu, and D. Pauly. 2004. Mapping global fisheries: sharpening our focus. *Fish and Fisheries* 5: 168–177.
- Watson, R., and T. Morato. 2013. Fishing down the deep: accounting for within-species changes in depth of fishing. *Fisheries Research* 140: 63–65.

NOTE

1. Cite as Lam, V. W. Y., A. Tavakolie, M. L. D. Palomares, D. Pauly, and D. Zeller. 2016. The *Sea Around Us* catch database and its spatial expression. Pp. 59–67 in D. Pauly and D. Zeller (eds.), *Global Atlas of Marine Fisheries: A Critical Appraisal of Catches and Ecosystem Impacts*. Island Press, Washington, DC.