Taxon distributions[5]

Maria-Lourdes.D. Palomares[a], William Cheung[b], Vicky Lam[a] and D. Pauly[a]

[a)] *Sea Around Us*, University of British Columbia, Vancouver, Canada.
[b)] Changing Ocean Unit, University of British Columbia, Vancouver, Canada.

Ecosystem-based fisheries management (EBFM, Pikitch *et al.* 2004) must include a sense of place, where fisheries interact with the animals of specific ecosystems. To be useful to researchers, managers and policy makers attempting to implement EBFM schemes, the *Sea Around Us* presents biodiversity and fisheries data in spatial form onto a grid of about 180,000 half degree latitude and longitude cells which can be regrouped into larger entities, e.g., the Exclusive Economic Zones (EEZs) of maritime countries, or the system of currently 66 Large Marine Ecosystems (LME) initiated by NOAA (Sherman *et al.* 2007), and now used by practitioners throughout the world.

However, not all the marine biodiversity of the world can be mapped in this manner; thus, while FishBase (www.fishbase.org) includes all marine fishes described so far (more than 15,000 spp.), so little is known about the distribution of the majority of these species that they cannot be mapped in their entirety. The situation is even worse for marine invertebrates, despite huge efforts (see www.sealifebase.org).

Scientific and common names

Before taxon distributions can be generated, the taxonomic 'validity' of a name needs to be verified, and all names standardized across all data sources being used. The names provided for the taxa included in the *Sea Around Us* catch data originate either from FAO or from other source material used by catch reconstructions (See Section #1 and Section #2), but were verified using FishBase for fishes and SeaLifeBase for non-fish taxa.

Common names, which is what most people know about most organisms, are provided in English, and increasingly also in other languages. FishBase provides common names in other languages for fish, covering nearly 200,000 different names in over 200 languages. FishBase also provides a rationale for the use of common names, and the way the names it contains were assembled.

Scientific names differ in various features, depending on whether they pertain to *species*, *genera*, *families*, *orders*, or broader taxonomic groups.

Species names always consist of two parts, a unique genus name (whose first letter is always capitalized) and a species epithet (whose first letter is never capitalized). Both components of the names should be written in italics whenever possible, i.e., *Gadus morhua* being the scientific species name for the Atlantic cod.

---

[5] Adapted from: Palomares MLD, Cheung WWL, Lam VWY and Pauly D 2016. The distribution of exploited marine biodiversity, *In:* Pauly D and Zeller D (eds.) *Global Atlas of Marine Fisheries: Ecosystem Impacts and Analysis*. Island Press, Washington, D.C.

The name of a ***genus*** (plural = *genera*) must be unique (i.e., there is no other such name in the entire animal kingdom) and its first letter is always capitalized. A genus can include one or several species, i.e. *Chanos* sp., or *Stolephorus* spp.. For more rules regarding the naming of species and genera, see www.fishbase.de/manual/fishbasespecies_of_fishes.htm

Families consist of one, or more commonly, several *genera*. Family names among animals always end in -*idae*, e.g. Gadidae (cods). Family names are not italicized, but always capitalized. Sometimes, 'common' names are derived from the scientific names of families, e.g. 'loliginids' for squids of the Family Loliginidae, but this usually leads to names that are little used, even when the family was based on a generic name, itself based on a (Latin) common name, e.g., '*Loligo*'. We have kept such names, however, if they occurred in the FAO catch database, in order to maintain as much compatibility as possible.

Orders consist of one or more families, and their names, in animals, end in -*formes*. Orders are not italicized but always capitalized. Thus, for example the Gadiformes include the families Gadidae (cods), Merluccidae (hakes), and others, all more closely related to each other than to, e.g., the herrings, sardines, etc. (the Clupeiformes).

The *Sea Around Us* data also include broader, but taxonomically ill-defined groups (e.g., 'miscellaneous marine fishes', also called 'marine fishes nei'[6] in FAO parlance), usually the result of suboptimal systems having been set up by various countries for collecting and reporting fisheries catch data. The *Sea Around Us* strives to disaggregate such data during the reconstruction process, i.e., to allocate them to the appropriate lower taxonomic levels, and we anticipate that the number of broad categories in the database, and especially the amount of catch they represent, will gradually decline.

Groups we report on besides 'taxa'

Because there are more than 2,000 species and other groups included in our global fisheries catches, we have decided to provide taxon specific data on our website for only a user-definable subset of the total number of individual taxa (plus a 'Others' group containing all other taxonomic entities combined), but we also provide data using two other types of aggregated groups for all catch.

The first is a general grouping of the catch by 12 broad groups that we call '**commercial groups**'. These are anchovies, herring-like fishes, perch-like fishes, tuna and billfishes, cod-like fishes, salmons and smelts, flatfishes, scorpion fishes, sharks and rays, crustaceans, mollusks, and 'other fishes and invertebrates'.

The other grouping is based partly on taxonomy, but mostly on habitat preferences, feeding habits, and maximum size, which define what we call '**functional groups**' as required for ecosystem modeling (e.g., Ecopath with Ecosim, Christensen *et al.* 2009). This grouping separates fish by where they live in the water column. *Demersal* animals that live on or are closely associated with the sea bottom are separated from those that live predominately in the

---

[6] 'nei' stands for 'not elsewhere included'.

water column or near the water surface (e.g., *pelagic*). *Benthopelagic* taxa refer to those that live and feed near the bottom as well as in mid-water or near the surface. Habitat separation is further described by depth zones, with *bathypelagic* and *bathydemersal* taxa referring to taxa living in the 1000-4000 m depth zone. Finally, we have separated out *reef associated* taxa as well as *sharks* and *rays*, *flatfishes*, and a few other individual groups. Most of these functional groups are further separated into those that are under 30 cm when at maximum length (e.g., small herring species), those 30 to 90 cm, and those over 90 cm (such as tunas), except for sharks, rays and flatfishes, which are grouped into two categories (small and medium versus large). Overall, we have defined 30 functional groups (Table 1). This grouping system, besides facilitating ecological studies, is useful for studying the impacts of fishing gears, as different functional groups tend to be impacted and targeted by various fishing gears differently.

Table 1. Functional groups as defined by the *Sea Around Us* for catch reporting and ecosystem modeling.

Small Pelagics (<30 cm)
Medium Pelagics (30 - 90 cm)
Large Pelagics (>=90 cm)
Small Demersals (<30 cm)
Medium Demersals (30 - 90 cm)
Large Demersals (>=90 cm)
Small Bathypelagics (<30 cm)
Medium Bathypelagics (30 - 90 cm)
Large Bathypelagics (>=90 cm)
Small Bathydemersals (<30 cm)
Medium Bathydemersals (30 - 90 cm)
Large Bathydemersals (>=90 cm)
Small Benthopelagics (<30 cm)
Medium Benthopelagics (30 - 90 cm)
Large Benthopelagics (>=90 cm)
Small Reef associated fish (<30 cm)
Medium Reef associated fish (30 - 90 cm)
Large Reef associated fish (>=90 cm)
Small to Medium Sharks (<90 cm)
Large Sharks (>=90 cm)
Small to Medium Rays (<90 cm)
Large Rays (>=90 cm)
Small to Medium Flatfishes (<90 cm)
Large Flatfishes (>=90 cm)
Cephalopods
Shrimps
Lobsters, crabs
Jellyfish
Other demersal invertebrates
Krill
Other taxa

Mapping distributions

We define as ‘commercial’ all marine fish or invertebrate species that are either reported in the catch statistics of at least one of the member countries of the Food and Agriculture Organization of the United Nations (FAO), or are listed as part of commercial and non-commercial catches (retained as well as discarded) in country-specific catch reconstructions (see Section #1 and Section #2). For most species occurring in the landings statistics of FAO, there were enough data in FishBase for at least tentatively mapping their distribution ranges. Similarly, most species of commercial invertebrates had enough information in SeaLifeBase for their approximate distribution range to be mapped. We discuss below the procedure we use for taxa that lacked sufficient data for mapping their distribution, which included only few taxa in the FAO statistics, but many from reconstructed catches, including discards.

In the following, we document how such mapping is done. Thus, this contribution presents the methods (improved from Close *et al.* 2006) by which all commercial species distribution ranges (totaling over 1,500 for the 1950-2010 time period) were constructed and/or updated, and consisting of a set of rigorously applied 'filters' that will markedly improve the accuracy of the *Sea Around Us* maps and other products.

The 'filters' used here are listed in the order that they are applied. Prior to the 'filter' approach presented below, the identity and nomenclature of each species is verified using FishBase or SeaLifeBase, the two authoritative online encyclopedia covering the fishes of the world and marine non-fish animals, respectively, and their scientific and English common names corrected if necessary. This information is then standardized throughout all *Sea Around Us* databases (see Section #4). Following the creation of all species-level distributions as described here, taxon distributions for higher taxonomic grouping, such as genus, family etc. are generated by combining each taxon-level's contributing components, e.g., for the genus *Gadus*, all distributions of species within this genus are combined.

Note that the procedures presented here avoid the use of temperature and primary productivity to define or refine distribution ranges for any species, even though these factors strongly shape the distribution of marine fishes and invertebrates (Ekman 1967; Longhurst and Pauly 1987). This was done in order to allow for subsequent analyses of distribution ranges to be legitimately performed using these variables, i.e., to avoid circularity.

Filter 1: FAO Areas

The FAO has divided the world's oceans into 19 statistical areas for reporting purposes (see Section #1). Information on the occurrence of commercial species within these areas is available primarily through (a) FAO publications and the FAO website (www.fao.org); and (b) FishBase and SeaLifeBase. Figures 1A and 2A illustrate the occurrence by FAO area of Florida pompano (*Trachinotus carolinus*) and silver hake (*Merluccius bilinearis*), i.e., examples representing pelagic and demersal species, respectively.

Filter 2: Latitudinal range

The second filter applied in this process is latitudinal ranges. The latitudinal range of a species is defined as the space between its northernmost and southernmost latitudes. This range can be found in FishBase for most fishes and in SeaLifeBase for many invertebrates. For fishes and invertebrates for which this information was lacking, latitudes were inferred from the latitudinal range of the EEZs of countries where they are reported to occur as endemic or native species, and/or from occurrence records in the Ocean Biogeographic Information System website (OBIS; www.iobis.org). Note, however, that recent occurrence records (from the 1980s onwards and known range extensions, e.g., of Lessepsian species) were not used to determine 'normal' latitudinal ranges, as they tend to be affected by global warming (Cheung *et al.* 2009).

A species will not have the same probability of occurrence, or relative abundance throughout its latitudinal range; it can be assumed to be most abundant at the center of its range (McCall 1990). Defining the center of the latitudinal distribution range is done using the following assumptions:

a) For distributions confined to one hemisphere, a symmetrical triangular probability distribution is applied, which estimates the center of the latitudinal range as the average of the range, i.e., [northernmost + southernmost latitude] / 2;

b) For distributions straddling the equator, the range is broken into three parts – the outer two thirds and the inner or middle third. If the equator falls within one of the outer thirds of the latitudinal range, then abundance is assumed to be the same as in (a). If, however, the equator falls in the middle third of the range, then abundance is assumed to be flat in the middle third and decreasing to the poles for the remainder of the range.

Figures 1B and 2B illustrate the result of the FAO and latitudinal filters combined. Both the Florida pompano and the silver hake follow symmetrical triangular distributions as mentioned in (a) above.

Filter 3: Range-limiting polygon

Range-limiting polygons help confine species in areas where they are known to occur, while preventing their occurrence in other areas where they could occur (because of environmental conditions), but do not. Distribution polygons for a vast number of species of commercial fish and invertebrates can be found in various publications, notably FAO's species catalogues, species identification sheets, guides to the commercial species of various countries or regions, and in online resources, some of which were obtained from model predictions, e.g., Aquamaps (Kaschner *et al.* 2008; see also www.aquamaps.org). Such polygons are mostly based on observed species occurrences, which may or may not be representative of the actual distribution range of the species.

Occurrence records assume that the observer correctly identified the species being reported, which adds a level of uncertainty to the validity of distribution polygons. Most often than not, experts are required to review and validate a polygon before it is published, e.g., in an FAO species catalogues. This review process is also important, notably for polygons that are automatically generated via model predictions such as Aquamaps. Note that for commercially important endemic species, this review process can be skipped as the polygon is restricted to the only known habitat and country where such species occurs.

For species without published polygons, range maps are generated using the filter process described here and compared with the native distribution generated in Aquamaps. Differences between these two 'model-generated' maps are verified using data from the scientific literature and OBIS/GBIF (i.e., reported occurrences, notably from scientific surveys). Note that FAO statistics, in which countries report a given species in their catch, can be used as occurrence records, the only exception being if the species was caught by the country's distant-water fleet.

Polygons are drawn based on the verified map (i.e., with unverified occurrences deleted). Additionally, faunistic work covering the high-latitude end of continents and/or semi-enclosed coastal seas with depauperate faunas (e.g., Hudson Bay, or the Baltic Sea) were used to avoid, where appropriate, distributions reaching into these extreme habitats. The results of this step, i.e., the information gathered from the verification of occurrences, are also provided to FishBase and SeaLifeBase to fill data gaps.

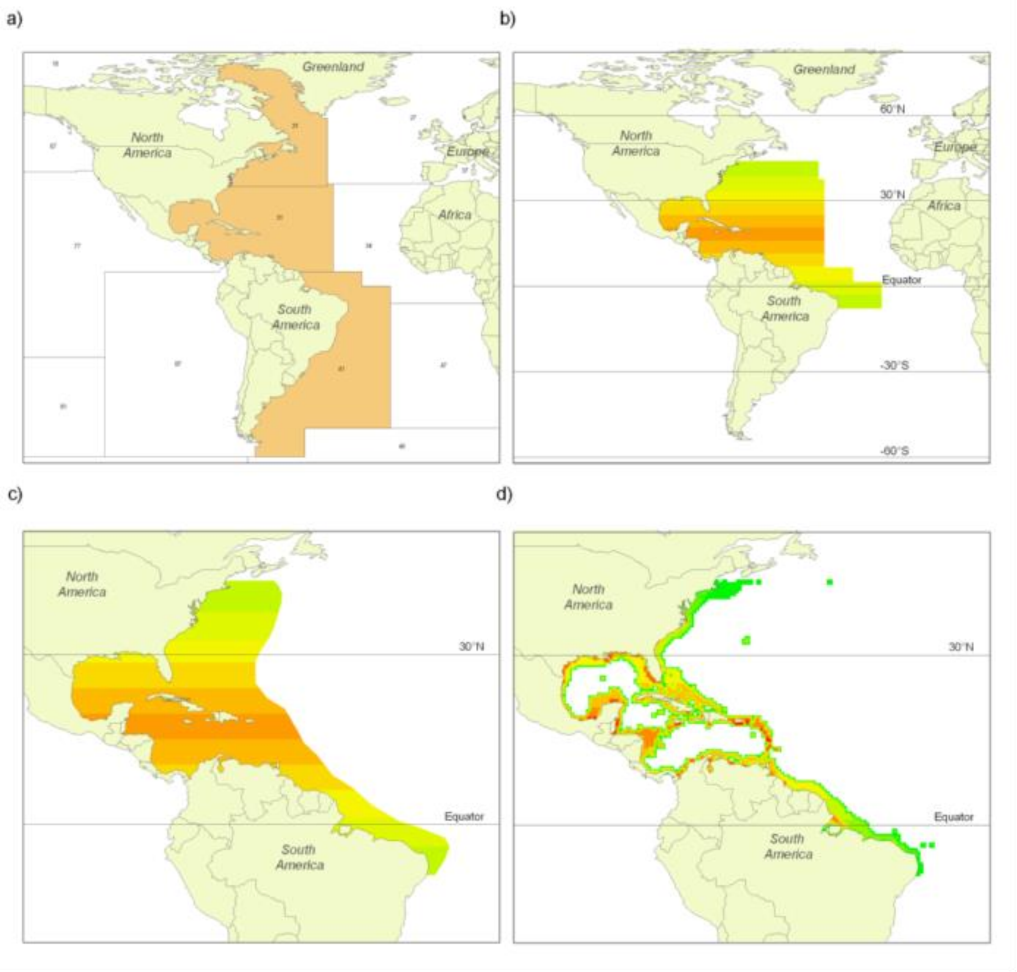

Figure 1. Partial results obtained following the application of the filters used for deriving a species distribution range map for the Florida pompano (*Trachinotus carolinus*): (**A**) illustrates the Florida pompano's presence in FAO areas 21, 31 and 41; (**B**) illustrates the result of overlaying the latitudinal range (43°N to 9°S; see Smith 1997) over the map in A; (**C**) shows the result of overlaying the (expert-reviewed) range-limiting polygon over B; and (**D**) illustrates the relative abundance of the Florida pompano resulting from the application of the depth range, habitat preference and equatorial submergence filters on the map in C.

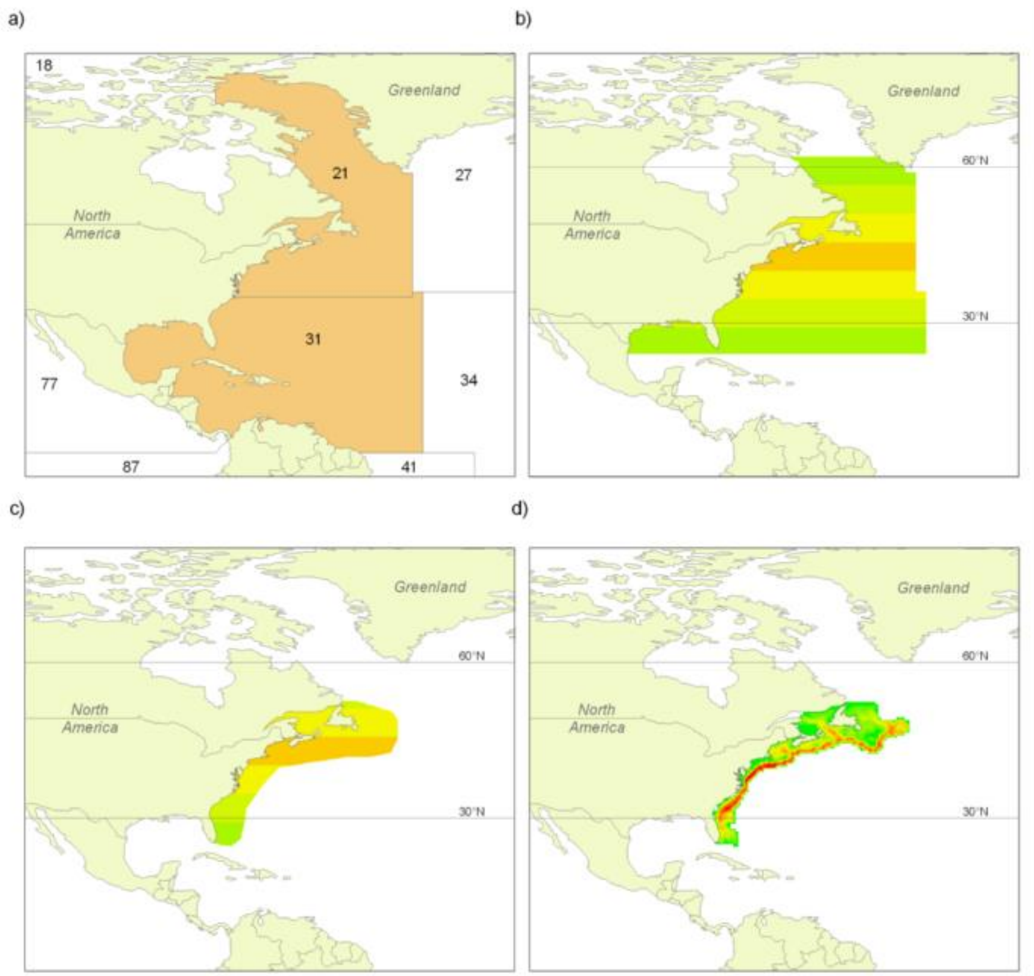

Figure 2. Partial results obtained following the application of the filters used for deriving a species distribution range map for the silver hake (*Merluccius bilinearis*): (**A**) illustrates the silver hake's presence in FAO areas 21 and 31; (**B**) illustrates the result of applying the FAO and latitudinal range (55°N to 24°N; see FAO-FIGIS 2001); (**C**) shows the result of overlaying the (expert-reviewed) range-limiting polygon over B; and (**D**) illustrates the silver hake's relative abundance resulting from the application of the depth range, habitat preference and equatorial submergence filters on the map in C.

All polygons, whether available from a publication or newly drawn, were digitized with ESRI's ArcGIS, and were later used for inferences on equatorial submergence (see below). Figures 1C and 2C illustrate the result of the combination of the first three filters, i.e., FAO, latitude and range-limiting polygons. These parameters and polygons will be revised periodically, as our knowledge of the species in question increases.

Note that because this mapping process only deals with commercially-caught species, the distribution ranges for higher level taxa (genera, families, etc.) were usually generated using the combination of range polygons from the taxa included in the higher-level taxon. Thus, the range polygons for genera were built using the range polygons of the commercial species that belong to the genus in question. Similarly, family-level polygons were generated from genus-level polygons, and so on. Latitude ranges, depth ranges and habitat preferences were expanded in the same manner. While this procedure will not produce the true distribution of the genera and families in question, which usually consists of more species than are reported in catch statistics, it is likely that the generic names in the catch statistics refer to the very commercial species that are used to generate the distribution ranges, as these taxa are frequently more abundant than the ones that are not reported in official catch statistics.

Filter 4: Depth range

Similar to the latitudinal range, the 'depth range', i.e., "[the] depth (in m) reported for juveniles and adults (but not larvae) from the most shallow to the deepest [waters]", is available from FishBase for most fish species and SeaLifeBase for many commercial invertebrates, along with their common depth, defined as the "[the] depth range (in m) where juveniles and adults are most often found. This range may be calculated as the depth range within which approximately 95% of the species biomass occurs" (Froese *et al.* 2000). Given this, and based on Alverson *et al.* (1964), Pauly and Chua (1988), and Zeller and Pauly (2001), among others, the abundance of a species within the water column is assumed to follow a scalene triangular distribution, where maximum abundance occurs at the top one-third of its depth range.

Filter 5: Habitat preference

Habitat preference is an important factor affecting the distribution of marine species. Thus, the aim of this filter is to enhance the prediction of the probability that a species occurs in an area, based on its association with different habitats. Two assumptions are made here:

a) That, other things being equal, the relative abundance of a species in a spatial ½ degree cell is determined by a fraction derived from the number of habitats that a species associates with in that same cell, and by how far the association effect will extend from that habitat; and
b) That the extent of this association is assumed to be a function of a species' maximum size (maximum length) and habitat 'versatility'. Thus, a large species that inhabits a wide range of habitats is more likely to occur far from the habitat(s) with which it is associated, while smaller species tend to have low habitat versatility (Kramer and Chapman 1999).

The maximum length and versatility of a species are classified into three categories, and it is assumed that a species can associate with one or more categories with different degrees of membership (0 to 1). A higher membership value means a higher 'probability' that the species is associated with that particularly category. The membership values are defined by a pre-specified membership function for each of the length and versatility categories (Figure 3). For example, the striped bass (*Morone saxatilis*) has a maximum length of 200 cm (total length). Based on the pre-defined membership function presented in Figure 3A, the striped bass has a large body size with a membership of 1. Note that there are maximum length estimates for all the exploited species used by the *Sea Around Us*, derived from FishBase and SeaLifeBase.

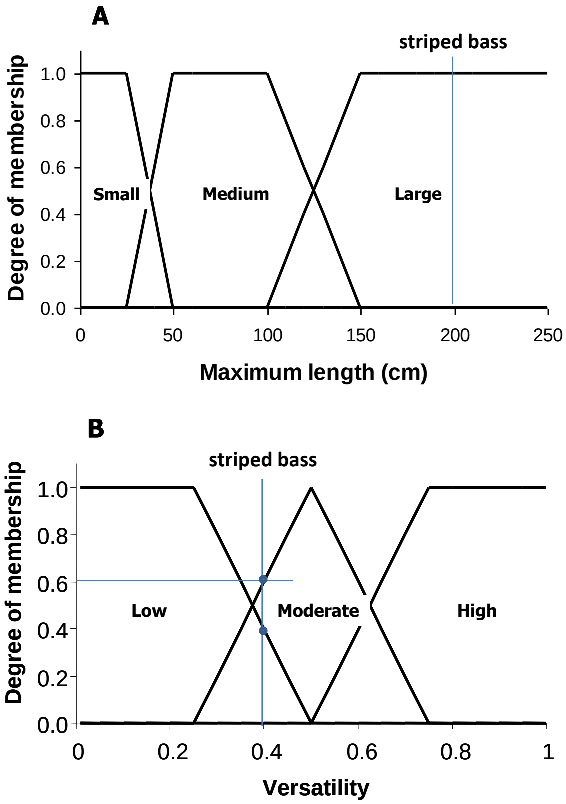

Figure 3. Fuzzy membership functions for the three categories of (**A**) maximum length and (**B**) habitat versatility of a species. Habitat versatility is defined as the ratio of the number of habitat types with which a species is associated to the total number of defined habitat types in Table 1. For example, the striped bass (*Morone saxatilis*) grows to a maximum total length of 200 cm (large body size; degree of membership = 1). It occurs in estuaries and ‘other habitats’ (2 of 5 defined habitats, i.e., versatility = 0.4, low to moderate degree of membership = 0.4-0.6).

The ability of a species to inhabit different habitat types, here referred to as ‘versatility’, is defined as the ratio between the number of habitats with which a species is associated to the total number of habitats as defined in Table 2. These habitats are categorized as ‘biophysical’ (i.e., coral reef, estuary, sea grass, seamount, other habitats), ‘depth-related’ (shelf/slope/abyssal), and ‘distance from coast’ (inshore/offshore). As species are generally specialized towards ‘biophysical’ habitats, this filter only takes those five habitats into consideration. Taking our example again, FishBase lists the following for the striped bass: “Inhabit coastal waters and are commonly found in bays but may enter rivers in the spring to spawn” (Eschmeyer *et al.* 1983). This associates the striped bass with estuaries and ‘other habitats’ (i.e., when it enters rivers to spawn). Given that the total number of defined biophysical habitats is five, and the striped bass is associated with two of those, then the versatility of striped bass is estimated to be 0.4 (i.e., 2/5). Finally, based on the defined membership functions shown in Figure 3B, the versatility of striped bass is classified as ‘low’ to ‘moderate’, with a membership of approximately 0.4 and 0.6, respectively.

Table 2. Habitat categories used here, and for which global maps are available in the *Sea Around Us*, with some of the terms typically associated with them (in FishBase, SeaLifeBase and other sources).

Categories	Specifications of global map	Terms often used
Estuary	Alder (2003)	Estuaries, mangroves, river mouth
Coral	UNEP-WCMC (2010)	Coral reef, coral, atoll, reef slope
Sea grass	Not yet available*	Sea grass bed
Seamounts	Kitchingman and Lai (2004)	Seamounts
Other habitats	–	Muddy/sandy/rocky bottom
Continental shelf	NOAA (2004)	Continental shelf, shelf
Continental slope	NOAA (2004)	Continental slope, upper/lower slope
Abyssal	NOAA (2004)	Away from shelf and slope
Inshore	NOAA (2004)	Shore, inshore, coastal, along shoreline
Offshore	NOAA (2004)	Offshore, oceanic

* The *Sea Around Us* is developing a global map of sea grass, which will be applied when available.

Determining habitat association

Qualitative descriptions relating the commonness (or preference) of a species to particular habitats (as defined in Table 1) are given weighting factors as enumerated in Table 3. Such descriptions are available from FishBase for most fishes and in SeaLifeBase for most commercially important invertebrates. Going back to our example, we thus know that the striped bass occurs in (and thus prefers) brackish water (i.e., estuaries), but enters freshwater (i.e., 'other habitats') to spawn. Given the weighting system in Table 3, estuaries is assigned a weight of 0.75 (usually occurs in) and 'other habitats' is given a weight of 0.5 (assuming a seasonal spawning period).

Table 3. Common descriptions of relative abundance of species in habitats where they occur and their assigned weighting factors. The weighting factor for ‘other habitats’ is assumed to be 0.1 when no further information is available.

Description	Weighting factor
Absent/rare	0.00
Occasionally, sometimes	0.25
Often, regularly, seasonally*	0.50
Usually, abundant in, prefer	0.75
Always, mostly, only occurs	1.00

* If a species occurs in a habitat, but no indication of relative abundance is available, a default score of 0.5 is assumed.

Maximum distance of habitat effect

Maximum distance of habitat effect (maximum effective distance) refers to the maximum distance from the nearest perimeter of the habitat which ‘attracts’ a species to a particular habitat. This is defined by the maximum length and habitat versatility of the species using the heuristic rule matrix in Table 4. Taking our example for the striped bass, with a ‘large’ maximum length (membership=1) and ‘low’ to ‘moderate’ versatility (membership values of 0.4 and 0.6), points to a ‘farthest’ maximum effective distance in Table 4. The degree of membership assigned to maximum effective distance is equal to the minimum membership value of the two predicates[7], in this example, 1 vs. 0.4 = 0.4 and 1 vs. 0.6 = 0.6. When the same conclusion is reached from different rules, the final degree of membership equals the average membership value (in this example, (0.4+0.6)/2=0.50).

The maximum effective distance from the associated habitat can be estimated from the 'centroid value' of each conclusion category, weighted by the degree of membership. The centroid values for ‘near’, ‘far’ and ‘farthest’ maximum effective distances were defined as 1 km, 50 km and 100 km, respectively. In our example, we obtained membership values of 0.4 for near (1 km) and 0.6 for farthest (100 km) maximum effective distance, respectively. This gives an estimate of (0.4*1 + 0*50 + 0.6*100)/(0.4 + 0 + 0.6) = 60.4 km (see Figure 4).

Table 4. Heuristic rules that define the maximum effective distance from the habitat in which a species occurs. The columns and rules in bold characters represent the predicates (categories of maximum body size and versatility), while those in italics represent the resulting categories of maximum effective distance.

Versatility	**Maximum body size**		
	Small	**Medium**	**Large**
Low	*Near*	*Near*	*Near*
Moderate	*Far*	*Far*	*Farthest*
High	*Far*	*Farthest*	*Farthest*

Estimating relative abundance in a spatial cell

Several assumptions are made to simplify the computations. First, it is assumed that the habitat always occurs in the center of a cell and is circular in shape. Second, species density (per unit area) is assumed to be the same across any habitat type; and that density declines linearly from the habitat perimeter to its maximum effective distance. Given these assumptions, the total relative abundance of a species in a cell equals the sum of abundance on and around its associated habitat, expressed as:

$$B'T = (\alpha j + \alpha j{+}1 \cdot (1 - \alpha j)) \cdot (1 - A) \qquad \ldots\ 4.1)$$

where $B'T$ is the final abundance, αj is the density away from the habitat from cell j, and A is the habitat area of the cell. The relative abundance resulting from the different habitat types is the sum of relative abundance, and is weighted by their importance to the species.

Although these assumptions on the relationship between maximum length, habitat versatility and maximum distance from the habitat may render uncertain predicted distributions at a fine spatial scale, this routine provides an explicit and consistent way to incorporate habitat considerations into distribution ranges.

---

[7] Predicate logic: a generic term for systems of abstract thought applied in fuzzy logic. In this example, the first-order logic predicate is “IF maximum weight is large”, and the second-order logic predicate is “AND versatility is moderate”. The resulting function, i.e., the conclusion category based on the predefined rules matrix in Table 3, is “THEN maximum effective distance is farthest”.

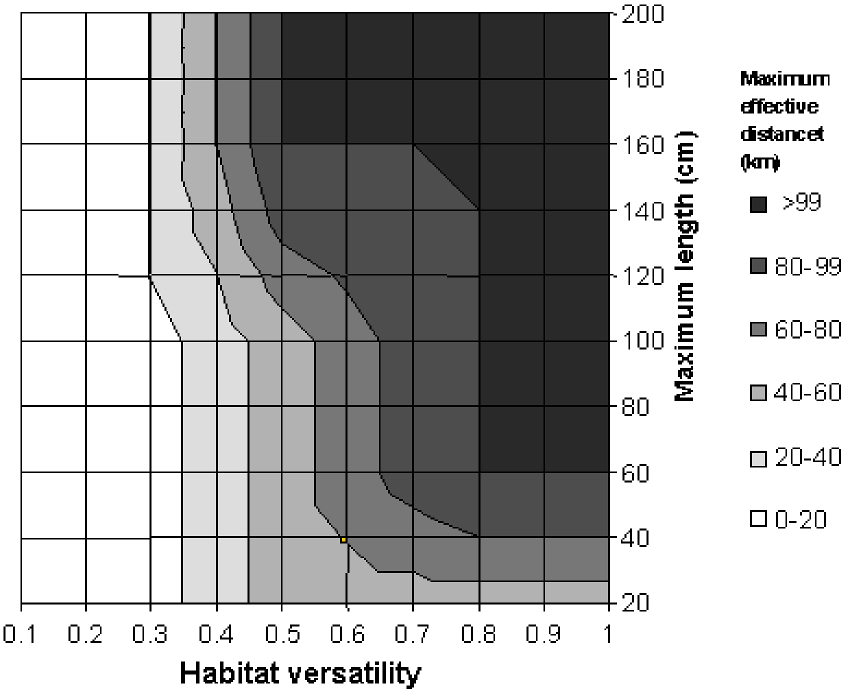

Figure 4. Maximum effective distance for striped bass (*Morone saxatilis*) estimated from the habitat versatility and maximum length of that species (see text).

Filter 6: Equatorial submergence

Eckman (1967) gives the current definition of equatorial submergence: "*animals which in higher latitudes live in shallow water seek in more southern regions archibenthal or live in shallow water seek in more southern regions archibenthal or purely abyssal waters […]. This is a very common phenomenon and has been observed by several earlier investigators. We call it submergence after V. Haecker [1906-1908] who, in his studies on pelagic radiolarian, drew attention to it. In most cases, including those which interest us here, submergence increases towards the lower latitudes and therefore may be called equatorial submergence. Submergence is simply a consequence of the animal's reaction to temperature. Cold-water animals must seek colder, deeper water layers in regions with warm surface water if they are to inhabit such regions at all.*" Equatorial submergence, indeed, is caused by the same physiological constraints which also determine the 'normal' latitudinal range of species, as described above, and it shifts due to global warming, i.e., respiratory constraints fish and aquatic invertebrates experience at temperatures higher than that which they have evolved to prefer (Pauly 1998, 2010).

Modifying the distribution ranges to account for equatorial submergence requires accounting for two constraints: (1) data scarcity; and (2) uneven distribution of environmental variables (temperature, light, food, etc.) with depth. FishBase and SeaLifeBase notwithstanding, there is little information on the depth distribution of most commercial species. However, in most cases, the following four data points are available for each species: the shallow end of the depth range

($D_{shallow}$), its deep end (D_{deep}) of the depth range, the poleward limit of the latitudinal range (L_{high}), and its lower latitude limit (L_{low}). If it is assumed that equatorial submergence is to occur, then it is logical to also assume that $D_{shallow}$ corresponds to L_{high}, and that D_{deep} corresponds to L_{low}.

Also, we further mitigate data scarcity by assuming the shape of the function linking latitude and equatorial submergence. Here, two parabolas (P) are used (Figure 5), one for the shallow limits of the depth distribution ($P_{shallow}$), and one for the deeper limits (P_{deep}), with the assumption that both $P_{shallow}$ and P_{deep} are symmetrical about the Equator. In addition, maximum depths are assumed not to change poleward of 60^0 N and 60^0 S. The uneven distribution of the temperature gradient can be mimicked by constraining $P_{shallow}$ to be less concave than P_{deep} by setting the geometric mean (D_{gm}) of $D_{shallow}$ and D_{deep} as the deepest depth that $P_{shallow}$ can attain. Three points draw the parabolas. In most cases, $P_{shallow}$ is obtained with $D_{60°N}=0$, $D_{60°S}=0$ and $D_{Lhigh}=D_{shallow}$, and P_{deep} with $D_{60°N}=D_{gm}$, $D_{60°S}= D_{gm}$ and $D_{Llow}=D_{max}$. If L_{high} is in the northern hemisphere and L_{low} is in the south, P_{deep} is drawn with D_{meep} at the Equator and conversely for the southern hemisphere. Finally, it is assumed that if a computed $P_{shallow}$ intercepts zero depth at latitudes higher than 60^0 N and/or lower than 60^0 S, then $P_{shallow}$ is recomputed with $D_{60°N}=D_{shallow}$, $D_{60°S}=D_{shallow}$ and $D_{Lhigh}=0$.

Figure 5 illustrates three cases of submergence based on different constraints. When this process is applied to a distribution based on latitudinal range and depth, but which did not account for submergence, these have the effect of 'shaving off' parts of the shallow-end of that distribution at low latitudes, and similarly, shaving off part of the deep-end end of the distribution at high latitudes. Also, besides leading to narrower and more realistic distribution ranges, this leads to narrowing the temperature ranges inhabited by the species in question, which is important for the estimation of their preferred temperature, as used when modelling global warming effects on marine biodiversity and fisheries.

The key outcome of the process described above consists of distribution ranges such as in Figure 6 for currently over 2,000 taxa, which can be viewed via the *Sea Around Us* website. They are also accessible via FishBase and SeaLifeBase (click '*Sea Around Us* distributions' under the 'Internet sources' section of the species summary pages). These distribution ranges serve as basis for all spatial catch allocation done by the *Sea Around Us* (Section #4), and we welcome feedback, i.e., suggested comments or corrections.

Predictions of distributions from the *Sea Around Us* algorithm are comparable in performance to other species modeling approaches that are commonly used for marine species (Jones *et al.* 2012). Specifically, AquaMaps (Kaschner *et al.* 2008), Maxent (Phillips *et al.* 2006) and the *Sea Around Us* algorithm are three approaches that have been applied to predict distributions of marine fishes and invertebrates. Jones *et al.* (2012) applied these three species distribution modelling methods to commercial fish in the North Sea and North Atlantic using data from FishBase and the Ocean Biogeographic Information System. Comparing test statistics of model predictions with occurrence records suggest that each modelling method produced plausible predictions of range maps for each species. However, the pattern of predicted relative habitat suitability can differ substantially between models (Jones *et al.* 2013). Incorporation of expert knowledge, as discussed above with reference to Filter 3, generally improves predictions, and therefore was given here particular attention.

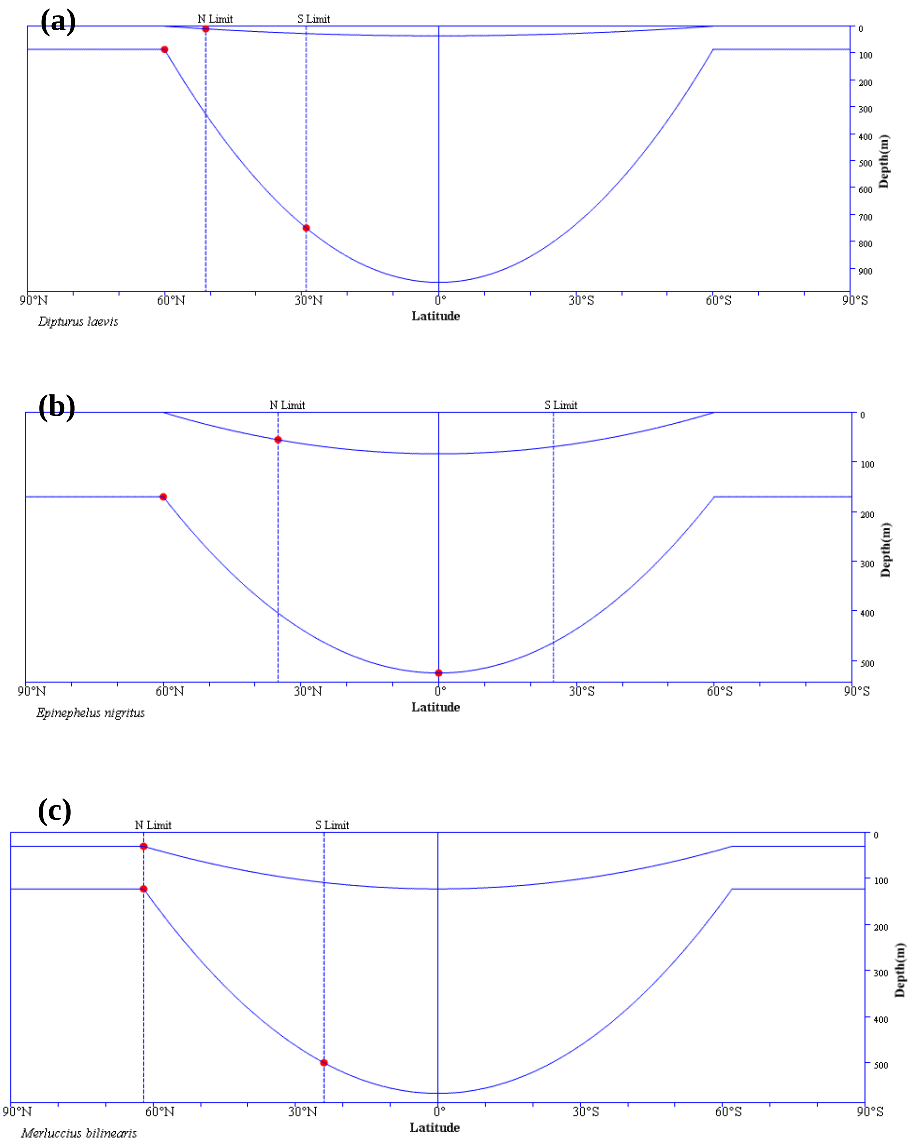

Figure 5. Shapes used to generate 'equatorial submergence', given different depth/latitude data: (A) Case 1: Barndoor skate (*Dipturus laevis*) – when the distribution range of the species is at lower latitudes than 60° N and/or S, the shallow parabola ($P_{shallow}$) is assumed to intercept zero at 60° N and S; (B) Case 2: When a distribution range is spanning the northern and southern hemispheres, as in the case of the Warsaw grouper (*Epinephelus nigritus*), the deepest depth of the deep parabola (P_{deep}) is at the Equator; (C) Case 3: Silver hake (*Merluccius bilinearis*), where the poleward limit of the latitudinal range (L_{high}) is at higher latitudes than 60° N and S.

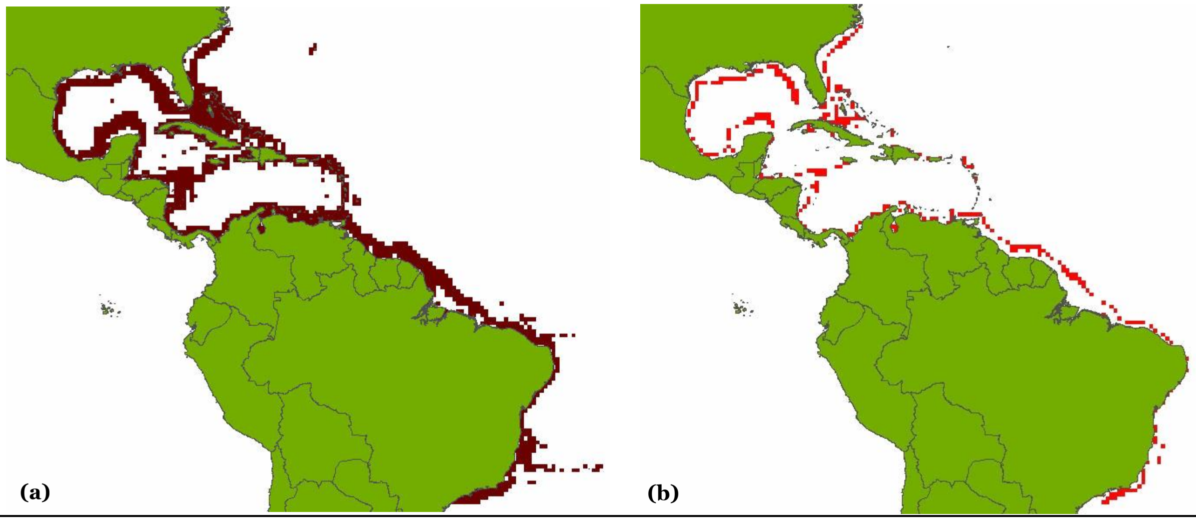

Figure 6. 'Equatorial submergence' has the effect of 'shaving off' areas from the distribution range of the Warsaw grouper, *Epinephelus nigritus*: **(A)** Original distribution; **(B)** Distribution adjusted for 'equatorial submergence'.

References

Alder J (2003) Putting the coast in the *Sea Around Us* Project. The Sea Around Us Newsletter, (15): 1-2.

Alverson DL, Pruter AL, Ronholt LL (1964) A Study of Demersal Fishes and Fisheries of the Northeastern Pacific Ocean. Institute of Fisheries, The University of British Columbia, Vancouver.

Cheung WWL, Lam VWY, Sarmiento JL, Kearney K, Watson R and Pauly D (2009) Projecting global marine biodiversity impacts under climate change scenarios. *Fish and Fisheries* 10: 235-251.

Christensen V, Walters C, Ahrens R, Alder J, Buszowski J, Christensen L, Cheung WL, Dunne J, Froese R, Karpouzi V, Kastner K, Kearney K, Lai S, Lam V, Palomares D, Peters-Mason A, Piroddi C, Sarmiento JL, Steenbeek J, Sumaila RU, Watson R, Zeller D and Pauly D (2009) Database-driven models of the world's large marine ecosystems. *Ecological Modelling* 220: 1984-1996.

Close C, Cheung WWL, Hodgson S, Lam VWY, Watson R and Pauly D (2006) Distribution ranges of commercial fishes and invertebrates, p. 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.

Ekman S (1967) *Zoogeography of the Sea*. Sidgwick & Jackson, London.

Eschmeyer WN, Herald ES and Hammann H (1983) *A field guide to Pacific coast fishes of North America*. Houghton Mifflin Company, Boston, USA 336 p.

FAO-FIGIS (2001) A world overview of species of interest to fisheries. Chapter: *Merluccius bilinearis*. Retrieved on 15 June 2005, from www.fao.org/figis/servlet/species?fid=3026. 2p. FIGIS Species Fact Sheets. Species Identification and Data Programme-SIDP, FAO-FIGIS.

Froese R, Capuli E, Garilao C and Pauly D (2000) The SPECIES Table. p. 76-85 *In:* R. Froese, Pauly, D. (eds.), *FishBase 2000: Concept, Design and Data Sources*. ICLARM, Los Baños, Laguna, Philippines.

Jones M, Dye S, Pinnegar JK, Warren R and Cheung WWL (2012) Modelling commercial fish distributions: prediction and assessment using different approaches. *Ecological Modelling* 225:133-145.

Jones MC, Dye SR, Pinnegar JK, Warren R and Cheung WWL (2013) Applying distribution model projections in the present for an uncertain future. *Aquatic Conservation: Marine and Freshwater Research* 23(5): 710-722.

Kaschner K, Ready JS, Agbayani E, Rius J, Kesner-Reyes K, Eastwood PD, South AB, Kullander SO, Rees T, Close CH, Watson R, Pauly D and Froese R (Editors) (2008) AquaMaps Environmental Dataset: Half-Degree Cells Authority File (HCAF) [see www.aquamaps.org].

Kitchingman A and Lai L (2004) Inferences of potential seamount locations from mid-resolution bathymetric data. p. 7-12 *In*: T. Morato and D. Pauly (eds.), *Seamounts: Biodiversity and Fisheries*. Fisheries Centre Research Reports 12(5), University of British Columbia, Vancouver.

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.