Aligning marine species range data to better serve science and conservation

Casey C. O'Hara1, Jamie C. Afflerbach1, Courtney Scarborough1, Benjamin S. Halpern1,2,3

Short title: "Aligning marine species range data"

1. National Center for Ecological Analysis and Synthesis, University of California, 735 State Street Suite 300, Santa Barbara CA 93101
2. Bren School of Environmental Science and Management, University of California, Santa Barbara CA 93106
3. Department of Life Sciences, Imperial College London, Silwood Park Campus, Buckhurst Rd, Ascot, West Berkshire SL5 7PY, United Kingdom

*Corresponding author:* Casey O'Hara, National Center for Ecological Analysis and Synthesis, University of California, 735 State Street, Suite 300, Santa Barbara CA 93101; (805) 892-2500; [ohara@nceas.ucsb.edu](mailto:ohara@nceas.ucsb.edu)

## Abstract

Species distribution data provide the foundation for a wide range of ecological research studies and conservation management decisions, yet most species ranges remain unknown, and existing range maps often suffer from data limitations and inconsistencies. AquaMaps and the International Union for Conservation of Nature (IUCN) are two distinct efforts to map marine species distributions at a global scale. Together these databases represent 24,637 species (92.9% within AquaMaps, 16.3% within IUCN), with only 2,279 shared species. Here we examine differences in predicted species ranges between the two datasets and find that these misalignments mainly result from divergent methodologies that introduce differing frequencies of commission and omission errors. We illustrate the scientific and management implications of these differences by repeating a global analysis of gaps in coverage of marine protected areas, and find significantly different results depending on how the two datasets are used. While we suggest methods to address specific errors in each spatial dataset, it remains essential to understand the implications of dataset differences for conservation planning and decision-making.

# Introduction

Knowing where species exist and thrive is fundamental to the sciences of ecology, biogeography, and conservation. This body of science has resulted in various compiled databases of species distribution and range maps which provide foundational information for understanding species diversity and extinction risk, predicting species responses to human impacts and climate change, and managing and protecting species effectively. No species range dataset can claim to represent the "truth" of any species' spatial distribution, but rather each offers its own distinct understanding of that latent distribution. These varying predictions of species presence and absence, driven by intent and methodology, can result in very different predictions of species presence; however, conservation managers and policy makers must base their actions on conclusions drawn from these imperfect datasets. It is critical to understand the differences in spatial range datasets and the implications of these differences for conservation research and decision-making.

The two most comprehensive, widely-used global-scale repositories that predict marine species ranges throughout the world's oceans are AquaMaps, which rely primarily on model predictions [1], and range data from the International Union for Conservation of Nature (IUCN), which rely primarily on expert opinion [2]. The two datasets are designed to serve different purposes: IUCN range maps are drawn to estimate extent of occurrence (EOO) [**???**], communicating range size as an [indicator/factor/criterion] for extinction risk assessments [**???**]; AquaMaps essentially communicate the area of occupancy (AOO) [**???**] of a species based on habitat suitability [**???**].

The fundamental differences between the datasets (EOO versus AOO, measure of range size versus distribution) suggest that the choice of one over the other should be carefully matched to the purpose for which it is to be used [**???** Other refs], and yet these datasets have been used in hundreds of studies and applications for a wide range of purposes, including assessing marine species status [3–5], evaluating global biodiversity patterns [6–9], predicting species range shifts [10], and setting conservation priorities [11]. In most of these cases, the implications of choosing one versus the other of these datasets is not evaluated or discussed, yet strong conservation and management conclusions are drawn from the results of the studies.

Here we focus not on how the two datasets should ideally be used; instead, in recognizing that each provides value for many aspects of conservation research, we highlight hidden assumptions and sources of error within each dataset, and explore the implications of choosing one over the other.

Uncertainties inherent in method and intent inevitably drive differences in the range predictions made by each dataset: extent of occurrence range data such as IUCN range maps will inevitably introduce commission errors (inaccurate indication of presence), while species distribution models such as AquaMaps are expected to introduce fewer commission errors at the expense of more omission errors (inaccurate indications of absence) [12]. Each type of error bears different implications for conservation goals: commission errors can result in prioritizing areas not relevant to conservation goals, while omission errors may result in protected area networks that fail to include important habitat and range [12,13].

To understand the implications of differences between the AquaMaps and IUCN datasets, we compare how each data source represents the global spatial and taxonomic distribution of the 24,637 marine species mapped by one or both datasets. Most notably, AquaMaps includes range maps for many more species (currently 22,889 species; 92.9% of total), such that most global analyses related to marine biodiversity to date have used AquaMaps (IUCN range map data exist for only 4,027 unique marine species). Yet the small overlap in mapped species between the two datasets means each provides unique value in taxonomic and geographic coverage.

Although relatively small in number, these overlap species present a unique opportunity to evaluate the two datasets overall. As understanding of a given species improves, we expect that range maps, regardless of method, would more closely approach the true distribution in predicting species presence and absence. Close alignment between two independent range maps can therefore indicate higher confidence in species presence or absence than a single map. For the 2,279 species (9.3% of total) mapped in both datasets, we examine how well the maps align in both spatial distribution and overall area.

Using this alignment framework, we explore methodological issues that can introduce errors, and suggest methods to improve confidence in species range predictions. In particular, we show how explicitly using depth as a constraint on coral ranges in IUCN range data dramatically mitigates a source of commission error, and highlight a side effect of AquaMaps methodology that results in potentially unexpected species range boundaries.

As a case study, we then reexamine a global analysis of gaps in protection afforded by marine protected areas (MPAs) [11] as a case study to explore the implications of selecting one data set over the other. Repeating the study using IUCN range data rather than AquaMaps data increases the apparent proportion of gap species from 1.4% to 6.4%, while also increasing the proportion of well-protected species from 1.5% to 2.9%. This shift highlights the implications of different data use decisions on our understanding of marine biodiversity status and protection.

While we have identified some potential issues in each dataset, we cannot simply recommend one data set over the other. Instead, we recommend simple methods to account for incidence of errors in each dataset, improving confidence in species range predictions to better inform conservation management and policy decisions.

# Methods and Analysis

## About the datasets

The IUCN publishes species range maps developed by species experts. These experts outline spatial boundaries that define the "limits of distribution" of a given species, based on observation records and informed by expert understanding of species' range and habitat preferences (Fig S1A).

In contrast, AquaMaps models species distribution based on envelopes of environmental preference, such as temperature, depth, and salinity; these preference envelopes are deduced from occurrence records, published species databases such as FishBase, and expert knowledge. The AquaMaps model overlays these environmental preferences atop a map of environmental attributes on a global 0.5° grid to determine suitable habitat, resulting in a "probability of occurrence" for each species (Fig S1B). To roughly constrain species ranges to appropriate georegions, the AquaMaps model uses Food and Agriculture Organization of the United Nations (FAO) Major Fishing Area [**???**] boundaries. Of the resulting maps, 1296 (5.7%) have been further refined through an expert review process [14,15].

**Comparison of taxonomic and regional distribution**: To examine the overall taxonomic distribution across the spatial datasets, we grouped species by taxonomic class and data source, and determined the proportion of each class represented in each dataset. To compare the spatial representation of the two datasets directly, we rasterized the IUCN species polygons to the same 0.5° grid as the AquaMaps species maps; we determined species presence within a grid cell as any non-zero overlap of a species polygon with the cell ***(Fig S2)***. For the AquaMaps dataset, we determined per-cell species count by including all species with non-zero probability of occurrence, to best approximate the "extent of occurrence" generally indicated by IUCN maps.

**Alignment framework for paired maps**: OR **Comparison of paired maps**: Using genus and species binomials as a matching key, we identified "paired map" species, i.e. the subset of marine species that have range maps in both IUCN and AquaMaps current native distribution. We used the taxize package [**???**] in R [**???**] to standardize species names and synonyms; for species with separate subpopulation maps in IUCN, we combined all subpopulations to create a single global population. For each of these paired map species, we determined species presence within each spatial cell for each dataset using the same criteria outlined above.

Overlaying paired distribution maps for each species, we defined and calculated *distribution alignment* and *area ratio* :

For each paired map species, and indicate the smaller and larger range representation (regardless of which dataset). represents the amount of overlapping area between the two datasets. Distribution alignment uses overlapping predictions of presence as means of identifying areas of higher confidence. Area ratio provides a proxy for frequency of commission and/or omission errors. [range area is one critical feature/use of IUCN used as a criterion to help define extinction risk; can AquaMaps provide useful information here as well?]

**(some heading here? to separate from alignment)** Corals provide a useful case study to examine differences in species range alignment. They are an intensely studied taxonomic group whose importance in supporting biodiversity is undisputed; their dependence on photosynthesis provides habitat-limiting life history traits, and they don't tend to move. Extracting data from the IUCN API [**???**] we identified depth limitations of each coral species mapped in the IUCN dataset to verify that none are found below the photosynthetic depth limit of 200 m. Using the same methodology as shown in Fig 2, we created a 200 m bathymetry raster from Natural Earth's 200 m bathymetry polygon [**???**] and masked our IUCN coral rasters to identify mapped coral presence below 200 m. The resulting maps were again compared to the AquaMaps rasters to examine distribution alignment and area ratio.

AquaMaps' reliance upon FAO zones to constrain species ranges seems likely to result in odd species range boundaries where ecologically suitable habitat intersects with human-defined borders. To explore the impact of these artificial constraints on species range predictions, we chose to focus on boundaries defined by longitude, as they seem less likely to coincide with ecological differences than boundaries defined by latitude. As an example, we identified all AquaMaps species whose eastern or western range limit between between 5° S and 25° N latitudes coincided with the large vertical boundary between FAO regions 71 and 77 at 175° W longitude.

The oceans abound with rare and poorly-understood species who offer sparse information on behavior and habitat preferences to inform models and expert opinion. To identify such data-poor species within the set of paired map species, we used the ***robis???*** and ***rgbif???*** packages [**???**] in R to identify known occurrences of each species; occurrences were averaged between the two occurrence databases as a proxy for data richness.

## Methods for MPA Gap Analysis case study

To assess the effectiveness of MPAs in protecting biodiversity, Klein et al. [11] compared the coverage of the global MPA network presented by the World Database on Protected Areas (WDPA) [16] to the species ranges described in the 2014 AquaMaps dataset [**???**]. For the primary analysis, the researchers defined species presence as 50% or greater probability of occurrence.

To reconstruct the primary analysis, we selected the subset of protected areas from the 2014 WDPA dataset classified as IUCN protected area management categories I-IV and spatially overlapping a marine area. The WDPA polygons and marine polygons were rasterized to 0.01° and then aggregated to AquaMaps native 0.5° cells, to calculate proportion of marine protected area within each cell. After verifying our results using the 2014 AquaMaps dataset, we updated the analysis to use the 2015 AquaMaps dataset, at a presence threshold of 50% (to compare to Klein et al. directly) and 0% (to better compare with IUCN spatial data). To analyze MPA coverage against IUCN spatial data, we extracted IUCN polygon weights per 0.5° cell for each species and compared against the protected area raster. Finally, we combined AquaMaps data (at 0% threshold) and IUCN data, using AquaMaps for the 2,279 overlapping species and again compared against the protected area raster.

All processing was completed using R statistical software [17], and all code and intermediate data are available on GitHub at <https://github.com/OHI-Science/IUCN-AquaMaps>.

# Results and Discussion

## Taxonomic and geographic coverage

[results] The two datasets have notably different taxonomic (Fig 1A) and regional (Figs 1B, 1C) coverage. AquaMaps encompasses a broader range of taxa than IUCN, as IUCN spatial data files are only available for select taxonomic groups that have been comprehensively assessed. While species numbers in both datasets peak in tropical latitudes near the equator, species counts for IUCN maps drop quickly beyond 30°N and 30°S, while AquaMaps includes distribution of species well into temperate latitudes. The longitude frequency plots show a slight shift in the IUCN dataset away from the Atlantic and eastern Pacific compared to AquaMaps.

**Fig 1.** (A) Number and proportion of species by taxa included in each dataset. Overlapping species are dominated by bony fishes (983 species, primarily tropical taxa) and corals (396 species). (B, C) Global marine species count per 0.5° cell according to (B) AquaMaps and (C) IUCN. The margin frequency plots show relative species count per cell at each latitude and longitude.

[ideas for discussion here? each provides information unavailable in the other; in addition to considerations about the nature of each dataset, the species and regions covered are an important consideration in which dataset to use]

## Distribution and range area alignment

[discussion] To explore differences in species distribution and range between the two datasets, we assessed the distribution alignment (how much of the smaller range falls within the larger range, i.e., where on the map) relative to the area ratio (ratio of smaller range area to larger range area, i.e., how much of the map) for each shared species (Fig 2A). Where two independent methods predict presence (or absence) of a species in the same location, we can assume a higher confidence in the accuracy of that prediction. Where one method predicts presence while the other predicts absence, we can investigate potential sources of commission and omission errors to refine our understanding of the species' range. Understanding these sources of error can provide guidance to improve predictions of species range even when only one source of spatial data is available.

**Fig 2.** (A) Distribution alignment (overlap of smaller range within larger) versus area ratio (the ratio of smaller range area to the larger range area) for 2,279 species included in both IUCN and AquaMaps datasets. The upper right quadrant (quadrant 1) comprises species whose maps largely agree (better than median value) in both spatial distribution and the extent of described ranges (n = 466; 20.1 %). The upper left quadrant (quadrant 2) comprises species whose maps agree well in distribution, but disagree in area (n = 687; 29.7 %). The lower right quadrant (quadrant 3) includes species for which the paired maps generally agree in range area, but disagree on where those ranges occur (n = 691; 29.9 %). The lower left quadrant (quadrant 4) indicates species for which the map pairs agree poorly in both area and distribution (n = 470; 20.3 %). (B) Alignment quadrant breakdown of species by taxonomic group.

[discussion] This analysis revealed a weak negative linear pattern, suggesting that increasing similarity in range area correlates with decreasing distribution alignment. AquaMaps tends to extrapolate species ranges into suitable areas beyond known occurrences, such that each additional unit of range predicted by AquaMaps will fall in different locations than an additional unit of range predicted using IUCN methodology. For species with dissimilar range areas, predicted distribution for the smaller range can more easily fall within the generous bounds of the larger range. For species with increasingly similar range areas, differences in methodology become more difficult to "hide," and the distribution alignment generally becomes poorer.

[results] The mean distribution alignment for species included in both datasets was 63%; the mean area alignment was 54.5%. By dividing the paired map species into quadrants based on these means, we highlight categories of relationships that help further explain this general pattern. Representative maps from each category are provided in the supporting materials (Fig S1).

[results/discussion] The upper right quadrant includes the species (n = 527) whose described ranges are above average in alignment of both spatial distribution and area. These species tend to be well-studied and include wide-ranging pelagic organisms such as marine mammals, tunas, and billfishes (Fig 2B). This result is not surprising, as species with very large ranges are likely to be more aligned regardless of methodology simply because their ranges span nearly the entire map.

[results/discussion] The area-mismatched ranges contained in the upper left quadrant (n = 709) include many species whose spatial distribution is similar, but where one range is notably larger than the other (For 88% of the species in this quadrant, the IUCN range is an average of 2.57 times larger than the AquaMaps range). This suggests a high rate of commission and/or omission errors by one or both datasets; further analysis is required to disentangle the source and type of error contributing to poor area alignment. The results of the coral analysis (described below) provide some insight.

[results/discussion] Species found in the lower right quadrant (n = 635) seem to represent cases of "two wrongs make a right." For these species, IUCN and AquaMaps both predict ranges extending far beyond the overlapping region, but the methodological differences result in very different extrapolations. Consequently, area ratios are close to 100%, though the poor distribution alignment indicates that one or both datasets are introducing significant errors. As above, further analysis is required to disentangle the causes of error.

[results] The lower left quadrant includes species (n = 443) where alignment is poor in both dimensions. Data-poor species are more common in this quadrant; indeed, the median number of species occurrence records (averaging occurrences from the Ocean Biogeographic Information System (OBIS) [18] and the Global Biodiversity Information Facility (GBIF) [19]) for this quadrant is 24 records, compared to a median of 97 records for species across the other three quadrants. In addition, the AquaMaps dataset offers a quality metric based on the number of unique cells containing valid occurrences; for this quadrant, the median "occurcells" is 11 compared to a median of 40 across the other three quadrants. Care should be taken when using distribution and range maps based upon fewer observations, as they clearly bear greater uncertainty; AquaMaps explicitly recommends using any of its maps generated with an "occurcells" count fewer than 10 [**???**].

## Coral depth exploration

[discussion] Noting from Fig 2B that corals dominate the upper-left "distribution-aligned" quadrant of Fig 2A (n = 237; 33.4% of all species in this quadrant), we chose to examine the effect of explicitly restricting IUCN ranges to depths based on species' life histories. Corals offer an excellent case study, due to their foundational role in biodiverse habitats, as well as their lack of mobility and reliance on photosynthesis. Ocean depth preference is explicitly included in AquaMaps models. While depth is recommended by the IUCN as a criterion for providers of range maps ("The limits of distribution can be determined by using known occurrences of the species, along with the knowledge of habitat preferences, remaining suitable habitat, elevation limited, and other expert knowledge of the species and its range." [**???** <http://www.iucnredlist.org/technical-documents/red-list-training/iucnspatialresources>]), it is not presented as a requirement, so we cannot take its inclusion for granted.

[results] Fig 3A shows aggregated ranges of the 463 coral species mapped in the IUCN dataset, with their ranges broken into proportional area deeper and shallower than 200 m. According to IUCN descriptions, none of these species is indicated to occur deeper than 200 m, and 94% are confined to waters shallower than 50 m; seven of the mapped species had no reported depth information. Clipping coral ranges to shallower than 200 m eliminated an average of 47.6% of the total predicted area while still allowing for a generous estimate of suitable habitat.

**Fig 3.** (A) Aggregate map combining ranges of the 463 coral species mapped in the IUCN dataset, showing raw ranges and ranges clipped to 200 m depth. (B) Alignment quadrant breakdown of paired map coral species using original data from IUCN and AquaMaps (as in Fig 2B) and the same species with IUCN ranges clipped to 200 m depth.

[results] In constraining coral ranges to shallow waters, we see a strong increase in the apparent alignment of species maps between IUCN and AquaMaps. Using the original quadrant definitions from Fig 2B, we see in Fig 3B a massive shift in 354 paired-map coral species due to constraining coral depths to 200 m. Membership in the "well-aligned" quadrant jumped from 22.4% to 76.2%, with a corresponding decrease in all other quadrants. See Fig SXXX to examine the shifts of individual species among the quadrants.

[discussion] We cannot know for certain the true distribution of each of these corals; certainly some commission errors result from satisfying the connectivity requirements of an EOO, and others may be due to experts taking a precautionary approach. But a sensible shift in method drastically decreases the likelihood of introducing commission errors, with little chance of introducing omission errors, greatly improving our confidence in the remaining reported distributions for most purposes. Note that this change applies just as readily to the IUCN coral maps that are not included in the paired map analysis, and likely to other reef-associated flora and fauna. While species depth preferences are an easy and consistent means of constraining range predictions, other conditions such as salinity and temperature could be cautiously used to refine the results of expert opinions.

## Georegional constraint exploration

[results] We identified 4,168 species whose equatorial ranges (between 5° S and 25° N) encounter a range limit at 175° W. This longitudinal boundary created a western limit for 512 species, and eastern limit for 3,656 species. Plotting the entire ranges of the 512 west-limited species (Fig 4) results in a very clear exclusion zone that matches perfectly with FAO region 71; while we did not focus on the boundary for region 57, a similar effect is readily apparent.

**Fig 4.** Composite map showing AquaMaps predicted ranges for 512 species whose equatorial range (between 5° S and 25° N) encounters a western limit exactly at 175° W, the boundary between FAO Major Fishing Areas 71 and 77. The range maps for these species conform exactly to the bounds of Area 71, as a result of AquaMaps using FAO areas as a rough georegional constraint on species ranges. The boundary of Area 57 creates an additional clear exclusion zone for most of these species, despite the fact that this was not used as a selection criteria.

[discussion] FAO Major Fishing Area boundaries provide a readily available method to roughly contrain AquaMaps predictions to appropriate ocean basins, thus eliminating a large source of potential commission errors and enabling rapid modeling of thousands of species ranges. However, these boundaries are defined for statistical purposes based on economic and political considerations rather than ecological considerations, and can result in odd species range predictions where otherwise suitable habitat is omitted where it encounters border between regions. While such an odd boundary would likely be obvious when inspecting the distribution of an individual species, the distinction is likely to be obscured when aggregating many species ranges as is typical for biodiversity or conservation studies. Applying an ecoregional spatial constraint, such as Marine Ecoregions of the World [**???**], may reduce commission errors on these questionable margins, boosting confidence in these predicted distributions; expert review, though time-consuming, is the most certain route to improving the quality of these maps.

## Case Study: MPA Gap Analysis

Klein et al. [11] compare the global distribution of species to the global distribution of marine protected areas to assess how well current MPAs overlap with species ranges and identify which species fall through gaps in protection. The study relied on the AquaMaps database, using a probability of occurrence threshold of 50% or greater, to determine species presence, and the World Database of Protected Areas to define zones of marine protection. They found that the global MPA network leaves 90.5% of marine species with less than 5% of their overall range represented within MPAs, and 1.4% of species have no protection at all (i.e., "gap" species).

We recalculated the amount of under-protected and gap species using either IUCN or AquaMaps data (using the most recent AquaMaps data and a 0% threshold to allow the most meaningful comparison to IUCN's "limits of distribution", Fig 4). We found a five-fold increase in the proportion of gap species (6.4% of species vs. 1.2%) and dramatically larger proportion of species with less than 2% of their range protected (73.2% of species vs. 47.7%). However, this comparison also indicates a larger proportion of well-protected species with greater than 10% of range protected (2.9% of species vs. 1.5%).

**Fig 5.** Percent of species range covered by MPAs based upon methods in Klein et al. (2015). Scenario 1 replicates the original results, measuring protected range of species in AquaMaps 2014 dataset, with a 50% presence threshold, against the 2014 World Database of Protected Areas, filtered for IUCN categories I-IV that overlap marine areas. Scenario 2 updates the results using AquaMaps 2015, showing very small changes despite the inclusion of an additional 5,545 species. Scenario 3, using 2015 AquaMaps data, drops the presence threshold to zero, showing an expected decrease in gap species, but also a decrease in species with 5% or greater protected range. Scenario 4 adds an additional 1745 species unique to IUCN, resulting in increases in gap species and species with less than 2% coverage. Scenario 5 examines species MPA coverage using only the IUCN dataset.

[discussion?]

[can this wrap up the results/discussion?] To achieve more comprehensive global coverage of species ranges these two datasets can be used together, understanding that the underlying differences complicate such direct comparisons. Major considerations include aligning the criteria used to determine "presence" (e.g. selecting an appropriate presence threshold for AquaMaps; *a presence threshold of 0% most closely approximates the "limits of distribution" criterion described by IUCN range maps* ***- not really discussed or backed up here***) and addressing mechanisms that introduce errors of commission and omission. ***as it stands, this paper really doesn't go into combining datasets and how to do it; omit this paragraph?***

# Conclusions

No spatial dataset can ever claim to know the "truth" of the whole of marine biodiversity. AquaMaps and IUCN range maps show strong agreement for many well-studied species, but substantial differences illustrate uncertainty in our understanding of spatial distribution for many others. Although many other approaches exist for species distribution modeling, these two are the largest approaches applied globally across a broad range of marine taxa. Method-driven differences in commission and omission errors produce clear and significant disagreement in species range descriptions between AquaMaps and IUCN datasets. Conclusions drawn from each of these datasets would paint dramatically different pictures of global marine biodiversity or the effectiveness of conservation management decisions. Identifying and addressing differences in these datasets will increase their utility for research and conservation actions.

Effective management and protection of marine species depends on a robust understanding of where species exist and where they do not; without this knowledge we risk wasting resources protecting low-value regions while missing opportunities to protect critical ones. By identifying the differences between these two important and fundamentally different marine species range datasets and understanding the likely mechanisms causing these discrepancies, we improve our ability to develop strategic and effective conservation policy that supports a resilient ocean ecosystem.

# Acknowledgments

We are thankful to Melanie Frazier and Julia Stewart Lowndes for their insightful comments on earlier versions of this manuscript, and to the National Center for Ecological Analysis and Synthesis for computation support. We acknowledge financial support from the Gordon and Betty Moore Foundation.

# References

1. Kaschner K, Rius-Barile J, Kesner-Reyes K, Garilao C, Kullander S, Rees T, et al. AquaMaps: Predicted range maps for aquatic species. 2015. Available: <www.aquamaps.org>

2. IUCN. The IUCN red list of threatened species. International union for the conservation of nature. 2015. Available: <http://www.iucnredlist.org>

3. Halpern BS, Longo C, Hardy D, McLeod KL, Samhouri JF, Katona SK, et al. An index to assess the health and benefits of the global ocean. Nature. 2012;488: 615–620. doi:[10.1038/nature11397](https://doi.org/10.1038/nature11397)

4. Halpern BS, Longo C, Lowndes JSS, Best BD, Frazier M, Katona SK, et al. Patterns and emerging trends in global ocean health. Tsikliras AC, editor. PLOS ONE. 2015;10: e0117863. doi:[10.1371/journal.pone.0117863](https://doi.org/10.1371/journal.pone.0117863)

5. Selig ER, Longo C, Halpern BS, Best BD, Hardy D, Elfes CT, et al. Assessing global marine biodiversity status within a coupled socio-ecological perspective. Guichard F, editor. PLoS ONE. 2013;8: e60284. doi:[10.1371/journal.pone.0060284](https://doi.org/10.1371/journal.pone.0060284)

6. Coll M, Piroddi C, Steenbeek J, Kaschner K, Ben Rais Lasram F, Aguzzi J, et al. The biodiversity of the mediterranean sea: Estimates, patterns, and threats. Bograd SJ, editor. PLoS ONE. 2010;5: e11842. doi:[10.1371/journal.pone.0011842](https://doi.org/10.1371/journal.pone.0011842)

7. Martin C, Fletcher R, Jones M, Kaschner K, Sullivan E, Tittensor DP, et al. Manual of marine and coastal datasets of biodiversity importance. United Nations Environment Programme; 2014 May.

8. Pimm SL, Jenkins CN, Abell R, Brooks TM, Gittleman JL, Joppa LN, et al. The biodiversity of species and their rates of extinction, distribution, and protection. Science. 2014;344: 1246752–1–1246752–10. doi:[10.1126/science.1246752](https://doi.org/10.1126/science.1246752)

9. Kaschner K, Tittensor DP, Ready J, Gerrodette T, Worm B. Current and future patterns of global marine mammal biodiversity. Bograd SJ, editor. PLoS ONE. 2011;6: e19653. doi:[10.1371/journal.pone.0019653](https://doi.org/10.1371/journal.pone.0019653)

10. García Molinos J, Halpern BS, Schoeman DS, Brown CJ, Kiessling W, Moore PJ, et al. Climate velocity and the future global redistribution of marine biodiversity. Nature Climate Change. 2015;6: 83–88. doi:[10.1038/nclimate2769](https://doi.org/10.1038/nclimate2769)

11. Klein CJ, Brown CJ, Halpern BS, Segan DB, McGowan J, Beger M, et al. Shortfalls in the global protected area network at representing marine biodiversity. Scientific Reports. 2015;5: 17539. doi:[10.1038/srep17539](https://doi.org/10.1038/srep17539)

12. Rondinini C, Wilson KA, Boitani L, Grantham H, Possingham HP. Tradeoffs of different types of species occurrence data for use in systematic conservation planning: Species data for conservation planning. Ecology Letters. 2006;9: 1136–1145. doi:[10.1111/j.1461-0248.2006.00970.x](https://doi.org/10.1111/j.1461-0248.2006.00970.x)

13. Jetz W, Sekercioglu CH, Watson JEM. Ecological correlates and conservation implications of overestimating species geographic ranges: *Overestimation of species ranges*. Conservation Biology. 2008;22: 110–119. doi:[10.1111/j.1523-1739.2007.00847.x](https://doi.org/10.1111/j.1523-1739.2007.00847.x)

14. Kaschner K, Watson R, Trites AW, Pauly D, others. Mapping world-wide distributions of marine mammal species using a relative environmental suitability (RES) model. Marine Ecology Progress Series. 2006;316: 2–3. Available: <http://www.vliz.be/imisdocs/publications/100462.pdf>

15. Ready J, Kaschner K, South AB, Eastwood PD, Rees T, Rius J, et al. Predicting the distributions of marine organisms at the global scale. Ecological Modelling. 2010;221: 467–478. doi:[10.1016/j.ecolmodel.2009.10.025](https://doi.org/10.1016/j.ecolmodel.2009.10.025)

16. IUCN, UNEP-WCMC. The world database on protected areas (WDPA). Cambridge, UK: UNEP-WCMC. 2014. Available: <www.protectedplanet.net>

17. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2016. Available: <https://www.R-project.org>

18. OBIS. Data from the ocean biogeographic information system. Intergovernmental oceanographic commission of UNESCO. 2016. Available: <http://www.iobis.org>

19. GBIF. Global biodiversity information facility (GBIF) memorandum of understanding. Global biodiversity information facility. 2010. Available: <http://www.gbif.org/resource/80661>