Aligning marine species range data to better serve science and conservation

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

*Author affiliation:*

1. National Center for Ecological Analysis and Synthesis, 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, 735 State Street, Suite 300, Santa Barbara CA 93101; (805) 892-2500; [ohara@nceas.ucsb.edu](mailto:ohara@nceas.ucsb.edu)

*Classification:* BIOLOGICAL SCIENCES: Sustainability Science

*Keywords:* Marine spatial planning, species distribution model, extent of occurrence, spatial ecology

## Summary

1. 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.
2. Here we examine differences in predicted species ranges between the two datasets by comparing alignment of distribution (predicted species range common to both datasets) and area (extent of range predicted by each dataset) to identify potential sources of misalignment.
3. We find that these misalignments mainly result from divergent methodologies that introduce differing frequencies of commission and omission errors. For example, species ranges predicted by IUCN range maps frequently overextend into unsuitably deep territory, as in the case of many coral and reef-associated species; similarly, the AquaMaps species distribution model often predicts species presence far afield from known occurrences when extrapolating based on limited observations.
4. **Policy implications:** We illustrate the scientific and management implications of these differences by repeating two recent applications - an assessment of global biodiversity within the Ocean Health Index and a global analysis of gaps in coverage of marine protected areas - and find significantly different results depending on how the two datasets are used. Until a single, highly accurate dataset of global marine species ranges becomes available, understanding the implications of dataset differences for conservation planning and decision-making remains essential.

# Introduction

Knowing where species exist and thrive is fundamental to the sciences of ecology, biogeography, and conservation, among many others. This knowledge provides foundational information for understanding species ranges and diversity, predicting species responses to human impacts and climate change, and managing and protecting species effectively.

One major outcome of this body of science is the various compiled databases of species distribution maps. Two global-scale repositories predict marine species ranges throughout the world's oceans – AquaMaps (Kaschner *et al.* 2015) and International Union for Conservation of Nature (IUCN) (IUCN 2015). These two spatial datasets have been used in hundreds of studies and applications for a wide range of purposes, including assessing marine species status (Halpern *et al.* 2012, 2015; Selig *et al.* 2013), evaluating global biodiversity patterns (Coll *et al.* 2010; Kaschner *et al.* 2011; Martin *et al.* 2014; Pimm *et al.* 2014), predicting species range shifts (García Molinos *et al.* 2015), and setting conservation priorities (Klein *et al.* 2015).

***sentence about SAUP - how can we rephrase the "two global-scale repositories" sentence to exclude it? "broad taxonomic coverage"? something about not focusing just on economically important species? (is that a valid assumption about SAUP)***

***or: "Several global-scale repositories predict ... In this paper we focus on two datasets that cover a broad taxonomic range: ..."***

Neither dataset claims to represent the "truth" of a species' spatial distribution, but rather each presents one understanding of the "truth" - one database relying primarily on expert opinion and the other on model predictions. Uncertainties inherent in the method and intent inevitably drive differences in the range predictions made by each dataset; significant differences could indicate dramatically different understandings of our marine ecosystems, with significant implications for policy and conservation recommendations. Inaccurate indications of presence or absence could lead to ineffective marine reserve systems and management plans (Rondinini *et al.* 2006; Jetz, Sekercioglu & Watson 2008). Importantly, biases in taxonomic or spatial coverage within a dataset could shift management and conservation actions away from places or species that are most in need.

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 marine species. 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).

***It seems fair to assume that/we can reasonably assume that*** range predictions from each dataset would be more closely aligned for well-understood species. For the 2,279 species (9.3% of total) mapped in both datasets, we examine how well the maps align, determine several issues that lead to misalignment between predicted species distributions, and suggest methods to improve confidence in species range predictions.

We then reexamine two recent marine biodiversity studies - an assessment of the status of global biodiversity within the Ocean Health Index (OHI) (Halpern *et al.* 2012, 2015) and a global analysis of gaps in protection afforded by marine protected areas (MPAs) (Klein *et al.* 2015) - as case studies to explore the implications of prioritizing one data set over the other. The results highlight possible consequences of different data use decisions on our understanding of marine biodiversity status and protection.

While we cannot recommmend one data set as more accurate than the other, we suggest methods to improve confidence in species range predictions to better inform conservation management and policy decisions.

***acknowledge or debunk other global data sets - e.g. SAUP?***

# Methods

\_About the datasets\_\_:

The IUCN publishes species range maps based on expert input of spatial boundaries of a given species' "limits of distribution" (IUCN 2015) - essentially a refined extent of occurrence (EOO), based on observation records and informed by expert understanding of species' range and habitat preferences. In contrast, AquaMaps models species distribution based on environmental preferences (e.g., temperature, depth, salinity) 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 0.5 degree grid, creating a global map of suitable habitat. The model masks habitat to certain FAO Major Fishing Areas as a rough cut of geographic range, resulting in a "probability of occurrence" for each species. Of the resulting maps, 1296 (5.7%) have been further refined under expert review (Kaschner *et al.* 2006; Ready *et al.* 2010). Studies using AquaMaps data frequently define "presence" by applying a probability threshold, e.g., previously, the Ocean Health Index defined presence as 40% or greater probability of occurrence (Halpern *et al.* 2012, 2015). In the 2016 OHI assessment, the presence threshold was eliminated, to better approximate the "extent of occurrence" generally indicated by IUCN maps, since the two datasets are used together.

***short paragraph about other datasets we didn't include, i.e. SAUP? BOTW?***

**Comparison of taxonomic and regional distribution**: To examine the overall taxonomic distribution across the spatial datasets (Fig. 1A), we grouped species by taxonomic class and data source, and determined the proportion of each class represented in each dataset. We defined "paired map" species as those with spatial data in both datasets.

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. SXXX)***. 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 ***(Fig. SXXX)***. We represented relative distribution of species count for each dataset by plotting average species count against latitude and longitude (Fig. 1B, 1C).

**Comparison of paired maps**: Using genus and species binomials as a matching key, we selected the "paired map" species - 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.* ***To avoid double-counting, we removed subpopulations and species aliases. Used the taxize package? which version?*** We determined species presence within each spatial cell using the same criteria outlined above.

Overlaying paired distribution maps for each species, we defined and calculated *distribution alignment* and *area ratio* and plotted these in Fig. 2A:

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 high confidence. Area ratio provides a proxy for frequency of commission and/or omission errors.*

Dividing the plot into quadrants based on mean values along each axis, we roughly categorized four qualities of alignment, and for each taxonomic group, based on IUCN taxonomic groupings, determined the proportion of paired map species that fell into each quality category (2B).

*Based on the results of the map alignment analysis, we chose to examine the effect of explicitly restricting IUCN ranges to depths based on species' life histories. We selected corals as a case study, due to their foundational role in biodiverse habitats, as well as their lack of mobility and general reliance on photosynthesis.* Using the IUCN's API [REF] to find life history information for all IUCN-mapped corals, we found that none of the paired-map corals is described as occurring deeper than 200 m; 94% are noted as being limited to 50 m or shallower. We created a 200 m bathymetry raster from Natural Earth's 200 m bathymetry polygon (**???**) ***(Fig. SXXX)*** and masked our IUCN coral rasters to exclude presence below 200 m. The resulting maps were again compared to the AquaMaps rasters; ***Fig. XXXX*** plots shifts in coral species alignment due to this depth clipping.

***To identify data-poor species, we used the robis??? and rgbif??? packages in R to identify known occurrences of each species; occurrences were averaged between the two bases as a proxy for data richness.***

**Methods for Ocean Health Index (OHI) case study:** The Species component of the OHI Biodiversity goal (SPP) uses IUCN Red List extinction risk as a proxy for a species' health. Each species is assigned a risk score based on IUCN Red List extinction risk category, from Least Concern (0.0) to Extinct (1.0) (Data Deficient and unassessed species are excluded). Species range is determined from its IUCN range map, where available, or its AquaMaps range map if an IUCN map is unavailable. For each assessed OHI region, an area-weighted mean extinction risk across all species is calculated, and rescaled to calculate a status score from 0 (mean extinction risk = 0.75) to 100 (mean extinction risk = 0) (Halpern *et al.* 2012, 2015).

*To highlight the differences in the datasets, we recalculated the OHI SPP status values using only one complete dataset at a time. For the AquaMaps dataset, we used a presence threshold of 0% to align with the OHI 2016 assessment. Fig. 3 compares these single-dataset scores to the 2016 OHI SPP status values. While the full OHI 2016 assessment includes the Birds of the World dataset (****???****) in addition to IUCN and AquaMaps, we omitted it from our calculations for simplicity.*

~To highlight the differences in the datasets, we selected the subset of species whose ranges are represented in both datasets, eliminating species found in only one or the other dataset. In the original OHI SPP analysis, IUCN data was used to determine ranges for these overlapping species; we calculated all regional SPP status scores for just these species as our baseline. We then recalculated the scores using AquaMaps ranges at a 40% presence threshold (as used in the original OHI analysis) and a 0% threshold (to better simulate the "limits of distribution" defined by IUCN maps) and compared them to our IUCN baseline.~

~We then modified the original OHI source code to recalculate the 2015 SPP status scores for all regions and all species to examine the effects of changing the AquaMaps presence threshold from 40% to 0% (all range included), and the effects of changing the data source for the overlapping species from IUCN to AquaMaps. We calculated the differences between these results and the published SPP status scores for 2015.~

**Methods for MPA Gap Analysis case study:** To assess the effectiveness of MPAs in protecting biodiversity, Klein et al. (2015) compared the coverage of the global MPA network presented by the World Database on Protected Areas (WDPA) (IUCN & UNEP-WCMC 2014) 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 covering 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 (R Core Team 2016), and all code and intermediate data are available on GitHub at <https://github.com/OHI-Science/IUCN-AquaMaps>.

# Results and Discussion

1. 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 species counts for AquaMaps remain robust 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.
2. *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) and reducing mechanisms that introduce errors of commission and omission.*

***from Methods*** (D) *The methodologies behind the creation of these datasets imply differences in prediction of species distribution due to errors of commission (falsely indicating species presence) and omission (falsely indicating species absence). Geographic range data such as IUCN range maps frequently introduce commission errors, while species distribution models such as AquaMaps will likely introduce fewer commission errors but more omission errors. 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 (Rondinini et al. 2006). By comparing maps resulting from IUCN and AquaMaps methodologies, we can identify and possibly address mechanistic causes for each type of error.*

(D/M) ***is this sufficiently described in methods? or revisit here?*** ***Both datasets strive for accuracy in different ways; the truth likely lies somewhere in between.*** To explore differences in species distribution and range between the two datasets, we plotted the distribution alignment (how much of the smaller range falls within the larger range, i.e., where on the map) against the area alignment (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.

1. This analysis revealed a general 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.
2. 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).
3. 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.
4. 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 or type of error contributing to poor area alignment. The results of the coral analysis (described below) provide some insight.
5. ***area ratio suggests similar commission/omission errors, but dist suggests otherwise*** 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. ***OMIT:*** *For these species, IUCN ranges frequently overextend into unsuitable depths, as in the case of many upper left quadrant species, while at the same time AquaMaps ranges often aggressively extrapolate presence into locations where IUCN predicts absence, likely introducing additional commission errors.* Consequently, area ratios are close to 100%, though the poor distribution alignment indicates that one or both datasets are introducing errors. As above, further analysis is required to disentangle the causes of error.
6. The lower left quadrant includes species (n = 443) where alignment is poor in both dimensions. ***present a solution here?*** 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) (OBIS 2016) and the Global Biodiversity Information Facility (GBIF) (GBIF 2010)) for this quadrant is 24 records, compared to a median of 97 records for species across the other three quadrants. It is not surprising that range maps based upon fewer observations bear greater uncertainty. ***OMIT:*** *When extrapolating from limited observations, the AquaMaps model often predicts species presence well beyond known occurrences, introducing commission errors; at the same time, IUCN range maps generally target known occurrences, possibly introducing omission errors for data-limited species.*

## New analysis: Coral maps (Distribution-aligned quads)

(D/M) 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 clipped the IUCN coral range maps to a 200 m bathymetry raster to examine the impact on distribution and area alignment. 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 clear that this cannot be taken for granted.

* Coral area plot

1. Fig. 3A shows the 463 coral species mapped in the IUCN dataset, with their ranges broken into proportional area deeper and shallower than 200 m. 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 conservative estimate of suitable habitat.

* Coral new quads

1. Masking coral ranges to shallow waters reduces a clear source of commission errors. In so doing, 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 and 3C a massive shift in 354 paired map 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.
2. We cannot know for certain the true distribution of each of these corals; much of the apparent over-prediction of coral ranges is likely due to very conservative and precautionary approach. But a small and sensible shift in method drastically decreases the introduction of likely commission errors, greatly improving our confidence in the remaining reported ranges. Note that this change applies just as well 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.

***What is missing here is something to show that AquaMaps ranges, clipped to FAO Major Fishing Areas, are probably introducing errors of commission as well*** - *the fish don't see that FAO boundary and suddenly decide not to cross; most obvious in the boundaries defined by longitudes. Can we show this with a couple of example maps? That wasn't received well for the corals, but I can't think of a clear and ambiguous way to analyze this problem.*

# Implications

Neither IUCN nor AquaMaps claims a lock on the "truth", and the small overlap in mapped species means each provides value in taxonomic and geographic coverage. Method-driven differences in commission and omission errors produce clear and significant disagreement in species range descriptions between AquaMaps and IUCN datasets. To examine the implications of these differences, we repeated two recent studies, varying only the selection of one data set over the other.

**Case Study: The Ocean Health Index**

The global Ocean Health Index (OHI) (Halpern *et al.* 2012, 2015), a composite index comprising ten sustainable benefits provided by a healthy ocean, uses species spatial distribution data and IUCN Red List conservation status to calculate biodiversity status (scored from zero to 100) for each of the world's 221 exclusive economic zones. To maximize the number of represented species, OHI gleans spatial distribution data from Red List-assessed species in both IUCN and AquaMaps datasets (n = 7,963), prioritizing IUCN data for the 2,026 species included in both sources.

***OMIT:*** *We calculated the OHI species status score under several scenarios to observe the impact of toggling the prioritized data set from IUCN to AquaMaps, and toggling the AquaMaps presence threshold from 40% to 0% (Fig. 3). Reducing the threshold increases the apparent range of a species, so a decrease in a region’s score under scenario 1 would indicate increased spatial representation of threatened species for that region. Shifting priority from IUCN to AquaMaps may increase or decrease a species’ apparent range, so a decrease in region score for scenario 2 may indicate a decrease in spatial representation of low-risk species, an increase in spatial representation of high-risk species, or a combination of both. Scenario three combines effects of scenarios 1 and 2.*

***OMIT:*** *Given that only 25.4% of Red List-assessed species are included in both datasets, it is surprising that changing the priority for overlapping species from IUCN maps to AquaMaps would result in substantial country-level score shifts as seen in scenarios 2 and 3. While the mean global score did not vary significantly from scenario to scenario, select countries gained up to 7.5 points while others dropped as many as 5 points. This result indicates that especially on a national or regional scale, an arbitrary change in how the two datasets are combined can result in a different assessment of species conservation status.*

**Case Study: MPA Gap Analysis**

Klein et al. (2015) 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%). Conclusions drawn from each of these datasets would paint dramatically different pictures of the protection afforded by our current global MPA network.

# Conclusions

AquaMaps and IUCN range maps show reasonable 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. Identifying and addressing differences in these datasets will increase their utility for research and conservation actions.

For IUCN range data, clipping ranges to known depth limits improves confidence in predicted ranges for coral species; similarly, depth may be a useful characteristic to consider for other reef-associated organisms. Other parameters such as salinity and temperature may provide useful constraints on other taxa. Simple but conservative rules of thumb will likely reduce commission errors without introducing substantial omission errors. For AquaMaps range data, dependent primarily on environmental and physical preferences and conditions, implementing area restrictions based on biogeographical criteria such as Marine Ecoregions of the World (Spalding *et al.* 2007), rather than political and economic criteria such as FAO Major Fishing Areas, would likely decrease commission errors and improve predictive power, especially for data-poor species. For either data set, maps based on few occurrences are more likely to bear high uncertainty in range predictions. Occurrence counts from external sources such as OBIS or GBIF can help identify relatively data-poor species; additionally, the "occurcells" attribute in the AquaMaps data set, which counts the number of half-degree cells used to generate the environmental envelope for each species, can be used in a similar manner.

***OMIT:*** *Many studies using AquaMaps data test the sensitivity of results by varying the probability of occurrence threshold to determine presence. This decision ultimately represents a tradeoff between errors of commission (low threshold) and omission (high threshold). Using AquaMaps in conjunction with IUCN can mitigate potential errors, while also increasing the taxonomic and spatial breadth of coverage, as long as the differences between the datasets can be reasonably minimized. In this case, a presence threshold of 0% most closely approximates the "limits of distribution" criterion defined by IUCN data providers.*

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 fundamental 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

Coll, M., Piroddi, C., Steenbeek, J., Kaschner, K., Ben Rais Lasram, F., Aguzzi, J., Ballesteros, E., Bianchi, C.N., Corbera, J., Dailianis, T., Danovaro, R., Estrada, M., Froglia, C., Galil, B.S., Gasol, J.M., Gertwagen, R., Gil, J., Guilhaumon, F., Kesner-Reyes, K., Kitsos, M.-S., Koukouras, A., Lampadariou, N., Laxamana, E., López-Fé de la Cuadra, C.M., Lotze, H.K., Martin, D., Mouillot, D., Oro, D., Raicevich, S., Rius-Barile, J., Saiz-Salinas, J.I., San Vicente, C., Somot, S., Templado, J., Turon, X., Vafidis, D., Villanueva, R. & Voultsiadou, E. (2010) The biodiversity of the mediterranean sea: Estimates, patterns, and threats ed S.J. Bograd. *PLoS ONE*, **5**, e11842.

García Molinos, J., Halpern, B.S., Schoeman, D.S., Brown, C.J., Kiessling, W., Moore, P.J., Pandolfi, J.M., Poloczanska, E.S., Richardson, A.J. & Burrows, M.T. (2015) Climate velocity and the future global redistribution of marine biodiversity. *Nature Climate Change*, **6**, 83–88.

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

Halpern, B.S., Longo, C., Hardy, D., McLeod, K.L., Samhouri, J.F., Katona, S.K., Kleisner, K., Lester, S.E., O’Leary, J., Ranelletti, M., Rosenberg, A.A., Scarborough, C., Selig, E.R., Best, B.D., Brumbaugh, D.R., Chapin, F.S., Crowder, L.B., Daly, K.L., Doney, S.C., Elfes, C., Fogarty, M.J., Gaines, S.D., Jacobsen, K.I., Karrer, L.B., Leslie, H.M., Neeley, E., Pauly, D., Polasky, S., Ris, B., St Martin, K., Stone, G.S., Sumaila, U.R. & Zeller, D. (2012) An index to assess the health and benefits of the global ocean. *Nature*, **488**, 615–620.

Halpern, B.S., Longo, C., Lowndes, J.S.S., Best, B.D., Frazier, M., Katona, S.K., Kleisner, K.M., Rosenberg, A.A., Scarborough, C. & Selig, E.R. (2015) Patterns and emerging trends in global ocean health ed A.C. Tsikliras. *PLOS ONE*, **10**, e0117863.

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

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

Jetz, W., Sekercioglu, C.H. & Watson, J.E.M. (2008) Ecological correlates and conservation implications of overestimating species geographic ranges: *Overestimation of species ranges*. *Conservation Biology*, **22**, 110–119.

Kaschner, K., Rius-Barile, J., Kesner-Reyes, K., Garilao, C., Kullander, S., Rees, T. & Froese, R. (2015) AquaMaps: Predicted range maps for aquatic species, <www.aquamaps.org>

Kaschner, K., Tittensor, D.P., Ready, J., Gerrodette, T. & Worm, B. (2011) Current and future patterns of global marine mammal biodiversity ed S.J. Bograd. *PLoS ONE*, **6**, e19653.

Kaschner, K., Watson, R., Trites, A.W., Pauly, D. & others. (2006) Mapping world-wide distributions of marine mammal species using a relative environmental suitability (RES) model. *Marine Ecology Progress Series*, **316**, 2–3.

Klein, C.J., Brown, C.J., Halpern, B.S., Segan, D.B., McGowan, J., Beger, M. & Watson, J.E. (2015) Shortfalls in the global protected area network at representing marine biodiversity. *Scientific Reports*, **5**, 17539.

Martin, C., Fletcher, R., Jones, M., Kaschner, K., Sullivan, E., Tittensor, D.P., Mcowen, C., Geffert, J., Bochove, J. van, Thomas, H., Blyth, S., Ravilious, C., Tolley, M. & Stanwell-Smith, D. (2014) *Manual of Marine and Coastal Datasets of Biodiversity Importance*. United Nations Environment Programme.

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

Pimm, S.L., Jenkins, C.N., Abell, R., Brooks, T.M., Gittleman, J.L., Joppa, L.N., Raven, P.H., Roberts, C.M. & Sexton, J.O. (2014) The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, **344**, 1246752–1–1246752–10.

R Core Team. (2016) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ready, J., Kaschner, K., South, A.B., Eastwood, P.D., Rees, T., Rius, J., Agbayani, E., Kullander, S. & Froese, R. (2010) Predicting the distributions of marine organisms at the global scale. *Ecological Modelling*, **221**, 467–478.

Rondinini, C., Wilson, K.A., Boitani, L., Grantham, H. & Possingham, H.P. (2006) Tradeoffs of different types of species occurrence data for use in systematic conservation planning: Species data for conservation planning. *Ecology Letters*, **9**, 1136–1145.

Selig, E.R., Longo, C., Halpern, B.S., Best, B.D., Hardy, D., Elfes, C.T., Scarborough, C., Kleisner, K.M. & Katona, S.K. (2013) Assessing global marine biodiversity status within a coupled socio-ecological perspective ed F. Guichard. *PLoS ONE*, **8**, e60284.

Spalding, M.D., Fox, H.E., Allen, G.R., Davidson, N., Ferdaña, Z.A., Finlayson, M.A.X., Halpern, B.S., Jorge, M.A., Lombana, A.L., Lourie, S.A. & others. (2007) Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *BioScience*, **57**, 573–583.