A comparative analysis of AquaMaps and IUCN species distribution maps

JA, CO, BH

0.1 Significance

This is important because of some reason or another. . . !!!

---

1 Introduction

Climate change, ocean temperature rise, ocean acidification, sea ice loss, human impacts from fishing and introduction of nonnative species are likely to significantly alter the global distribution of marine species - but how well do we really understand the current distribution? Understanding species abundances and distributions is a critical component to assessing global biodiversity and detecting patterns and trends through time.

AquaMaps and IUCN marine species distribution maps are both global-scale spatial data collections that have been used to analyze marine species status (Halpern et al. 2012, Selig et al. 2013), global biodiversity patterns and trends (other references), predicting range shifts (Molinos et al. 2015), and setting conservation priorities (Klein et al. 2015). The two data collections ostensibly describe the same information, but significant differences in methodology, intent, and taxonomic and geographic sampling could lead to dramatically different understandings of our marine ecosystems.

To better understand the similarities and differences between the two data collections, we examined three fundamental questions:

- How do the two spatial data collections compare in their representation of global spatial distribution of species diversity?
- How do the two spatial data collections compare in their representation of species diversity across taxonomic groups?
- For species represented in both collections, how closely do the spatial distribution maps align?

The Ocean Health Index and MPA Gap Analysis study both provide case studies examined the implications of alternate data use scenarios on two large-scale studies of global biodiversity and conservation.

2 Understanding each dataset separately

The IUCN publishes species distribution maps for as collections of shapefiles for a range of taxonomic groups, including both terrestrial and marine organisms. Species spatial distributions are determined by experts based upon known occurrences of the species and expert understanding of range and habitat preferences. These maps are typically based on species occurrence records from databases such as ***OBIS*** and ***GBIF***; IUCN provides general guidelines to facilitate consistency in how the maps are generated. The IUCN presents these maps as 'limits of distribution'; they are meant to indicate that a species probably occurs within the defined polygon, but are not meant to indicate that the species is evenly distributed within the polygon. While the polygons roughly define regions of presence/absence, additional attributes provide information on extant/extinct ranges, native/introduced ranges, and seasonality.

IUCN publishes spatial distribution map sets only for taxonomic groups that have been "comprehensively assessed," i.e. in which at least 90% of the species within the taxonomic group have been evaluated. While this mitigates sampling bias within taxa, it also means that entire taxonomic groups remain unavailable until they have met this threshold of comprehensive assessment. As of December 2015, IUCN had published species distribution maps for nearly 4500 marine species across 24 taxonomic groups. (For this analysis, we did not consider IUCN range maps for bird species, as those data are hosted separately by BirdLife International.)

AquaMaps (Kaschner et al. 2013) offers species range maps based on modeled relative environmental suitability. For each assessed species, data from occurrence records such as OBIS and GBIF, published species databases such as FishBase, and expert knowledge are used to generate bioclimatic envelopes for sea surface temperature, salinity, depth, productivity, sea ice concentration, and distance to land. The species' suite of bioclimatic envelopes is compared to a map of environmental attributes on a 0.5 degree global grid, creating a global map of probability of occurrence, with cell values ranging from 0 to 1.00. These maps can then be reviewed by experts to fine-tune the parameters to better match known occurrences. As of December 2015, AquaMaps current native distribution maps have been produced for nearly 23,000 species, though only 5,600 of these species have been evaluated for the IUCN Red List. *See (Kaschner et al. 2013) for more detailed methods.*

---

3 Methods

3.1 Analysis of taxonomic and regional inclusion

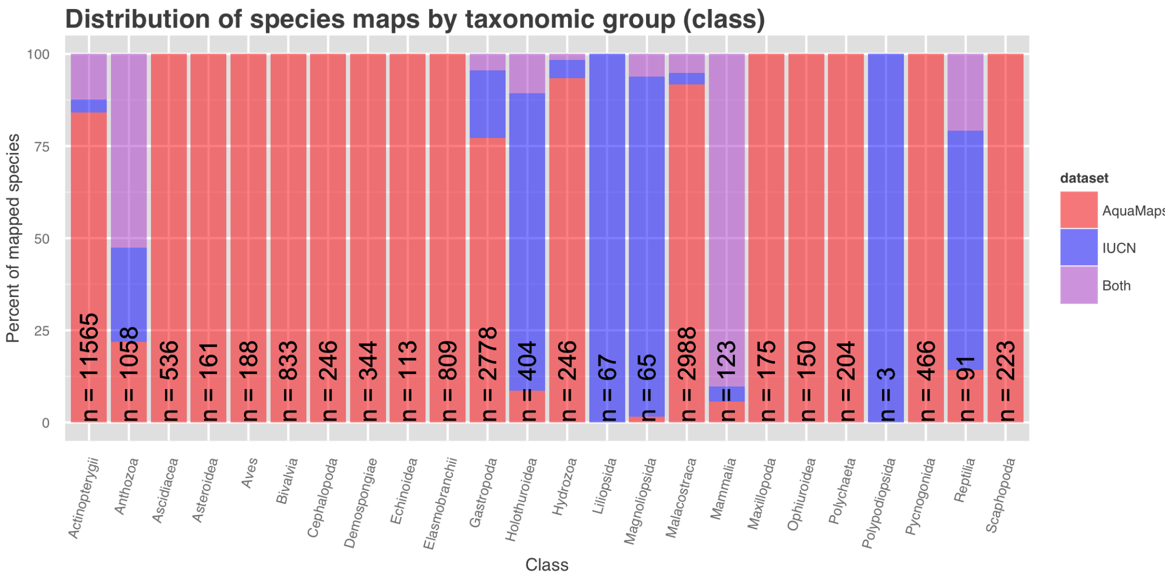
3.1.1 Overlap between assessed species

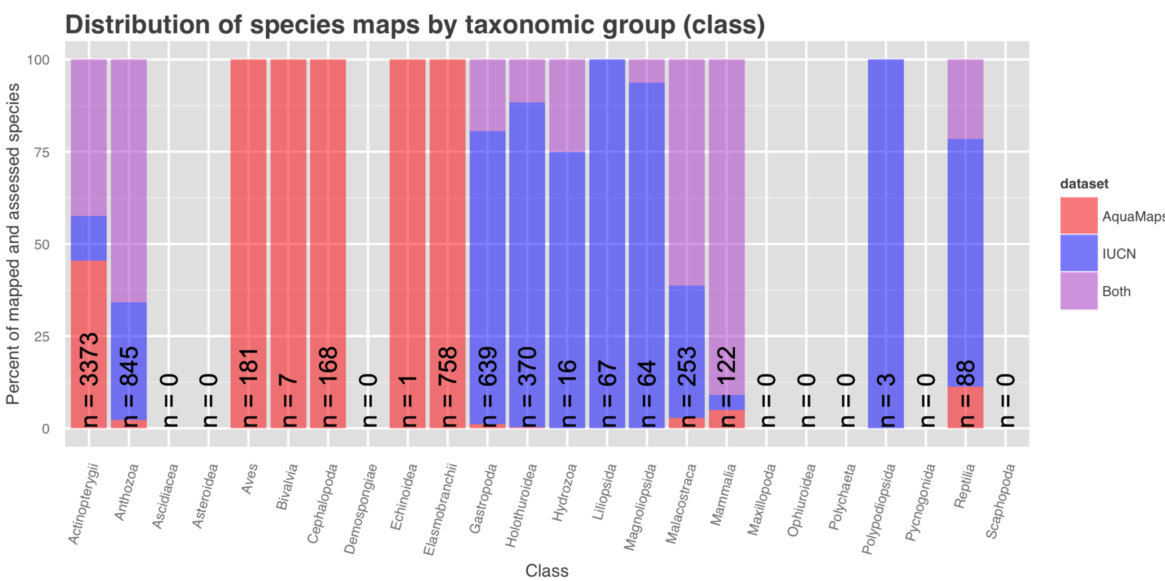
Due to differences in both methodology and intent, there are significant differences in the coverage afforded by the species distribution maps published by AquaMaps and those published by the IUCN. This is immediately clear from an examination of the number of species maps offered by each source, and the small proportion of species with maps from both sources.

Map set	n_{total}	n_{Red_List}
AquaMaps	22889	5246
IUCN	4343	all
Both	2455	all

3.1.2 Taxonomic distribution between data collections

Studies of species richness and biodiversity depend on sampling a representative cross section of ecosystems, and thus benefit from a broad representation across taxonomic groups. To examine the overall taxonomic distribution across the spatial data collections, we grouped species by taxonomic class and data source (IUCN, AquaMaps, or both), and examined the proportion of each class represented in each data source category. Since biodiversity studies often rely on both spatial distribution and indicators of species health, we further divided the data by whether species had been evaluated for the IUCN Red List of Threatened Species.



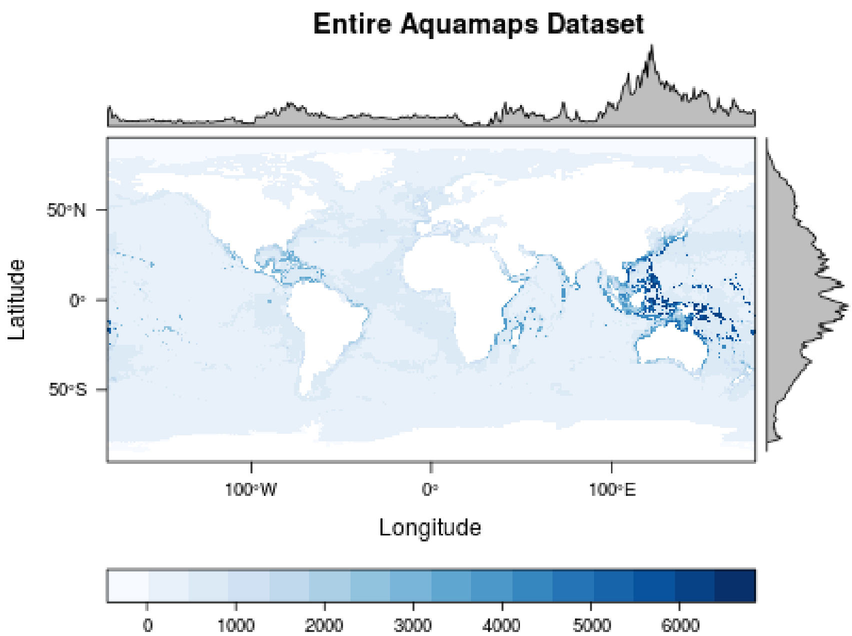


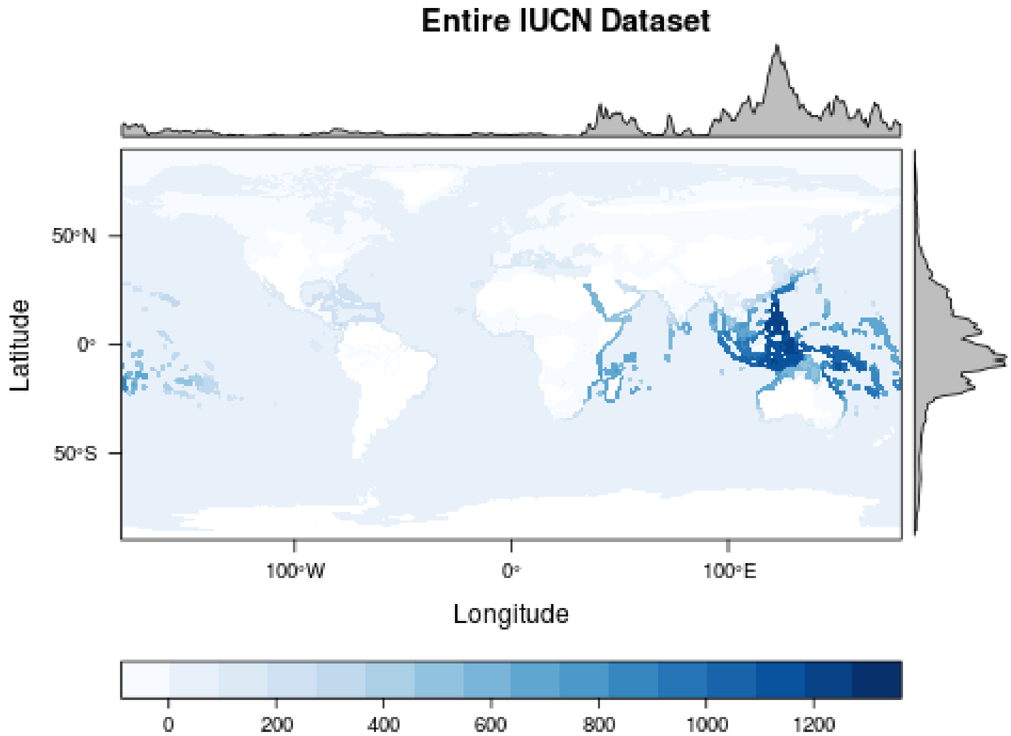
The IUCN only releases spatial data sets when a taxonomic group (on the scale of order or family) has been comprehensively assessed. As such, spatial data for many taxonomic classes are unavailable, and within a class, the assessed sub-groups may not represent the entire class. For example, as of this writing, IUCN has released no spatial data for class Elasmobranchii (cartilaginous fishes including sharks and rays); and while IUCN offers a large number of maps within class Actinopterygii (ray-finned bony fishes), the available maps include only a few primarily tropical taxonomic sub-groups, such as wrasses, damselfish, butterflyfish, tunas, and billfishes, but are missing economically important subgroups including salmon, rockfish, and clupeids. However, IUCN's criterion of comprehensive assessment greatly reduces the risk of sample bias within the bounds of the assessed taxonomic groups.

The release of AquaMaps distribution maps is not limited to comprehensively-assessed taxa, and maps are available across a much larger range of taxonomic classes; however, there is no guarantee that the list of species included within each class is a representative cross section of the entire class.

Some global-scale studies of biodiversity, such as the Ocean Health Index, require both spatial distribution information and non-spatial information such as extinction risk, population trend, or native/non-native status provided by IUCN Red List assessments. While all species represented in IUCN distribution maps have also been evaluated for Red List inclusion, we found that only 22.9% (5246 of 22889) of AquaMaps species have been similarly evaluated, severely limiting the catalog of available species maps for studies of biodiversity and ecosystem health.

Broad taxonomic inclusion is critical for understanding distribution and health of ecosystems at any spatial scale, but global and large regional studies of species diversity also require representative sampling across regions. To compare the spatial representation of the two data collections directly, we first rasterized the IUCN species polygons to the same 0.5° cells as the AquaMaps species maps; species presence within a cell was determined by any non-zero overlap of a species polygon with the cell, and species richness per cell was simply the sum of the species present. For the AquaMaps data collection, we determined per-cell species richness by summing all species with non-zero probability of occurrence. We represented relative distribution of species richness for each data collection by plotting average species count against latitude and longitude. To highlight the differences between the two data collections, we also examined the relative difference in per-cell species richness against latitude and longitude.





The balance of species represented in IUCN spatial data skews toward tropical latitudes compared to the distribution represented in AquaMaps distribution maps. This may be driven by the fact that the taxonomic distribution of IUCN distribution maps focuses heavily on coral reef-associated species (see fig. XXX). While both data collections indicate highest species richness in the Coral Triangle and western Indian Ocean, the AquaMaps data collection shows a relatively larger representation of species in the Atlantic, Caribbean, and eastern Pacific than does the IUCN data collection.

Each data collection offers spatial distribution information for large numbers of species. However, the data collections vary in terms of taxonomic coverage and regional coverage. For spatial assessments of biodiversity, the choice of one data collection over the other is likely to create significantly different results. For studies confined to a narrow range of taxa or to a narrow spatial scale, one data collection may offer an advantage over the other in the number of species maps available. For global scale biodiversity studies, however, the selection of one data collection over the other will entail tradeoffs in spatial coverage, taxonomic breadth, and taxonomic depth. ***find a reference that describes what makes a "good" data collection for global biodiversity, e.g. OHI*** *or maybe species richness vs diversity vs "health" or whatnot*

3.2 Analysis of spatial alignment by species

3.2.1 Defining spatial alignment between the two data collections

For species described in both spatial data collections, we would expect to see some spatial correlation, both in the global pattern of species distribution and the extent of species range. Large discrepancies in distribution and range area could indicate species that require further study to consolidate expert knowledge and improve species distribution models.

Using genus and species binomials to identify paired maps, we selected the subset of marine species that have range maps in both IUCN and AquaMaps current native distribution (n = 2234). Assigning "presence" values using the same methodology above, we found that for 69.6% of species, the IUCN distribution maps indicated a larger species range than the range indicated by AquaMaps.

Overlaying distribution maps from the two data sets for a given species, we calculated two dimensions of spatial alignment: *distribution alignment*, which we defined as the ratio of the smaller range that overlapped with the larger range; and *area alignment*, which we defined as the ratio of the smaller range to the larger range.

$$Match_{dist} = \frac{A_{overlap}}{A_{larger}} * 100\%$$

$$Match_{area} = \frac{A_{smaller}}{A_{larger}} * 100\%$$

For a species whose distribution is well understood, and whose range is described in both data collections, we would expect to see a value near 100% for each dimension of alignment, indicating near-total inclusion of the smaller range within the larger (*distribution alignment*) and very similar extent of range (*area alignment*).

For each of the 2234 species whose range is described by both data collections, we calculated the distribution alignment and the area alignment, and plotted the species map pairs on a scatter plot of distribution alignment vs area alignment (Figure 1).

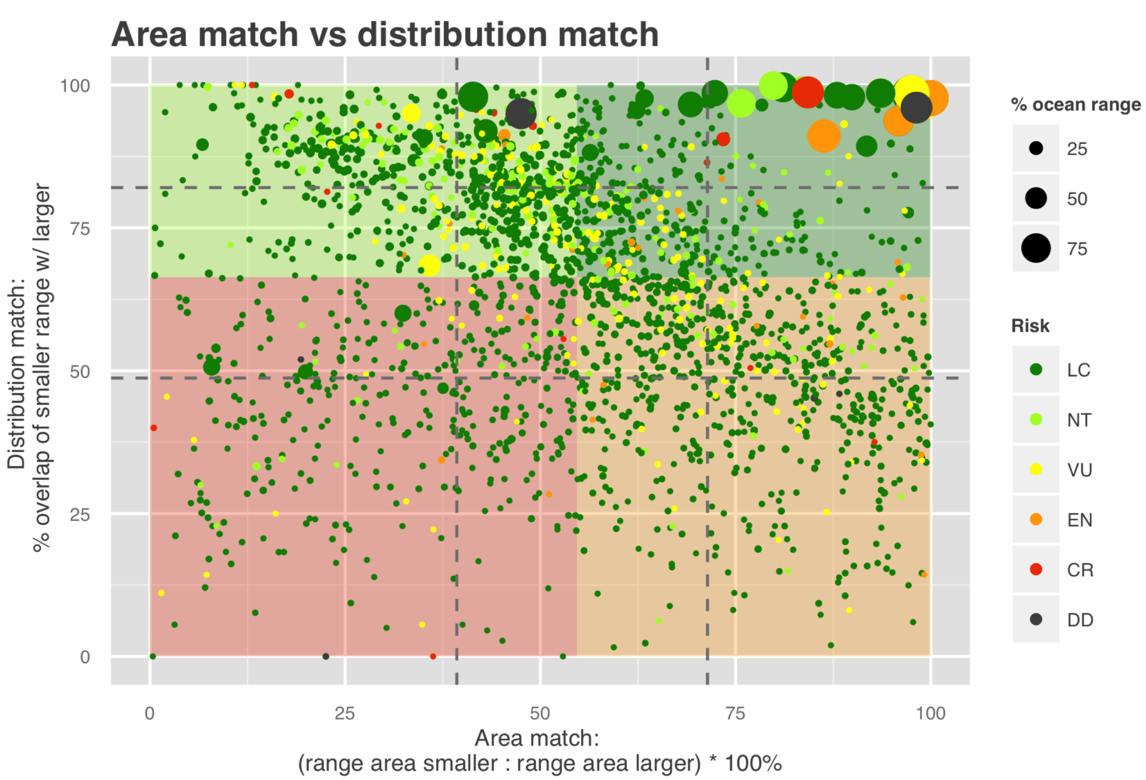


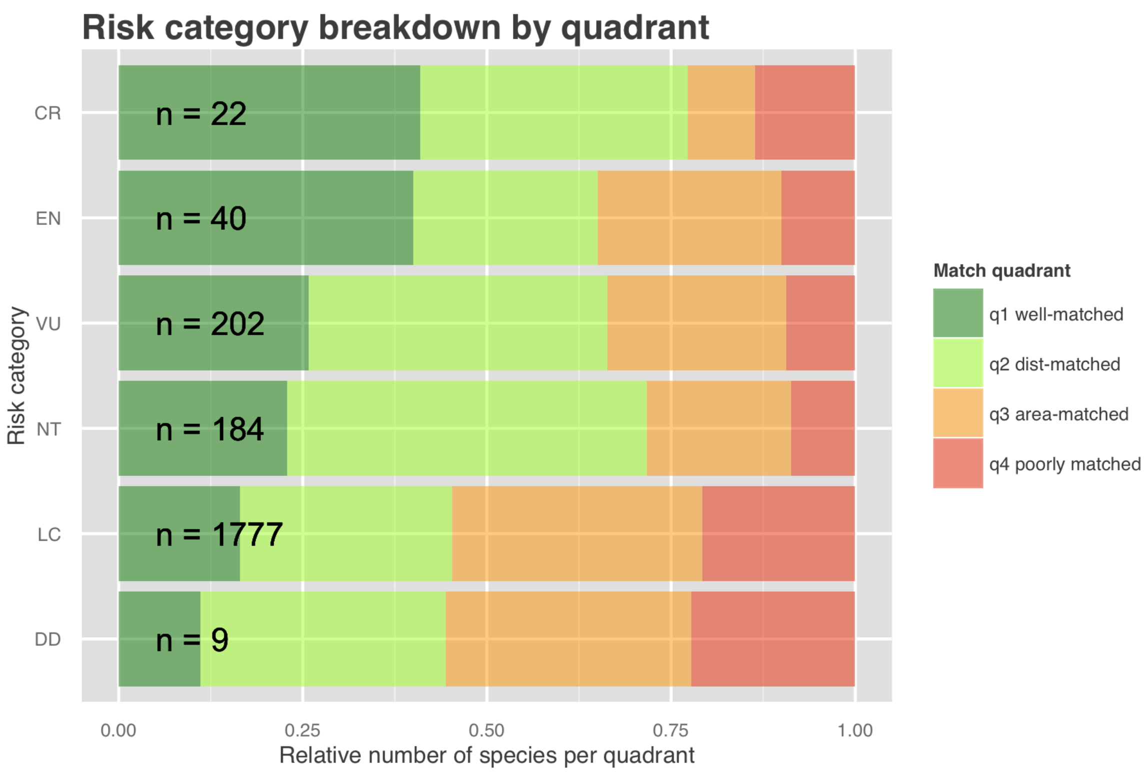
Figure XXX: Distribution alignment vs area alignment for 2234 species mapped in both Aquamaps and IUCN species distribution maps.

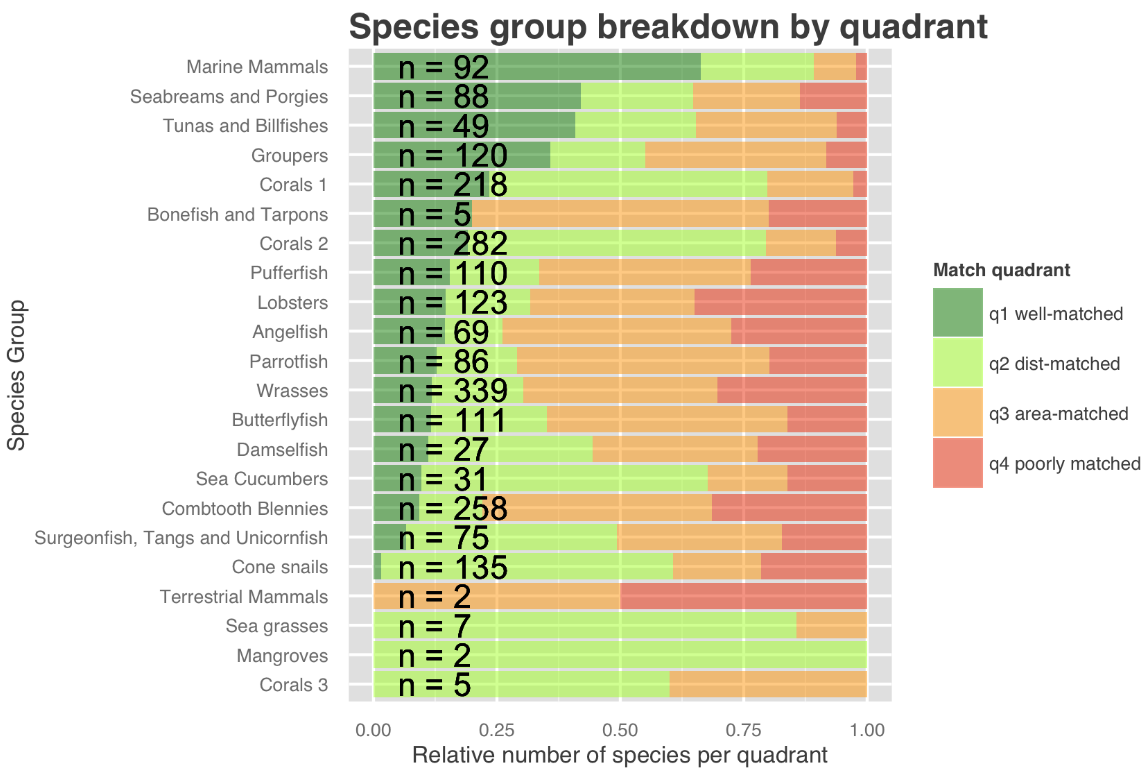
We divided the plot into quadrants based on the median values for each axis to differentiate between four different qualities of alignment:

Description	Distribution match	Area match
q1: Well-aligned	>= 66.4 %	>= 54.7 %
q2: Distribution-aligned	< 66.4 %	>= 54.7 %
q3: Area-aligned	>= 66.4 %	< 54.7 %
q4: Poorly-aligned	< 66.4 %	< 54.7 %

caption?

The upper right quadrant (quadrant 1) comprises species whose maps largely agree (better than median value) in both spatial distribution and the area of described ranges. For these species, the expert-drawn range map and the bioclimatic envelope model produce similar predictions of species distribution. (n = 407; 18.2 %) The other quadrants each suggest a different story. Species map pairs that fall within the upper left quadrant (quadrant 2: well-matched in distribution but poorly-matched in area) indicate species in which the range of the smaller map falls generally within the larger map, but the larger map may include more generous buffers or may include areas unrepresented in the smaller map. (n = 708; 31.7 %) *The lower right quadrant (quadrant 3: well-matched in area but poorly matched in distribution) includes species for which the two maps generally agree in range area, but disagree in where those ranges occur. Disagreement in distribution seems to indicate a greater issue than disagreement in area extent, so we categorized map pairs in this quadrant as being more problematic than map pairs in either of the upper quadrants. (n = 706; 31.6 %)* The lower left quadrant (quadrant 4: poorly matched in both dimensions) indicates species for which the map pairs fail to agree in both area and distribution. (n = 413; 18.5 %)





Breaking down the quadrants by IUCN extinction risk categories, species with higher extinction risk tend to be better matched between the two data collections; this is not surprising, as these at-risk species are likely to face closer expert scrutiny than species ranked as Least Concern.

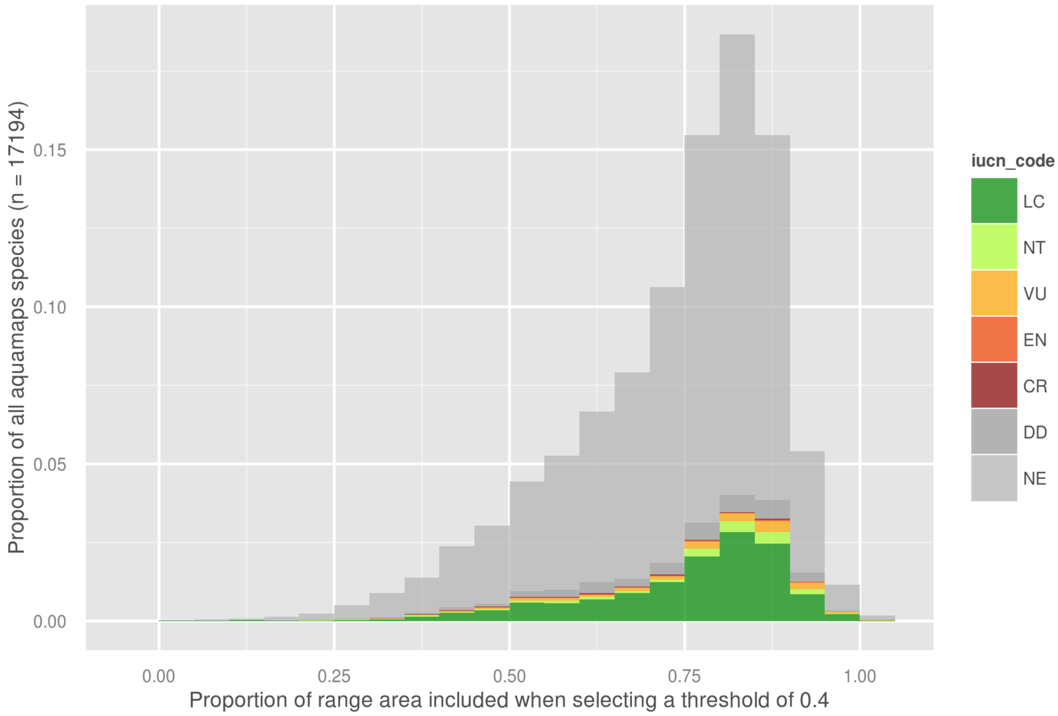
When broken down by taxonomic group, we found that certain taxonomic groups were far more likely than others to be spatially well-matched; in particular, wide-ranging pelagic organisms such as marine mammals, tunas, and billfishes were well-matched (quadrants 1 and 2), . Coral species are predominantly found in quadrant 2, indicating that while general global distribution trends seem to be consistent between the two data collections, predicted areas are not well-matched; AquaMaps accounts for seafloor depth as a limiting factor, while IUCN maps for coral species often include wide buffers that would extend far beyond appropriate depth for coral habitat, and thus likely overestimate the actual distribution.

Other hypotheses/case studies to highlight other possible mechanistic differences?

3.3 AquaMaps: Effect of changing "presence" threshold on area representation

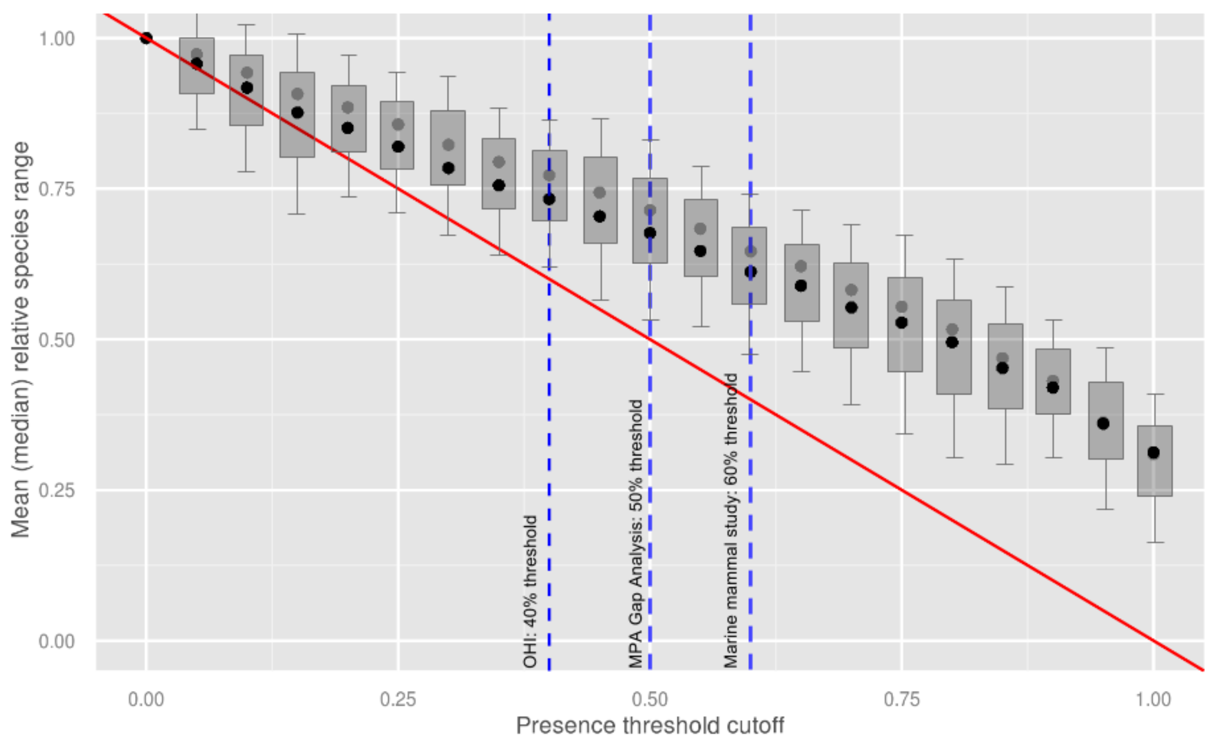
Why should anyone care about AquaMaps thresholds. ***Identify several studies that used different thresholds - OHI, MPA gap analysis, AquaMaps mammal distribution*** to justify this micro-analysis

what would it mean if there were non-linearities or generally linear trend?



add a mean/median vline on this? do we need this figure? it looks cool though... why broken down by risk category? Ditch NE category since most studies probably care about the species health as well as distribution?

Identify several studies that used different thresholds - OHI, MPA gap analysis, AquaMaps mammal distribution to justify this micro-analysis



Effect of selecting different thresholds on total "present" area - linearish, no big steps or drops to indicate sudden change in tradeoff. Do regression - shallow slope indicates that sensitivity is less than 1:1.

---

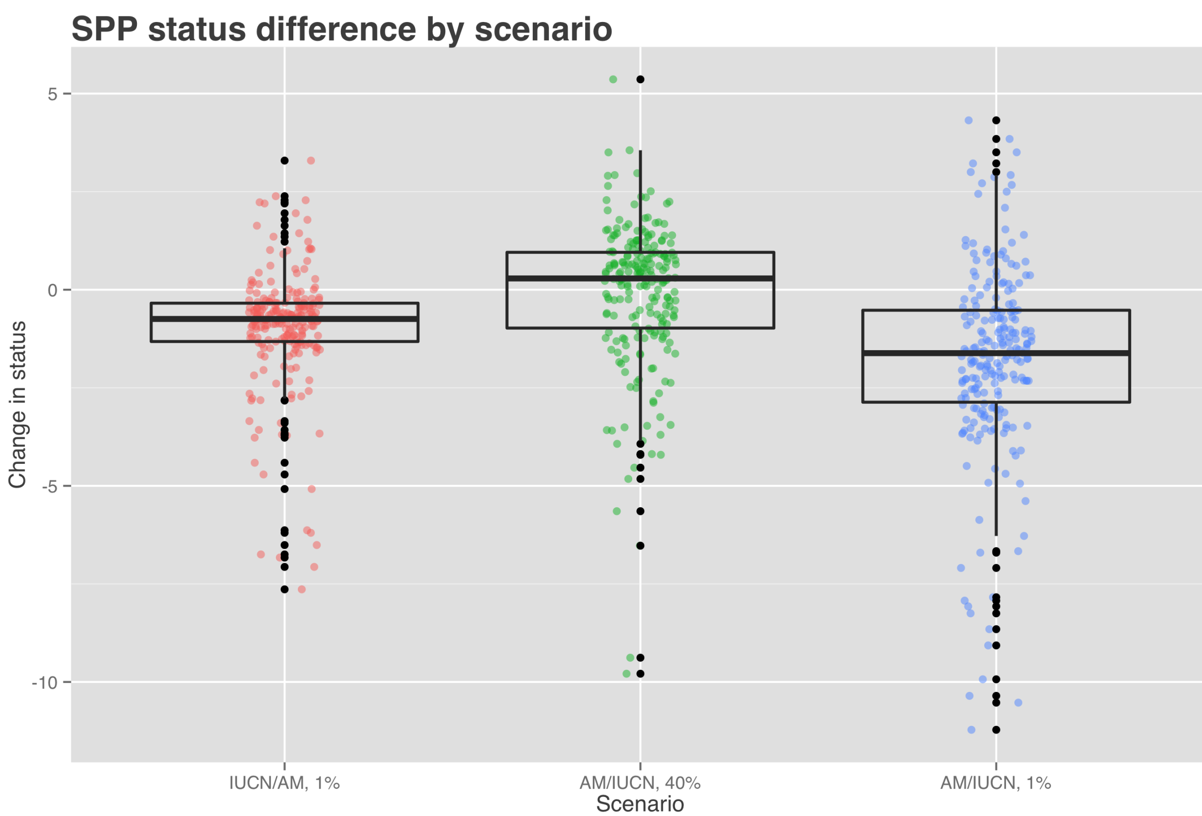
4 Implications

4.1 Application to OHI

The global Ocean Health Index (Halpern et al. 2012), an index made up of 10 goals, utilizes both of these datasets to inform the Biodiversity goal. The biodiversity goal is made up of two parts; Species and Habitats. As it is currently calculated, the Species subgoal uses the IUCN range map data and IUCN Red List conservation status to calculate an area-weighted mean species status in each of 221 exclusive economic zones. For those species that are are represented within the IUCN range maps, AquaMaps distribution maps were used. Species with no spatial data in either data collection were excluded, as were species with insufficient information to determine conservation status (including species listed as not evaluated or data deficient).

This decision to use IUCN data preferentially to AquaMaps data was made based on the fact that experts are developing the IUCN maps and AquaMaps data is based off a computer model, which in some cases is then consulted by an expert and appropriately altered. ***is this really the reasoning behind the decision? let's get the real reason in here***

Figure [X] shows the change in status score for the Species Subgoal within the global Ocean Health Index under three different scenarios.



Scenario	Priority data source	AquaMaps presence threshold
Scenario 0 (current)	IUCN	>= 40%
Scenario 1	IUCN	> 0%
Scenario 2	AquaMaps	>= 40%
Scenario 3	AquaMaps	> 0%

Scenario 1 shows the effect of reducing the threshold for AquaMaps presence. This will always increase the apparent range of any AquaMaps species in the analysis. The slight decrease in average score suggests increased spatial representation of species at greater risk of extinction.

Scenario 2 shows the effect of prioritizing AquaMaps data over IUCN, while maintaining the same presence threshold. This will have different effects depending on the species; in general, AquaMaps ranges are smaller than IUCN ranges, so we could predict an average decrease in range for these overlapping species. *The slight bump in mean score indicates that a slight increased representation of low-risk species counterbalances a decreased representation of high-risk species.* ***can we really make any generalizations here? not really anything useful without further analysis***

Scenario 3 shows the combination of the two previous scenarios. The larger decrease in score indicates that the ***hmm gotta think about this a little more - come back to it***

4.2 Application to MPA Gap Analysis

yeah still gotta do something on this one

---

5 Conclusion

---

6 References

Key AquaMaps publications:

Kaschner, K., D.P. Tittensor, J. Ready, T Gerrodette and B. Worm (2011). Current and Future Patterns of Global Marine Mammal Biodiversity. PLoS ONE 6(5): e19653. PDF

Ready, J., K. Kaschner, A.B. South, P.D Eastwood, T. Rees, J. Rius, E. Agbayani, S. Kullander and R. Froese (2010). Predicting the distributions of marine organisms at the global scale. Ecological Modelling 221(3): 467-478. PDF

Papers based on AquaMaps: ***Use these for section 4?***

Jones, M.C., S.R. Dyeb, J.K. Pinnegar and W.W.L. Cheung (2012). Modelling commercial fish distributions: Prediction and assessment using different approaches. Ecological Modelling 225(2012): 133-145. PDF

Coll, M., C. Piroddi, J. Steenbeek, K. Kaschner, F. Ben Rais Lasram et al. (2010). The biodiversity of the Mediterranean Sea: estimates, patterns, and threats. PLoS ONE 5(8): e11842. PDF

Martin C.S., Fletcher R., Jones M.C., Kaschner K., Sullivan E., Tittensor D.P., Mcowen C., Geffert J.L., van Bochove J.W., Thomas H., Blyth S., Ravillious C., Tolley M., Stanwell-Smith D. (2014). Manual of marine nd coastal datasets of biodiversity importance. May 2014 release. Cambridge (UK): UNEP World Conservation Monitoring Centre. 28 pp. (+ 4 annexes totalling 174 pp. and one e-supplement). PDF

7 References

García Molinos, Jorge, Benjamin S. Halpern, David S. Schoeman, Christopher J. Brown, Wolfgang Kiessling, Pippa J. Moore, John M. Pandolfi, Elvira S. Poloczanska, Anthony J. Richardson, and Michael T. Burrows. "Climate Velocity and the Future Global Redistribution of Marine Biodiversity." Nature Climate Change advance online publication (August 31, 2015). doi:10.1038/nclimate2769.

Halpern, Benjamin S., Catherine Longo, Darren Hardy, Karen L. McLeod, Jameal F. Samhouri, Steven K. Katona, Kristin Kleisner, et al. "An Index to Assess the Health and Benefits of the Global Ocean." Nature 488, no. 7413 (August 30, 2012): 615–20. doi:10.1038/nature11397.

Klein, Carissa J., Christopher J. Brown, Benjamin S. Halpern, Daniel B. Segan, Jennifer McGowan, Maria Beger, and James E.M. Watson. "Shortfalls in the Global Protected Area Network at Representing Marine Biodiversity." Scientific Reports 5 (December 3, 2015): 17539. doi:10.1038/srep17539.

Selig, Elizabeth R., Catherine Longo, Benjamin S. Halpern, Benjamin D. Best, Darren Hardy, Cristiane T. Elfes, Courtney Scarborough, Kristin M. Kleisner, and Steven K. Katona. "Assessing Global Marine Biodiversity Status within a Coupled Socio-Ecological Perspective." PLoS ONE 8, no. 4 (April 11, 2013): e60284. doi:10.1371/journal.pone.0060284.

- *IUCN reference:*
  - https://www.conservationtraining.org/course/view.php?id=217&lang=en
  - http://www.iucnredlist.org/technical-documents/red-list-training/iucnspatialresources