A comparative analysis of AquaMaps and IUCN species distribution maps

JA, CO, BH

0.1 Main message

Understanding the differences in marine species range maps from IUCN and AquaMaps, and the implications of using one or both.

Species range maps are useful in conservation and management, as well as understanding the spatial distribution of movement and overlap of species in different habitats globally. Marine species prove more difficult to map based on their inherent nature of being difficult to observe and track. Two main sources for marine species are the IUCN and AquaMaps. These two datasets differ in methodology and final products, yet the differences between the two are not well documented *(though Aquamaps, and probably IUCN, each probably have some validation studies)*. AquaMaps uses a computer model based on preferential environmental parameters to predict species locations and probability of occurrence, while the IUCN uses expert insight to assign spatial boundaries around known locations of marine species. Approximately 1,900 species are found in both datasets and there are very large differences between these maps. Depending on data usage, choosing one of these datasets over the other can lead to potentially large differences in results as exemplified here by the Ocean Health Index as an example.

---

1 Introduction

can we set up a compelling story line here?

Mapping species abundances and distributions is a critical component to understanding global biodiversity and detecting patterns and trends through time. There have been multiple efforts to map terrestrial, freshwater and marine species from local to global scales. These pieces of information inform our understanding of taxonomic assemblages, biogeographical limits to species ranges, suitable habitat for species.

These two spatial datasets have been used to analyze species status globally (Halpern et al. 2012, Selig et al. 2013), global biodiversity patterns and trends, predicting range shifts (Molinos et al. 2015),and setting conservation priorities (Klein et al. 2015). Yet there have been no comprehensive comparisons of these two datasets to identify the differences and provide insight into how using one or the other can influence the outcome of an analysis.

Differences in methodology (model-driven vs expert knowledge-driven); differences in intent (what does it communicate); differences in represented species and regions

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

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

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

- sensitivity analysis to AquaMaps presence threshold - diff reports use diff thresholds; what is the likely effect? is there a "correct" threshold?
  - this seems like a sub-analysis to justify choice of 0% threshold etc.
  - Should this just be in supplemental info? This could be important though...

2 Understanding each dataset separately

2.1 IUCN species distribution maps

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 dependent on observation records from databases ***such as GBIF other examples?***, and 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 12,000 marine species across 26 taxonomic groups, though fewer than 4,500 of these have been formally assessed under the IUCN Red List of Threatened Species.

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

2.2 AquaMaps standardized range maps of marine species

AquaMaps (Kaschner et al. 2013) offers species range maps based on modeled relative environmental suitability. For each assessed species, data from observations and expert knowledge are used to generate bioclimatic envelopes for sea surface temperature, salinity, depth, productivity, and where applicable, sea ice cover and/or 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*** (or ***relative environmental suitability?*** (see Kaschner et al 2006 and 2011)), with cell values ranging from 0 (unsuitable) to 1.00 (highly suitable). These maps can then be tweaked by experts altering the parameters and regenerating the modeled map. As of December 2015, nearly 23,000 species have been mapped, though only 5,600 of these species have been evaluated for the IUCN Red List. ***Expert mode to further refine models?*** See (Kaschner et al. 2013) for more detailed methods.

- ***how is expert insight used? how frequently? new data set includes "reviewed" field*** expert insight seems to be used to generate the environmental suitability envelope for species, to provide sightings info that use used in training the model, and to exclude species from regions of known absence; but the model generates the maps. However, some species maps are reviewed by experts (approved? changed? what happens in the review)

2.3 Comparison between AM and IUCN

- Differences in intent: capture all observations and be inclusive as possible, vs model suitable habitats
- Differences in methodology - expert-defined range maps (over-predicting) vs sightings (under-predicting) vs models: from **Kaschner et al 2011**: *Similar to the trade-offs of different habitat prediction modeling approaches [11], these two methods lie on* **opposite ends of a spectrum from potentially over-predicting expert-derived (range maps) to potentially underpredicting (empirical sighting surveys) range sizes**
  – are there other sources (non-IUCN and AM sources) that can back this up?
- Differences in the interpretation: probability/RES, vs. presence "probably within this polygon, but not everywhere within" - apples to oranges.

---

3 Analysis of taxonomic and regional inclusion

3.1 Overlap between assessed species

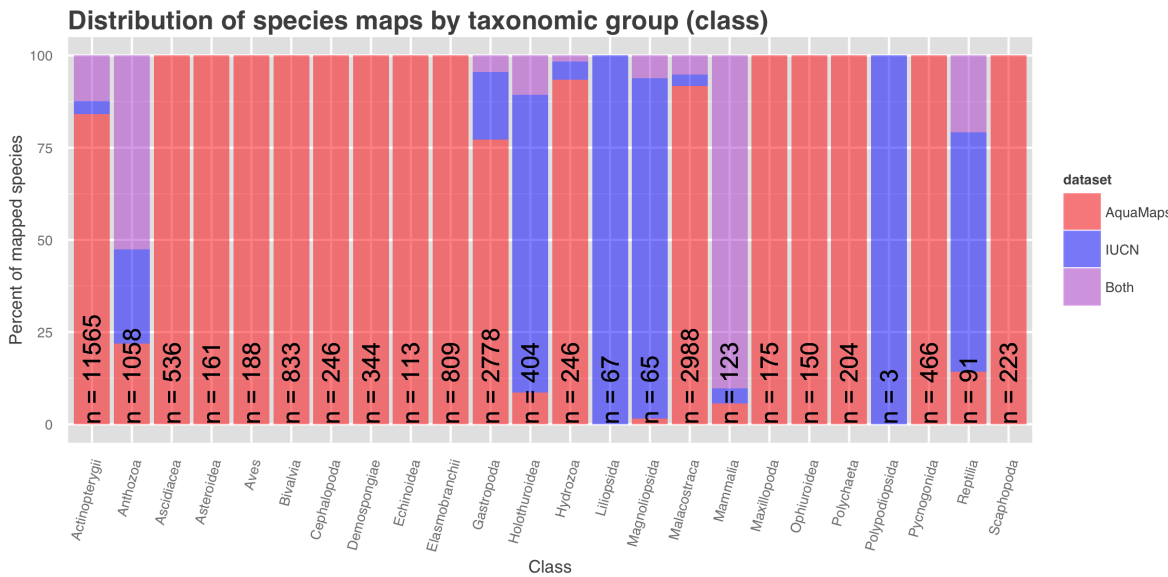
Due to differences in both methodology and intent, there are significant differences in the coverage afforded by the species distribution maps published by AquaMaps and those published by the IUCN. This is immediately clear from an examination of the number of species maps offered by each source, and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

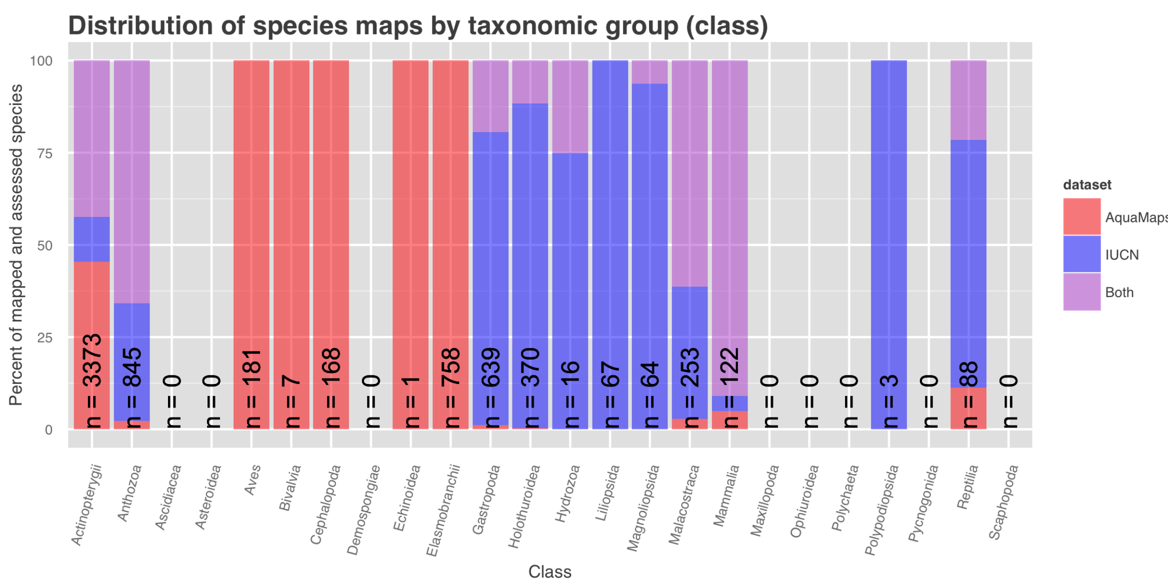
Map set	n_{total}	n_{Red_List}
AquaMaps	22889	5246
IUCN	4343	all
Both	2455	all

Should we break it down by "reviewed" Aquamaps as well?

The AquaMaps species list includes all species available as of December 2015. The IUCN species list includes all species maps available directly from the IUCN Red List website (http://www.iucnredlist.org/technical-documents/spatial-data) as of December 2015, filtered to include only species with habitat listed as "marine." For this analysis, we did not include IUCN range maps for bird species, as those data are hosted separately by BirdLife International.

3.2 Taxonomic distribution between data sets





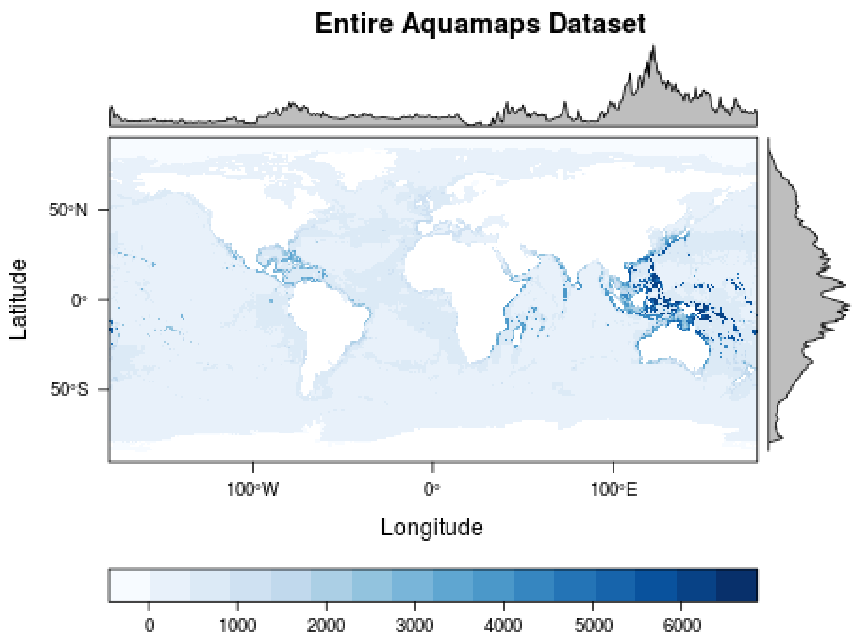
The IUCN only releases spatial data sets when a taxonomic group (on the scale of order or family) has been comprehensively assessed. As such, spatial data for many taxonomic classes are not available, and within a class, the assessed sub-groups may not represent the entire class. For example, as of this 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. However, IUCN's criterion of comprehensive assessment greatly reduces the risk of sample bias within the bounds of the assessed taxonomic groups.

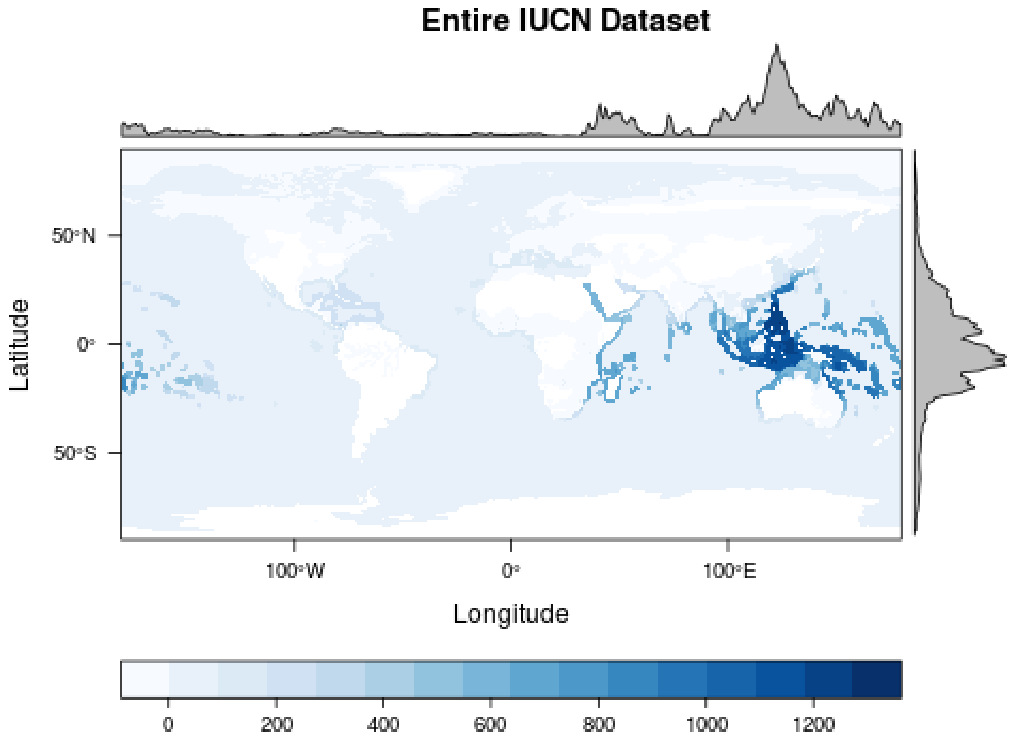
The release of AquaMaps distribution maps is not limited in the same way, and maps are available across a much larger range of taxonomic classes; however, there is no guarantee that the list of species included within each class is a representative cross section of the entire class. ***How do they choose their species***

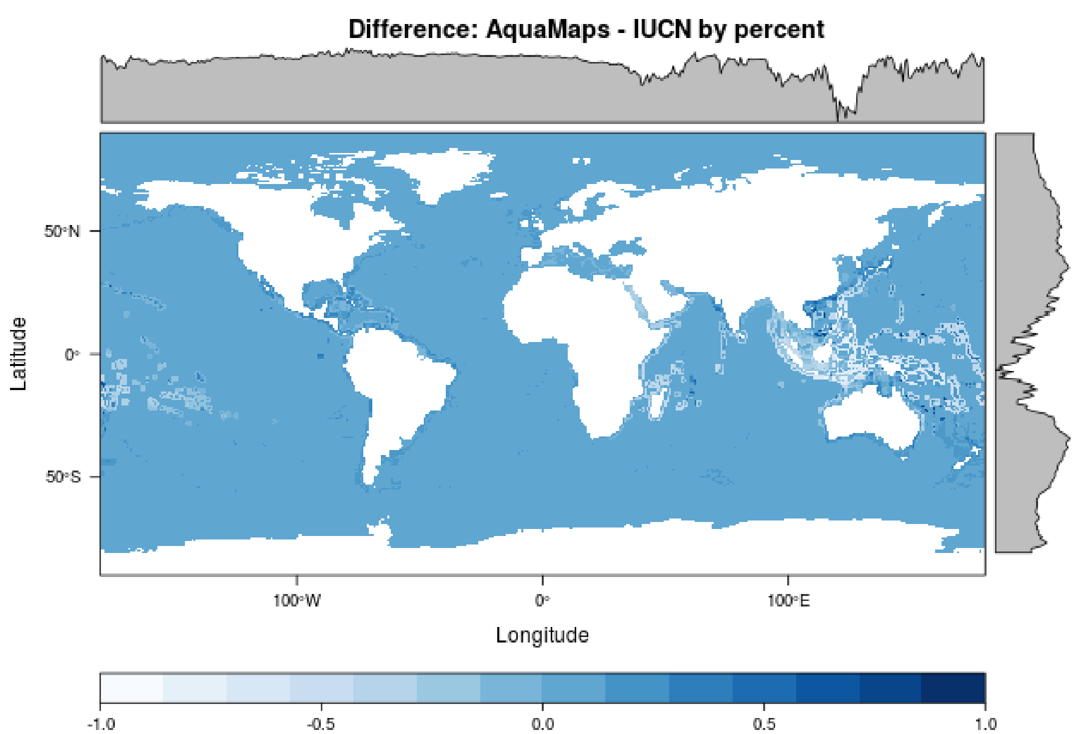
to assess?

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 species diversity and ecosystem health.

3.3 Spatial global distribution between data sets




Entire IUCN Dataset
50°N
0°
50°S
Latitude
100°W
0°
100°E
Longitude
0
200
400
600
800
1000
1200


Difference: AquaMaps - IUCN by percent
50°N
0°
50°S
Latitude
100°W
0°
100°E
Longitude
-1.0
-0.5
0.0
0.5
1.0

The balance of species represented in IUCN spatial data skews toward tropical latitudes compared to the distribution represented in AquaMaps distribution maps. (The tropical skew of IUCN species is also evident from an examination of the taxonomic groups included in the IUCN map set; see Fig. XXX). While both data sets indicate highest species richness in the Coral Triangle and western Indian Ocean, the AquaMaps data set shows a relatively larger representation of species in the Atlantic, Caribbean, and eastern Pacific than does the IUCN data set.

Each data set offers spatial distribution information for large numbers of species. However, the data sets vary in terms of taxonomic coverage and regional coverage. For spatial assessments of biodiversity, the choice of one data set 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 set 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 set over the other will entail tradeoffs in spatial coverage, taxonomic breadth, and taxonomic depth. ***find a reference that describes what makes a "good" data set for global biodiversity, e.g. OHI*** *or maybe species richness vs diversity vs "health" or whatnot*

---

4 Analysis of spatial alignment by species

4.1 Defining spatial alignment between the two data sets

AquaMaps and IUCN marine species distribution maps ostensibly describe the same information, and while differences in methodology and input data will inevitably lead to variation between range maps for any given species, we would expect to see some spatial correlation.

To facilitate our analysis, we selected the subset of marine species that have range maps in both IUCN (filtering out ranges identified as extinct) and AquaMaps current native distribution (n = 2234), and used genus and species binomials to identify paired maps. We then rasterized the IUCN shapefiles to match the Aquamaps half-cell raster. We flattened each map by assigning to each cell a simple "presence/absence" value:

- For IUCN data, we defined "presence" within a cell as any non-zero overlap with a range map polygon.
- For AquaMaps data, we defined presence as any cell with non-zero probability ***(or RES?)***.

- In both cases, these criteria for presence will maximize the number of cells included.

Using these criteria, 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 match, which we defined as the ratio of the smaller range that overlapped with the larger range; and area match, 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 sets, 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).

4.2 Comparing spatial alignment between the two data sets

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

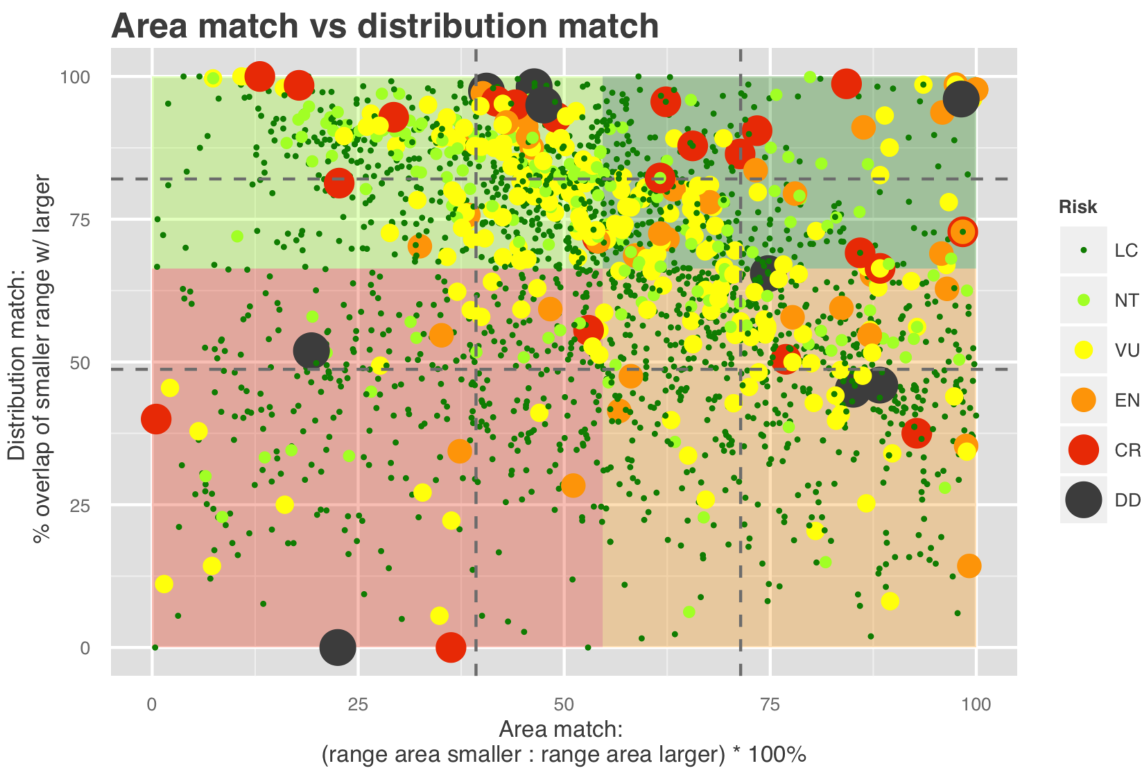


Figure XXX: Distribution alignment vs area alignment for 2234 species mapped in both Aquamaps and IUCN species distribution maps.

ditch the species identifiers here? or highlight ones that are used in discussion of possible mechanistic explanations for differences?

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 %

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 %)

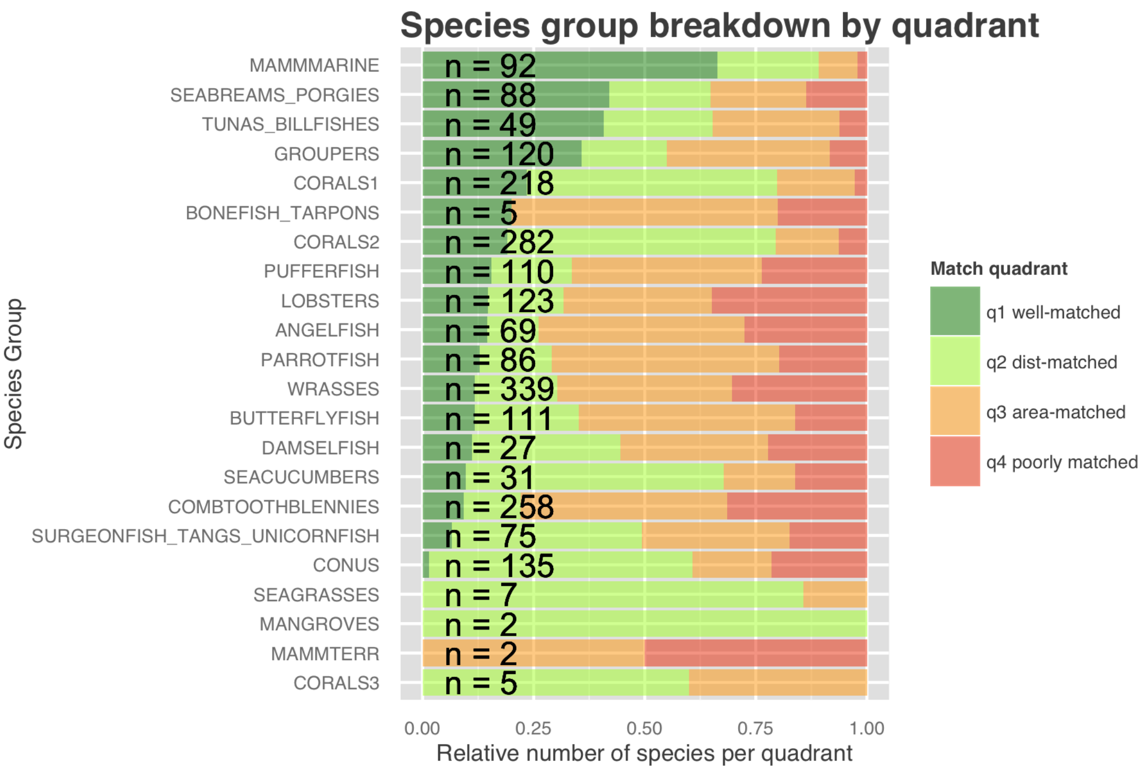


Figure 2 divides the 2234 species map pairs into taxonomic groups as classified by the published IUCN range maps, and examines the breakdown of each taxonomic group by quadrant.

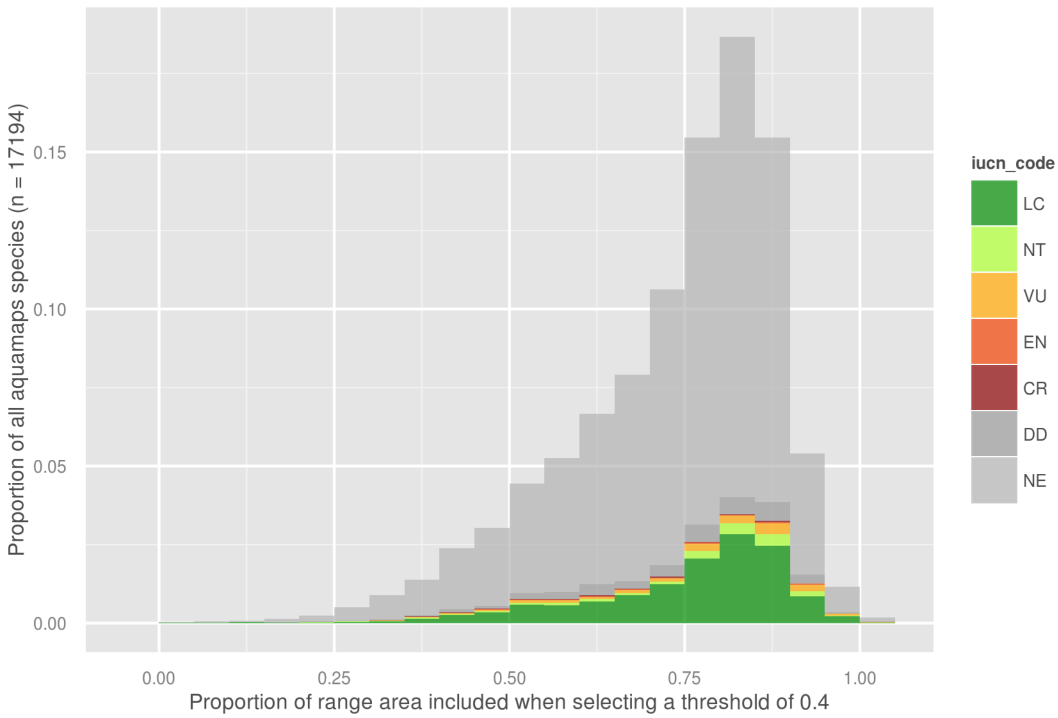
4.2.1 Mechanistic explanations for some of the variation?

We were not surprised to find that map pairs for marine mammals, tunas, and billfishes were well-matched (quadrants 1 and 2); many of these species are global-ranging pelagic organisms, with some ranges encompassing nearly the entire ocean.

Map pairs for corals, most of which fall into quadrant 2, **hypothesis... check a bunch of maps to see if this is a real trend** *generally agree well in distribution, but in many cases IUCN ranges include buffers that extend much farther from coastlines than AquaMaps modeled ranges, which use seafloor depth as a limiting factor.*

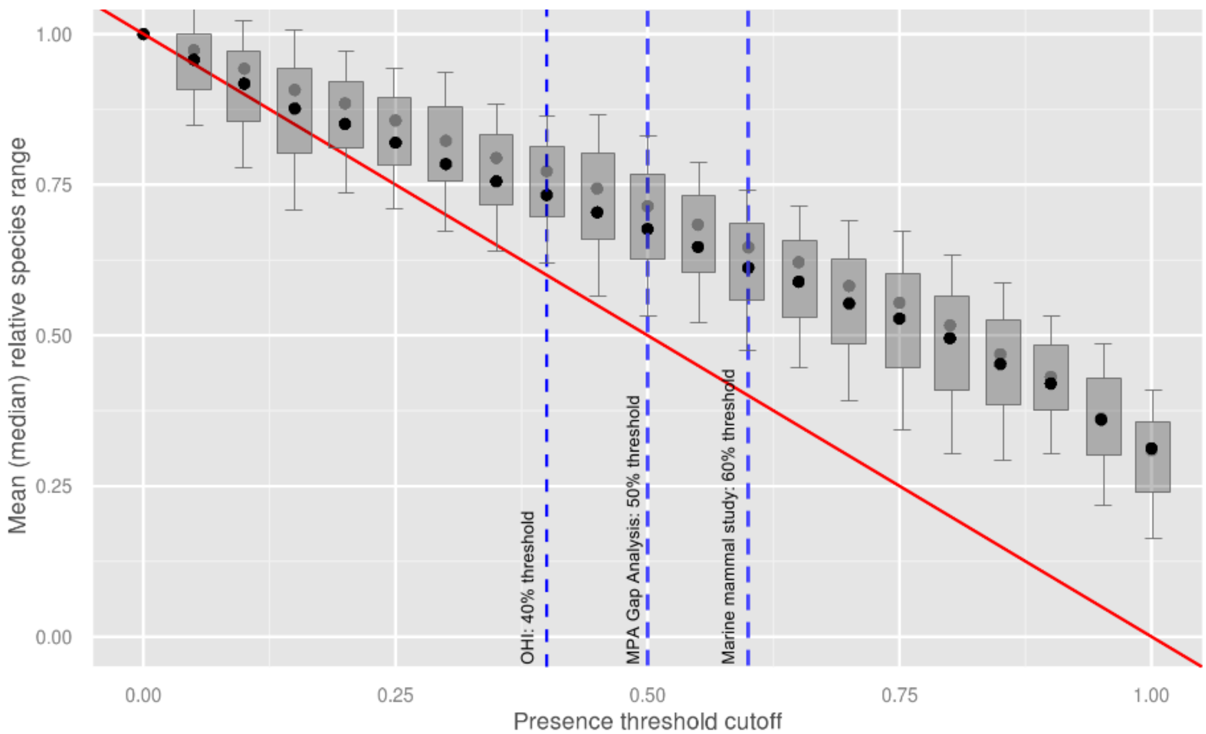
Other hypotheses/case studies to highlight possible mechanistic differences?

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



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.

---

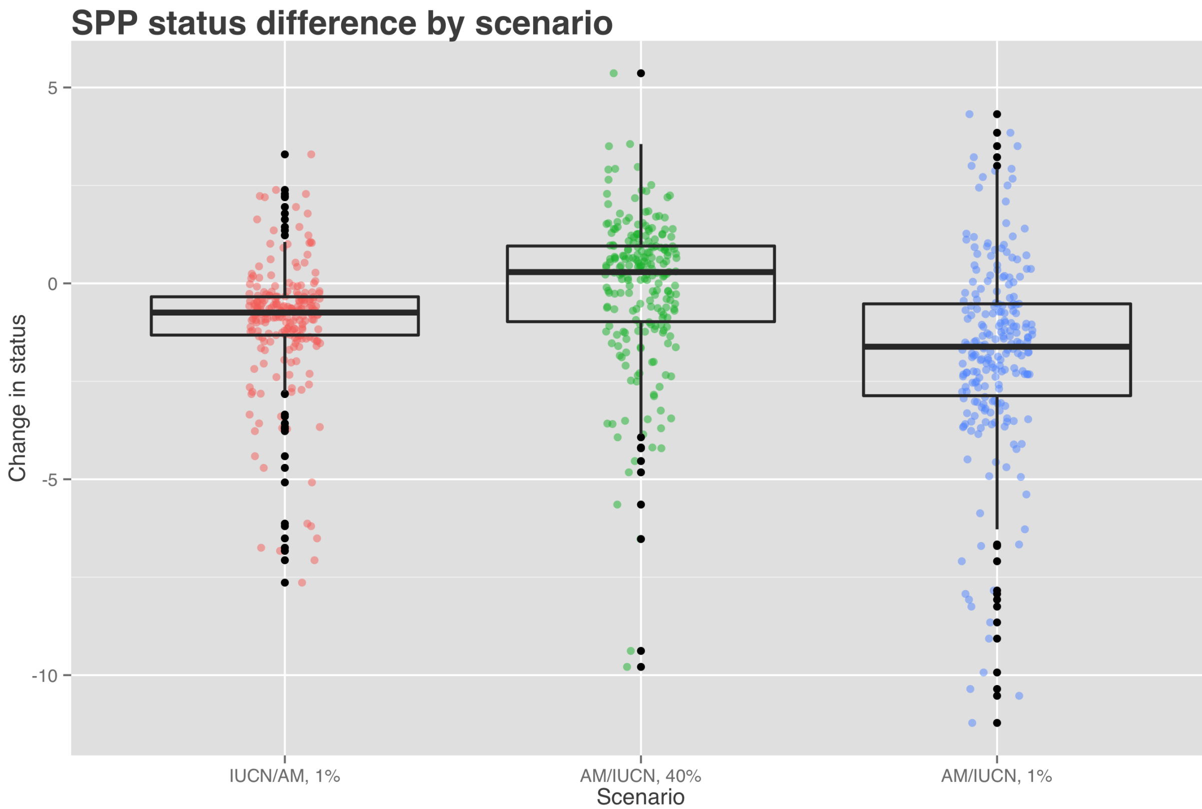
6 Implications

6.1 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 set 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***

6.2 Application to MPA Gap Analysis

yeah still gotta do something on this one

7 Conclusion

8 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

9 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.