REVIEWER 2

General Comments

The dataset documented in this manuscript is a really impressive achievement, and constitutes an extremely valuable resource for the marine biodiversity research and policy communities in Aotearoa NZ and worldwide. The use of data-driven SDMs corroborated with expert judgement has resulted in a really high quality set of species-level maps, and the methodology behind these is generally sound and clearly described. So overall I am very positive about this manuscript. I have a few comments about its structure and content, and some suggestions about maximising access and uptake.

We thank the reviewer for their constructive comments and overall endorsement of the work. As for reviewer 1, changes have been made to the text following suggestions by both reviewers and line numbers in the response refer to the resubmitted manuscript. All changes to the text are provided as track changes. Detailed responses are provided below.

Specific Comments

I like the infographic (fig 1), and hesitate to make suggestions that would overload it with too much information, but I think it would be possible to include a little more info on the different biotic groups - e.g. in addition to number of taxa, maybe also include number (and maybe time-span) of surveys and sampling points?

We believe that adding this information in the graphic may make the graphic too cluttered. However, we acknowledge the importance of this information and have added this to the figure caption, which now reads (L104 - 111):

"Figure 1 Infographic illustrating the development of the atlas described in this study. Presence and absence (p/a) records for four taxonomic groups were combined with over 20 spatially explicit environmental variables and used to run ensembled SDMs of Boosted Regression Trees and Random Forest models==. The number of unique spatial occurrences available for modelling ranged between 50-13926 unique spatial occurrences for demersal fish (collected from 1979 – 2016), 36-339 unique spatial occurrences for reef fish (collected from 1986 – 2014), 70-10804 unique spatial occurrences for subtidal invertebrate (collected between 1896 – 2019) and 50-422 unique spatial occurrences for macroalgae taxa (collected between 1896 - 2018).== The models were statistically validated using best practice procedures and evaluated by taxonomic experts to further appraise model accuracy. These assessments were incorporated within the metadata of each layer and uploaded, along with the layers themselves, to the Zenodo data portal. World imagery basemap utilised on inset (ESRI 2022)."

L120-125: I would find it useful here to provide an illustration (maybe in the supplementary material) of different cases - e.g. where model performance statistics and expert assessment agree / disagree particularly well. This would I think help to make this process a bit more intuitive.

This is a very useful suggestion. We now point to this illustration in the main body of the text (L133 - 134):

"(as an illustration, see examples of where model performance statistics and expert assessment agree and disagree in the supplementary materials 1)."

and provide these examples in the supplementary materials 1 (L195 – 204):

"In most cases, statistical and expert validations were congruent. For example, the predicted distributions of kahawai (Arripis trutta – demersal fish), the erect branching deep-water coral genus Solenosmilia sp. (subtidal invertebrate); the wrasse Pseudolabrus miles (reef fish); and the New Zealand bull kelp (Durvillaea antarctica, macroalgae) were all considered to be "accurate" / "very accurate" by experts and had "excellent" statistical model validation scores. In contrast, there were some taxa for which the expert evaluation scores were much lower ("inaccurate" / "somewhat inaccurate") than the statistical evaluation score ("good"). For example, the pacific salmon (Oncorhynchus tshawytscha, demersal fish), the bryozoan Genus Figularia (subtidal invertebrates), the wrasse Coris sandeyeri (reef fish), the red algae Champia novae-zelandiae (macroalage). Care must be taken when using predictions where there is discrepancy between statistical and expert evaluation (with expert evaluation assumed to be more accurate)."

Figure 2: There is a reasonably large body of evidence now showing the the rainbow colour palette has a number of shortcomings in terms of accessibility and perceptual biases (for a short recent overview see Westaway 2022 https://doi.org/10.5194/gc-5-83-2022). Although the palette used in these maps is not the 'classic' rainbow, it shares some of its characteristics, and I would strongly encourage the authors to consider changing their default colour palette, or at least to provide options.

We acknowledge the reviewers' legitimate concerns about colour palettes but in this case we feel that the interval classification means that at least perceptual bias may not be an issue. The visualisations of these layers are provided on the geoportal hosted and managed by the New Zealand Department of Conservation and we are not in a position to request these be changed. We believe the colour palette used is adequate for interested parties to explore patterns of distribution and should a more detailed view of the data be warranted we would encourage that they download the layers (from the linked Zenodo database) and use their preferred colour palette.

L140: The efforts to make the data available are commendable, and for many users the Zenodo archive plus the interactive GIS will be perfectly adequate. As a biodiversity data scientist, however, the thing that would really open up these datasets would be programmatic API access. I do not think this is essential for this release of the dataset, but it is something that I would encourage the authors to consider as a future development (and maybe to discuss that briefly in the ms). As a committed R user I typically look to ROpenSci (http://ropensci) for implementations of this kind of functionality, and ideed it appears that the 'deposits' package (https://docs.ropensci.org/deposits/articles/deposits.html) does already provide means to access data in Zenodo, so it may not be that significant a task to provide some example code or even to package the atlas up - this could really trigger higher uptake among various research communities.

We thank the reviewer for this suggestion. We agree that this would be a valuable addition and now refer to this as a future development in the L265 - 267. We will explore incorporating this functionality in future iterations of this work.

"Furthermore, future iterations of databases may also seek to include programmatic API access to commonly used statistical software (e.g, R statistical software) to facilitate use from biodiversity scientists"

Supplementary materials

It's not clear from the description of the data, but are the raw occurrence data also deposited in any biodiversity data aggregators (OBIS would be the obvious choice)? If not, and if this is feasible, they would constitute a valuable addition.

Some of the data are freely available (e.g., demersal fish) whereas others are not (e.g., subtidal invertebrates from the NIWA invert which contains commercially sensitive data). However, the underlying sample data can be available upon reasonable request (now stated L 6 in supplementary materials 1)

It would be useful to provide a little more information about absences. My understanding is that for both fish datasets and for the invertebrates, absence data is available (i.e., the assumption is that all species from the relevant group were identified at each sampling event, so the absence of a species indicates that it was looked for but not found). Is that correct? And if so, it's not clear why absences needed to be generated using 'target group background data' (from L100) - is it not possible to infer that the surveys where a species was not recorded is an absence, and to use all of these? I may have misunderstood but it would be useful to provide a little clarification. I understand that there is presence-only data for macroalgae - this is clearly explained.

The reviewer raises a good point as this was not clear in the text. We now elaborate on the different datasets to clarify this at L100:

"To estimate taxonomic distributions, BRT and RF models require locations of both presences (occurrence records) and absences. For biotic groups demersal fish and reef fish, the consistent methods to collect occurrence data means that where a taxon was not sampled, we assume that taxon was absent. For biotic group subtidal invertebrates, the same assumption was made but split by gear type (to account for differences in sampling efficiency). For biotic group macroalgae, given the differences in sampling methods used to collect occurrence data, we used 'target-group background data' (Phillips et al., 2009) as absences (referred to here as relative absence), i.e., a location where a different taxon to that being modelled was recorded (Stephenson et al., 2020). In practice different labelling of 'absences' does not affect the modelling approach but illustrates differences in the certainty of the absences (and therefore the outputs). Absence (and relative absences for macroalgae) were generated for each taxon from occurrences within taxonomic groups (i.e., demersal fish absences were generated from demersal fish occurrence records). The location of absences and relative absences was required to be at least 1 km from presence data and the number of absence and relative absence data was set to be equal to the number of presences (following best practice outlined in Aiello-Lammens et al. (2015) and Barbet-Massin et al. (2012))."

Do you have any thoughts about how reproducible the species distribution modelling process is, given that there is a reasonable amount of model tuning and analyst judgement involved? There have been some efforts to improve the reproducibility of SDMs (e.g. Golding et al. 2017, https://doi.org/10.1111/2041-210X.12858) but I am unsure what the current best practice is considered to be.

We believe that we provide detailed methods which would allow other researchers to reproduce our work given the same data. Furthermore, the model tuning is mostly automated and therefore there wouldn't be much scope for analyst judgment. The idea with these spatial distributions is that they

may be periodically updated as and when new data is available. In that sense, these data may "change" every 5 – 10 years but this will be due to differing datasets used in the modelling. In line with this, the expert judgement may vary should different experts be used in future, but in this work we had access to a large number of number experts who arrived at consensus. If this process were repeated, we do not think their assessments would vary greatly.

Technical Comments

Figure 1: Zonodo should be Zenodo

Changed – thank you

L129: This should refer to Figure 2, not Figure 1

Changed - thank you