Summary:

This manuscript describes an original method for integrating different datasets into a multispecies occupancy model, and showcases this modeling framework using the case study of dolphin-trawling interactions in the Mediterranean Sea (hence making the analogy between trawlers and another species). More specifically, the authors developed an integrated Bayesian model (NIMBLE-based; code available online) which, fitted to detection/non-detection data, estimates how environment affects where trawling activites and dolphins occur and co-occur while implicitly accounting for the potential impact of each component on the distribution of the other. Taking advantage of the Bayesian approach, the multispecies co-occupancy model makes use of all available information by making very different surveys (onboard and aerial) comparable through an estimate of their respective sampling efficiency. Using the multispecies occupancy model in a predictive way, the authors can then map the co-occurrence of dolphins and actively fishing trawlers in the Gulf of Lion, highlighting the potential of the approach in the context of by-catch management.

Main comments:

I appreciated reading this manuscript, which is pretty straightforward and clear, well illustrated, and goes with helpful complementary information and documented code. The ecological modeling work described in the manuscript is of great quality. The authors use the appropriate tools to answer the main ecological questions and show a fine understanding of the assumptions underlying each step of the modeling procedure. The framework they developed is transferable to other case studies and contexts and may have concrete applications in the near future for assessing and managing fishing-wildlife interactions.

The main asset of this work is the modeling framework itself, both with respect to (i) the way it maximizes the amount of information that can be incorporated and (ii) the originality of its application to fishing-protected species interactions by considering the trawlers as a second "species" of the occupancy model to account for the potential interdependence of their distribution.

As partially highlighted by the authors in their Introduction, to date, many studies mapping the distribution of interactions between fishing and protected species rely on data biased towards one of the components involved in these interactions that are sometimes combined together and rarely in an adequate way: reported by-catch, observers onboard fishing boats, or simply fishing effort distribution for the former, megafauna-focused survey for the latter. Although surveys and monitoring programs able to provide information on both megafauna and fishing remain limited nowadays, we can expect that the increasing use of new autonomous monitoring systems (e.g., drones or gliders) and the development of tools for the analyses of the data they collect (~ automatic analyses through machine learning) will favour the acquisition of such data in the near future. In that context, the present modeling approach looks particularly promising.

On a minor and purely methodological aspect, I also liked how the model was assembled. To couple the occupancy and the detection modules of the model, the authors hacked a very popular tool in marine ecology for building species distribution modeling: generalized additive models (and associated mgcv package). While Bayesian integrated models are less accessible and mastered by a smaller part of the community, the framework proposed in this article and the available supplementary material might help researchers transition or get more familiar with such types of frameworks.

The results of this model application are slightly frustrating as it seems that (I'm putting it provocatively), the conclusion is that "dolphins and fishing boats co-occur where they occur": the spatial patterns in the probability of occurrence of dolphins/trawlers are very close and are similar to that of their co-occurrence. It suggests that, probably due to some modeling choices, data constraints, and the reality of fishing-dolphin interactions at that scale of time and space, interactions between dolphins and trawling activity are poorly structuring their distribution in comparison with environment. Nonetheless, this case study remains helpful in showcasing the potential of this approach.

For all these reasons, I think this work is definitely valuable for the ecological modeling community, with perspectives for management applications. Once that said, I have some reservations regarding several points.

My main reservation is related to several modeling choices made in the present study that would require a better justification and whose implications regarding the outcomes of the study should be discussed. A first example is the temporal and spatial resolution of the data and model. I would expect that the manuscript describes more basic choices regarding, for instance, grid-cell size or the seasons covered. All these choices are probably driven by the amount of data available but need to be justified. The data description in the present version of the manuscript does not help: there is no information about the temporal coverage, and information on the spatial coverage is only provided in the Annexes. Maybe interspecies interaction effects on co-occurrence may be stronger with a finer resolution..? All of this should be discussed. Another example is the choice of the covariates included in the distribution model. To explain the presence or absence of dolphins and fishing, the authors chose one environmental covariate only (and a spatial tensor), i.e. depth. I know that even complex habitat models built for odontocetes often identify depth as one of the primary variables structuring their distribution. Still, this should at least be justified with some references. Also, due to technical limitations or fuel coats, most of the trawling activity is generally restricted to the continental shelf, 0-250m, and in smaller extent, the continental slope (trawlers are operating deep in this study but it remains difficult for me to analyze since the "trawlers" are not defined in the M&M; see specific comments). As a consequence, integrating only depth could limit the pertinence of the model. It will likely show a depth effect at the scale of the whole study area, which includes deep waters not frequented by fishing, while interesting patterns in shallower waters, i.e., where most of the fishing activity occurs, may be missed; Especially since depth has a linear effect.

Another important reservation is about some choices the authors made regarding the manuscript's content.
A complementary work was conducted to characterize how the integrative aspect of that approach was crucial. One of the annex documents compares the outputs of the models when integrating one dataset only and both datasets. I find this part relatively interesting and, given (i) the brevity of the manuscript and (ii) the number of statements regarding the interest of integrative approaches, I think it would make sense to move it into the main manuscript. It is just a suggestion, though.
More importantly, the Introduction emphasized the importance of accounting for the role of interspecies when studying their distributions. As far as I understand, this aspect motivated the authors to adopt the multispecies co-occurrence modeling approach. However, no output, figure or analysis allows us to assess the benefits of this approach. I think the authors should find a way to highlight it better (this is, to me, even more important than discussing the value of data integration). From what we can see of the model outputs, the advantage of the co-occupancy model over 2 separate monospecies approaches may be

relatively limited in the present case study. I am not asking the authors to run the model in a monospecies configuration (unless it is straightforward!). Nonetheless, the manuscript would benefit from an assessment (even qualitative) of the added value of the multispecies aspect. At least, I would expect much more about this topic in the Discussion.

Last, I strongly recommend the authors to carefully check their text as I found a lot of typos should not be so numerous at that stage (especially those detectable with Word or other automatic correctors). Some sentences were not clear in different places in the manuscript; I tried to flag them in my specific comments.

Note that I reviewed the code provided by the authors but didn't take the time to run the model by myself.

Conclusion

For the reasons mentioned above, I consider the present manuscript highly valuable and worth publishing. However, I suggest it is revised to tackle the issues mentioned in my reservations before any acceptance.

Specific comments:
Summary

L22 – *The* possibility (ability?)

Intro

L.27 – the fact that it is especially affected by anthropogenic [not "anthropic"] is not only resulting from its "seaway function" – reformulate

L.29 – reference for "increasing interactions"?

L31 – there are probably some experts in the co-authors, but in my mind, "depredation" is the fact that these predators feed on fish caught (or at least localized and targeted) by fishers. So "foraging behind trawlers" IS predation (~~leads to depredation~~), right? Then, OK, depredation behavior can occasion by-catch.

L34 – "depredation pressure" is hard to quantify. Maybe mentioning "depredating behavior" is sufficient..?

L37 – "in multiple locations worldwide" → "worldwide"?

L38 - "Following mortality events that have been documented about bottlenose dolphins (Manlik et al. 2022), interactions have raised conservation concerns and mitigation measures trialed thus far have not proven effective (Snape et al. 2018, Bonizzoni et al. 2020)" → "Following documented bottlenose dolphins mortality events (Manlik et al. 2022), interactions have raised conservation concerns and mitigation measures tested to date have not proven effective (Snape et al. 2018, Bonizzoni et al. 2020)"

L41-44 – Bunch of different approaches listed here; try to be more specific and quickly display the advantages and drawbacks of each method

L47 – be more specific (this isn't the first sentence of the intro); try something like "….conservation conflicts and is, therefore, particularly strategic in the case of the mammal by-catch issue"?

L51 – "needs"

L47-49 – Maybe I would put just after this sentence the reference to Pirotta (L43): it is step 0 of interactions mapping. It would support your statement about the need for multisp co-occurrence model.

L58 – "underestimation of species distribution" what do the authors mean?

L62 – "to combine multiple datasets into" → "combining multiple datasets within"?

M&M

L72 – Data paragraph: maybe the authors could add a short description of the study area to highlight its relevance regarding the dolphin-fisheries interactions

L76 – maybe first the megafauna (main focus), then human activities?

L77 – "detections and non-detections data" → "detection /non-detection data"?

L80 – Onboard survey?

L82 – Ony detections here? Or non-detections too?

L83 – harbours

L83 – Trawlers: which type (all? Pelagic? Demersal? Which size?)? We want to be sure that it is pertinent with interactions that have been observed in the literature

L86 – "as the transect length (in km) of each monitoring program for each grid-cell" → "as the total length (in km) of transects conducted in each grid-cell by each monitoring program"?

L87 – Using depth as the only covariate should be, to my mind, justified.

L81-87 – We have here information about the spatial coverage of the dataset – but we don't have anything about the temporal resolution/coverage.

L89-97 – ok, helpful clarification

L112 – I guess X and Y are long/lat of the center of the grid-cells?

L114 – Specify the family chosen + the type of smoother

L114 – Based on the study area map, I have the feeling that correlation may be strong between the covariate, depth, and the spatial smoother, e.g., X positively correlated with depth, and Y negatively correlated with depth. I guess the authors have already checked this but could they discuss this in this paragraph? Is there a correlation, is it strong etc… ? If there is a substantial correlation, what are the implications for the inferred probabilities of co-occurrence?

L119 – "winter, spring, summer, and autumn" – first time that authors mention seasons, should be described in the Data part.

L136-157 – I find this part a bit laborious. I am actually not sure that the 16 feet long list of "observation events" is helpful here. Maybe it would be ok to transfer it to an Annexe, together with a more detail explanation of how the theta matrix is built; and replace L135-154 with smthg like "These four probabilities could then be used to explain the simultaneous detection or undetection of each species by each survey, consisting 16 observation "events" (= (2 species) ^ (2 detection status) ^ (2 surveys))." I let the authors think about it.

L159 – I would instead write "multiple datasets/data types integration."

L166 – 80% is not consistent with what is indicated in the caption → which one is correct?

Results

L169 – Would it be possible to compare the relative importance of the depth effect vs spatial tensor? The most important results are displayed, but people may be curious about it.

L171 – keep the greek letter for psi

L171 – detail but help clarify the paper: dolphins on Fig 1.3 should be in green to respect the color choices of Fig1.2 (if the authors choose to include figures from the Annexe in the main, they should also check this aspect)

L177-178 – May also be put in the M&M: data limitation issue should be mentioned prior modeling

L179 – "that trawlers only or that ~~trawlers~~ only" dolphins?

L182 – "Although~~,~~ dolphins"

L186 – Just a visual thing: proximity between the 2 upper and the 2 lower panels is confusing when reading the figure. Maybe consider putting more space between the titles of lower panels and upper panels

Discussion

L203-207 – One point that is worth commenting on is also the impact of the grid size that has been selected.

L213 – "Supplementary materials": I would not be shocked to read more about the outcomes of this analysis in the Discussion.

L221 – campaigns → surveys

Annexes

In the script: please define all variables, even the intermediary variables – can be done in an intro to your code. Quickly showing the structure/content of the input data would also help.

Supplementary Material Annexe 2 overlaps a lot with the M&M paragraph of the manuscript – except for the data part, which is so short that it can be transferred into the M&M.