Reviewer Responses for "Assessing the Population-level Conservation Effects of Marine Protected Areas" (previously "The Regional Effects of MPAs")

Daniel Ovando Jennifer E. Caselle Christopher Costello Olivier Deschenes
Steven D. Gaines Ray Hilborn Owen Liu

Contents

1 Reviewer Responses

This is a re-submission of MS# 2019-17437 (formerly titled "The Regional Effects of Marine Protected Areas"). We thank the two reviewers and the Editor for their constructive responses. We have fundamentally restructured the paper to place much greater emphasis on the results from the Channel Islands Marine Protected Area (MPA) network. This is to our knowledge the first empirical assessment of the population-level effects of a large and iconic MPA network on a wide assemblage of commercially and economically important species. Our inability to estimate a clear population-level effect in such a well-enforced and well-studied location is we feel an important "null-result" both for scientists and MPA managers to be aware of. As such we hope that it warrants publication in PNAS.

While our results are more focused on the empirical results, we explicitly link our results to the large body of MPA theoretical work, using it to help illustrate why we might expect such a null-result and help managers understand when they might be able to detect the population-level effects of MPA networks. As MPAs are increasingly used as a core tool of marine resource management, our results provide a useful update to our empirical expectations for MPAs, that to date have been based almost entirely on inside-outside MPA response ratios.

As pointed out by the reviewers, many of the more theoretical points we make about the importance of fleet dynamics and the challenges of response ratios have been made in the past. However, we feel there is substantial value in revisiting those discussions in the context of a concrete empirical example, to provide scientists and practitioners alike a cohesive reference to observe and understand the full ramifications of these past findings on our practical ability to monitor and manage marine protected area networks. We have tried to make clearer that the theoretical section is a reflection of broad body of previous work.

We would like to call out two particular changes to the paper. We felt that term "regional effects" was needlessly vague and causing confusion. We have therefore re-framed our results around "population-level" effects, a more easily defined and we hope understandable metric. In particular, we feel that this framing

places these results in clearer contrast to prior works such as Caselle et al. 2015 that examined inside-outside MPA effects, as opposed to population-level effects.

We have also replaced our empirical results with a slightly simplified model fit to aggregated data (as opposed to the raw transect-level fit reported before). The updated results are we feel much clearer to understand and equally empirically valid, and have the added benefit of allowing for a broader range of sensitivity analyses (all included in the Supplemental Information). They are also much easier for other users to reproduce. The transect-level results reported in the prior draft are still available in the Supplemental Information.

We hope you find the updated version of our manuscript improved thanks to the inclusions of the careful points raised by the last round of reviews of this paper.

1.1 Reviewer 1

- *I think that greater emphasis needs to be placed up front, abstract, statement of significance, introduction..perhaps even title, that this model and paper only addresses fish biomass... not other 'conservation outcomes'. You do define this but not until page 2, line 148-154. I recommend moving this up.*
  - We apologize for not making this clearer in our earlier draft, and have included this key caveat in the abstract, statement of significance, and introduction (Line 77)
- *I was surprised that there was no discussion about enforcement being a potential confounding factor. Currently the paper assumes that the Channel Islands MPAs are fully enforced, without poaching. This may or may not be the case and should be stated as such. You do address well other system wide impacts like water temperature (lines 557-8, 691+). Perhaps enforcement could be added to the discussion around line 711-13.*
  - This is indeed a critical point that we have endeavored to correct. We have included enforcement under our summary of key enabling conditions for effective MPAs (line 109), as well as the discussion (line 406). While we do not have published data on strength of enforcement within the MPAs, anecdotal evidence suggests it is relatively high, as well as evidence from the mainland California MPAs (Zellmer et al. 2018) indicates it is high. In addition, the presence of a clear increase in response ratios within the Channel Islands provides evidence that the MPAs are large enough and sufficiently well-enforced as to provide meaningful protection within their borders (Figure 1).
- *Your discussion of the model also emphasizes that stock status prior to MPA designation is very key to outcomes; in the SI most of the species included for the Channel Island review have unknown status—which influences your discussion, and may better be described.*
  - We make clearer that we iterate over unknowns such as pre-MPA stock status in our description of the Channel Islands simulations (line 132), and the implications of pre-MPA stock status in the discussion (Line 383, Figure 5).

1.2 Reviewer 2

- *Yes, these are important questions but they are not new to ecology, economics, and fishery science. You are not the first to ask them and point out their importance; see articles at bottom for some examples. What are you doing to answer them that is better than all of the other papers that come before you? And why is your case study an exemplary place to answer the questions? The conclusions on line 709-728 are well-known in the MPA literature.*
  - We have fundamentally restructured the paper based around the importance of the Channel Islands as a case study in measuring MPA effectiveness. We thank the reviewer for pointing out gaps in our literature, and we have incorporated the studies they recommend. The theoretical points we raise have indeed been pointed out in the literature before, and we feel our new manuscript makes

that clearer. Specifically, we now use our simulation analysis to demonstrate when we might be able to detect population-level MPA effects, rather than focusing on the expected effect sizes as a function of traits that have been previously raised in the literature for

  - The Channel Islands MPAs are an iconic network that has served as a model for protected areas around the world. It is also extremely well-studied and well-enforced. While previous studies have looked the effectiveness of the Channel Islands MPAs within their borders, this is the first study in the Islands (and globally of which we are aware) to attempt an estimate of the effect of the MPAs at the population scale (both inside and outside of the MPAs themselves). That we are unable to estimate a clear population-level effect in even this system is an important result that we feel warrants publication in PNAS. As MPAs are increasingly turned to tools for ocean conservation, our results provide an important update and caution to expectations based solely on inside-outside comparison studies.

- *I would also be more cautious regarding your conclusion about your model predicts what should be expected in the data analysis. It could be simply luck that there is agreement between the two.*

  - We have made the language clearer that this concordance could be simple coincidence, and is not in any way proof of the validity of our results (Line 209) However, we feel that it is critically important to ground our theoretical results in theoretical predictions. In addition, as we mention in the discussion, we encourage MPA scientists and stakeholders to undertake such simulation modeling to consider the plausibility of their results. As the reviewer says, empirical results that match theoretical predictions are not proof that the empirical methods are robust, but they do provide am additional and important line of evidence: empirical results that fail to match theoretical predictions need to be carefully considered.

- *What about using the simulation model to demonstrate the validity of the technique used to identify the MPA effect used in the paper, which was pioneered in publication in 38? Or to understand whether the implication of the statement on line 539, that is, species ranges exceed the Island and your sampling domain, is important for measuring the impacts of MPAs outside of their boundaries? In sum, in its current form, I would move all of the simulation model to the SI (lines 131-252) and mention in the discussion section that these results are consistent with the predictions of a stylized single species model.*

  - We have greatly expanded the simulation validation of our method, which was previously included in the supplementary material. Figure 4 of the paper now demonstrates the validity (and shortcomings) of the model under a variety of plausible conditions. We would note that this method is related to that described in Caselle et al. (2015) (previously reference #38), but is not the same. Our method explicitly estimates the net difference in the targeted and non-targeted biomass densities pre-and-post MPA as an estimate of the mean population-level effect of the MPAs both inside and outside their borders. Caselle et al. 2015 used trends in targeted and non-targeted species inside and outside of MPAs to illustrate the effects of the MPAs within their borders. We stress this difference in an effort to make clear that our results and methods are not simply an update to Caselle et al. 2015, but address a fundamentally different question. We hope we have made this clearer in the text, e.g. on Line 89, Line 103, and Line 153.
  - With regards to the specific question about the possibility of the species range exceeding the sampling domain, this would be reflected as a reduction in the percent of the population protected by the MPAs: a 25km^2 MPA may provide near total protection for a sedentary species with brood-rearing, but be insignificant to a migratory tuna. To that end, our assessment of model performance as a function of MPA effect size (Fig.4, Fig.5) addresses the question of how MPA size relative to species range affects our models performance. All else being equal the smaller a portion of a population's range is protected by MPAs, the smaller the expected effect size, and by extension the more challenge the estimation model will have.
  - We apologize that we did not present our prior simulation work in a sufficiently compelling manner, and failed to differentiate our results from the prior literature. As suggested by the reviewers and the editor, we have restructured the paper to place much greater emphasis on the critical results

from the Channel Islands. However, we feel that the simulation results do provide vital context to our results, and provide tangible recommendations as to under what circumstances we might be able to estimate population-level effects of MPAs. The reviewer correctly points out that the individual components of our model are not themselves novel, and that the importance of fleet dynamics and fish dispersal have been previously discussed. As such as have greatly reduced the attention these results receive, except as a vehicle for the explicit consideration of the feasibility of estimating population-level MPA effects. In addition, while these broad points have been made in the literature before, we feel there is value in cohesively demonstrating the ramifications of these diverse pieces of MPA theory (that generally are looked at only one or a few at a time) in once place, both through our Figure 5 and the accompanying web application. We hope that these modifications make clear the contribution of the simulation modeling.

- *The discussion on how to detect changes outside the MPA (lines 254-347) follows very closely the discussion in Ferraro et al (#41 in your list). It seems like you can remove most of this text and point the reader to that paper*
  - We have greatly reduced the amount of text addressing response ratios, and restructured the response-ratio analysis to be linked explicitly with the empirical response ratios observed in the Channel Islands (in response to comments from the Editor), in Figure 1. We feel this strategy makes much clearer how the results of this paper expand beyond the findings of Caselle et al. 2015, and why the results of Caselle et al. 2015 are evidence of within-MPA effects, but not of the question of population-level effects addressed in this manuscript. Given that much of the MPA literature relies on response ratios as the empirical evidence for the effectiveness of MPAs, we feel that it is critically important to directly address the challenges of response ratios as measures of population-level effects in this paper with an empirical example.
- *To me the potential contribution of the paper starts on line 350, which is way too late in a paper of this size*
  - We have restructured the paper so that it now focuses on the Channel Islands case study, and as suggested have moved this discussion up to the introduction (Abstract, Significance Statement, Line 77).
- *Is part of assumption b) (line 372) that the dispersal characteristics of the fished and non-fished species need to be similar? I could envision a scenario where the two sets of species have different dispersal characteristics, non-fished have smaller dispersal ranges than fished species, for example. How does this impact your identification strategy? If the trends are similar regardless of the mechanistic process by which it occurs, does it matter for the empirics? More discussion on both assumptions (a) and (b) and what they imply about the biological mechanisms seems warranted.*
  - The estimation model does not depend on the dispersal characteristics of the fished and non-fished species per-say. The model is operating at the level of biomass densities. As such, even if non-targeted species have a much bigger range than the targeted species (or even the sampling region), so long as the parallel trends assumption holds within the sampled portion of the non-targeted species range, the identification strategy may be valid. Similarly, the model is agnostic as to the root causes of the parallel trends, so long as they exist. However, if for example the parallel trends are dependent on a particular temperature regime which during the study is disrupted (as is likely the case in the later years of our data), then the model may be reliable for the early years but not later. We have made this point clearer in the paper (Line 276, Line 450), and made clear that the lack of a clear effect in the recent years is likely due to an emerging violation of the parallel trends assumption, and not to a failure of the MPAs per say. However, this is still an important result, since it suggests that neither the response ratios nor the difference-in-difference method give us a robust estimate of the population-level effects of the Channel Island MPAs. As we stress in the paper, that even in a network such as the Channel Islands we are left without a clear empirical effect at the population level has critical ramifications for the design and monitoring of MPAs globally.

- *The use of visual data surveys has, of course, problems that introduce errors in dependent and independent variables that could be correlated with features of the fished and non-fished species, e.g., size distribution. How can you be sure that this is not driving your results? Could you use the simulation model to investigate this?*

  - Visual survey data are indeed challenging, but also often the only form of data available for MPA analysis. Our base model operates at the level of total biomass density of targeted and non-targeted species, at the site level. This is the finest spatial scale at which the sampling regime would be likely to provide an estimate of the variable in question, since while any individual transect is unlikely to observe all of the targeted / non-targeted, but aggregating across transects provides a better chance of this. These totals are derived from average biomass densities of individual species across multiple transects within each site, which are then summed across species to achieve the total mean biomass density at the site level. The averaging process acts to smooth out some of the sampling variation, and reflects a broader discussion about the right degree of smoothing in spatio-temporal data.
  - How might this affect our results? So long as sampling errors are consistent over time, they should not bias the mean or the trends of our results (e.g. if we systemically under/over estimate the biomass of some species). There is no doubt that we are underestimating the true magnitude of the credible intervals though. For example, the model assumes constant length-to-weight relationships over time, whereas in reality there is clearly variation both among individuals and over time. Accounting for this should add more noise to our data, further complicating the ability of our model to achieve precise estimates. Complicated relationships such as the MPAs themselves affecting the length-to-weight relationship would indeed skew our results, but we feel that is beyond the scope of this or any other empirical MPA study of which we are aware to address.
  - We have run our model at a wide array of resolutions, from the updated aggregated form we use here to a complete hierarchical framework that attempts to control for errors in the individual observation process (Fig.4, Fig.S39). The fundamental results remains the same: a general upwards trend, followed by a downward trend, with the clearest signal in the years around 2010. We would actually be more concerned about the robustness of our results if we had found a precise but small effect size: such an effect could easily disappear if we accounted for errors in the data (e.g. through allowing for variations in the length-weight relationship). Our relatively uncertain findings instead suggests that it is unlikely that we would get fundamentally different conclusions by including more error; we would simply be even less certain than the uncertain state the model estimates we are in. We have expanded the simulation testing of the method a bit to help illustrate how violations of the model assumptions play out, and how they are related to the magnitude of the effect size (Fig. 4). The full suite of alternative resolution models can be found in Figs.S16-26 (all of which show the same general result), and a a simulation test of our model incorporating factors such as differing diver skill can be found in Fig.S39 (which shows that the model is capable of performing under such circumstances).

- *(The answer to your question on line 709 that appears on lines 714-719 has been known in the literature on MPAs for over 20 years; see, e.g., papers by Martin Smith below. Given that your question is what "your results imply about the future of marine science", my answer is not very much. Again, you did a rigorous empirical case study that should be published. I am however finding it difficult to see a path to publication in PNAS.)*

  - We hope that our restructuring around the novel empirical results resolves this concern. While we agree that framing our results around the future of marine science would be a step too far, our goal is to ask what our results imply for the future of MPA science more specifically, which we hope is less of a stretch.
  - Given the increasing role of MPAs in conservation strategies, the lack of a clear population-level effect in a place like the Channel Islands should be known and taken into account by MPA practitioners around the globe. We agree that we framed our simulation results poorly in previous drafts, but feel that they do compliment the empirical results to help managers understand why

the Channel Islands results may not be surprising, and when they might expect to be able to detect and effect, and when not. See section 2.1.2 ("When can we detect the effects of MPAs") and our updated conclusions, in particular the paragraph beginning on line 357. Works such as the papers by Martin Smith suggested by the reviewer are indeed critical to our understanding of MPAs, and we have taken steps to place our results in the context of such work. In particular, we advocate for more works such as Smith, Zhang, & Coleman 2006 for estimating the fishery effects of MPA networks (which differs from the population-level effects estimated here)

- *Examples of relevant literature, you should expand your net to find other relevant literature to better distinguish your contribution*

  – We thank the reviewer for pointing out these important omissions. We have included each of these references and their key messages in their appropriate portions of the paper.

2 References

Caselle, Jennifer E., Andrew Rassweiler, Scott L. Hamilton, and Robert R. Warner. 2015. "Recovery Trajectories of Kelp Forest Animals Are Rapid yet Spatially Variable Across a Network of Temperate Marine Protected Areas." *Scientific Reports* 5: 14102.

Zellmer, Amanda J., Heather Burdick, Ivan Medel, Daniel J. Pondella, and Tom Ford. 2018. "Aerial Surveys and Distribution Models Enable Monitoring of Fishing in Marine Protected Areas." *Ocean & Coastal Management* 165 (November): 298–306. https://doi.org/10.1016/j.ocecoaman.2018.08.027.