Reviewer Report

Letter from the editor

There are some rather loose statements in the ms, starting with the Abstract (see also Rev 1) where it is implied that a 'well-managed' fishery should result in no effect in a reserve. This is false. Most 'well-managed' fisheries aim to bring B0 to around 50% of unfished biomass, so reserves should on average contain double the density of fished areas. There is one unsubstantiated statement (line 81) of good enforcement, but could it not be just as likely that there is fishing activity in these reserves? Poaching is an equally valid explanation as any examine din the Discussion.

As per Reviewer 1, much more clarity is needed in Methods. It is unclear what data are being tested and what the sampling design was, especially the spatial arrangement of controls. The modeling approach was also confusing to me - you have categorical responses so why 'multiple regression'? Your reporting indicates a log-linear model (especially the use of lambda) which is probably the most appropriate structure for fish data.

There is also some fairly lazy scholarship here: e.g. BACI designs originate in papers other than those cited on line 112-113, and were demonstrated in papers well before those in 114-115, and again in 133-134. The Moland et al. paper may not be a good example anyway since their sampling was spatially confounded. Discussion of spatial comparisons and before-after studies seems to imply nobody has ever used monitoring time series before. I am also confused by lines 250-252 where it is implied that this does study does not in fact have fished controls in the design.

Recommendation: This ms requires major revision.

Reviewer 1

Recommendation: Substantial revisions

Q1 - Please summarize the main findings of the study.

The manuscript provides an analysis of ecological and economic fishery indicators across TURF-reserves in three Mexican communities, finding no substantial changes in target species or associated fisheries, or in value or quantity of landings. This is not necessarily surprising and the likely reasons are discussed. This study provides an example (adding to an existing literature) of community based reserves that have local support but do not necessarily deliver substantial fishery benefits.

Q2 - Please highlight the limitations and strengths.

The question at the heart of this manuscript is an interesting one because the combination of TURFS and reserves offers the possibility of securing rights for fishers while also implementing conservation areas, thereby potentially avoiding some of the conflicts associated with exclusion from fishing grounds that can arise with MPA designation. However, I didn't feel that the paper fully lived up to its potential in addressing this question (see Q3). Beyond the key findings as described above (Q1), there is a lot of additional material in the manuscript that didn't seem to add much to the key message. If the focus on TURF-reserves is the key area of novelty, I think this should be further emphasised and explained – see more detailed comments below.

Q3 - Please provide your detailed review report to the editor and authors (including any comments on the Q4 Check List):

This review focuses on the more substantial points, and I have not commented on minor issues.

ABSTRACT

- I found the abstract somewhat misleading after reading the full ms because it implies that the findings include data on the level of support among fishing communities (L14), but this data is not presented in the results. In line with this the last sentence of the abstract seems to be speculation, and is not supported by evidence presented in the manuscript.

INTRODUCTION

- The introduction is brief and not easy to follow. It could be usefully expanded to develop the argument for this study and demonstrate the novel elements. My suggestion would be to expand it to more clearly explain what TURFS are, and what benefits TURF-reserves are expected to deliver in comparison to TURFS or reserves alone.
- I found the introduction confusing in parts because some statements say there is little evidence around TURFs, while others seem to refer to documented benefits. There is quite a lot of switching between references to studies of TURFS, reserves, and TURF-reserves. A clearer explanation of the existing science would help convey the knowledge gap. One general issue is that there are several places in the ms where it was difficult to distinguish between theoretical/expected outcomes or relationships, and those supported by evidence from this or other studies – it would be helpful if the authors could clarify these areas.
- In the third paragraph it wasn't clear to me whether the authors see TURF-reserves as synonymous with community-based marine reserves, or if not what the differences are. As the authors note, who these rights are conferred and recognised by is an important consideration and needs further explanation. In line with this the authors should perhaps revisit the title, which refers to community-based marine reserves, but not to TURFs.
- I suggest moving the specific details of Mexican TURFs to the case study section to provide context for the study sites, and focus on the broader issues in the introduction.

METHODS

- In the study area section more information would be useful on the wider context (some of this is in the introduction

– e.g. the TURF system in Mexico and governance arrangements for this system and the reserves within it).

- The claims that the communities are representative and that insights could be generalised to elsewhere in the world could be toned down. Some evidence that the communities share similarities with other fishing communities is useful, but the authors themselves suggest in the discussion that wide generalisations are not possible.
- More information on data collection would help the reader to have greater confidence in the data and make the study more replicable – e.g. numbers of surveys/data points, location of control sites and counterfactuals, processes of site selection, etc. Some of this could be added to Table 1 (or perhaps to supplementary info?).
- With regard to the evaluation of SES – it wasn't clear to me what the purpose of this evaluation was, and the data sources (and analysis of these) seemed sparse. There is little explanation of why these

factors are important in relation to TURF-reserves, and only minimal reference to the original work that led to this framework.

- My assumption is that the statistical analysis is appropriate to the data and questions. This is not really my area of expertise, though as someone who is reasonably conversant with statistical modelling I did find it difficult to follow the explanation of the analysis.

RESULTS

- Results are concisely presented. The key findings are clear but because the figures only show effect size it's not easy to get a feel for the data and its variability across sites and controls.
- The presentation of governance information is brief. It seems that the purpose is to assess whether the characteristics of the system are conducive to effective local management of the resource, but this is not very informative given the depth of analysis – the same broad descriptions are presented for all three sites for many of the variables. I would question the value of this somewhat 'tick box approach' to SES systems across a small number of case studies. What are the questions driving this analysis and what does it add to the paper? It seems that much of this could be presented as a contextual summary of the study sites rather than as analysis.

DISCUSSION

- The key message seems to be that there are no significant changes in the species targeted for protection, nor other fishery-related species or fishery economic benefits. There are a number of plausible explanations presented that are all well established in the MPA literature (reserve size, duration, quality of existing fisheries management and environmental variability). While I appreciate the importance of publishing negative results in the MPA literature, it's a shame that some of these factors weren't anticipated and explored in more detail.
- I found the second paragraph of the discussion difficult to follow – the key points here could be more clearly explained as they seem to point to the novelty of the statistical analysis.
- The latter part of the discussion focuses on the additional potential benefits of reserves including the positive perceptions of fishers noted in another study. My view is that these benefits are overemphasised here, given that there is no data in the manuscript to support them. Furthermore the potential counter-arguments (e.g. that resources invested in these reserves might be put to better use) are not considered, which leads to a fairly one-sided view.
- It would be worth discussing fully the limitations of the data e.g. presumably the lack of data on fishing effort restricts interpretation of the time series, and the indicators representing socio-economic benefits of the MPAs are limited

Q4 - Please comment on the methods, results and data interpretation. If there are any objective errors, or if the conclusions are not supported, you should detail your concerns.

Please see Q3

Q5 - Check List

Is the English language of sufficient quality?

- Yes

Is the quality of the figures and tables satisfactory?

- No

Does the reference list cover the relevant literature adequately and in an unbiased manner?

- Yes

Are the statistical methods valid and correctly applied? (e.g. sample size, choice of test)

No answer given.

Are the methods sufficiently documented to allow replication studies?

- No

Are the data underlying the study available in either the article, supplement, or deposited in a repository? (Sequence/expression data, protein/molecule characterizations, annotations, and taxonomy data are required to be deposited in public repositories prior to publication) (yes/no/not applicable)

- No

Reviewer 2

Recommendation: Moderate revisions

Q1 - Please summarize the main findings of the study.

This paper assesses the effectiveness of 3 TURF-reserves using a Before-After-Control-Impact design. The results indicate that the reserves have not increased the density of targeted species. Besides the biological analysis, the authors also evaluate the social-ecological system at each study community based on selected indicators. They conclude that in order to maximise biological effectiveness, community-based TURF reserves need to have full community support as well as meet other reserve design criteria, such as appropriate size and location.

Q2 - Please highlight the limitations and strengths.

Strengths: The study evaluates both the biological as well as social and governance system of marine reserves, which is not addressed simultaneously in many studies assessing marine reserve effectiveness.

Limitations: The connection between the SES and biological evaluations need to be strengthened. At the same time, the SES provides a static picture of governance and resource conditions, whereas the biological assessment reflects changes through time. To bring these two components together, there is a need to provide information about how the SES has changed through time. Also the SES analysis lacks evaluation criteria. This inhibits a full picture of the combined biological and social contribution of TURF reserves.

See further comments in Q3 below.

Q3 - Please provide your detailed review report to the editor and authors (including any comments on the Q4 Check List):

- This paper assesses the effectiveness of 3 TURF-reserves using a Before-After-Control-Impact design. The results indicate that the reserves have not increased the density of targeted species. Besides the biological analysis, the authors also evaluate the social-ecological system at each study community based on selected indicators. While it is encouraging to see the authors consider the social and governance aspects of marine reserves, it appears that this is secondary to the biological assessment. As it is presented now, I don't see the added insights that the SES analysis brings to the study. The effectiveness of reserves are primarily assessed based on the biological indicators only. This is due to the lack of evaluation criteria for the SES component. I'd suggest using some rating system to rate each of the SES indicators. E.g., for RS4.1 stock status – range of possible values would be overfished, underfished, etc. Where would each study site fall along the gradient of underfished to overfished stocks?
- By doing this you can assess each site in terms of biological performance and SES conditions. This then allows you to establish a stronger connection on how biological and social factors together act to contribute to better protection and effectiveness of marine reserves. This to me, is where the study can make a contribution to the literature.
- Further to my comment above, my understanding of this paper is that the authors are making an argument for TURF-reserves. However the rationale for this does not come out clearly in the paper. There needs to be stronger justification provided for why the authors expect the 18 TURF-reserves to perform better than i) non-TURF community based marine reserves, and ii) TURF fisheries (but not necessarily marine reserves). Without this clarification, there is nothing to distinguish this study from others which assess the fishery impacts of marine reserves.
- Does fishing occur inside the marine reserves, or outside the marine reserves? If the fishing occurs inside the marine reserve then please clarify that the marine reserves are NOT no-take marine reserves.
- Were the fisheries at each study site managed as TURFs prior to the establishment of the marine reserves?
- L67. What is a triple bottom line evaluation of the effectiveness of community based marine reserves? Please clarify triple bottom line and what the criteria for this are.

Methods

- Please provide the size of each of the reserves.
- Reserve size is an important factor affecting ecological outcomes, especially when comparing comparatively mobile vs non-mobile species (as noted by the authors in the Discussion). Why was the effect of reserve size not assessed?
- Are the control sites managed with TURFs even though they are non-reserves?

The socioeconomic analysis needs to address these points:

- Provide the ex-vessel prices of lobster and invertebrates at each study site.
- Did you control for difference in prices between sites?
- Would observed differences in revenue inside/outside reserves be due to difference in landings or due to price differences?

Results

- Please provide results for biological and socioeconomic indicators - Average density/biomass, landings and revenue for lobster, invertebrates and fish by year in and outside marine reserves (control sites). If

space is an issue then include these results in the Supplementary Materials.

- What was the species and community effect on biological and socioeconomic indicators, respectively?
- Table 2: Some variables need further clarification.
- RU5 – Number of targeted species is the variable, but this isn't specified in the narrative. In addition to noting that lobster is the main target species please enter the total number of targeted species.
- A1 – Number of fishers. I assume that the numbers listed refer only to the number of fishing co-op members. Are there other non-coop fishers in the fishing communities? If yes, what is the ratio of co-op to non co-op fishers? The point of this question is to see how the benefits of TURFs are distributed among the entire community, not just within the fishing co-op.
- A3 – Level of isolation. The narrative says that fishing grounds and reserves are away from dense urban centers – can you provide the distance from each study site to the nearest urban centre (km).
- GS6.2 Which rules (larger min catch size, lower quotas etc) pertain to each site?
- GS9.1 Social monitoring – Do fishing coops at all 3 sites have the same monitoring and enforcement regime in place?
- GS9.2 Biophysical monitoring – Are all the listed monitoring activities performed at all 3 sites?
- GS 10.1 Graduated sanctions – Again, which of the penalties are relevant to each site?
- As it is presented now, the governance variables are a bit unclear as to which type of rules/regulations are practiced at each of the three sites. I recognise that this is because these variables are more qualitative in nature. However to make it clearer, the Narrative column could be split into 3 - one column per study site, and the relevant information entered under each site column.
- Table 2. As per my comment about rating the SES indicators, add a column in between the "Variable" and "Narrative" column to describe the range of potential values/conditions for each "Variable". E.g., for Governance system Social monitoring can range from "Strong" to "Weak". You could attach points to a gradient of "Strong" to "Weak" to produce a quantitative assessment, or enter a qualitative assessment. Then each site will either be assessed with points or a qualitative descriptor. The main point is to provide the reader with a better sense of what a "strong" or "weak" governance system looks like, and where each site falls along this gradient. This is missing from the current table.
- Fig 2B and 2D. Why are there big fluctuations in fish biomass from Years 7-10, but density stays pretty much the same for the same time period?

Discussion

- L238-240: "Our analysis of economic data supports this hypothesis. . . .reserve implementation". I don't understand what the authors are trying to convey in this sentence. What are the expected short-term costs associated to the first years of reserve implementation? Please clarify.
- L208-209 – The authors state that "Understanding the socio-ecological context . . . .provide insights as to why this happens", but then don't provide further elaboration. The reader is not getting any insight as to why the communities continue to support the reserves. Even though this is presumably covered in the provided reference to Ayer et al. 2018 a brief explanation is required here. Please add this to the Discussion.
- Clearly state the lessons learned. L274 simply says that "Lessons learned from these cases can guide implementation. . . " But what are these lessons?
- L277 the authors state that "Community-based marine reserves might have more benefits. . . " Are the authors referring to the reserves assessed in this study, or to community based marine reserves in general? If referring to this study, the conclusion is not supported by the results, as the authors did not compare community-based to non-community based reserves. Need to clarify this point.

- L268-269 “Furthermore, the lack of effectiveness observed in these reserves should not be generalizable to other reserves established under the same legal framework”. This appear to contradict L99-100 which says that “it is safe to cautiously generalize our insights to other similar reserves. . . ”
- L196. “. . . the benefits of conservation directly benefit the members of the fishing cooperatives. . . ” But what about benefits to the community as a whole? Going back to my question about number of actors in Table 2 – I assume that the invertebrate fisheries are high value fisheries. As such, it’s important to consider how the benefits of TURF-reserves are distributed among the entire community. Please provide some discussion about the distribution of fishery benefits – do the number of TURF fishers make up only a small proportion of total active fishers in the community? Or are all fishers in each community able to participate in the TURFs?
- The Discussion is focused mainly on ecological factors which may have affected the effectiveness of the TURF-reserves. What about potential changes in social-economic and governance conditions over the assessed time period which may have affected the reserve? For instance was there a breakdown in enforcement which may have led to illegal fishing, thereby resulting in lower density of target species? What about market drivers that may have motivated increased fishing pressure?
- Elaboration on these factors will strengthen the authors’ attempt to bring in the SES framework to this primarily biological analysis.

Minor comments:

Figure 2 caption: first line (IN; red circles) “circles” is spelled wrong.

Q4 - Please comment on the methods, results and data interpretation. If there are any objective errors, or if the conclusions are not supported, you should detail your concerns

- These comments are covered in Q3 above.

Q5 - Check List

Is the English language of sufficient quality?

- Yes

Is the quality of the figures and tables satisfactory?

- Yes

Does the reference list cover the relevant literature adequately and in an unbiased manner?

- Yes

Are the statistical methods valid and correctly applied? (e.g. sample size, choice of test)

- Yes

Are the methods sufficiently documented to allow replication studies?

- Yes

Are the data underlying the study available in either the article, supplement, or deposited in a repository? (Sequence/expression data, protein/molecule characterizations, annotations, and taxonomy data are required to be deposited in public repositories prior to publication) (yes/no/not applicable)

- Yes