Recommendation by the Subject Editor (anonymous):  
This paper was read by two reviewers, both of whom saw value in the paper for the Ecography audience. However, both had concerns about the presentation and accessibility of the manuscript as currently written, which will take some hard work to address. A number of the statistical details need to be presented more clearly (and some potential statistical issues related to estimability and AIC addressed), and the overall framing presented more clearly to non-fisheries audiences. Ecography's readership works in a wide range of ecosystems, and this paper should be accessible to that range.  
  
Reviewer(s)' Comments to Author:  
Reviewer: 1  
  
Comments to the Author  
I was excited to read this manuscript as it presents a new way to detect shift in species distributions, which is clearly an important topic. Overall, I like the approach presented and think it will be a good contribution. I also think it fits well with Ecography’s goals and scope. As you see below, however, I do have some reservations with the approach. In addition, I think many of the details are missing, which makes it hard to fully assess the validity of the method and the quality of the results.  
  
Major comments:  
  
1. In the methods section, most of the model details are lacking. I appreciate that part of the reason was potentially to keep the description general, but by doing so it makes it really hard for readers and reviewers to assess the validity of the method and nearly impossible to reproduce the method on their on data. Thus, I think the manuscript needs to present the full models (with all of the equations).  I think it would make sense to present the full model for the fish data in details in the main text, thus including the observation equation, the link function, the distribution used for the random fields, the correlation structure (Mattern?), etc.  I would also write down the full model for the simulation, but this could be placed in the supplementary material. I realize that the code is available, but this is not sufficient, the model must be described. Describing the model is essential to understand things like the parameters stated in Table S1 and the results presented in Fig. S1.

**We appreciate the desire of the reviewer and other readers to have these details readily available. We have revised the methods in the main text to include all relevant details and full equations. We have also added details regarding the simulations to the appendix as requested.**  
  
2. I think these kinds of spatial random fields model suffer from Identifiability issues even without observation error and looking at the equation definitely make me fear that this is the case. The simulation studies should explore changes in the magnitude of all variance parameters (sigma\_0, sigma\_{0,trend} in addition to sigma\_E and phi).

**There is an existing literature exploring these issues in models with the same essential structure, but without the spatial trend component. We modified the text to point the reader to these studies. Furthermore, to address the question in the context of this specific model, we performed additional simulations as recommended to determine how the magnitude of the spatial variance (sigma\_O) and spatial trend variance (sigma\_0\_trend) influence model performance. These results were in line with our intuition and are described in a new appendix figure (Fig. S2) along with additional text in the results of the main manuscript.**  
  
3. I’m not a 100% sure what the theta is in the simulation results (e.g. in Fig 3). My understanding is that Z\_s will have multiple values for a simulation (one for each point s), so is theta the mean Z\_s or are you taking each Z\_s value of a simulation to be independent (then why not just call theta Z\_s and \hat{Z}\_s?). I don’t feel that either is great nor terrible, but knowing what it is definitely essential.

**What is shown in the figure are the distributions of location-by-location comparisons. We understand the confusion and have revised the text in the figure caption to clarify.**  
  
4. My understanding is that AIC can be pretty poor at selecting between random effects due to parameter boundary problems. It would be worth using the simulation to see whether you can select between the model with and without the spatial trend. This could be done by fitting the model with and without the trends to your simulations and see whether you can recover the good model. You may need to add a model version without a spatial trend.

**We have now generated data with and without a spatial trend and compared models using AIC. We were able to recover the correct model and have summarized these results in a new figure (Fig. S3).**

5. Have you fitted the model with SST as a covariate and compared it with the trend model? It looks like some of the patterns are likely driven by changes temperature and it would help when understanding the patterns and support some of the statements in the discussion.

**We have not included such covariates because the purpose of the analysis was to provide the best description of the spatially-explicit temporal trends rather than ascribe them to specific dynamic environmental drivers. We would argue that this is out of scope and would need to be tackled in a separate paper. The purpose of the latent variable modeling approach here is to account for these unmodeled variables. We’ve already noted in the introduction and methods that other covariates could be included and cited literature where others have explored the costs and benefits of doing so. To further emphasize this, we have added text to the methods with additional citations and a sentence describing the reasoning for the latent variable approach described here.**  
  
6. L217-225: it’s only after reading the results and really digging into the figures that I understood what you are comparing. My understanding now is that what is done by others is calculate the COG either coastwise/for the whole range of species (which you call coarse-scale), or by sub-regions (which we could call med-scale?) - in your study there are 3 sub-regions divided by Point Conception and Cape Mendocino (you could specific these in a map and refer to it, see comment below for Fig 2). You are proposing that looking at the spatial trend (Z\_s, which you call fine-scale) is much better. Could you please clarify this.

**We revised the final paragraph of the introduction to more explicitly set up the contrast from coarse- to fine-scale indicators, noting which approach applies to each and clarifying that this is a major objective of the study.**   
  
7. Similarly here, it took me a while to understand the clustering algorithm and its goal. I think it needs to be clarified here and likely needs its own paragraph (separated from the COG). I might be wrong, but my understanding is this is a non-spatial clustering technique that is used, and that the only spatial component is the latitude. Wouldn’t you want to use a spatial clustering algorithm? Or something that allows to identify hotspots and coldspots? My intuition is that this would allow you to better identify the shifts inland that are discussed in the discussion. Here the cluster are really restricted by latitude and the patterns of the clusters are not particularly striking. I wonder (not sure if it makes sense), whether it would be worth plotting the change in COG for the clusters?

**The reviewer is correct in thinking this is a nonspatial clustering technique, and we clarify this in the text. We do not use a spatial clustering technique because the spatial autocorrelation is already accounted for in the local trend model. Certainly, there are many ways that one could attempt to summarize the results of the local trend predictions more coarsely, and we have modified the text to acknowledge this point and emphasize that this is simply meant to serve as a qualitative summary of the spatial structure of the local trend. As for “hot” and “cold” spots, we already do this in Figure 4 by identifying which clusters had anomalously higher or lower values of the spatial trend. Regarding the influence of latitude on the cluster patterns, we also experimented with clustering the local trend alone without latitude and the resulting patterns were quite similar (yet slightly patchier in space). Finally, we also experimented with computing the COGs for each cluster rather than the biogeographic boundaries, and patterns were similar, yet we did not include the analysis in this form because we were concerned about potential confounding or “double-counting” of the influence of latitude in the clustering and in weighting densities by latitude (as is done to compute the COG).**  
8. I don’t find Fig 4 compelling. Almost none of the clusters matched the bottom line, and only a few matched the top line. What are you defining as near? I woonder if adding area boundaries to the second column of Fig 5 would better show your results? Or maybe the clusters are not really getting at these boundaries?  
  
Minor comments:  
  
The title is not really selling the paper. It feels like you are comparing metrics that are already available, but in fact you are proposing a new method.

**We revised the title to emphasize that we are introducing a novel approach.**

L147: I think it’s a bad idea to call it spatial trend. You are trying to get a temporal trend? I think generally spatial trend is understood to be a systematic (and generally smooth) change in the mean value of a variable over an area. Here it’s just a random field no (and from Fig 5 it not particularly smooth and systematic)? And the main interest is that it demonstrates how the other random field change through time.

**To avoid confusion, we define this additional random slope field as the spatially explicit temporal trend, but have introduced a new term to replace “spatial trend” as the shorthand for this term, by replacing this phrase with “local trend” throughout.**  
  
L153: I realize that the model presented is likely a state-space model, especially when we include the observation error, but nowhere is it referred as such (and actually the observation equation is never presented), so this connection is not clear.  Worth mentioning above that it is a state-space model. Also, would it be a state-space if there were no observation error?

**There is no explicit observation model. However, even without an observation model, it is still a state space model because the random effects are estimated at the knot locations and predicted to the data locations. However, we do not think there is a benefit to describing the model using state-space model terminology as it could also be labeled as a mixed effects model (as we describe it in the text) or random effects model and using state-space terminology may confuse the matter for some readers.**   
  
L170-182: Could you add 1-2 sentences describing the survey effort.

**We added additional details regarding the survey methods and data, as requested.**

I would number all of the equations. It helps to refer to them.

**Added equation numbers as requested.**

Fig. 2 Please add boundaries around the 3 sub-regions and/or large lines at the two breaks, so the sub-regions are clear.

**We did not include lines separating these boundaries in this figure because at this point in the main text, there has not been a reference to these subregions. However, we do include labels for the locations that serve as latitudinal boundaries separating the region and clearly explain this in the figure legend.**  
  
Fig 3.  why is  sigma=1 (which according to table S1 was explored) is not presented?   
This is a really really small detail, but it would help quickly understand the results (without going back to the table S1), if you had 0.01 rather than 0.00 on the y-axis.

**We have clarified which levels of the parameters were evaluated in the figure caption to address both of these comments. We did not show the results for sigma=1 because they differed little from sigma=0.75, thus we simplified the figure to focus on the main gradients/thresholds. We added text to the figure caption describing this reasoning.**  
  
Fig.5 Why the mean density value over all years? Why not the value at t 0 and final time, to show the trends. Even if it’s not striking, it might make the point that just comparing the predicted value at time 0 and last year is less powerful than looking at the trend estimated. The map projection is also strange.

**We chose the mean density over all years because this can be interpreted as the “weight” on the local trend, serving as a basis for discussing how distributions change relative to their average distribution. We consider this a more important point for explaining distribution shifts than simply pointing out that the trend is more interesting than a comparison of distributions at the initial and final time (a point which was already made more rigorously by comparing post-hoc regression to estimate local trends rather than estimating the trend within the model). As for the map projection, we agree that it is slightly distorted but prefer this projection as it makes the prediction surfaces easier to see, rather than being extremely laterally compressed.**

Reviewer: 2  
  
Comments to the Author  
Thank you for giving me the opportunity to review the manuscript “Comparing metrics of species distribution change across spatial scales with spatiotemporal models”.  The authors present a new modeling approach to account for spatial and temporal variability in species density and distribution.  I find the title somewhat misleading as I don’t see this as a comparison but rather the introduction of a new model.  Although I am familiar with species distribution models, I’m afraid I found the paper hard to follow and some of the figures completely indecipherable.  I fear the utility of the paper may be lost on someone unfamiliar with fisheries.

**We revised the title to highlight our development of a novel approach. We also performed a major reorganization and revision of the introduction to broaden the scope and further emphasize relevance to fields beyond fisheries. The same modeling approach can be applied to many other forms of biological and environmental data. We have also revised the figures, figure captions, and methods to clarify the main points they are meant to communicate.**

As a terrestrial species biologist I found myself looking for terrestrial analogies and trying to understand the methods in that context.  The authors cite Yackulic et al 2013 who provide a review of MAXENT and presence only data.  I wonder if a better comparison would be to the extensive literature on spatially explicit mark recapture (SECR) that easily incorporates temporal and spatial variation in species density.  I’m obviously biased and perhaps this is an unfair criticism but I wonder if the paper could be made more broadly approachable by adding comparisons to spatial mark recapture and or terrestrial systems and surveys?

**We added additional text and citations to the introduction and discussion to provide more direct analogies to established terrestrial approaches to species distribution modeling and multiple data types.**

If I understand correctly a major limitation in the model is that it requires a linear trend in density and distribution.  Given the introduction I assumed that the model would take into account species density rather than distribution to better track population trend (i.e. a shift in distribution may suggest a decline in population that is not supported if density is accounted for).  I’m left wondering what happens if a decline is not systematic and why this is better (as stated on line 170).

**We are unclear on what the referee is asking here, but it seems to arise from some slight misunderstandings. Indeed, the model does predict population density at each location and does not make any strict assumptions about the temporal trend in abundance of the whole population. To rectify similar potential misunderstandings, we have worked to improve clarity throughout, including a more explicit explanation of what is meant by references to a linear response. To clarify, a linear model in this context is one where the response is a linear function of the coefficients. However, this does not mean that all response variables must have a linear relationship with covariates. We also note that if nonlinear responses in many local areas are present in the data, one could adapt the model to account for this (now noted at end of first methods section).**

Lines 34-37 – You provide examples of the ranges of things that might be have distributions we want to track, but you don’t anything about what might cause distribution shifts.  Might also be beneficial to add a sentence regarding what causes distribution shifts.

**We added some text in this first paragraph of the introduction to clarify the mechanisms by which distributions change.**

Line 40 – “However, when reliable population density data are available, distribution shifts are better quantified by spatial predictions of population size” This statement might benefit from the support of a citation.

**While there is not necessarily a great citation that demonstrates this directly, it can be understood to be true based on first principles. We modified the text to make the link between these sentences and subsequent paragraph stronger, where we have references that demonstrate how the hypothesis that the distribution of abundance within a species range is greatest at the center and declines smoothly toward the range edge (the abundant-center hypothesis) has been debunked (Sagarin and Gaines 2002, Sagarin et al. 2006). We also added additional text to this first paragraph of the introduction to explain why population size data is inherently richer than say, presence/absence data.**

Line 55 – I found this transition a little distracting.  I don’t find the choice of appropriate scale any more or less germane to this question than any other in ecology so am not sure this is necessary.

**We added a citation to better emphasize why scale is a particularly important issue in SDMs.**

Line 109 – This is the first use of SDM in the manuscript so the acronym needs to be defined.

**Defined SDM as species distribution model upon this first use, as requested.**

Line 149 – Figure 1, like all of the figures in the manuscript excluding figure 2, is a bit of a challenge.  Judging from the model is it necessary that the trend be linear from year to year?  I’m not sure what it is the authors are trying to convey here in the figure.

**To clarify the message, we added another axis label to the figure to indicate that the numbers represent time steps. In addition, we added more text to the figure caption and methods to better explain the main point. With regard to the question about linear assumptions, please see our response to the referee’s general comments. Additional information regarding this approach has already been published in (Anderson and Ward 2019).**

Figure 4 – I find it almost impossible to decipher anything from this figure.

Figure 5 – labeling the North American continent may make it more readily apparent that each tile represents the west coast.  What do the colors in column 2 represent?  There is no legend.

**The study region was established in figure 2, so we’d hope that the reader would recall this. We did not include a legend for colors in column 2 because they simply represent different clusters of the spatial trend (as labeled at the top of this column), and since species differ in their number of clusters, a single legend would be inadequate or confusing. However, we do include a description of what these colors represent in the figure caption.**

**Literature Cited in Response**

Anderson, S. C., and E. J. Ward. 2019. Black Swans in Space: Modeling Spatiotemporal Processes with Extremes. Ecology 100:e02403.

Sagarin, R. D., and S. D. Gaines. 2002. The ‘abundant centre’ distribution: to what extent is it a biogeographical rule? Ecology Letters 5:137–147.

Sagarin, R. D., S. D. Gaines, and B. Gaylord. 2006. Moving beyond assumptions to understand abundance distributions across the ranges of species. Trends in Ecology & Evolution 21:524–530.