

Editor and reviewer comments (we provide below the full context of the two reviewers' comments) are in *italics*, while our responses are in regular text.

Response to editor's comments:

We are happy to hear that the three reviewers recognized the value of our work and that the manuscript is ready for publication. We have worked to fix the last small points raised by the reviewers, and we would like to thank them again for their careful reading. Furthermore, we also particularly appreciate the really interesting perspective on the policy implications from M. Authier, and...

Response to reviewers' comments:

Reviewer 1

The authors have addressed all of my concerns and I believe the MS is (nearly) ready for publication. I suggest one more pass for copy editing to fix small mistakes like inconsistencies for example in "figure" "Figure" and "e.g." "e.g.",

We thank the reviewer for their positive comments regarding our manuscript. We have worked to correct these inconsistencies.

Reviewer 2

All of my major comments were sufficiently addressed in this revision of the manuscript. I appreciate that the supplement is now referenced in the main document (I missed it on my last round of review). I have a few editorial comments for the main text and a few minor comments about the supplement, but I feel this manuscript is almost ready for publication.

...

Manuscript Minor comments:

Abstract: potential typo in second-to-last sentence. Should "where forecasting a natural output" be "where forecasting is a natural output"?

...

Line 39: should "ignore" be "ignores"

Thanks for catching this, it is fixed now.

Line 51: "hypotheses" or "a hypothesis" rather than "hypothesis"

Corrected.

Line 138: there appears to be an extra "in" in "in only in"

Corrected.

Line 215: comma after parameters

Added.

A couple typos in Box e.g., "calibrating submodes separately allows to avoid" missing "us"?

...

Supplement Minor comments:

Although the supplement works through an example far simpler than the types of research questions discussed in the main text, it serves as a reasonable (and accessible) example of how the workflow

could be implemented. There is a brief discussion towards the end of the supplement about the "scalability" of the proposed workflow, but it leaves something to be desired. Have the authors applied this workflow (or pieces of it) to examples that are more complex that could be cited, or is that future work?

...

First green box: typo in RQ.. "What is global trend..." should read "what is the global trend..."

Thank you, corrected.

Step 2: so assuming populations are exchangeable? Is this reasonable?

We have edited the text here to clarify this assumption:

"In this first hierarchical model, we thus assume that populations are exchangeable—an assumption we might refine later."

Step 2: priors on $\sigma_{\alpha,sp}$ and $\sigma_{\beta,sp}$ do not seem appropriate, these parameters have positive support, so should these be half-normals (folded normals)?

...

Stan model not included so could not be assessed (model1_nc2.stan), I recommend sharing for transparency.

...

Figure 15, unclear what the posterior predictive distribution is of, this detail could be added for transparency. Could also add many posterior predictive datasets to the same plot and then overlay the observed data to facilitate easier comparisons between $y\text{-tilde}$ and $y\text{-obs}$.