Response to Reviewers

We thank the editor for the thoughtful comments to improve our manuscript. Our responses to each comment are included below in blue.

Comments to the author:
Tracked changes version
Lines 42-44: This sentence structure is confusing me.

This sentence has been deleted and the prior sentence now reads as the following: " More recently, weighting based solely on skill has given way to weighting based upon both skill and independence in recognition of differences in skill between regions and variables and the lack of independence between GCMs resulting from common bases in model structure (Massoud et al. 2019, 2020a; Sanderson et al. 2015b, 2017; Knutti, 2010; Knutti et al. 2017)."

Lines 50-60: This is coming across as though you're trying to justify your own work. I'd rather see these citations raised in the context of scientific discovery and how the studies have advanced the field.

This paragraph (Lines 49-56) now reads as: "Several studies have examined the effect of model weighting on the outcome of climate change projections from multiple ensembles. For example, in Massoud et al. (2019), the authors utilized information from various model averaging approaches to evaluate an ensemble from the Coupled Model Intercomparison Project Phase 5 (CMIP5; Taylor et al., 2012), finding that Bayesian Model Averaging (BMA) reduced the error by one third and constrained the uncertainty to 20-25% of the raw ensemble for projections of atmospheric river frequency. Massoud et al. (2020a) found that BMA constrained the uncertainty in precipitation projections over the contiguous United States (CONUS) to be a third of that in the original ensemble. In Wootten et al. (2020a), the authors found that ensemble weighting can change dramatically when weighting schemes are applied to statistically downscaled ensembles compared to a raw GCM ensemble."

Lines 62-70: You're kind of just listing these studies and what they did. I'd like to see conclusions that motivate why you chose the problem that you're tackling in the present study.

This paragraph (Lines 58-77) is now written as the following: "Other studies have applied model weighting to a certain variable or to multiple variables and went on to investigate climate change impacts for other variables (e.g., energy and hydrologic cycles). For example, Knutti et al. (2017) extended the weighting scheme of Sanderson et al (2015a; 2017) to projections of Arctic September temperatures and sea ice, finding that the uncertainty could be constrained by the scheme while noting the proposed weighting scheme is one of several that could be used for multiple applications. The National Climate Assessment had previously considered weighting based only on commonly used climate variables (e.g., precipitation and temperature, Wuebbles et al., 2017), but discussions to use additional variables, such as equilibrium climate sensitivity, are currently ongoing. Other studies have calculated weights based on metrics in one domain (e.g. globally) and then applied them to projections for another domain (e.g. North America or Europe) (Massoud et al., 2019). However, these studies are rare, as are studies providing comparisons of various weighting schemes. Examples of these studies include Shin et al. (2020), Brunner et al. (2020a), and Kolosu et al. (2021). Shin et al. (2020) suggested that researchers may provide results from several weighted ensembles to capture the uncertainties of future changes, but did not explore weighting strategies beyond different weighting schemes. Brunner et al (2020a) found that the region can influence the agreement between approaches to constrain uncertainty in the CMIP5 multimodel ensemble. Finally, Kolusu et al. (2021) focusing on a water-related decision context in Africa, finds that projected risk profiles were less sensitive to the weighting schemes used. Such studies as in these examples tend to focus on the sensitivity associated with one to a few components of a multi-model weighting strategy. No prior study (to the author's knowledge) offers a comprehensive cross-comparison of the sensitivity resulting from the choices of the domain, variable, weighting scheme, and ensemble that comprise multi-model weighting strategies. In addition, the primary focus of these studies are continental regions although climate projections are now being used by regional and local organizations for climate impacts assessments and climate adaptation with additional modeling efforts."

Lines 81-83: I like that you're pointing out the novelty of the study. Please also explain why it's important to do this. That is, there are previous studies that have looked at individual pieces of this problem. Why is it important to look at them all together?

The following sentences are added after this sentence (Lines 90-92) to address this: "Prior studies have examined some of these dimensions individually, but the comprehensive experimental matrix used here allows the comparison of model weighting results based on all these dimensions. This is important because there could be high sensitivities in the estimated model weights based on how the weighting strategy is formulated."

Lines 86-87: I'd like to see a sentiment like this earlier in the paper. Saying that we don't have reliable local information for adaptation planning (and that it's kind of difficult to get, hence motivating your study) really points out the problem.

The following has been added at Lines 75-77: "In addition, the primary focus of these studies are continental regions although climate projections are now being used by regional and local organizations for climate impacts assessments and climate adaptation with additional modeling efforts."

Lines 102-105: I like this, but I'd also like to see something a little punchier. Perhaps a sentence after this one saying something like "That is, we want to figure out under what circumstances the projections are sensitive to weighting and why." I think you're getting at this in the final sentence of this paragraph (lines 107-109), but it could be clearer.

To address this, the final sentence of the paragraph (Lines 116-118) now reads as: "Our purpose in this study is not to address the skill of the multi-model weighting strategies in future projections, but rather to assess under what circumstances the projections are sensitive to multi-model weighting strategies and why."

Lines 122-123: It's good to include this, but also CMIP5 is a perfectly legitimate ensemble to use. Communities making adaptation plans aren't necessarily always going to use the latest and greatest, nor should they - imagine if they had used the CMIP6 archive before we figured out why some models had erroneously high climate sensitivities. A flexible method that can work on CMIP5 (or even CMIP3) is valuable.

Thank you for pointing this out. To make this point we have added the following in section 2.2 at Lines 132-133– "That said, the weighting schemes used here are applicable also to other ensembles such as CMIP6 and CMIP3. Therefore, the findings of this study are generalizable to other ensembles."

Lines 130-132: (I'm including this here because I'm reading in order. If you address this later, please ignore this comment.) I agree with getting a high signal-to-noise ratio. It might be worth a comment in the conclusions section about using other scenarios. For example, if you used RCP2.6, would your uncertainty ranges go down, meaning that the amount of climate change has a direct bearing on the sensitivity of the weighting scheme? (This also has implications for the time period that one is using - there may be less uncertainty for mid-century vs late-century projections.)

Thank you for pointing this out. To address this point, we have added the following in Section 4.3 (Lines 509-512): "In this study, we focused on the sensitivity under RCP 8.5 to maximize the effects observed from different weighting strategies. Given the smaller change signals under other RCPs it is possible that the sensitivities observed here have a lesser magnitude under other RCPs. Considering this component is another aspect that could be explored in future work."

Line 444: Did you mean that this has _not_ been explored? Otherwise I'd like to see some citations.

Thank you for pointing this out. You are correct, we meant to say "has not been explored." This is corrected in the text.