

We thank the referee for the invested time to assess our manuscript and for pointing out specific issues to improve upon. We reply individually to all specific points below in blue but start with the general intended changes and comments.

We agree that the two lines of inquiry in the initial manuscript, (i) being the intercomparison of emulators and (ii) being the use cases of emulated climate information for the IAM's emission scenarios, were not sufficiently motivated and structured to pursue one consistent goal. In the initial manuscript the intercomparison appeared only in the final figure item, for one region and without comparison to any ESM, while the uses cases relied throughout on the output of a single emulator. The initial manuscript did not make clear how these two aspects belong together. We intend to address this by making the intercomparison the basis throughout the manuscript, so that both parts provide results to the same overarching question: How much does the choice of regional emulator, and which ESMs it emulates, change (i) the regional climate outcome and (ii) the socioeconomic consequences derived from it?

We intend to make four major changes to address this. Later referred to as C1 to C4.

C1: The intercomparison becomes the central element.

Former Fig. 6 (one region, one scenario, three emulators, no ESMs) is removed and replaced by two new figure items built on the common core set of emulations (C3). New Fig 1 shows the distribution (boxplots) for each AR6 region of annual mean temperatures from all three emulators and for the seven CMIP6 models behind the common core set (C3). New Fig 2 focusses on the upper percentiles (exceedance frequencies and absolute values) of annual temperature, for the same regions and again all three emulators & CMIP6 ESMs and additionally contrasts each emulator on the common core ESM set with the emulator output based on its full ESM set. The new Fig 2 responds directly to RC1's point that agreement at the centre of the distribution is of limited interest, and its second element responds to RC1's point on lines 318–320 by measuring the effect of ESM selection.

C2: The use cases are recomputed to integrate all three emulators.

The regional emulator use cases become part of the intercomparison. Former Figs. 3, 4 and 5 (the scenario-to-climate outcome space, exposed population, and exposed GDP against mitigation cost) were previously computed from MESMER output alone. They will be recomputed and shown for each of the three emulators now based on upper percentile “tas” instead of “TXx”. This aims to bridge the two major inquiries of the manuscript. (i) The differences in climate alone of the emulated tas established in the two new figures 1 & 2 (C1), and (ii) what those differences mean for derived climatic consequences of emission scenarios like GDP or population exposure.

C3: The shared core protocol will be described explicitly.

Our experiment did run a shared core: Identical SCM-derived global-mean temperature trajectories, the same five probability levels of the same SCM ensembles, a common scenario set, a core set of seven emulated ESMs, with annual mean near-surface air temperature delivered by all three emulators and all output harmonised to a common 2.5° grid. The initial submitted manuscript never said so explicitly and only stated the overall emulated variables, ESMs, etc. The revision describes the core protocol explicitly and separately from the extended set of ESMs and variables each emulator contributed beyond it.

C4: The manuscript is restructured so that the intercomparison comes first.

The revised section order that follows from this is: 1 Introduction; 2 Design and protocol (C3); 3 Intercomparison of emulated regional climate (C1, new Figs. 1–2, replacing former Fig. 6); 4

What the choice of emulator and ESM set implies for scenario assessment (C3, former Figs. 3–5, all three emulators); 5 Towards endogenous links (former Sect. 3.1, shortened); 6 Discussion and conclusions; 7 Methods.

RC2

The paper “Advancing the Scenario-to-Climate Link with Spatial Emulators in the FASTMIP Pilot Experiment” by Windisch and colleagues is an interesting paper that argues for the usefulness of spatial emulation and compares three spatial emulators with different properties among each other (PRIME, MESMER and STITCHES).

I think that the paper includes very valuable information about (i) the intercomparing of three (spatial) emulators as well as (ii) why spatial emulators are useful in the IAM-SCM-ESM-pipeline. However, after reading the paper, I was left with the feeling that both of these lines of results are unsatisfying as neither of them is the focus of the paper. Therefore, I wondered why they appear jointly in the same manuscript while both are not worked out fully so that they represent a standalone paper (maybe the IAM-SCM usefulness is closer to a standalone contribution in my view). In that regard my main criticism aligns with the criticism of reviewer #1 and I would ask the authors to either separate these two contributions off into two individual pieces, or very clearly argue why they have to belong to the same paper (if this is possible).

One more specific point on this:

Between sections 3 and 4 I felt that there was the largest break in the narrative of this paper and (at the latest there) I got confused of whether the paper discusses the IAM-->SCM pipeline or intercompares spatial emulators. From reading the title of the paper, I expecting a study that leans more towards section 4 but section 3 is larger.

This is my main point as I think the rest of the paper is relatively well written and at this point in time, I only have very few minor comments:

We agree with the reviewer that in the initially submitted manuscript the two lines of inquiry, the intercomparison and the use cases of spatial emulation, were not sufficiently integrated.

We intend to restructure (C4) the manuscript by adding a fully new intercomparison first (C1). The two new figure items and assessments will intercompare the three emulators and CMIP6 output with boxplots of the emulated and modelled warming across AR6 regions. New figure 1 will use the common core of seven ESMs for the intercomparison. New figure 2 will more closely look at the upper percentiles and further contrasts the results obtained from the common core set of seven emulated ESMs with the results from each emulator’s full ESM set. This new intercomparison is intended to provide the common basis for understanding what aspects of the regional climate information depend on these choices.

This will be followed by the potential use cases of regionally emulated climatic consequences (C2, former figures 3 to 5, scenario-to-climate outcome space, exposed population, and exposed GDP). These three analyses will move from being based on a single emulator to all three emulator contributing, which will consequently integrate the consequences of emulator type and emulated ESM choice into the space of socioeconomic exposure and potential use cases for emission scenarios. The intention is that the first part (C1) establishes how the choice of emulator and of the ESMs being emulated affects the regional climate information obtained from spatial emulation, while the second part (C2) asks what these differences imply when this information is used to derive climatic consequences of emission scenarios.

Hence, the restructured (C4) and new assessments (C1 & 2) intend to show that knowledge of how emulator choice and the emulated ESM set affect the resulting regional climate is key to understanding and interpreting the computed consequences. In this way, we intend the intercomparison and the use cases to form a single line of inquiry: How much does the choice of regional emulator and the choice of which ESMs are emulated affect the computed consequences for emission scenarios by IAMs?

1. Line 88-90: The authors are right, this is all existing pathways and routinely used. Can the authors provide some key references where these standard procedures can be found in literature?

We thank the referee for pointing out the missing reference here. We intend to add Kikstra et al., 2022 ("The IPCC Sixth Assessment Report WGIII climate assessment of mitigation pathways: from emissions to global temperatures", <https://doi.org/10.5194/gmd-15-9075-2022>) which provides an overview of the climate-assessment procedure using scenarios of IAMs in conjunction with simple climate models.

2. Overall, I believe the connection to other MIPs (such as ScenarioMIP) could be exploited even more. I understand that running the ScenarioMIP database might be too much for this work but at least putting it into context would make sense I believe

We agree that the connection to the former ScenarioMIP endeavor could be made stronger. In changes described in C1, we intend to now directly compare emulated and simulated data of the SSP-RCP framework's ESM output of the previous ScenarioMIP effort leading to CMIP6. A similar coordinated emulator effort for the current ScenarioMIP cycle leading to CMIP7 is underway but will not be made part of this manuscript.

3. From all the examples in section 3, I only found section 3.2 (a more comprehensive scenario-to-climate outcome space) the one that would be integrable most easily into the intercomparison/uncertainty part (section 4) of the paper.

We thank the referee for this suggestion. As stated in the general changes C2, we intend to make all the examples except fire weather (scenario-to-climate outcome space, exposed population, and exposed GDP against mitigation cost) part of the intercomparison by switching from using TXx as only computed by MESMER to "tas" computed by all three emulators. Further, the former section 4 will be replaced by changes of C1 towards a more comprehensive intercomparison between emulators and to simulations of CMIP6 ESM runs.