ISIMIP3b paper review comments

Reviewer 2:

Summary

The paper provides a general description of the climate-relative forcings, based on climate simulations performed during CMIP6, and improvements made for realistic representation of various climate forcings in the third simulation round of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP3b). The paper serves as a reference to perform impact model simulations using the forcing specifications described here, which can be used to evaluate the impact of climate forcings in different scenarios.

This is not a particularly long paper since most of it consists of tables and lists, but it is riddled with references pointing to incorrect tables and especially non-existent ones, whose contents cannot be found anywhere on the manuscript. As such, I was not able to comprehensively review this manuscript and recommend that the manuscript go through another round of revision.

Reply: Thanks so much for the careful review. There were indeed some flaws regarding table references and we apologize for that. We corrected these and are happy to receive further comments.

Comments

L93: I know that the description has been given in the first paper for different sectors being considered; however, I'd like to see a concise summary of these sectors in the introduction, especially because too many references have been made here about how this work influences different sectors, but it is difficult to see what they are exactly.

Reply: Thanks for pointing this out. We added the following to the respective sentence in the introduction: "... provides a common scenario framework for cross-sectorally consistent climate impact simulations. Currently, operational simulation protocols exist for the following sectors: Agriculture, Biomes, Energy, Fire, Food security and nutrition, Groundwater, Labour, Lakes global, Lakes regional, Fisheries and marine ecosystems global, Fisheries and marine ecosystems local, Peatland, Permafrost, Water global, Water regional. Additional protocols for Coastal systems, Regional forests, Temperature-related mortality, health indicators, Terrestrial biodiversity and Water quality sectors are under development."

L144: This sentence needs some proof-reading. The scenarios were chosen to "capture a wide range of possible futures", then "the availability of climate model simulations"? Perhaps the authors meant to "utilize" the availability or something along those lines?

Reply: Thank you very much for the hint. The sentence has been changed to: "The selection of ISIMIP3b scenarios (see **Figure 1**) was generally driven by i) the aim to capture a wide range of possible futures from low to high emission scenarios and, provide of a long baseline simulation

assuming pre-industrial climate conditions that allows for a robust estimation of reference return levels of extreme events and ii) the availability of climate model simulations."

L155: "Only" sounds odd here. Perhaps the authors meant "eventually"?

Reply: Yes, the text has been adjusted accordingly.

L157: Wrong quotation mark.

Reply: Thank you very much! The text has been changed accordingly.

L158: I have two issues with this whole discussion on the plausibility of different scenarios. The first is that it has been claimed that the scenarios were chosen to "capture a wide range of possible futures" (L144), but the only thing that is discussed here is how some of the SSP scenarios are unrealistic. I do agree that some SSP scenarios are unlikely [1], but I do not see anything particularly wide regarding the choices. I do not think all future predictions must include extreme cases, but I am just not fully convinced of the stated goal of the ISIMIP3b scenarios. The second is that it is not clear what the authors consider as the "business as usual" scenario. What do the authors actually think of the "business as usual" scenario, now that they have refuted almost all scenarios used in ISIMIP3b? My main concern is that the authors leave a number of open questions regarding the choice of plausible scenarios.

Reply: The selection of the scenarios is a community-driven process constrained by the availability of climate model simulations (multi-GCM ensemble per scenario) and socio-economic background information (such as land use patterns, populations and GDP data etc. additionally required as 'Direct Human Forcing' for the ISIMIP3b, group III impact model simulations that will be introduced in an upcoming paper). These criteria have made CMIP6-ScenarioMIP our reference point for the selection. The 'Tier 1' ScenarioMIP scenarios comprise only four scenarios: SSP5-8.5, SSP3-7.0, SSP2-4.5, and SSP1-2.6 where 'Tier 1 spans a wide range of uncertainty in future forcing pathways important for research in climate science, IAM, and IAV studies, while also providing key scenarios to anchor experiments in a number of other MIPs (see last column in Table 2)' (O'Neill et al. 2016)). When saying that we aim 'to capture a wide range of possible futures' we refer to selecting the highest and the lowest scenario from the CMIP6-ScenarioMIP ensemble. SSP3-7.0 was added in a second step after a discussion among the sectoral coordinators and at the ISIMIP workshops in response to e.g., (Hausfather and Peters 2020). However, it is important to note, that none of the four ScenarioMIP scenarios is considered a 'business as usual scenario' and within ISIMIP do not do so and have never done so either. The critique by Hausfather et al., does not refer to the SSPs but the high emissions of RCP8.5. We modified the paragraph to highlight the clear linkage to CMIP6-ScenarioMIP and to underline that there simply is no 'business as usual scenario' and that there was not even the intention to design one within ScenarioMIP. The paragraph now reads:

'The selection of the scenarios is a community-driven process constrained by the availability of climate model simulations (multi-GCM ensemble per scenario) and socio-economic background information (such as land use patterns, populations and GDP data etc. additionally required as 'Direct Human Forcing' for the ISIMIP3b, group III impact model simulations that will be introduced in an

upcoming paper). These criteria have made CMIP6-ScenarioMIP the reference point for the selection (O'Neill et al. 2016). The selection of ISIMIP3b scenarios (see **Figure 1**) from the four ScenarioMIP Tier 1 scenarios was additionally driven by the aim to capture a wide range of possible futures from low to high emission scenarios and to provide of a long baseline simulation assuming pre-industrial climate conditions that allows for a robust estimation of reference return levels of extreme events. This is why the original selection comprised the pre-industrial baseline ('picontrol'), the historical simulations ('historical'), SSP1-2.6 representing the 'low end of the range of future forcing pathways in the IAM literature' (O'Neill et al. 2016), and SSP5-8.5 representing the 'high end of the range of future pathways in the IAM literature' (O'Neill et al. 2016). Given recent mitigation efforts, some estimates of recoverable coal reserves, and decreasing prices for renewable energies the emissions underlying SSP5-8.5 have been criticised for being unplausibly high (Hausfather and Peters 2020). Based on these discussions, the 'medium to high end of the range of future forcing pathway' SSP3-7.0 (O'Neill et al. 2016) has been added to the ISIMIP3b scenario set-up. While this scenario is described as 'average no climate protection policy' by (Hausfather and Peters 2020), we highlight that we explicitly do not consider it as a 'business as usual scenario' and that this was not the framing within ScenarioMIP either. Instead SSP3-7.0 is based on rather extreme assumptions about land use changes and aerosol emissions e.g. leading to a scaling of precipitation with global mean temperature that diverges from the scaling identified in the other scenarios (Shiogama et al. 2023). In addition, SSP5-8.5 is explicitly kept in the ISIMIP3b ensemble as its particularly strong warming signal allows testing to what degree the simulated impacts of climate may scale with global mean temperature, which could allow for a translation of impacts to other emission scenarios. In addition, even under lower emission scenarios, global warming levels as the ones reached under SSP5-8.5 in 2100 will eventually be reached later in time as long as emissions are not reduced to zero. These impacts of high warming levels would not be captured when only considering lower emission scenarios ending in 2100.'

L205: To be honest, I cannot keep track of all the errors when it comes to references to figures and tables in this paper. Here, "Table 11" should be the correct one. I have also found cases where the authors refer to tables that simply do not exist. The manuscript needs an overhaul.

Reply: Thanks, we reviewed all table and figure references and corrected all errors.

L228: This should be "to include".

Reply: True, thanks.

L229: I believe it should be "impact distribution".

Reply: We think of extreme value distribution as a certain type of statistical distribution. To clarify that we have modified the sentence in the following way:

'In order to allow for the fitting of extreme value distributions such as Gumble or Generalized Extreme Value (GEV) distributions to e.g., annual maximum river discharge to estimate reference 100 year return levels….'

L237: I believe the authors are referring to "Table 2". I will not comment further about this, but always directly refer to the table being discussed, not "the table".

Reply: Correct, thanks.

L298: I do not think a comma is needed here.

Reply: Corrected, thanks.

L309: "Too sparse"?

Reply: Indeed, thanks.

L333: Again, wrong table.

Reply: Changed to Table 2.

L453: I think it is unusual to only write the variable names. The authors did write "sea level pressure (psl)" at one point, and I think this should be the convention. Write both the variables names and the actual names of the properties being referenced so the readers will not have to go back to the table every time they encounter a variable name, at least for the first time.

Reply: Done

L479: Table 16 does not exist.

Reply: Thanks, should be Table 8.

L512: I do not find it particularly significant that the mean ECS is exactly the same as that of CMIP6, as it is a subset of CMIP6 ensemble. However, one important aspect of the estimation of ECS in CMIP6 is that the variability in the estimates of ECS is still very large [2], [3]. I would like to see a detailed discussion on the variability of ECS, especially on how various factors influence the estimate of ECS in each GCM. For example, the five GCMs chosen for this study seem to produce short-wave and long-wave radiative effects from clouds that are consistently close to median RMSD of the CMIP6 ensemble, and I would assume that as a result, the variability should be much smaller than the whole CMIP6 ensemble. Is it really the case? If not, what is driving the variability?

Reply: It's true that the RMSD in these two variables is relatively similar across our five models, and somewhat lower than in the CMIP6 multi-model median (Fig. 2). This is no coincidence, since the relative RMSD shown in Fig. 2 was used as one of several criteria for selecting models for ISIMIP3b; as explained higher up in the same section. The five selected GCMs therefore "by construction" tend to have relatively low RMSD across the set of variables tested in Fig. 2.

However, Fig. 2 shows the (dis-)agreement of models with historical observations, in terms of spatiotemporal variability in different variables; it does not show the variability itself. Therefore, the fact that the five selected models have low RMSD compared to the CMIP6 median does not imply that they exhibit low variability, in terms of cloud radiative effect. Moreover, even though

cloud-related processes appear to play an important role in explaining differences in ECS across models, these relationships are complex and differences cannot be easily attributed to individual processes (Zelinka et al. 2020; Meehl et al. 2020). Most importantly, the two variables in question - lwcre and swcre - measure the *short-term* cloud radiative effect, which has no correlation with the *long-term* cloud feedback that defines ECS (Chao, Zelinka, and Dessler 2024). Thus, the results shown in Fig. 2 are useful for assessing model skill compared to historical observations, but do not allow for any conclusions about the spread in ECS between different models. Generally, different models could achieve a similar agreement with observations through various different combinations of processes and feedback strengths; which is one reason why model estimates of ECS still diverge so strongly. The five models selected here reflect this divergence in estimates, while all being in relatively good agreement with historical observations, which in our opinion makes them a reasonable choice for studying climate impacts (noting that data availability additionally constrained model selection).

We have expanded our discussion of ECS in the climate models, in response to this comment as well as a related comment by reviewer 1, as follows:

"The five GCMs provide a good representation of both the mean and the range of the full CMIP6 multi-model ensemble ECS. According to (Meehl et al. 2020), the CMIP6 multi-model mean ECS is 3.7°C, which is precisely met by the mean ECS of the five ISIMIP3b GCMs. The transient climate response (TCR) of 2.0°C is also precisely met. This provides an improvement over ISIMIP2b, in the sense of the selected GCM subset reflecting the statistics of the larger CMIP ensemble. In ISIMIP2b the mean ECS for the full CMIP5 was 3.2°C compared with a mean ECS of 3.72°C for the four ISMIP2b GCMs (see Table S1 and S2 in (Jägermeyr et al. 2021)). The ISIMIP3b ensemble includes three models with below-average ECS (GFDL-ESM4, MPI-ESM1-2-HR, MRI-ESM2-0) and two models with above-average ECS (IPSL-CM6A-LR, UKESM1-0-LL) (see **Table 4**). In line with their ECS values, we find GFDL-ESM4 and UKESM1-0-LL to project the weakest and strongest global warming, respectively, under any future scenario considered (see **Figure 3**). Under SSP5-8.5, the global mean near-surface temperature in 2100 is about 3°C larger in UKESM1-0-LL than in GFDL-ESM4. Under SSP1-2.6, the projections are about 1.5°C apart. The ensemble mean warming of the ISIMIP3b CMIP6 models is significantly higher than the warming of the ISIMIP2b CMIP5 models, across global land area by an average of 0.3°C, but over the main breadbasket cropland regions by more than 0.5°C between 1983–2013 and 2069–2099, under both SSP1-2.6 and SSP5-8.5 (Table S1 in (Jägermeyr et al. 2021). This is in line with the higher median ECS in CMIP6 compared to CMIP5; indeed, some CMIP6 models have an ECS above the assessed likely (2.5°C to 4°C) and very likely (2°C to 5°C) ranges in the IPCC's sixth assessment report (AR6) (Forster et al. 2021). The reasons for these higher estimates of ECS are complex, with cloud feedback processes playing an important role (Zelinka et al. 2020). While the plausibility of the very high ECS estimates has been questioned, recent studies indicate CMIP6 models with high ECS tend to simulate cloud properties better than low ECS models (Bock and Lauer 2024); also, unaccounted natural variability may have biased the IPCC's assessed ranges somewhat low (Watanabe et al. 2024; Liang, Gillett, and Monahan 2024).

The ISIMIP3b ensemble reflects the spread in ECS of the overall CMIP6 ensemble, with two models above the AR6 likely range and one of these (UKESM1-0-LL) above the very likely range. The strong warming response of these models should be kept in mind when conducting ISIMIP3b-based impacts studies. However, depending on the region and variable of interest, the high ECS does not necessarily

have any bearing on the magnitude or realism of projected regional impacts, and any further selection of models should not be based solely on ECS but on the models' suitability for the impacts variables in question (Swaminathan et al. 2024). In many applications, results can be harmonized by describing the simulated impacts in terms of global mean temperature changes instead of time for the different emission scenarios."

L518: Table 12 does not exist. At this point, there are too many errors involved in references to its own tables. I will not mention them from now on, but I do find it very odd that there are a number of references to tables that simply do not exist in the manuscript.

Reply: Yes, we wanted to point to Table 4 "Characteristics of CMIP6 climate models used in ISIMIP3b". The text has been modified accordingly.

L548: What do you mean that the new approach "improves the spatial variability"?

Reply: A simple interpolation from coarser resolution data to high resolution underestimates the spatial variability within the grid cells. In the new approach this is avoided by adding additional variability at the target resolution. We have modified the sentence to:

"Compared to ISIMIP2b, where climate model output was first spatially interpolated to the target resolution and then bias-adjusted, the new approach avoids the associated underestimation of the spatial variability at the target resolution (Lange 2019)."

L577: From this line on, references appear with parentheses where they should not.

Reply: Has been adjusted.

L612: Why was the 5% threshold chosen? What was the reasoning behind this specific number?

Reply: We used the 5% significance threshold because it is a widely accepted convention in statistics, corresponding to a 95% confidence level and roughly ±2 standard deviations (±1.96) under the normal distribution. This level is commonly applied in many areas of research as it balances the risk of false positives with maintaining sufficient statistical power, whereas stricter (1%) or more lenient (10%) thresholds are typically reserved for specific contexts.

L679: "Artifacts"

Reply: Thanks, corrected.

L722: I am not sure what I am supposed to find out in Table 4, which does not include any specifications of forcings or GCM data. This could be another case of incorrect table reference, but I cannot find the relevant information anywhere on the submitted manuscript.

Reply: We wanted to point to Table 6 where the GCM data used to force the TC models are listed. The text has been adjusted accordingly.

L783: I am fairly certain that this is against GMD policy.

Reply: We have revised the Data and code availability section and implemented an automatic process into the ISIMIP repository to retrieve the data upon signing a non-commercial license. We addressed the issue in detail in the answer to the related chief editors comment.

L787: What is the point of having the "Datasets" column if they are all the same? Surely you could have just added a note just for the surface temperature, or write it down in caption?

Reply: Besides others, the paper serves as a look-up document for modelers performing climate impact simulations in line with the ISIMIP protocol. To make information as easy to find as possible and to avoid confusion we decided to stick to a common and systematic table layout also in cases where this may result in very short tables, or repeated table entries.

Bibliography

Bock, Lisa, and Axel Lauer. 2024. "Cloud Properties and Their Projected Changes in CMIP Models with Low to High Climate Sensitivity." *Atmospheric Chemistry and Physics* 24 (3): 1587–1605.

Chao, Li-Wei, Mark D. Zelinka, and Andrew E. Dessler. 2024. "Evaluating Cloud Feedback Components in Observations and Their Representation in Climate Models." *Journal of Geophysical Research Atmospheres* 129 (2). https://doi.org/10.1029/2023jd039427.

Forster, P., T. Storelvmo, K. Armour, W. Collins, J-L Dufresne, D. Frame, D. Lunt, et al. 2021. *Chapter 7: The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity. Climate Change 2021: The Physical Science Basis*.

Hausfather, Zeke, and Glen P. Peters. 2020. "Emissions – the 'business as Usual' Story Is Misleading." Nature Publishing Group UK. January 29, 2020. https://doi.org/10.1038/d41586-020-00177-3.

Jägermeyr, Jonas, Christoph Müller, Alex C. Ruane, Joshua Elliott, Juraj Balkovic, Oscar Castillo, Babacar Faye, et al. 2021. "Climate Impacts on Global Agriculture Emerge Earlier in New Generation of Climate and Crop Models." *Nature Food* 2 (11): 873–85.

Lange, Stefan. 2019. "Trend-Preserving Bias Adjustment and Statistical Downscaling with ISIMIP3BASD (v1.0)." *Geoscientific Model Development* 12 (7): 3055–70.

Liang, Yongxiao, Nathan P. Gillett, and Adam H. Monahan. 2024. "Accounting for Pacific Climate Variability Increases Projected Global Warming." *Nature Climate Change* 14 (6): 608–14.

Meehl, Gerald A., Catherine A. Senior, Veronika Eyring, Gregory Flato, Jean-Francois Lamarque, Ronald J. Stouffer, Karl E. Taylor, and Manuel Schlund. 2020. "Context for Interpreting Equilibrium Climate Sensitivity and Transient Climate Response from the CMIP6 Earth System Models." *Science Advances* 6 (26): eaba1981.

O'Neill, Brian C., Claudia Tebaldi, Detlef P. van Vuuren, Veronika Eyring, Pierre Friedlingstein, George Hurtt, Reto Knutti, et al. 2016. "The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6." *Geoscientific Model Development* 9 (9): 3461–82.

Shiogama, Hideo, Shinichiro Fujimori, Tomoko Hasegawa, Michiya Hayashi, Yukiko Hirabayashi, Tomoo Ogura, Toshichika Iizumi, Kiyoshi Takahashi, and Toshihiko Takemura. 2023. "Important Distinctiveness of SSP3–7.0 for Use in Impact Assessments." *Nature Climate Change* 13 (12): 1276–78.

Swaminathan, Ranjini, Jacob Schewe, Jeremy Walton, Klaus Zimmermann, Colin Jones, Richard A. Betts, Chantelle Burton, et al. 2024. "Regional Impacts Poorly Constrained by

Climate Sensitivity." *Earth's Future* 12 (12). https://doi.org/10.1029/2024ef004901.
Watanabe, Masahiro, Sarah M. Kang, Matthew Collins, Yen-Ting Hwang, Shayne McGregor, and Malte F. Stuecker. 2024. "Possible Shift in Controls of the Tropical Pacific Surface Warming Pattern." *Nature* 630 (8016): 315–24.
Zelinka, Mark D., Timothy A. Myers, Daniel T. McCoy, Stephen Po-Chedley, Peter M. Caldwell, Paulo Ceppi, Stephen A. Klein, and Karl E. Taylor. 2020. "Causes of Higher Climate Sensitivity in CMIP6 Models." *Geophysical Research Letters* 47 (1). https://doi.org/10.1029/2019gl085782.