Huber et al. ESD-2014-23 point-by-point responses

Referee #1

General: The manuscript provides a brief review of previous research to quantify the impacts of climate change, and describes the specific contributions of the ISI-MIP project, including how the design of the framework addresses specific challenges in modeling impacts, adaptation, and vulnerability. Overall the manuscript represents a contribution to the climate change science literature, in particular with regards to the review of existing multi-sector analyses, and the descriptions of how the ISI-MIP frame-work is providing insights regarding the relative importance of uncertainties related to structure and design of models and other uncertainty sources that are more commonly explored in the literature. I believe that the theme and content of the manuscript are appropriate regarding the scope of the Journal, and therefore potentially suitable for publication following major revision.

Specific: Page 723, lines 24-26: It is worth clarifying that a "comprehensive assess-ment" of climate change impacts would be the ideal type of information to inform mitigation decisions, but that analyses which are less than comprehensive (yet still containing a rich set of information) can be policy relevant. Given the large gaps which still exist in the science community's ability to produce such a comprehensive assessment, this paper should not suggest that effective policies cannot be designed without this ideal level of information.

We agree with the referee that adequate mitigation decisions can be based on analyses that are less then comprehensive. We have chosen a more careful wording in the revised version of the manuscript (ll. 45-50) to avoid any misunderstandings.

Page 724, lines 12-16: "Assessing the vulnerability of human and natural systems to climate change is not possible without accounting for the interactive effects of ..." Again, this wording seems to strong here, and suggests a situation of the perfect being the enemy of the good. One could argue that some climate change impacts can be reasonably modeled in isolation, such as impacts on individual animal species, to tell policy-relevant stories. Other sectors obviously deserve attention to integration, and perhaps more importantly, connectivity on key inputs/assumptions.

We have changed the wording in the revised version of the manuscript (ll. 58-64), expressing that accounting for interactive effects is highly desirable but that studies investigating specific climate impacts in isolation are of course also valuable.

Section 2: The review of existing model-based assessments is a worthwhile endeavor, and if improved will provide a useful set of information to the research community. However the current draft mischaracterizes some of the projects, and care should be taken to appropriately represent what others are doing in the field.

One issue that should be clarified at the start of the section is that the review is based on the description of modeling projects as described in the papers which have been published on the efforts. But due to the delays in the publication process, which creates a gap or delay between reported and actual project status, and because these papers may not reflect the totality of any one project (as some papers just focus on one aspect or sector of the broader project), there is a risk that the information presented in this review is not complete.

We think that most readers will be aware of this shortcoming, which is relevant to any literature review. Nevertheless, in order to appreciate the many ongoing integrative efforts in climate impacts research that are not yet reflected in publications we explicitly mention the inevitable incompleteness of information in the revised version of the manuscript (ll. 139-140).

Another issue is that the categories of model-based assessments developed by the authors do not provide for clean fits for some of the efforts. For example: The CIRA project does not fit perfectly in section 2.1 (several sectors, one model) for the following reasons: 1) over 20 different partial equilibrium, bottom-up models are used to estimate impacts; 2) some sectors have multiple models to analyze structural uncertainties (e.g., three electric power system models, three agricultural yield models) – although inter-model comparison is clearly not a focus of CIRA to the extent of ISI-MIP; and 3) there is integration, or at least linkages, across some sectors(e.g., agriculture, water, energy). As the lines separating the categories established by the authors appear to become blurred

Indeed, some of the mentioned projects would fit into several of the established categories. Nevertheless, we think that the proposed categories (i.e., the impacts integration matrix; Fig. 1) are a useful tool to summarize the current state of integrated assessments of climate change impacts. In the revised version of the manuscript, we now state explicitly that subcomponents of the presented projects may be placed into a different category (ll. 140-143). This is particularly relevant for projects such as CIRA, which contain some elements of model intercomparison, without them being a focus of the project though (ll. 152-154).

Similarly, PESETA uses multiple partial equilibrium models to estimate impacts, reports these impacts at the sector level, and then feeds the information into a single CGE to capture welfare effects in an economic framework (i.e., it is not a project investigating several sectors using one model). Also, the description of the PESETA project should be updated to account for the recently released PESETA II report. Additional sectors have been modeled, along with other methodological improvements, which may or may not be worth mentioning here.

We discuss the PESETA project as an approach that investigates several sectors within a consistent framework using one bottom-up model *per sector*. It is important to note that the relevant heading "Several sectors, one model" should be interpreted as referring to both integrated multi-sectoral models and multi-sectoral assessments of climate impacts, using one model in each sector. We clearly distinguish these two options at the beginning of section 2.1 (ll. 126-137). Also, the revised version of the manuscript now includes information on the more recent second phase of the PESETA project (ll. 145-150).

Page 729, lines 19-20: It is worth noting that there are different levels of integration, and the authors should be clear regarding what has been done in the ISI-MIP project. For example, multi-sector projects can conduct initial linkages between models (e.g., water availability for hydrologic models informing irrigation availability in crop models). A more in-depth effort would entail cross-model convergence to equilibrium, which typically requires many coordinated runs of each model involved. Finally, integration at the extreme end would entail the creation of an integrated assessment modeling framework to dynamically link the sectors.

As pointed out in the previous response we discuss two of these levels of integration in the beginning of section 2.1: first, dynamically linking sectors by creating an integrated model, and second using offline simulations of different sectoral models (which entails at least some harmonization of the input). Between these two extreme ends, there are obviously intermediate steps of increasingly integrated models/model runs. In the revised version, we have added a sentence to be clearer of what level of sectoral integration has been pursued in ISI-MIP (ll. 192-194).

Page 729, lines 26-27: My sense is that this is not an accurate statement, as there are sectoral models to analyze a number of these impacts, particularly for the EU, US, and Japan. I agree that these sectors have not received as much attention when compared to agriculture, water, and energy, but the modeling platforms do exist – even if they are not region-specific and there are not ten of them to do an inter-comparison exercise.

We agree that for some of these sectors models exist. However, these are generally employed at the regional scale only (EU, US, Japan as the referee points out), and there are not yet enough models to undertake a meaningful comparison. As we understand it, our statement on p.729, l. 26-27 ("For some of these areas, not even one global-scale model exists yet, let alone ensembles of comparable models.") does not contradict the referee's and our assessment of the state of the art.

Pages 731, line 13 through page 732, line 22: These paragraphs are very interesting, and this work will be of great interest to the impacts community. This is clearly a strength of the ISI-MIP approach.

Thank you.

Page 733, line 25: Another limitation is that intercomparisons cannot be undertaken until there are a sufficient number of models to analyze jointly. In the interim, individual models can still provide policy relevant information, even if the structural uncertainties associated with impact model development have not yet been fully explored.

We agree that individual models can provide policy relevant information, as long as the specific assumptions of the model are communicated (see revised version of the manuscript ll. 403-404). However, to our mind there is no such thing as an insufficient number of models for an intercomparison exercise. Comparing two models can make sense depending on the research questions posed. On the other hand, a very large model ensemble can still be too small in order to capture the conceivable structural uncertainty.

Referee #2

General: This paper provides a useful overview of the state of the art in climate change impacts modelling with particular reference to the question of uncertainties within impacts projection. It provides reasonable coverage of the methods that have been hitherto employed to address this important topic. It continues to describe a recent initiative to break new ground in this area and provides a useful summary of the methodology and some of the new findings. It can therefore be said to present a novel concept (that of the new model comparison process) and continues to suggest future useful research ideas, and reaches important conclusions about where research in this area needs to be improved.

The topic which the paper addresses, that of uncertainties in climate change impacts modelling, is of critical importance owing to the status of the UNFCCC climate change negotiations and the upcoming review of the global temperature rise target. The topic is of relevance to the Earth System Dynamics Discussions.

Proper credit is given to related work, all of the references should be retained. The title and abstact are appropriate. The paper is clearly presented and well written,and the language is fluent and precise. As this paper is largely a review and opinion piece, describing the state of the art reached a result of a number of existing publications, and suggesting further ways forward, questions 4 to 6 asked of peer reviewers are not relevant to this paper. The figures used to illustrate the text are useful and should be retained.

One or two remarks made in the paper are questionable and should be reworded – for detail see below. There is one key issue that should be detailed further, see below.

P734 line 9-10. This sentence is questionable and needs to be reworded. It also prompts me to highlight an issue which is not discussed in the paper and needs to be highlighted and clarified. Delete 'modellers tend to adjust their models'.

I think the concept you should be conveying here is that model intercomparison needs to do the following (i) Allow identification of errors in input data, so that this can be excluded (ii) Isolate the influence of model structure, which requires harmonisation of model input data. The output of such a comparison does produce more convergence of output. (iii) However, once errors and structural issues have been explored/addressed, one then still has to go back and explore the total uncertainty, which results from the sum of uncertainties in input data and uncertainties in structure. Obviously, where there is genuine uncertainty in input data, a model intercomparison that then ignores this uncertainty is biased. In ISIMIP I think that of necessity input data has been harmonised? Rather than saying that modelers have been adjusting their data, implying malpractise, say that of necessity input data has to be harmonised in order to tease out those output differences that are dependent only on model structure (i.e. (ii) above). Then, when one returns to assessing the uncertainty in model output, one has to put BACK the diversity in knowledge of the input data in order to encompass the full range of uncertainty (i.e. (iii) above). It is important to highlight whether this has or has not been done in the ISIMIP papers published so far, in other words, whether the range of uncertainties presented is actually lower than it should be due to artificial harmonisation of input data in order to isolate the effect of model structure only.

Of course, if separately, you think that modelers have been deliberately adjusting their models to produce a mean outcome, then that is quite a serious inditement of (presumably particular) scientists integrity and I would recommend that you steer clear of making such statements which would be extremely difficult to substantiate and likely to open up an unhelpful debate about scientific consensus about climate change in general that could undermine the work of large numbers of extremely meticulous and upright scientists.

We did not want to imply any malpractice of modelers, intentionally adjusting their models so as to follow the ensemble mean. Yet, we agree with Knutti (2010) that a tendency towards consensus can be a potential shortcoming of model intercomparison projects. Based on his experience as a leading scientist of CMIP, he has termed this phenomenon an 'element of social anchoring'. In the revised version of the manuscript we have deleted "modelers tend to adjust their models" as suggested and cite Knutti (2010) more directly, using his wording (ll. 393-396). (In fact, in the previous version of the manuscript we accidently cited Sanderson and Knutti, 2012 on this point.)

Concerning the referee's more general statements about dealing with uncertainties in model intercomparison projects, we mention on several occasions that ISI-MIP relied on harmonized input data (e.g., ll. 193; ll. 197-202). The identification of errors was part of the process of providing data to participating modeling groups. The simulation output was then investigated in two ways as suggested by the referee: Some studies focused on structural uncertainties of impact models basing their analysis on a selected input dataset (i.e., one of the five GCMs and one RCP scenario considered in ISI-MIP) (ll. 376-384). Other studies investigated "total uncertainty" combining uncertainties in input data (stemming from different RCPs, different GCMs, and bias-correction) with structural uncertainties from impact models (ll. 333-340). However, as we point out in the manuscript, this assessment of total uncertainty is strongly dependent on the considered model ensembles, both with regard to input (climate model) and output (impact model) data (ll. 358-360).

P733 lines 19-25 Remove the citation to Tavoni & Tol, this argument does not make any sense – just because only a few people have estimated something doesn't mean it is necessarily underestimated. It just means the numbers are less certain. In fact, economic costs of mitigation are probably overestimated because of lack of proper incorporation of economic gains resulting from investment in new technologies, and the incorporation of assumptions that the economy is in perfect equilibrium in many models. Secondly, cobenefits (such as energy security and improved health effects) were not included in the AR5 and these influences would seem to be far more important- effectively reducing the costs of mitigation by a large amount. There is plenty of evidence for these processes in IPCC AR5 and citations therein.

The citation to Tavoni and Tol (2010) was meant to provide an example of a potential shortcoming of model intercomparison, rather than be a comment on the question whether economic costs of stringent mitigation were underestimated in the IPCC AR 4. In the revised version of the manuscript, we have reformulated the entire paragraph, so as to make a much more general statement on the importance of communicating model assumptions to policy makers (ll. 402-413). We now cite Tavoni and Tol (2010) together with another paper (Knopf et al. 2012), presenting their argument as part of a controversial debate about whether AR4 results were biased due to the selection of specific models in the underlying EMF ensemble.

Minor comments: P728 line 20-25 The millions at risk approach was implemented, not proposed. Please change 'proposed' to 'implemented'. I am not convinced that the population scenarios used therein were inconsistent – this study used SRES scenarios. Whilst these have since been updated, it does not follow that they were inconsistent. Such a statement needs backing up with a citation where this has been conclusively demonstrated.

We agree that the millions at risk approach implemented by Parry et al. 2001, based on the SRES scenarios, is not a good example of a cross-sectoral synthesis that relied on partly inconsistent scenario input. In the revised manuscript, we cite two other review-type studies that summarized impacts as a function of global mean temperature rise (ll.215-219). These studies relied on a huge literature basis with the trade-off that some inconsistencies in the underlying data could not be avoided.

P729 line 1. The discussion of the hotspots work should emphasize that this paper does not definitively identify the areas which are the most affected by climate change in the world, because it does not include all impacts sectors, and also it is very difficult to decide how to 'weight' different levels of impacts in different sectors. Rather, these hotspots perhaps show where interactions between climate change impacts upon different sectors will be most likely to manifest themselves.

To our mind, it is impossible to "definitively identify the areas which are the most affected by climate change in the world". Any hotspot map will be dependent on specific assumptions, e.g., about what is considered severe change in the different sectors considered. However, we agree that the Piontek et al. study is only a first step towards a more general hotspot map, which would need to rely on data from more impact sectors and optimally account for adaptation potentials. We mention this point in the revised version of the manuscript (ll. 225-229).

P729 line 28 to P730 line 12. Consider moving this paragraph which seems out of place here.

This paragraph discusses the integration of global, regional, and local models and therefore belongs to the section entitled "Integrating impacts projections across sectors *and scales"*. In the revised version we have added "spatial" in the section title, and we have changed the beginning of the paragraph (l. 264). It should now be obvious that we turn from discussing sectoral integration to discussing the integration across different spatial scales.

P731 line 11-21 Mention the debate over whether it is appropriate to weight GCMs, including whether their ability to represent current climate is related to their ability to represent future change.

We have added a sentence in the revised manuscript (ll. 305-307) to express this specific caveat with regard to the weighting of GCM output.