Paper:
Challinor, AJ, Martre, P, Asseng, S, Thornton, P and Ewert, F (2014) *Making the most of climate impacts ensembles.* Nature Climate Change, 4. 77 - 80.

http://dx.doi.org/10.1038/nclimate2117

Making the most of climate impacts ensembles

Andy Challinor, Pierre Martre, Senthold Asseng, Philip Thornton, Frank Ewert

Increasing use of regionally and globally oriented impacts studies, coordinated across international modelling groups, promises to bring about a new era in climate impacts research. Coordinated cycles of model improvement and projection are needed in order to make the most of this potential.

Climate impacts ensembles, usually comprising multiple impact models, are a promising tool for projecting future crop productivity [e.g. 1] and increasing coordination between international modelling groups, evident in Model Intercomparison Programs (MIPs), is producing high-profile multi-model studies [e.g. 2]. An increasing number of these studies are global in extent, while model accuracy and data quality are often better at local to regional scales. Here, we explore the implications of this trend for the design and coordination of future studies. We develop recommendations based on the assertion that a single model intercomparison study, if it is to avoid being unwieldy, can focus on either projecting impacts, or on model evaluation and model intercomparison, but not both. Further, we assess the suitability of global *vs.* regional studies for achieving each of these aims. Whilst our analysis is presented for agriculture, it applies to a range of climate impacts. We define global studies as those with full global coverage, and regional studies as those with limited geographic extent, such as a country or province. We also include in the latter the modelling of specific fields (i.e. local studies), since the ultimate aim of local-scale studies is often to draw conclusions for the region. The models used for the different studies may be the same, although in practice they often differ in complexity.

The value of multi-model impacts assessments in quantifying uncertainty is increasingly well documented [3]. However, we cannot simply take our cue from the larger body of work on climate ensembles, since impacts ensembles are different: they involve calibration towards a small subset of variables which may depend on the output variable of interest (e.g. crop yield), as opposed to seeking to reproduce a broad set of properties of a closed system. For example, crop models are usually used primarily to simulate yield, which is only one of the many aspects of crop growth and development. Climate models, in contrast, are assessed on their representation of rainfall, temperature, wind (jet streams, monsoon circulations), ocean properties and a host of other physical properties Assessment of multiple properties of impact models is less advanced. This is not least due to differences in model structure constraining the identification of comparable properties, and difficulties in obtaining adequate data, particularly at regional scales. Whilst these problems are not insurmountable, the relative lack of progress means that, crop models are prone to often unknown compensation of errors, making on-going assessment of causal relationships in our impacts models particularly important. The same issue arises in other impacts models, for example modelling hydrology [4] and tree distribution [5] .Unpicking this compensation of errors is intractable in practice, since it involves separating calibration from tuning, and we do not often have the data to do this adequately. Thus a model can never be truly ‘validated’ for future use, only continually evaluated in the light of the most recent data. Thankfully, model evaluation is becoming increasingly coordinated amongst model groups, and increasingly sophisticated. Progress has been facilitated by greater international coordination, e.g. through the Agricultural Model Intercomparison and Improvement Project (AgMIP) [6].

The issue of compensation of errors is illustrated by a recent inter-comparison of 27 wheat simulation models, where parameter calibration led to a greater improvement in yield error than for any other variables, including leaf area index, harvest index and cumulative evapotranspiration [2]. Figure 1 presents further evaluation of the calibration procedure conducted for that study. There is no clear relationship between the total number of genotypic parameters – which can be taken as a proxy of model complexity – and the relative error of either harvest index or grain yield (panels 1a and 1c). This result suggests that the models have more degrees of freedom than can be constrained by experimental data. Subsequent calibration of the models using experimental data led to significant reduction of model error, although this improvement (y axis in panels 1b and 1d) was generally greater for yield than for harvest index, suggesting some compensation of errors. However, there was no relationship between the number of calibrated parameters and the reduction of model error (Fig. 1b); i.e. no evidence of model over-tuning. Detailed comparisons of a range of model variables are needed if we are to determine the nature of the compensation of errors – i.e. the extent to which the models are getting the right answer for what is, in part at least, the wrong reason. Multi-variable impacts studies could facilitate assessments of crop sustainability (through e.g. nitrogen and water use) and crop quality (through e.g. grain protein or mycotoxin concentration).

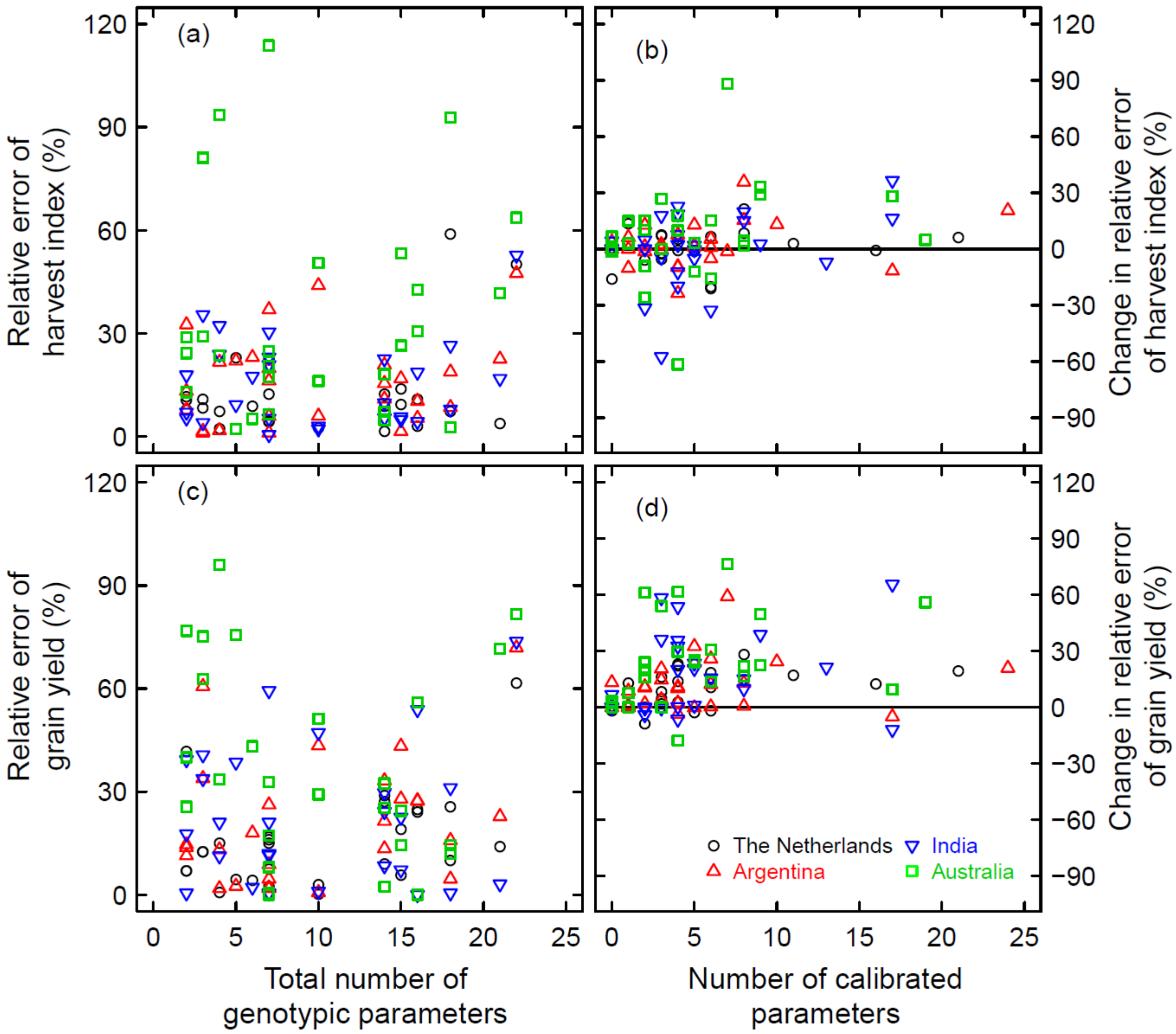

Figure 1. Relationship between models relative errors and the number of genotypic and calibrated parameters for 27 wheat crop simulation models in The Netherlands, Argentina, India and Australia. Relative error of (a) harvest index and (c) grain yield, versus the total number of genotypic parameters. Also shown are the changes in relative error, due to model calibration, of (b) harvest index and (d) grain yield, versus the number of calibrated parameters. The experiments and the simulation protocols were developed by AgMIP and are described in [2].

The above analysis of compensation error and model tuning has an important implication: it illustrates the value of exploring differences – and in particular the *reason* for differences – between models. Thus, as has been argued for climate ensembles, a focus on narrowing uncertainty – i.e. seeking consensus – is too limited [7, 8]. The differences matter. Good quality observational data sets, for both agriculture and climate, are critical in determining and unpicking these differences [9]. Integrated analyses of model chains (e.g. from climate to bio-physical impact to economic consequences) are also important if model differences are to be understood. Impacts modellers should not be insular, but rather should recognise the benefit of engaging with the climate modelling community and with those who model other sectors (e.g. the Inter-Sectoral Impact Model Intercomparison Project, ISI-MIP).

A second implication of model error and potential over-tuning is the need to be clear about how much can be achieved in any one study. In particular, how much of the globe can an ensemble study reliably assess? Global studies have the advantage of employing a consistent set of assumptions and therefore producing projections that are consistent and directly comparable, thus facilitating benchmarking across a range of environments. However, the datasets used to drive global models need to be both global in coverage and consistent, thus limiting the pool of available data. Consequently there will certainly be areas over which more comprehensive and reliable data are available. Furthermore, specific impacts models developed and/or evaluated using such data typically perform better than a global assessment does over the same region. This can be due to calibration and/or the incorporation of key processes (e.g. heat stress during anthesis) and interactions (e.g. between canopy temperature and transpiration) relevant to that region but not yet incorporated into global assessments.

There is a question, then, of whether projections of climate change impacts are better made by ensembles that are global or regional in scope. The former can produce a single consistent and global evaluation – a significant advantage given the importance of quantifying uncertainty. However, we do not have data at the global scale to determine whether or not models are getting the right answer for the right reason (c.f. Fig. 1). Thus there is a danger that global simulations will appear robust, because of their consistency, whilst in fact lacking valuable regional specificity.. The end result in this case is that both future research agendas and policy may be misinformed. Whilst there is currently no evidence of such misinformation, in-depth analyses have rarely been performed.

Clearly global and regional studies each have advantages and disadvantages. This invites two questions: How can MIPs be designed in a way that makes the most of both types of study? And how can the need for projections – generating and synthesising results based on impact scenarios – be balanced with the need for model intercomparison and improvement?

Ensuring that detailed process research feeds through into improved projections requires strategic planning of MIPs that takes into account the different drivers of the component research. At the global level, policy makers are interested in the use of state-of-the-art models and methods to produce probabilistic projections of climate impacts. The ongoing increase in the use of land surface models for crop simulation [e.g. ref. 10] is underpinning progress in this area. At the regional level, it is critical to ensure that improved projections enable adaptation and deliver improved livelihoods. This focus has led to an increase in outcome-orientated research for development, for example within the CGIAR.

The different drivers of global and regional research lead to different – if somewhat overlapping – communities. Modelling and international coordination strategies need to be carefully thought through if we are to make the most of the full range of models and researchers, rather than deepen the existing divide. Figure 2 draws on the reasoning above to conceptualise effective coordination of MIPs. This conceptualisation recognises that different MIPs might – indeed probably *should* – have different aims. Effective coordination involves the exploration of synergies between different MIPs and associated international programmes. For example, work that is focussed on livelihoods could make use of the broader global research base. Such multi-level analyses, at least at the regional-to-local scale, have been shown to support development outcomes that are consistent across different types of production systems [e.g. 11].

The conceptualisation in Figure 2 also emphasises model improvement, underpinned by inter-comparison, as an important aim for MIPs. Thus studies should contain explicit statements on the assumptions made in the modelling, and they should report discrepancies in addition to agreement. Detailed modelling studies and experimental data are needed in order to understand response mechanisms and ensure they are included in models [e.g. 12]. Studies should also assess multiple variables, for example, nitrogen, water use and crop quality and yield. This not only provides stronger constraints on models, it also facilitates assessments of crop sustainability and crop quality.

At the heart of Figure 2 is the concept of coordinated cycles of model improvement and multi-model projection. Quantifying the effect of model improvement on predictive skill will help to focus research efforts. For example, globally-oriented studies can incorporate key processes that have been identified by regional studies. Coordination efforts need to be underpinned by international data and modelling strategies, such as those being developed by Future Earth [13]. The projection phase of a MIP should be inclusive, by using systematic inter-comparison of impacts studies in order to synthesise knowledge and utilise new insights. For example, multi-model impacts studies [14] or meta-analyses [15] can be used to develop response functions, and associated uncertainties. These response functions quantify change in crop yield as a function of the changes in local temperature resulting from climate change.

Finally, a critical component of coordinated MIPs is a full treatment of uncertainty, which goes beyond impacts models. Integrated treatments of climate and impacts model chains not only provide more accurate assessments of uncertainty; they can also lead to improved ways of presenting uncertain information [16,17]. Such methods contrast with the scenario-driven approaches that are still prevalent in impact MIPs. The recommendations we have outlined above are summarised in Box 1.

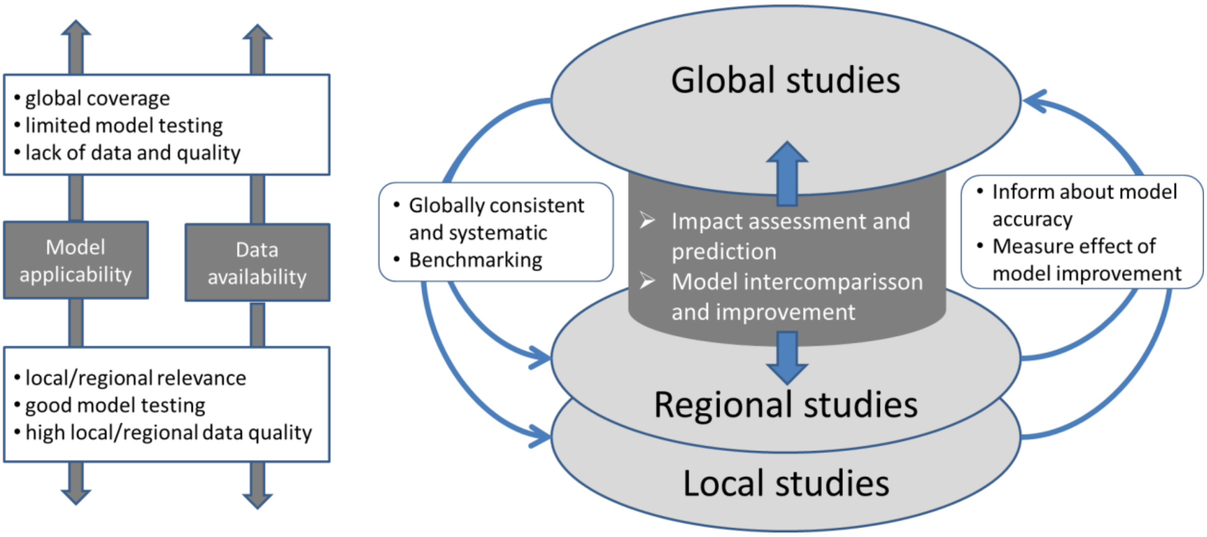

Figure 2: Proposed coordinated cycle of model improvement and projection, based on recognition of the strengths and weaknesses of global, regional and local studies. Effective use of studies with different geographical domains is contingent on coordination within and across Model Intercomparison Programmes.

Box 1: recommendations.

- Recognition of the separate but linked strategies of different Model Intercomparrison and Improvement Programmes.
- Increased assessment of multiple variables within single impacts studies, e.g. nitrogen, water use, crop quality and yield.
- Design and implementation of coordinated cycles of model improvement and multi-model projection (Fig. 2).
- Use of systematic intercomparison of impacts studies to synthesise knowledge
- Full treatments of uncertainty, which go beyond impacts models and include relationships between climate and its impacts

References

1. Challinor, A. J., M. Stafford Smith, P. K. Thornton (2013). Agro-climate ensembles: emerging tools for quantifying uncertainty and informing adaptation. Agricultural and Forest Meteorology 170, Pages 2-7.
2. Asseng, S., et al. (2013). Uncertainty in simulating wheat yields under climate change. Nature Climate Change 3: 827-832
3. Rötter, R. P., T. R. Carter, J. E. Olesen and J. R. Porter (2011). Crop–climate models need an overhaul. Nature Climate Change, 1, 175-177, doi:10.1038/nclimate1152
4. Beven, K. (2006). A manifesto for the equifinallity thesis. Journal of Hydrology 320, 18-36.
5. Cheaib, A., Badeau, V., Boe, J., Chuine, I., Delire, C., Dufrêne, E.,François, C., Gritti, E.S., Legay, M., Pagé, C., Thuiller, W., Viovy, N. & Leadley, P. (2012) Climate

change impacts on tree ranges: model intercomparison facilitates understanding and quantification of uncertainty.Ecology Letters 15 533–544
6. Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburne, P., Antle, J.M., Nelson, G.C., Porter, C., Janssen, S., Asseng, S., Basso, B., Ewert, F., Wallach, D., Baigorria, G., Winter, J.M., 2013. The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. Agricultural and Forest Meteorology 170, 166-182.
7. Knutti, R. (2010) The end of model democracy? An editorial comment. Clim Change 102(3-4):395-404.
8. Knutti and Sadlace, (2013) Robustness and uncertainties in the new CMIP5 climate model projections. Nature Climate Change 3, 369–373 doi:10.1038/nclimate1716
9. http://ccafs.cgiar.org/node/304
10. Osborne, T. M., D. M. Lawrence, A. J. Challinor, J. M. Slingo, T. R. Wheeler (2007). Development and assessment of a coupled crop-climate model. Global Change Biology 13 169-183
11. Rufino MC, Thornton PK, Ng'ang'a SK, Mutie I, Jones PG, van Wijk MT, Herrero M (2013). Transitions in agro-pastoralist systems of East Africa: impacts on food security and poverty. Agriculture, Ecosystems and Environment (in press).
12. Lobell, D.B.. G.L. Hammer, G. McLean, C. Messina, M.J. Roberts, and W. Schlenker. 2013. The critical role of extreme heat for maize production in the United States, Nature Climate Change, 3, 497–501
13. http://www.icsu.org/future-earth
14. Rosenzweig, C., J. Elliott, D. Deryng, A.C. Ruane, A. Arneth, K.J. Boote, C. Folberth, M. Glotter, N. Khabarov, C. Müller, K. Neumann, F. Piontek, T. Pugh, E. Schmid, E. Stehfest, H. Yang, and J.W. Jones, (2013). Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. Proceedings of the National Academy of Sciences.(in press).
15. Challinor et al., NCC in review. NCLIM-13020143A
16. Vermeulen, S. J., A. J. Challinor, P. K. Thornton, B. M. Campbell, N. Eriyagama, J. M. Vervoort, J. Kinyangi, A. Jarvis, P. Läderach, J. Ramirez-Villegas, K. J. Nicklin, E. Hawkins and D. R. Smith (2013). Addressing uncertainty in adaptation planning for agriculture PNAS 2013 ; published ahead of print May 14, 2013, doi:10.1073/pnas.1219441110
17. Joshi, M., R. Sutton, J. Lowe, and D. Frame, 2011: Projections of when temperature change will exceed 2 °C above pre-industrial levels. Nature Climate Change, 1(8), 407-412.