Dr Aideen Foley
Department of Geography, Environment and Development Studies
Birkbeck College
Malet Street
London WC1E 7HX

December 10, 2015

MS# esd-2015-35
Revised for *Earth System Dynamics*

"Climate model emulation in an integrated assessment framework: A case study for mitigation policies in the electricity sector"

Responses to reviewers – December 10, 2015

To whom it may concern:

We would like to thank the reviewers and the editor for their thoughtful comments of our manuscript. We have adopted the reviewers' suggestions, including contextualising our results relative to the policies explored in the IPCC AR5.

We think that the manuscript has been greatly improved by these revisions and we hope that you will now find it suitable for publication.

Our point-by-point responses to comments are detailed on the following pages.

Yours sincerely,

Aideen Foley

Dr Aideen Foley

Response to Reviewer #1

1. **In section 3.1 where the PLASIM_ENTSem is discussed, it would be useful to know how the relationship between CO2 concentrations and global temperature rise compares with the CMIP5 simulations reported in IPCC AR5 WGI The Physical Science Basis (IPCC 2013). The earlier part of the paper demonstrates that the relationship between emissions and concentrations is very similar to that of the relationships found in VanVuuren et al. 2011. Hence it would useful also to have this analogous information about PLASIM_ENTSem.**

   The following remarks on work carried out in Holden et al (2014) have been added:
   'The response of PLASIM_ENTSem to RCP forcing was analysed in Holden et al (2014, Figure 6); in all four scenarios, the emulated ensemble distribution was found to compare favourably with the multi-model CMIP5 ensemble.'

2. **In section 3.2, at the end on page 1291 the authors need to provide information about the baseline scenario in the main text – in particular whether its emissions are similar to those which others have published as consistent with RCP8.5, or are they higher or lower? Do they grow more/less rapidly at different times during the 21st century? When it comes to the results, text making comparison of the trends in CO2 concentrations and temperatures with the RCPs would be useful – so that one can relate the outcomes of the scenarios being explored to these.**

   The following description of the baseline scenario has been added to section 3.2:
   'The baseline scenario extends current policies in the energy sector to 2050. It assumes no additional technology subsidies worldwide, feed-in tariffs in some EU countries, and carbon pricing in the EU. Figure 3 illustrates that the emissions associated with this scenario are of a similar magnitude as emissions associated with RCP 8.5, but following a more linear trajectory.'

3. **In the conclusion, the authors need to put their results in context of existing work on mitigation policy by:**

   - detailing the latest IPCC AR5 figures estimating the contribution that the electricity sector makes to the total CO2 emissions; and also the total GHG emissions…This is a key factor in assessing the significance or otherwise of the results in terms of global mitigation policy considerations and in determining the extent to which these results might suggest more pessimistic outcomes for mitigation (in terms of reducing warming) than IPCC AR5 WGIII (IPCC 2014, Mitigation of Climate Change).

   - discussing how their results compare with the IAM model ensemble database of IPCC AR5 WGIII. These are mainly outputs of IAMs which use simple climate model emulators, MAGICC6 or others. Obviously the IAM database mitigation scenarios represent mitigation in many sectors not just the electricity sector, but comparison could still be made in terms of the level of GtC removed, to see if the relationship between emissions, concentration and temperature in the WGIII database differs from that in this paper. It would be useful to specify precisely which feedback mechanism is responsible for the difference.

   - making a comparison with the level of decarbonisation in the electricity sector in the IPCC AR5 WGIII database: are you simulating similar levels of decarbonisation by similar to dates to many of these scenarios, or do your scenarios examine greater rates of decarbonisation than are explored in this database?

   - if a rigorous comparison with the database is extremely time consuming and thus beyond the scope of the paper, statements along these lines could instead be made by making approximations based on reading from Figures in the IPCC AR5 assessment report, and by expert judgement.

   - please make clearer if the policy relevance of the paper lies rather in pointing out the inadequacy of policies that focus on the electricity sector alone in reaching the 2C target; rather than in suggesting that mitigation policy will be less effective than as stated in IPCC 2014 because of the inclusion of non-linear dynamics that the IAMs underpinning the database don't include. It would be very helpful to understand which of these points you are trying to make - either or both.

   - It would be useful to add some background about whether there is some possibility or not that real world policies might be in danger of focusing on the electricity sector whilst leaving the other sectors to their own devices. You could discuss whether mitigation in this sector cheaper than in the other sectors for example.

The following text addresses these comments, by contrasting the types of policies explored in this paper with those of the AR5, and explicitly stating the added value of the non-linear dynamics included in this framework; the failure of electricity sector only policies in this paper is in spite of non-linear feedbacks, which are expected to promote decarbonisation:

'Even the most successful mitigation strategy considered here results in warming of above 3.5°C by 2100, a level of warming which Parry (2009) notes could result in substantial harmful impacts, including risks of water shortage and coastal flooding. As such, in a context where the global electricity sector is decarbonised by 90%, further emissions reductions must be achieved in other sectors (e.g. transport and industry) to enable CO2 concentrations to remain below 450~ppm, and correspondingly, global warming below 2°C (Meinshausen et al., 2009).

The latest IPCC AR5 notes that in 2010, the energy supply sector accounted for 35% of total GHG emissions, therefore there is scope for reductions to be achieved in other sectors. For instance, policy options explored by Luderer et al (2012) which keep CO2 concentrations below 450~ppm, using the IMACLIM-R and ReMIND-R models , include mitigation in the transportation sector to reduce energy demand. However, the IPCC AR5 notes that based on scenario analysis, sectors currently using liquid fuel may be more costly, and therefore slower, to decarbonize than electricity. Additionally, it is worth noting that the most successful mitigation scenarios explored in the IPCC AR5, which lead to CO2 eq concentrations in the range of 430-480 ppm by 2100 (approximately equivalent to RCP 2.6) feature large-scale, long-term application of carbon dioxide removal (CDR) technologies, in addition to large emissions reductions (IPCC, 2014).

This analysis, focusing on the effectiveness of mitigation policies in the electricity sector, therefore highlights the danger of focusing mitigation efforts on this single sector, where the cost of decarbonisation is lower; not only are such efforts insufficient to maintain global warming below 2°C, but additionally, the heterogeneous distribution of climate impacts globally will need to be addressed.

Furthermore, the inadequacy of electricity sector to solve the emissions problem is in spite of the fact that the inclusion of non-linear feedbacks on technology uptake is expected to promote decarbonisation in our model, compared to the equilibrium models in the IPCC AR5 database, which may not capture the complexities of real-world human behaviour in mitigation decision-making (Mercure et al. 2015).'

4. **Detailed comments Page 1285 lines 5-7 Note that the RCPs are not emission scenarios but concentration pathways – suggest that you edit the the phrase 'RCP emission scenarios' to read 'emissions consistent with the RCP pathways that are reported in VanVuuren et al 2011' or similar.**

   Text has been changed to read: 'The coefficient ranges were chosen to span emissions consistent with the RCP pathways'

5. **Page 1285 lines 6-7 and 14-16 Justify the choice of values for E1, E2 and E3 – how do I go from Moss et all or Van Vuuren et al. to derive these? Similarly for R1, R2, R3. Page 1287 lines 1-9. Link this to the IPCC AR5 treatment of uncertainty in the terrestrial carbon sink, as Holden et al 2013a presumably is not based on IPCC AR5?**

   The motivation for these ranges has been expanded on, as below:
   'The E1 and R1 coefficients define the 2100 CO2 emissions and non-CO2 radiative forcing respectively. The ranges for these coefficients have been chosen to encompass (and exceed) the ranges of 2100 forcing in Moss et al (2010). The range of input values for the training dataset needs to be wide in order to avoid extrapolation when using the resulting emulator. The maximum E1 = 30 gives 2100 CO2 emissions of E0+E1=39.166GTC, which compares to RCP8.5 emissions of 28.817GTC. Maximum radiative forcing of R0+R1=10.619Wm-2 was allowed to greatly exceed RCP estimates (maximum 1.796Wm-2) in order to allow the potential application of the emulator to extreme non-CO2 forcing scenarios.'

   In section 2.4, we direct the reader to the evaluation of uncertainty in section 2.6.
   'We evaluate the resulting emulated uncertainty through a comparison with CMIP5 in Section 2.6.'

6. **Page 1288 lines 5-10. Should there be illustrations or tables to support statements about how the simulated ensemble mean and the emulated ensemble mean compare in section 2.5 in the SM?**

As the manuscript already contains several images, we determined that the inclusion of an illustration of the comparison of simulated mean and ensemble mean did not add value to the paper. We have additionally cited the RSQ value for an additional comparison of emulated and simulated values.

7. **Page 1291 line 4. Please detail the baseline assumptions here and how does the scenario compare with other analysts' emissions for RCP8.5?**

   See response to comment 2.

8. **Page 1292 line 11. This paragraph refers to 'the' mitigation scenario – aren't there several? In which is there 90% decarbonisation?**

   This line has been amended to 'mitigation scenarios', as there are indeed multiple scenarios.

9. **Page 1295. Line 5 suggest insert 86' before 'ensemble members' to clarify**

   This clarification has been made in the revised text.

10. **Page 1295 line 23. Suggest reword. The statistical performance of the pattern scaling seems to be generally quite good really, so I would rephrase this to say that assumptions of pattern scaling may perform less well, rather than saying 'especially likely to break down', or say what % error you think there might be that your method can improve upon.**

    The language has been changed to:
    '…the assumptions of pattern scaling may not be optimal when applied to strong mitigation scenarios.'

11. **Page 1296. See my comments about what is missing from the conclusion. If the journal allows you may want to insert a separate discussion prior to the conclusion where this comparison is made.**

    Discussion in the conclusions section has been extended to address these points.

<u>**Response to Reviewer #2**</u>

1. **1278:17. "in response to" – perhaps "associated with" is better.**
   M1278:16-20. I do not feel the paper ultimately establishes this for the case that is studied…. So this sentence needs to be re-cast so as to reflect the reality of the results.

   The sentence has been rephrased as follows:
   'Our approach also highlights the regional temperature and precipitation patterns associated with the global mean temperature change occurring in these scenarios…'

2. **1279:10. GCM not defined (nor is AOGCM later – a systematic checking of acronyms would be useful)**
   1282:2. ESM?

   These acronyms have now been defined.

3. **M1282:5 non-CO2 radiative forcing is very ambiguous especially as we are told very late in the paper (1295:2) that the model lacks aerosol forcing (but then 1284:22 says aerosols are included). There is a need for greater clarity.**

   The description has been clarified as follows:
   'The emulator takes a time series of anthropogenic carbon emissions and non-CO2 radiative forcing (stemming from CH4 N2O, halocarbons, and other forcing agents including O3 and aerosols) as inputs and provides a time series of atmospheric CO2 concentration as output.'

4. **M1282:11. For clarity, is the full GENIE-1 simulator here the same as the GENIE-1 ESM referred to on line 1 of this page. A consistent terminology would help the reader.**

   References to the GENIE-1 simulator have been replaced with GENIE-1 ESM.

5. **M1284:21-22. First if all climate forcings (even aerosols) are represented as perturbers of long-wave radiation, then important characteristics of the effect of, for example, aerosols (especially for precipitation changes) are lost (see 1295:2), and any spatial influence of the forcing on the response is lost, if a globally-uniform modification is applied (again see 1295:2), as this is not even the case for CO2 alone. This would be a potentially significant limitation to the model, especially for the application here, and this needs to be spelt out as a caveat more clearly.**

   We have added the following clarification under the section 'The GENIEem carbon cycle model emulator':
   'In the integrated assessment framework developed here, the time series of anthropogenic carbon emissions is provided by E3MG-FTT, while non-CO2 forcing data is derived from global timeseries of forcing data obtained through the RCP Database. As such, GPem emulates high-dimensional climate outputs as a function of scalar model inputs. We note that certain forcings, such as aerosol forcing, are characterised by complex spatial patterns and so would benefit from an approach in which the inputs are also high-dimensional. However, incorporating such forcing into the emulator framework would involve coupling an aerosol model to PLASIM-ENTS in order to build an ensemble of simulations and a subsequent emulator, which is beyond the current scope of this work.'

6. **Second, methodologically I do not understand the apparent permanence of the modification of the outgoing longwave radiation. Radiative forcing changes the top of the atmosphere radiative budget in only a transient perturbation – the climate system responds (via warming) to eradicate the perturbation in radiative budget (and so globally it returns to zero in an equilibrium situation). If the forcing is applied as a permanent modification of the LW budget, then how does the longwave budget re-adjust following a warming? I could understand this more if the emissivities (which appear to be used in the Fanning and Weaver model) were instead modified.**

   Adding a constant number to either side of any equilibrium relationship will result in a different equilibrium state.
   The radiative balance expressed most simply, in 0D, is: S(1-alpha)/4=cT^4
   where T is temperature, S solar radiation and alpha planetary albedo.
   Applying a perturbation P as we do gives: S(1-alpha)/4=cT^4 + P
   This finds equilibrium at a different temperature.
   To clarify, we have amended the phrase 'globally uniform modification to' to ' globally uniform additional term in'.

7. **M1285:1:16. If I apply Equation (3) with the stated parameters I generate some very strange time profiles of forcing… Is Equation (3) wrong (I note in Holden and Edwards that the 0.5 embraces R1, R2 and R3 rather than just R1 here)? Not being a Chebyshev expert, I was also confused by the R3 parameter; lists I see in text books etc have 4x**3 - 3x, but perhaps this is what is meant by "modified" here?**

   There was an error in transcribing this equation. As the reviewer suggests, the 0.5 should embrace R1, R2 and R3. The modified Chebyshev parameters are arrived at through linear decomposition of the first three Chebyshev polynomials. This has been made explicit.

8. **1285:17 and 1285:21. I didnt understand what "were reproduced three times" and what "successfully" means. Could you clarify?**

   This step has been clarified as follows:
   'The 86 parameter sets were replicated three times, and each of these three 86 parameter sets was combined with different future emissions profiles to produce a 258-member ensemble.'

   "Successfully" is this context is redundant, and the sentence has been amended to:
   '257 simulations completed; in the remaining simulation, input parameters led to an unphysical state and ultimately, numerical instability.'

9. **1286: 11. Is this time-series of concentrations or emissions?**

This is a reference to the output timeseries of CO_2 concentrations, and the text has been amended to reflect this: 'Each individual simulated CO_2 concentration time series can thus be well approximated as a linear combination of the first four components, scaled by their respective scores.'

10. **1291:2 and 1291:11. I wasn't clear whether Figure 3 was emissions just from the power sector, or the different electricity scenarios on the total CO_2 emissions. I guess the latter, as I could not see a 90% reduction on Figure 3.**

    Figure 3 refers to total CO_2 emissions, so that the scenarios explored in the paper (which refer to the electricity sector) can be compared with RCP scenarios, which cover all sectors and land use. The figure caption has been updated to reflect this.

11. **1291:25-28:. These sentences seem contradictory – one says the appropriate non-CO_2 RCP is chosen, but it then says that RCP8.5 non-CO_2 is used for all scenarios. Which is it?**

    RCP8.5 non-CO_2 forcing is used for all scenarios as the mitigation scenarios explored here lack a suitable analog in the RCP database. This is described in the text as follows:

    'As FTT:Power-E3MG does not simulate non-CO_2 radiative forcing, we select the RCP that best matches the CO_2 concentrations associated with the baseline scenario(RCP 8.5) and force GENIEem with the non-CO_2radiative forcing associated with that RCP. The RCP 8.5 non-CO_2 radiative forcing was applied to all scenarios as the RCPs lack a suitable analog to the CO_2 concentrations associated with the power sector mitigation scenarios examined in this work."

12. **M1292:5-15. The implication here is that non-CO_2 here means just methane and nitrous oxide? Is that correct? If so, what is the implication of just considering these non-CO_2 gases rather than the wider mix including the short-lived pollutants?**

    Please refer to response to comment 3. CH_4 and N_2O are specifically mentioned here as they are emissions we expect might be impacted by the mitigation pathways explored, but they are not the only non-CO_2 forcings considered. The text added in response to comment 3 clarifies this.

13. **1292:7. In principle, the correct application of equation (7) to obtain equivalent CO_2 is to sum the forcings before calculating the equivalent CO_2. There is a hint in the next sentence that the CO_2 seen by the model is the sum of equivalent CO_2's calculated individually for actual CO_2 and non-CO_2 forcings. Perhaps the difference is negligible, but it would be worth clarifying.**

    Forcings are summed before calculating equivalent CO_2 (see text below):
    '…GENIEem ensemble CO_2 concentrations are converted to radiative forcing following:
    $F=5.35 \ln(CO2=280)$ W m^{-2}
    RCP 8.5 non-CO_2 forcing is added to this time series to give total radiative forcing, which is converted to equivalent CO_2 using the previous relationship. '

14. **1294:5. "due to the effect of non-CO_2 forcing" – does this mean via the carbon-cycle feedbacks in the model? I was unsure.**

    This sentence refers to the fact that although the mitigation policies explored lead to reductions in CO_2, the combination of remaining CO_2 forcing and non-CO_2 forcing still have a warming effect. This has been clarified as follows:
    'While the mitigation policies explored generate reductions in CO_2 emissions from the energy sector, due to the effect of non-CO_2 radiative forcing on climate, combined with remaining CO_2 emissions, CO_2 concentrations continue to increase in mitigation scenarios.'

15. **1295:12-14. Indeed, but this is not what is implied in the abstract, which is altogether more tantalising.**

    See response to comment 1. The abstract has been amended in response to the reviewer's comment.

16. **1295:8. … I do not think a strengthening of the Hadley Circulation is needed to generate this pattern. They emerge from the differences in water vapour amount in the atmosphere that follows (assuming fixed relative humidity) from the warming – in the absence of a circulation change, you still amplify the precipitation fields as more water is available in the convergence zones to condense. See e.g. http://dx.doi.org/10.1175/JCLI3990.1**

    The text has been amended to incorporate the reviewer's suggestion:
    'Generally, areas that experience a significant increase/decrease in precipitation under scenario iv (i.e. larger than 1 mm day $^{-1}$ experience even greater extremes under scenario I, which can be attributed to differences in water vapour amount in the atmosphere due to warming (Held and Soden, 2006); precipitation fields are amplified as more water is available in the convergence zones to condense.'

17. **1295:21. "demonstrates" – I think "suggests" is safer. I suspect that in the CMIP5 simulations it is the short-lived forcings that are important in modulating the precipitation pattern in the scenarios which are not CO2 dominated, but the model here cannot represent this.**

    The language has been changed accordingly with the reviewer's suggestion.

18. **1295:25. I am not familiar with the literature on the climate effects of mitigation in the electricity sector, but I would be surprised if there were not several studies using simpler model frameworks already. I might have expected some discussion in the conclusions about what has been learnt here which goes beyond these studies. If no such studies exist, it may be worth stressing this, as it would render this paper more original.**

    Some information on this can be gained from the IPCC WG3 database. See response to reviewer #1 comment 3. As most integrated assessment studies involve policies across multiple sectors, there are insights to be gained from focusing on individual sectors. We note in the text added that the electricity sector has been judged to be a less difficult sector to decarbonise, compared to those that use liquid fuel. By illustrating the inadequacies of policies that focus only on this sector, particularly in a framework that features non-linear dynamics that should favour decarbonisation, we highlight the risk of focusing mitigation efforts on sectors that are easier to decarbonise.

Manuscript prepared for Earth Syst. Dynam.
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 10 December 2015

Climate model emulation in an integrated assessment framework: A case study for mitigation policies in the electricity sector

A.M. Foley[1*], P.B. Holden[2], N.R. Edwards[2], J.-F. Mercure[1], P. Salas[1], H. Pollitt[3], and U. Chewpreecha[3]

[1]Cambridge Centre for Climate Change Mitigation Research, Department of Land Economy, University of Cambridge, 19 Silver Street, Cambridge, CB3 9EP, UK
[2]Environment, Earth and Ecosystems, Open University, Milton Keynes, MK7 6AA, UK
[3]Cambridge Econometrics Ltd, Covent Garden, Cambridge, CB1 2HT, UK
[*]now at Department of Geography, Environment and Development Studies, Birkbeck, University of London, 32 Tavistock Square, London, WC1H 9EZ, UK

Correspondence to: A.M. Foley (a.foley@bbk.ac.uk)

Abstract.

We present a carbon cycle-climate modelling framework using model emulation, designed for integrated assessment modelling, which introduces a new emulator of the carbon cycle (GENIEem). We demonstrate that GENIEem successfully reproduces the CO_2 concentrations of the Representative Concentration Pathways when forced with the corresponding CO_2 emissions and non-CO_2 forcing. To demonstrate its application as part of the integrated assessment framework, we use GENIEem along with an emulator of the climate (PLASIM-ENTSem) to evaluate global CO_2 concentration levels and spatial temperature and precipitation response patterns resulting from CO_2 emission scenarios. These scenarios are modelled using a macroeconometric model (E3MG) coupled to a model of technology substitution dynamics (FTT), and represent different emissions reduction policies applied solely in the electricity sector, without mitigation in the rest of the economy. The effect of cascading uncertainty is apparent, but despite uncertainties, it is clear that in all scenarios, global mean temperatures in excess of 2°C above pre-industrial levels are projected by the end of the century. Our approach also ~~reveals the diverse~~ highlights the regional temperature and precipitation patterns ~~that could occur regionally in response to~~ associated with the global mean ~~temperatures associated with~~ temperature change occurring in these scenarios, enabling more robust impacts modelling and emphasising the necessity of focussing on spatial patterns in addition to global mean temperature change.

1 Introduction

Integrated assessment modelling can be used to explore the climatic consequences of particular climate mitigation policy scenarios. However, most integrated assessment models (IAMs) do not directly utilise sophisticated coupled Atmosphere Ocean General Circulation Models, such as those employed in the Coupled Model Intercomparison Project Phase 5 (CMIP5: Friedlingstein et al., 2013), to represent the climate and carbon cycle. Due to the large computational resources they require, the direct use of such models within IAMs is not feasible.

Instead, many IAMs have used simple mechanistic models to represent the carbon cycle. One such simplified carbon-cycle/climate model is MAGICC6 (Meinshausen et al., 2011a), which is calibrated against higher complexity models from the Coupled Carbon Cycle Climate Model Intercomparison Project (C4MIP), to emulate the atmospheric CO_2 concentrations of those models. Schaeffer et al. (2013) used MAGICC6 to derive probability distributions for radiative forcing, which drive a simple climate model that projects ~~GCM~~ global mean temperature response by linearly scaling the CO_2 step experiment response of 17 CMIP5 ~~GCM~~ General Circulation Model (GCM) $4\times CO_2$ simulations. Such approaches can be used to generate large ensembles quite quickly; for instance, MAGICC6 has been used to generate a 600-member perturbed parameter ensemble (Schaeffer et al., 2013) of CO_2-equivalent concentration and global-mean surface-air temperature change projections.

It has been suggested that a conceptual advantage of this approach is that the mechanistic model fit adds some confidence when extrapolating beyond the training data (Meinshausen et al., 2011a). A limitation of simplified mechanistic models is that they may contain a high level of parameterization. For example, the Meinshausen et al. (2011a) carbon cycle calibration procedure uses global mean temperature as a proxy for changes in patterns of temperature and precipitation. These drivers of change in the carbon cycle would be explicitly represented in a more sophisticated model.

To represent regionally varying patterns of climatic change, as opposed to global mean temperature change, many IAM studies have used pattern-scaling (e.g. IMAGE: Bouwman et al., 2006). This computationally inexpensive technique linearly relates regional climatic change, derived from stored GCM ensembles such as those generated in CMIP5, to global mean temperature change, simulated using a simplified model, so that the regional response to many emissions scenarios can be computed quickly (e.g. Cabré et al., 2010). Simple pattern scaling assumes that the climate response is spatially invariant (with respect to time and forcing), and therefore cannot capture aspects which may be sensitive to the greenhouse gas (GHG) concentration pathway (O'Neill and Oppenheimer, 2004; Tebaldi and Arblaster, 2014). Tebaldi and Arblaster (2014) cite a number of instances where it is liable to break down, in particular for scenarios with strong mitigation or less mean temperature change. Recent advances in pattern-scaling have considered the effects of different forcing components; for example, with the most recent iteration of MAGICC-SCENGEN, the effects of aerosols can be estimated for some climate parameters by generating patterns specific to these emissions [1].

[1]MAGICC/SCENGEN user manual, p. 2: http://www.cgd.ucar.edu/cas/wigley/magicc/UserMan5.3.v2.pdf

The ~~AOGCM~~ Atmosphere-Ocean General Circulation Model (AOGCM) ensembles used in pattern scaling are usually multi-model ensembles (MMEs). Such ensembles consist of simulations from different models, and are neither a systematic nor random sampling of potential future climates (Tebaldi and Knutti, 2007). Similarities between models may lead to a lack of independence amongst ensemble members (Foley et al., 2013), complicating the interpretation of the ensemble as a whole (Knutti et al., 2013).

Perturbed physics ensembles (PPEs) offer a more systematic sampling of potential future climates, but embedding a PPE approach into an IAM framework requires a computationally fast climate model. In this context, statistical emulation of complex models is a useful alternative. For example, Castruccio et al. (2014) constructed a statistical climate model emulator using simulations performed with the Community Climate System Model, version 3 (CCSM3), in which statistical models are fitted to temperature and precipitation for 47 subcontinental-scale regions. Such an approach is suitable for applications requiring annual temperatures of specific regions, but is less appropriate when climate impacts within regions are to be considered. Carslaw et al. (2013) apply a similar approach to the grid-cell level. However, such an approach requires many emulators, and correspondingly, computational resources. Furthermore, the global emulation may not be self-consistent, as the individual emulators do not utilise the correlations between grid cells.

In this paper, we demonstrate how model emulation using singular vector decomposition (SVD) can be used within an IAM framework to generate perturbed physics ensembles, systematically capturing uncertainty in the future climate state while also providing insight into regional climate change. We introduce the GENIEem-PLASIM-ENTSem (GPem) climate-carbon cycle emulator, which consists of a statistical climate model emulator, PLASIM-ENTSem, to represent climate dynamics (Holden et al., 2014a), and a new carbon cycle emulator GENIEem. Compared to a simple mechanistic model, the purely statistical GENIEem does not impose a predefined functional structure, allowing the emulator to capture more of the behaviour of the underlying simulator, and notably providing a representation of the parametric uncertainty of the simulator. Although parametric uncertainty of MAGICC itself can be investigated (Meinshausen et al., 2009), this is distinct from representing the parametric uncertainties and associated non-linear feedbacks in the underlying simulator. Similarly, compared to pattern-scaling, the more complex statistical approach used in PLASIM-ENTSem enables a representation of spatial uncertainties due to parametric uncertainties in the underlying model. The use of SVD to decompose spatial patterns of climate parameters makes PLASIM-ENTSem computationally efficient, compared to techniques in which statistical relationships are developed for each grid-cell.

We demonstrate how these emulators can be applied in an IAM framework to resolve the regional environmental impacts associated with policy scenarios by coupling GPem to FTT:Power-E3MG, a non-equilibrium economic model with a technology diffusion component. Our work builds on that of Labriet et al. (2013) and Joshi et al. (2014 subm.) who also derived IAMs from economic and energy technology system models coupled to PLASIM-ENTSem.

2 The GENIEem carbon cycle model emulator

The carbon cycle model emulator GENIEem is an emulator of the GENIE-1 ~~ESM~~ Earth System Model (ESM) (Holden et al., 2013a) (i.e. a statistical model that approximately reproduces selected outputs from the full GENIE-1 ~~simulator~~ESM). The emulator takes a time series of anthropogenic carbon emissions and non-CO_2 radiative forcing (stemming from CH_4, N_2O, halocarbons, and other forcing agents including O_3 and aerosols) as inputs and provides a time series of atmospheric CO_2 concentration as output. ~~Uncertainty in the carbon cycle is captured by emulating 86 possible futures, each with a different set of GENIE-1 parameter inputs~~

In the integrated assessment framework developed here, the time series of anthropogenic carbon emissions is provided by E3MG-FTT, while non-CO_2 forcing data is derived from global timeseries of forcing data obtained through the RCP Database.[2] As such, GPem emulates high-dimensional climate outputs as a function of scalar model inputs (Holden et al., 2015) . We note that certain forcings, such as aerosol forcing, are characterised by complex spatial patterns and so would benefit from an approach in which the inputs are also high-dimensional. However, incorporating such forcing into the emulator framework would involve coupling an aerosol model to PLASIM-ENTS in order to build an ensemble of simulations and a subsequent emulator, which is beyond the current scope of this work. ~~The 86 parameter sets cover a wide range of values, reflecting a large assumed structural uncertainty.~~

2.1 GENIE-1 description

The full GENIE-1 ~~simulator~~ ESM comprises the 3-D frictional geostrophic ocean model GOLDSTEIN (Edwards and Marsh, 2005) coupled to a 2-D Energy Moisture Balance Atmosphere based on that of Fanning and Weaver (1996) and Weaver et al. (2001), and a thermodynamic-dynamic sea-ice model based on Semtner (1976) and Hibler (1979). Ocean biogeochemistry is modelled with BIOGEM (Ridgwell et al., 2007), coupled to the sediment model SEDGEM (Ridgwell and Hargreaves, 2007). GENIE-1 is run at 36×36 spatial resolution (≈10×5 degrees on average) with a ≈1 day atmospheric time step, and 16 depth levels in the ocean. Vegetation is simulated with ENTSML (Holden et al., 2013a), a dynamic model of terrestrial carbon and land use change (LUC) based on the single plant functional type model ENTS (Williamson et al., 2006). ENTSML takes time-varying fields of LUC as inputs. Each simulation used to build the emulator is a transient simulation from 850 AD through to 2105. Historical forcing (850 to 2005 AD), including changing land use, is prescribed as described in Eby et al. (2013). Future forcing (2005 to 2105) is defined by a CO_2 concentration

[2] Data available via the RCP Database at http://tntcat.iiasa.ac.at:8787/RcpDb

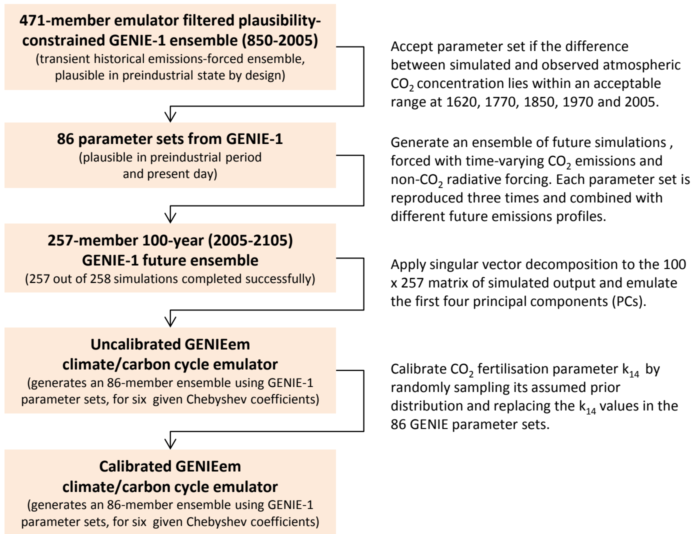

Figure 1. Schematic describing the construction of GENIEem.

time series and a non-CO_2 radiative forcing time series, both represented by polynomials (see section 2.1.2). The LUC mask is held fixed from 2005, as capturing LUC-climate-carbon feedbacks in the emulator would require high dimensional inputs, a significantly more complex ensemble design and emulation challenge. The future forcing due to LUC is instead subsumed into the CO_2 concentration (LUC emissions) and non-CO_2 radiative forcing (LUC albedo).

The configuration is the same as that applied in the Earth system model of intermediate complexity (EMIC) intercomparison project (Zickfeld et al., 2013). Due to its reduced complexity, GENIE-1 is a good choice for performing the many simulations required to build an emulator.

2.2 GENIE-1 parameter set selection

Construction of GENIEem is summarised in (Figure 1). To build the carbon cycle emulator, a subset of the 471-member emulator filtered plausibility-constrained parameter sets described in Holden et al. (2013b) is used. Each of these 471 parameter sets was previously applied to a CO_2 emissions-forced transient historical simulation (850 to 2005 AD). They comprise experiments 1 and 2 of Holden et al. (2013a). In addition to emissions forcing, these simulations were forced by non-CO_2 trace gases, LUC, anthropogenic aerosols, volcanic aerosols, orbital change and solar variability, as described in Eby et al. (2013).

The 471 parameter sets are constrained to be plausible in the preindustrial state by design (Holden et al., 2013b). However, they are not constrained to be plausible in the present day as neither the anthropogenic carbon sinks nor the LUC emissions are calibrated. Additionally, these 471 parameter sets are known to contain members that display numerical instabilities (Holden et al., 2013a).

In order to identify useful parameter sets, we apply a filter to this transient historical ensemble. A parameter set is accepted as plausible if the difference between simulated and observed atmospheric CO_2 concentration lies within an acceptable range at each of five time points, 1620, 1770, 1850, 1970 and 2005 AD:

$$|CO_2(t) - CO_2^*(t)| < \sqrt{\epsilon_0^2 + \epsilon_t^2} \quad (1)$$

where CO_2(t) and CO_2^*(t) are simulated and observed atmospheric CO_2 concentration, evaluated at each time slice t, and the acceptable errors ϵ_0 and ϵ_t relate to the preindustrial spin-up state and to the transient change. The time points span the preindustrial period and are not associated with volcanic eruptions as these can lead to an unrealistic carbon-cycle response in GENIE due to the single layer soil module (Holden et al., 2013a).

The ϵ_0 term dominates the acceptable error during the preindustrial era and is designed to reject any simulations that exhibit numerical instability. It is set equal to 2 standard deviations (9 ppm) of the 471-member spin-up ensemble. The ϵ_t term is given by 0.22×(CO_2^*(t)-280) ppm. This term dominates the acceptable error in the post-industrial era and is designed to reject simulations that exhibit an unreasonable strength for the CO_2 sink. It approximately limits the range of acceptable uncertainty to the inter-model variance of the multi-model C4MIP ensemble (Friedlingstein et al., 2006), assuming that the range of simulated CO_2 change across the C4MIP ensemble scales linearly with simulated CO_2 change relative to preindustrial (280 ppm). Eighty-six parameter sets satisfied this constraint at all 5 time points.

2.3 GENIE-1 ensemble design

These 86 parameter sets from the full GENIE-1 ~~simulator~~ ESM were used to generate an ensemble of future simulations (2005 to 2105) forced with time-varying CO_2 emissions and non-CO_2 radiative forcing. Each simulation was continued from its respective transient historical simulation. Radiative forcing was applied as a globally uniform ~~modification to~~ additional term in outgoing long-wave radiation to capture the combined effects of non-CO_2 trace gases, aerosols and LUC on global temperature. The LUC mask was fixed at the 2005 distribution, but effects of future land use changes are accounted for, albeit approximately, in the applied radiative forcing and emissions anomalies.

To capture the range of possible future forcing we followed the approach of Holden and Edwards (2010). The CO_2 emissions profile is represented as:

$$E = E_0 + 0.5[E_1(t+1) + E_2(2t^2 - 2) + E_3(4t^3 - 4t)] \quad (2)$$

where t is time, normalised onto the range -1 to 1 (2005 to 2105). The coefficient ranges were chosen to span ~~the RCP emission scenarios~~ emissions consistent with the RCP pathways (Moss et al., 2010):

E_1 = -30 to 30 GtC yr^{-1}, E_2 = -15 to 15 GtC yr^{-1}, E_3 = -15 to 15 GtC yr^{-1}. The 2005 emissions E_0 = 9.166 GtC yr^{-1}. Note that Eq. (2) is strictly a linear combination of Chebyshev polynomials such that the first two terms give the linear increase in emissions; we refer to the coefficients henceforth as 'Chebyshev coefficents'.

The non-CO_2 radiative forcing profile is also represented by a linear combination of modified Chebyshev polynomials:

$$R = R_0 + 0.5[R_1(t+1) + R_2(2t^2 - 2) + R_3(4t^3 - 4t)] \quad (3)$$

These Chebyshev coefficients are varied in the ranges R_1 = -10 to 10 Wm^{-2}, R_2 = -5 to 15 Wm^{-2}, R_3 = -5 to 5 Wm^{-2}. The 2005 non-CO_2 radiative forcing R_0= 0.619 Wm^{-2}.

The E_1 and R_1 coefficients define the 2100 CO_2 emissions and non-CO_2 radiative forcing respectively. The ranges for these coefficients have been chosen to encompass (and exceed) the ranges of 2100 forcing in Moss et al. (2010) ; for emulator training we apply wider ranges than we expect to apply in order to ensure the emulator is never used under extrapolation.

The maximum E_1 = 30 gives 2100 CO_2 emissions of E_0+E_1=39.166 GtC, which compares to RCP 8.5 emissions of 28.817 GtC. Maximum radiative forcing of R_0+R_1=10.619 Wm^{-2}) was allowed to greatly exceed RCP estimates (maximum 1.796 Wm^{-2}) in order to allow the potential application of the emulator to extreme non-CO_2 forcing scenarios.

The 86 parameter sets were ~~reproduced~~ replicated three times, and each of these three 86 parameter sets was combined with different future emissions profiles to produce a 258-member ensemble. To achieve this, the six coefficients were varied over the above ranges to create a 258-member Maximin Latin Hypercube design, using the maximinLHS function of the lhs package in R (R Development Core Team, 2013). 257 simulations completed~~successfully~~; in the remaining simulation, input parameters led to an unphysical state and ultimately, numerical instability.

2.4 Construction of GENIEem

The emulation approach closely follows the dimension reduction methodology detailed in Holden et al. (2014a). We have an ensemble of 257 transient simulations of the coupled climate-carbon system, incorporating both parametric uncertainty (28 parameters) and forcing uncertainty (6 modified Chebyshev coefficients). For coupling applications we require an emulator that will generate the annually resolved evolution of CO_2 concentration through time (2006 to 2105). The simulation outputs were combined into a (100×257) matrix Y, and SVD was performed on the matrix

$$Y = LDR^T \quad (4)$$

where L is the (100×257) matrix of left singular vectors ("components"), D is the 257×257 diagonal matrix of the square roots of the eigenvalues and R is the 257×257 matrix of right singular vectors ("component scores").

We retain the first four components, which together explain more than 99.9% of the ensemble variance. Each individual simulated CO_2 concentration time series can thus be well approximated as a linear combination of the first four components, scaled by their respective scores. Each set of scores consists of a vector of coefficients, representing the projection of each simulation onto the respective component. As each simulated time series is a function of the input parameters, so are the coefficients that comprise the scores. So each component score can be viewed, and hence emulated, as a scalar function of the input parameters to the simulator.

Emulators of the first four component scores were derived as functions of the 28 model parameters and the 6 concentration profile coefficients. These emulators were built in R (R Development Core Team, 2013), using the stepAIC function (Venables and Ripley, 2002). For each emulator, we first built a linear model from all 34 inputs allowing only terms that satisfy the Bayes Information Criterion (BIC). BIC-constrained stepwise addition of quadratic and cross terms was then performed, allowing only inputs present in the linear model.

While the variance in emulator output is dominated by the Chebyshev forcing coefficients, uncertainty for a given forcing scenario is generated through emulator dependencies on GENIE-1 parameters. The most important of these is the CO_2 fertilisation parameter, k_{14}, describing the uncertain response of photosynthesis to changing CO_2 concentrations. To use the emulator, we constrain k_{14} using the calibration of Holden et al. (2013a), to better quantify the uncertainty associated with the terrestrial sink. We evaluate the resulting emulated uncertainty through a comparison with CMIP5 in Section 2.6.

We approximate the prior as a normal distribution with mean 500 ppm and standard deviation 150 ppm, following the base posterior of Holden et al. (2013a). We sampled values at random from this distribution and replaced the k_{14} values in the 86-member training parameter set. Then, to generate a perturbed parameter ensemble of emulated futures, the emulation is performed for each of the resulting 86 parameter sets.

2.5 Validation of GENIEem

To validate the emulator, we apply leave-one-out cross-validation, which involves rebuilding the emulator 257 times with a different simulation omitted and comparing the omitted simulation with its emulation. The proportion of variance V_T explained by the emulator under cross-validation is given by:

$$V_T = 1 - \sum_{n=1}^{257}\sum_{t=1}^{100}(S_n,t - E_n,t)^2 \,/\, \sum_{n=1}^{257}\sum_{t=1}^{100}(S_n,t - \bar{S}_t)^2 \qquad (5)$$

where $S_{(n,t)}$ is the simulated CO_2 concentration at time t in left-out ensemble member n, $E_{(n,t)}$ the corresponding emulated output and $\bar{S}_t$ is the ensemble mean output at time t. V_T measures the degree to which individual emulations can be regarded as accurate (Holden et al., 2014a)

The cross-validated root mean square error of the emulator is given by:

$$RMSE = \sqrt{\sum_{n=1}^{257}\sum_{t=1}^{100}\frac{(S_n,t-E_n,t)^2}{25700}} \tag{6}$$

The proportion of variance explained by the emulator under cross-validation is found to be 96.8%, and the cross-validated root mean square error of the emulator is 34 ppm.The ensemble distribution of cross-validated emulator error does not exhibit any significant trends as a function of the forcing, being approximately distributed about zero, independently of the final CO_2 concentration. This suggests that the emulator errors are likely dominated by describing parametric uncertainty with low order polynomials, and so would be randomly distributed across a perturbed parameter emulated ensemble. To test this we performed a simulation ensemble forced by RCP8.5. The simulated ensemble mean of 2100 CO_2 = ~~1040~~ 990 $\pm$ ~~99~~ 92 ppm. This compares to the emulated ensemble mean of ~~1032~~ 975 $\pm$ ~~79~~ 73 ppm with the same forcing. The R^2 value for emulated versus simulated output is 74.5%. The emulator explains 74% of the variance in 2100 CO_2 across the RCP 8.5 simulation ensemble, demonstrating that the parametric uncertainty is reasonably well approximated.

Given that the RCP estimate is 936 ppm, this data appears to show that the emulator and simulator overstate the RCP 8.5 concentration in the median. However, the reason for this is is that this validation did not use the CO_2 fertilization prior, which is applied to the emulator to constrain the predictions.

2.6 Evaluation of GENIEem using RCPs

To further evaluate the emulator's performance, we consider GENIEem's response to forcing by Representative Concentration Pathways (RCPs; Van Vuuren et al., 2011). For each RCP, CO_2 emissions, non-CO_2 radiative forcing and CO_2 concentrations are provided by Meinshausen et al. (2011b).[3] GENIEem is run using Chebyshev coefficients derived by fitting Eqs 1 and 2 to RCP CO_2 emissions and non-CO_2 radiative forcing data. Emulated CO_2 concentrations are compared to the CO_2 concentrations corresponding to that RCP in the RCP Database. For RCP 8.5, we also compare the emulator range with the CMIP5 ensemble range of CO_2 concentrations for that RCP.

GENIEem median CO_2 concentrations are generally well centred on the RCPs (Figure 2). The RCP profiles were derived assuming carbon cycle rates that were calibrated to the median of the C4MIP models. This good agreement is therefore not imposed, but is desirable as it suggests that the ensemble of GENIE-1 parameter sets is not significantly biased with respect to C4MIP. The full

[3]Data available via the RCP Database at http://tntcat.iiasa.ac.at:8787/RcpDb

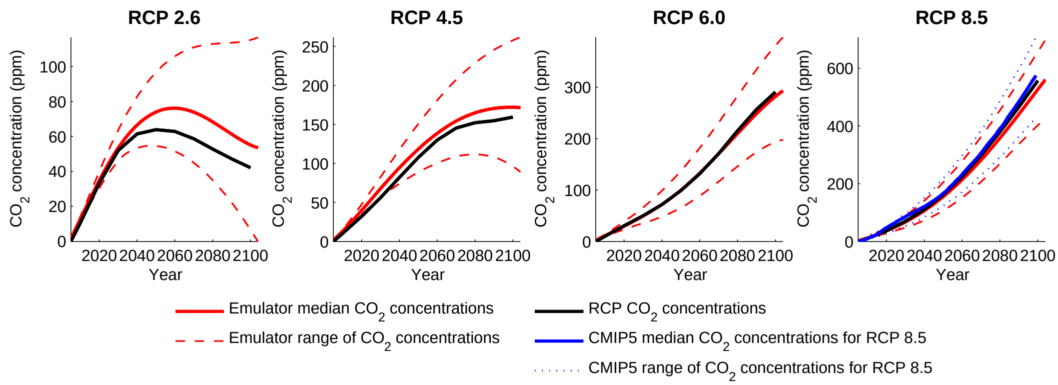

Figure 2. Carbon cycle emulator output compared with RCP data, for the four RCPs 2.6, 4.5, 6.0 and 8.5.Anomalies are relative to 2005. For RCP 8.5, CMIP5 data is presented as a reference.

range of 2105 emulated CO_2 concentrations under RCP 8.5 forcing is 806 to 1076 ppm. When forced with the same RCP, 11 CMIP5 models simulate a range of 795 to 1145 ppm by 2100 (Friedlingstein et al., 2013), demonstrating that the emulator can reproduce existing estimates of the carbon cycle uncertainty. In a related analysis, the ensemble mean and variance were shown to be easier to emulate than individual simulations (Holden et al., 2014a). The emulator's capacity to capture the CMIP5 simulation ensemble suggests that this is also the case here.

For RCP 2.6, the difference between the RCP value and the emulator median reaches about 15 ppm. One possible explanation for this is the formulation of land use change. When land use is changed in GENIE, soil carbon evolves dynamically to a new equilibrium. Therefore, although the LUC mask is held fixed after the transient 850-2005 AD spin-up, there are ongoing land-atmosphere fluxes in the future (2005-2105) due to historical LUC. Since the RCP emissions data used to force GENIEem already include the contribution from soil carbon fluxes, the inconsistency of approaches is liable to lead to a net additional forcing while the historical contribution decays. These residual emissions would be most significant in RCP 2.6 because other emissions are lowest in this scenario, potentially contributing to the excess concentrations in the emulation of RCP 2.6. This difference could be reduced by using a more sophisticated treatment of the forcing inputs that separated fossil fuel and land use carbon emissions, with land use emissions calculated from spatially explicit scenarios based on above-ground carbon change, as in Houghton (2008).

3 Application of GPem in an IAM framework

To demonstrate the utility of emulation within an integrated assessment framework, we describe how GENIEem, along with PLASIM-ENTSem has been used to explore the climate change implications of four of the policy scenarios for the electricity sector, as presented in Mercure et al. (2014). GPem is coupled to FTT:Power-E3MG, which combines a technology diffusion model with a non-equilibrium economic model. Mercure et al. (2014) emphasises the policy instruments that can be applied to decarbonisation of the global energy sector, and analysis of climate impacts is limited to mean surface temperature anomalies. Here, we extend that work to illustrate the regional patterns of climate variability associated with different policy scenarios, and discuss these results in the context of "dangerous climate change" (Jarvis et al., 2012).

3.1 The climate model emulator: PLASIM-ENTSem

PLASIM-ENTSem is an emulator of the GCM PLASIM-ENTS; both simulator and emulator are described by Holden et al. (2014a). The GCM consists of a climate model, PLASIM (Fraedrich, 2012), coupled to a simple surface and vegetation model, ENTS (Williamson et al., 2006), which represents vegetation and soil carbon through a single plant functional type. PLASIM has a heat-flux corrected slab ocean and a mixed-layer of a given depth, and a 3D dynamic atmosphere, run at T21 $\sim$5 degree resolution. It utilises primitive equations for vorticity, divergence, temperature and the logarithm of surface pressure, solved via the spectral transform method, and contains parameterizations for long and short-wave radiation, interactive clouds, moist and dry convection, large-scale precipitation, boundary layer fluxes of latent and sensible heat and vertical and horizontal diffusion. It accounts for water vapour, carbon dioxide and ozone.

As an emulator of PLASIM-ENTS, PLASIM-ENTSem emulates mean fields of change for surface air temperature and precipitation well, while emulations of precipitation underestimate simulated ensemble variability, explaining $\sim 60-80\%$ of the variance in precipitation (compared to $\sim 95\%$ for surface air temperature) (Holden et al., 2014a).

The response of PLASIM-ENTSem to RCP forcing was analysed in Holden et al. (2014a, Figure 6) ; in all four scenarios, the emulated ensemble distribution was found to compare favourably with the multi-model CMIP5 ensemble.

3.2 Policy scenarios and emissions profiles

FTT:Power is a simulation model of the global power sector (Mercure, 2012), which has been coupled to a dynamic simulation model of the global economy, E3MG (Mercure et al., 2014)[4]. These models are described in greater detail in the supplementary information of this paper. Policies within the electricity sector drive the uptake or phasing out of types of generators, leading to different CO_2 emission profiles (Figure 3).

Here we consider four scenarios, a subset of the ten scenarios explored in Mercure et al. (2014). Scenario *i* is the no-climate-policy baseline. The baseline scenario extends current policies in the energy sector to 2050. It assumes no additional technology subsidies worldwide, feed-in tariffs in some EU countries, and carbon pricing in the EU. Figure 3 illustrates that the emissions associated

[4] www.4cmr.group.cam.ac.uk/research/FTT/fttviewerhttp://www.4cmr.group.cam.ac.uk/research/FTT/fttviewer

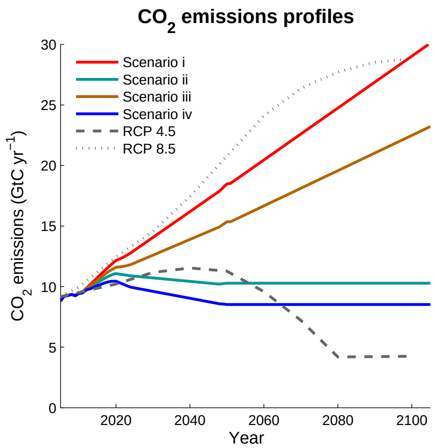

Figure 3. Total CO_2 emissions associated with four different electricity sector-only policy scenarios. Total CO_2 emissions associated with two RCPs are shown for reference. (Note that the RCP scenarios cover all sectors and land use.)

with this scenario are of a similar magnitude as emissions associated with RCP 8.5, but following a more linear trajectory.

Scenario *ii* introduces carbon pricing, which rises to 200-400 2008\$/tCO_2. Scenario *iii* explores the use of carbon pricing, along with technology subsidies and feed-in tariffs in the developed world only. Finally, scenario *iv* uses carbon pricing, along with technology subsidies and feed-in tariffs to incentivise decarbonisation, and also includes regulations to ban the construction of new coal power plants in China if not equipped with Carbon Capture and Storage; this policy set decarbonises the global electricity sector by 90% (relative to 1990 emissions) by 2050.

3.3 Coupling procedure

As FTT:Power-E3MG runs until 2050, emissions for 2050-2105 are estimated using a linear best-fit trend, except in the case of successful mitigation scenarios, where such an approach could lead to implausible emissions reductions by 2105. In these scenarios, the emissions in $PgCy^{-1}$ reached in 2050 were assumed to remain constant beyond 2050 (i.e. in these scenarios, it is assumed that by 2050, the energy sector has decarbonised as much as can be incentivised under the specified policies).

Chebyshev coefficients are calculated to provide least squares fits to each emissions profile produced by FTT:Power-E3MG. If we conservatively assume that any error in emissions due to differences between the FTT:Power-E3MG emissions profile and the corresponding Chebyshev curve has an infinite lifetime in the atmosphere, the accumulated error does not exceed 4.5 ppm in any scenario over the period 2005-2105, well within the 5th-95th percentiles of GENIEem.

As FTT:Power-E3MG does not simulate non-CO_2 radiative forcing, we select the RCP that best matches the CO_2 concentrations associated with the baseline scenario (RCP 8.5) and force GENIEem with the non-CO_2 radiative forcing associated with that RCP. The RCP 8.5 non-CO_2 radiative forcing was applied to all scenarios as the RCPs lack a suitable analog to the CO_2 concentrations associated with the power sector mitigation scenarios examined in this work. Values for Chebyshev coeficients are calculated and these three coefficients, together with the three CO_2 emissions coefficients, are the inputs to GENIEem.

This approach maintains comparability across the different scenarios, although we expect some small reductions in CH_4 and N_2O in the mitigation scenarios, due to a reduction in leaks of these GHGs from drilling. Representations of these GHGs in E3MG-FTT are not sufficiently detailed to provide forcing data for GPem, but reductions in fuel use-related CH_4 and N_2O emissions of around 10-15% by 2050 in the mitigation ~~scenario~~ scenarios can be inferred. After 2050, we expect a stabilisation at this new level, as the sectors involved have decarbonised by 90%, producing a reduction in forcing of roughly 0.1 Wm^{-2} (relative to total forcing of 7.3 to 8.3 Wm^{-2} in the baseline and 5.3 to 6.2 Wm^{-2} in the mitigation scenario, accounting for carbon cycle uncertainty). This small reduction in forcing is well within the uncertainty bounds of GENIEem.

Climate-carbon feedbacks are emulated entirely within GENIEem. No climate information is passed from PLASIMem to GENIEem. PLASIM-ENTSem takes inputs of both actual CO_2 (for CO_2 fertilization) and equivalent CO_2 (for radiative forcing). Chebyshev coefficients are calculated to provide least squares fits to the median and 5th-95th percentiles of the GENIEem ensemble CO_2 concentrations; these coefficients, therefore, correspond to actual CO_2 concentrations. Chebyshev coefficients for equivalent CO_2 are also calculated, corresponding to combined CO_2 and non-CO_2 forcings. To determine these coefficients for equivalent CO_2, the median and 5th-95th percentiles of the GENIEem ensemble CO_2 concentrations are converted to radiative forcing following:

$$\Delta F = 5.35 \ln(CO_2/280) Wm^{-2} \tag{7}$$

RCP 8.5 non-CO_2 forcing is added to this time series to give total radiative forcing, which is converted to equivalent CO_2 using the previous relationship. Chebyshev coefficients for equivalent CO_2 are fitted to the resulting time series.

Thus, PLASIM-ENTSem is forced with three sets of six coefficients (three actual CO_2 and three equivalent CO_2 each for the median and 5th-95th percentiles of the GENIEem ensemble).

We calculate the median warming of the PLASIM-ENTSem ensemble based on the 5th and 95th percentiles of the GENIEem ensemble. These bounds, therefore, illustrate parametric uncertainty of the carbon cycle model alone.

We also calculate the median and 5th-95th percentiles of warming of the PLASIM-ENTSem ensemble from the median GENIEem ensemble output. These bounds reflect parametric uncertainty in the climate model alone.

Finally, we calculate the 5th percentile of warming from the PLASIM-ENTSem ensemble based on the 5th percentile of CO_2 concentration from the GENIEem ensemble, and the 95th percentile of warming from the PLASIM-ENTSem ensemble based on the 95th percentile of CO_2 concentration from the GENIEem ensemble. This third set of bounds reflects warming uncertainty due to parametric uncertainty in the climate model and the carbon cycle model, computed under the assumption that GENIEem and PLASIM-ENTSem projections are perfectly correlated, i.e. that states exhibiting the greatest CO_2 concentration in GENIEem correspond to states exhibiting greatest warming in PLASIM-ENTSem. Many carbon cycle processes are affected directly by changes in temperature, or by variables which covary with temperature (Willeit et al., 2014), so while such a correlation is not absolute, there is a motivation for this approach.

4 Results

4.1 GPem mean warming under policy scenarios

We applied GPem to determine the atmospheric CO_2 concentrations and mean global temperature anomalies associated with different mitigation policiies applicable to the energy sector. ~~Due~~ While the mitigation policies explored generate reductions in CO_2 emissions from the energy sector, due to the effect of non-CO_2 radiative forcing on climate, combined with remaining CO_2 emissions, CO_2 concentrations continue to increase in mitigation scenarios (Figure 4). Figure 4 also illustrates the temperature anomalies associated with each of the scenarios. Modelled anomalies are relative to the model baseline, 1995-2005. Therefore historical warming, estimated at ≈ 0.6°C in 2000 (IPCC, 2013) is added to give anomalies relative to the preindustrial period. While there is no scenario in which temperature stabilises by 2100, in scenario *iv*, the rate of warming remains roughly constant, while in scenario *i*, the rate of warming appears to increase towards the later half of this century. The effect of cascading uncertainty is apparent (Jones, 2000; Foley, 2010), leading to large uncertainty bounds for temperature projections.

4.2 GPem regional climate under policy scenarios

Figure 5 illustrates the 2095-2105 December-February and June-August warming anomalies associated with scenario *i* and *iv*, presenting the median and 5th/95th percentiles of the PLASIM-ENTSem ensemble outputs calculated independently at each grid point. These emulated ensembles are forced with GENIEem median CO_2 concentrations for the respective scenario, giving an indication of the range of PLASIM-ENTSem parametric uncertainty associated with the projection. It is evident that the warming associated with the baseline scenario would be partially offset under the mitigation sceanrio. However, certain hotspots of warming are apparent even under the 5th percentile projection. In both scenarios, there is cooling in south-east Asia in summer, which likely arises due to a

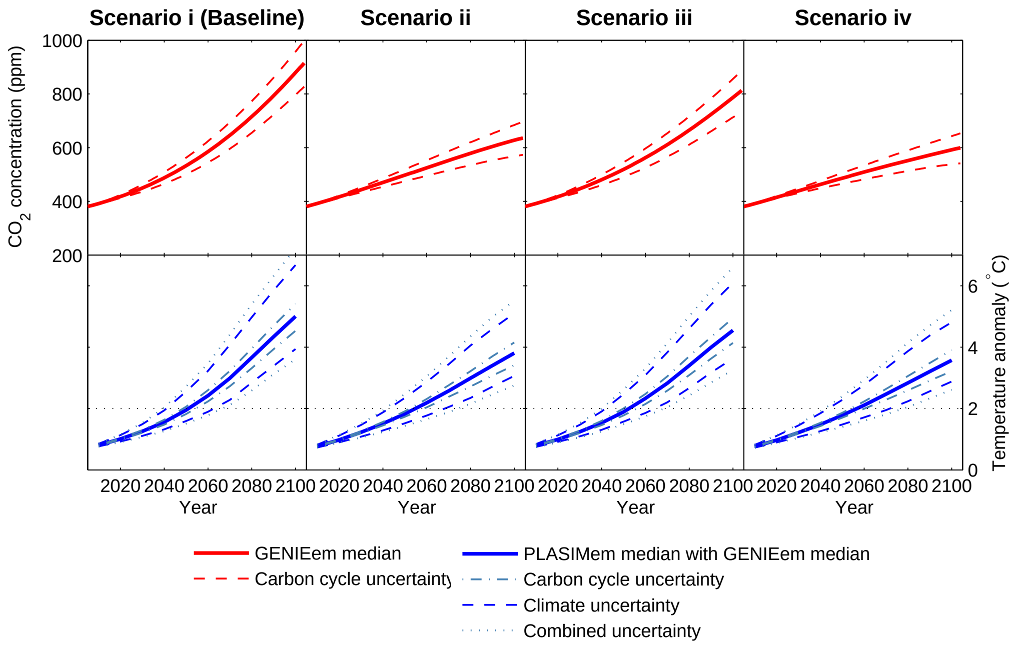

Figure 4. *Top*: Median CO_2 concentrations for scenarios *i* (baseline), *ii*, *iii* and *iv*, simulated by GENIEem, with uncertainty bounds (GENIEem 5th/95th percentile). *Bottom*: Median temperature anomalies relative to preindustrial conditions for scenarios*a* (baseline), *d*, *i* and *j*, simulated by PLASIM-ENTSem using median GENIEem CO_2 concentrations. Uncertainty bounds are based on carbon cycle uncertainty (PLASIMem median with GENIEem 5th/95th percentile), climate uncertainty (PLASIMem 5th/95th percentile with GENIEem median), and combined uncertainty (PLASIMem 5th/95th percentile with GENIEem 5th/95th percentile). The 2°C target, described as 'the maximum allowable warming to avoid dangerous anthropogenic interference in the climate' (e.g Randalls, 2010), is also illustrated by the grey dashed line.

strengthening of the monsoon in PLASIM-ENTSem. However, Holden et al. (2014a) note that this signal may not be robust as the model lacks aerosol forcing.

Figure 6 illustrates the mean 2095-2105 December-February and June-August precipitation patterns associated with scenario *i* and *iv*, along with the proportion of the 86 ensemble members simulating increased precipitation in each case. Generally, areas that experience a significant increase/decrease in precipitation under scenario *iv* (i.e. larger than $\pm$ 1 mm ~~/yr~~day^{-1}) experience even greater extremes under scenario *i*, which can be attributed to ~~strengthened Hadley circulation in a warmer world~~differences in water vapour amount in the atmosphere due to warming (Held and Soden, 2006) ; precipitation fields are amplified as more water is available in the convergence zones to condense. Plotting the proportion of ensemble members that project increasing precipitation shows that in most regions of the world, there is high agreement between ensemble members on the direction of change for precipitation.

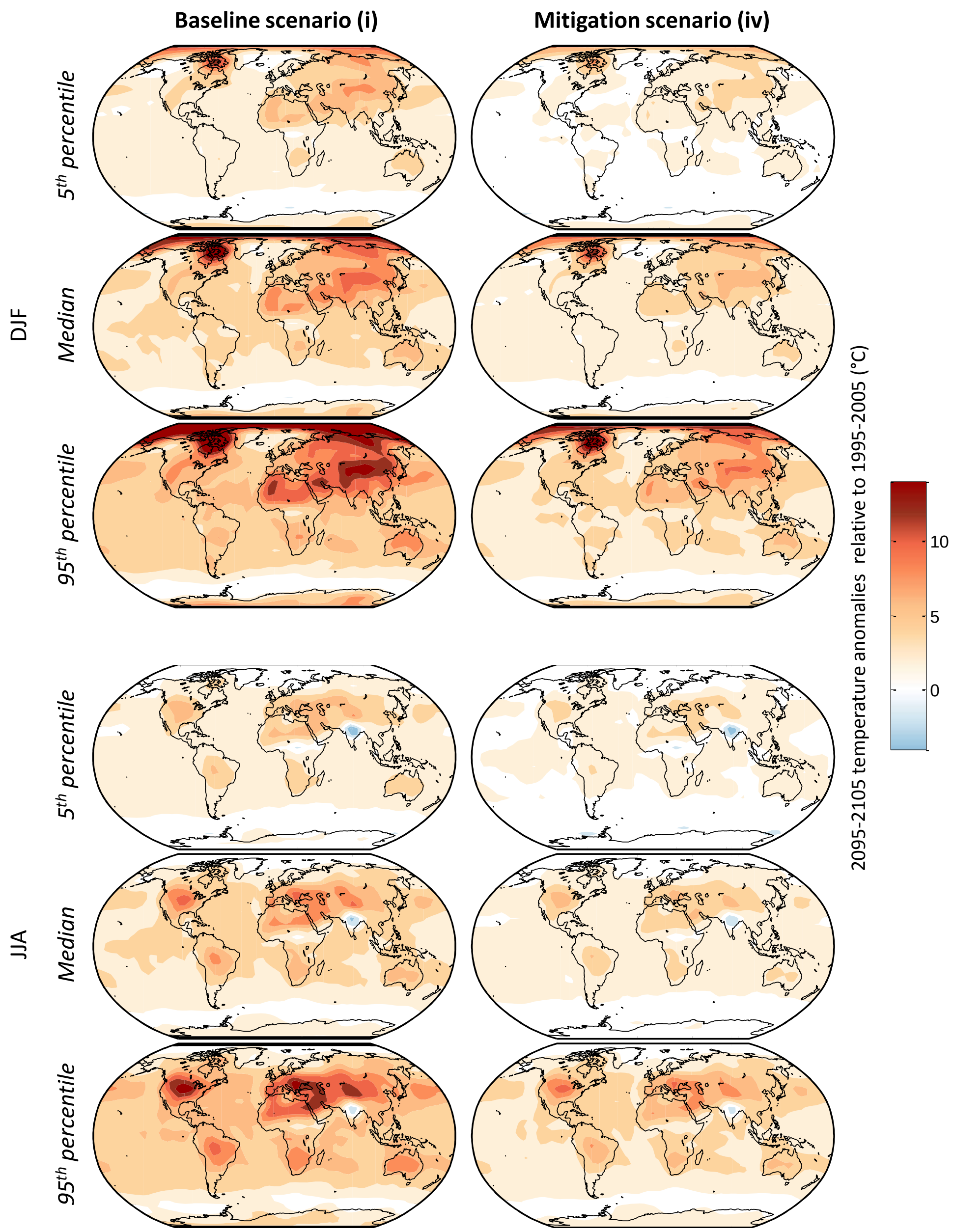

Figure 5. 2095-2105 temperature anomalies relative to 1995-2005 for DJF and JJA under the baseline scenario *i* (right) and the mitigation scenario *iv* (left). The 5th, 50th and 95th percentile of the PLASIM-ENTSem ensemble are calculated independently at each grid point. The PLASIM-ENTSem ensembles are forced with GENIEem median CO_2 concentrations for that scenario.

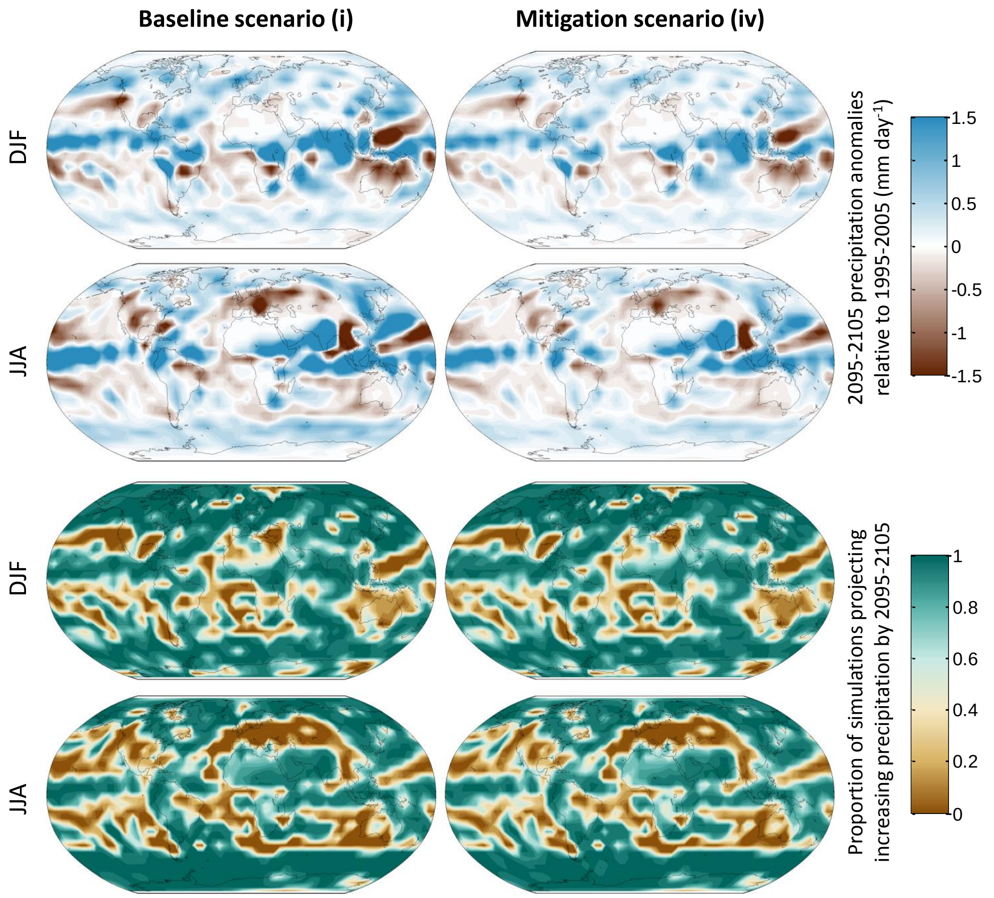

Figure 6. 2095-2105 precipitation anomalies (ensemble means) relative to 1995-2005 under the baseline scenario *i*, and the mitigation scenario *iv* (top) and proportion of ensemble members simulating increased precipitation (bottom).

Precipitation patterns are similar for the two scenarios presented (r=0.99), suggesting that a simple pattern scaling approach would have sufficed in the particular example considered here, at least for estimation of the ensemble mean field. However, Tebaldi and Arblaster (2014) considered correlations between the averaged precipitation anomaly fields (2090-1990) of the CMIP5 multi-model ensemble when forced with different RCPs; the lowest correlation (0.85) was between ensembles forced with RCP 2.6 and RCP 8.5, while a correlation of 0.97 was found between RCP 4.5 and RCP 8.5. Applying our emulation framework yielded correlations of: 0.89-0.93 (RCP 2.6, RCP 8.5) and 0.97-0.98 (RCP 4.5, RCP 8.5), depending on season. This comparison ~~demonstrates~~ suggests that the emulation framework captures non-linear feedback strengths that are comparable to those found in a high-complexity high-resolution multi-model ensemble and, furthermore, that the assumptions of pattern scaling ~~are especially likely to break down~~ may not be optimal when applied to strong mitigation scenarios.

5 Conclusions

We have described and validated a new carbon cycle model emulator, GENIEem, and applied it along with PLASIM-ENTSem to demonstrate the utility of statistical model emulation in an IAM setting. The climate-carbon cycle emulator GPem was used to examine atmospheric CO_2 concentration, mean global temperature anomolies, and spatial temperature and precipitation response patterns resulting from CO_2 emission scenarios associated with various mitigation scenarios for the electricity sector.

Even the most successful mitigation strategy considered here results in warming of above 3.5°C by 2100, a level of warming which Parry et al. (2009) notes could result in substantial harmful impacts, including risks of water shortage and coastal flooding. As such, in a context where the global electricity sector is decarbonised by 90%, further emissions reductions must be achieved in other sectors (e.g. transport and industry) to enable CO_2 concentrations to remain below 450 ppm, and correspondingly, global warming below 2°C (Meinshausen et al., 2009).

The latest IPCC AR5 notes that in 2010, the energy supply sector accounted for 35% of total GHG emissions (IPCC, 2014) , therefore there is scope for reductions to be achieved in other sectors. For instance, policy options explored by Luderer et al. (2012) which keep CO_2 concentrations below 450 ppm, using the IMACLIM-R and ReMIND-R models, include mitigation in the transportation sector to reduce energy demand. However, the IPCC AR5 notes that based on scenario analysis, sectors currently using liquid fuel may be more costly, and therefore slower, to decarbonize than electricity. Additionally, it is worth noting that the most successful mitigation scenarios explored in the IPCC AR5, which lead to CO_2 equivalent concentrations in the range of 430-480 ppm by 2100 (approximately equivalent to RCP 2.6) feature large-scale, long-term application of carbon dioxide removal (CDR) technologies, in addition to large emissions reductions (IPCC, 2014). This analysis, focusing on the effectiveness of mitigation policies in the electricity sector, therefore highlights the danger of focusing mitigation efforts on this single sector, where the cost of decarbonisation is lower; not only are such efforts insufficient to maintain global warming below 2°C, but additionally, the heterogeneous distribution of climate impacts globally will need to be addressed.

Furthermore, the inadequacy of electricity sector to solve the emissions problem is in spite of the fact that the inclusion of non-linear feedbacks on technology uptake is expected to promote decarbonisation in our model, compared to the equilibrium models in the IPCC AR5 database, which may not capture the complexities of real-world human behaviour in mitigation decision-making (Mercure et al., 2015) .

The 2°C warming threshold is often a focal point of climate mitigation policy and scholarship, and is indeed useful as a guiding principle (e.g. Den Elzen and Meinshausen, 2006; Oberthür and Roche Kelly, 2008; Shindell et al., 2012). However, it is also vital to consider the complex temperature and precipitation patterns that could occur, lest a focus on the global mean temperature result in regional climate impacts being overlooked. Furthermore, consideration must be given to how to adapt to diverse regional climate change, should this target not be met (Parry et al., 2009). Applying the GPem framework yields a more systematic representation of uncertainty in future regional climate states, when compared the pattern-scaling approaches that are based on "ensembles of opportunity" (Stone et al., 2007).

While uncertainty associated with carbon cycle and climate modelling in this framework are accounted for through the use of ensembles, it is still possible that the actual future climate state may fall outside the simulated range. Uncertainty associated with emissions profiles is more difficult to quantify as these depend, ultimately, on human decision-making. Therefore many policy contexts should be modelled in order to find out which ones effectively lead to desired outcomes.

Acknowledgements. We acknowledge the support of D. Crawford-Brown. We thank F. Babonneau, for providing code to generate Chebyshev coefficients, P. Friedlingstein for provision of CMIP5 data and M. Syddall, for his advice regarding data visualisation.This work was supported by the Three Guineas Trust (A. M. Foley), the EU Seventh Framework Programme grant agreement n 265170 'ERMITAGE' (N. Edwards and P. Holden), the UK Engineering and Physical Sciences Research Council, fellowship number EP/K007254/1 (J.-F. Mercure), Conicyt (Comisión Nacional de Investigación Científica y Tecnológica, Gobierno de Chile) and the Ministerio de Energía, Gobierno de Chile (P. Salas), and Cambridge Econometrics (H. Pollitt and U. Chewpreecha).

References

Bouwman AF, Kram T, Klein (eds) (2006) Integrated modelling of global environmental change. An overview of IMAGE 2.4. Netherlands Environmental Assessment Agency, Bilthoven, the Netherlands

Cabré MF, Solman S, Nuñez M (2010) Creating regional climate change scenarios over southern South America for the 2020s and 2050s using the pattern scaling technique: Validity and limitations. Climatic Change 98(3-4):449–469

Carslaw K, Lee L, Reddington C, Pringle K, Rap A, Forster P, Mann G, Spracklen D, Woodhouse M, Regayre L, et al. (2013) Large contribution of natural aerosols to uncertainty in indirect forcing. Nature 503(7474):67–71

Castruccio S, McInerney DJ, Stein ML, Liu Crouch F, Jacob RL, Moyer EJ (2014) Statistical emulation of climate model projections based on precomputed gcm runs. Journal of Climate 27(5):1829–1844

Den Elzen M, Meinshausen M (2006) Meeting the EU 2°C climate target: Global and regional emission implications. Climate Policy 6(5):545–564

Eby M, Weaver A, Alexander K, Zickfeld K, Abe-Ouchi A, Cimatoribus A, Crespin E, Drijfhout S, Edwards N, Eliseev A, et al. (2013) Historical and idealized climate model experiments: an intercomparison of Earth system models of intermediate complexity. Climate of the Past 9:1111–1140

Edwards NR, Marsh R (2005) Uncertainties due to transport-parameter sensitivity in an efficient 3-d ocean-climate model. Climate Dynamics 24(4):415–433

Fanning AF, Weaver AJ (1996) An atmospheric energy-moisture balance model: Climatology, interpentadal climate change, and coupling to an ocean general circulation model. Journal of Geophysical Research: Atmospheres (1984–2012) 101(D10):15,111–15,128

Foley A (2010) Uncertainty in regional climate modelling: A review. Progress in Physical Geography 34(5):647–670

Foley A, Fealy R, Sweeney J (2013) Model skill measures in probabilistic regional climate projections for Ireland. Climate Research 56(1):33–49

Fraedrich K (2012) A suite of user-friendly global climate models: Hysteresis experiments. The European Physical Journal Plus 127(5):1–9

Friedlingstein P, Cox P, Betts R, Bopp L, Von Bloh W, Brovkin V, Cadule P, Doney S, Eby M, Fung I, Bala G, John J, Jones C, Joos F, Kato F, Kawamiya M, Knorr W, Lindsay K, Matthews H, Raddatz T, Rayner P, Reick C, Roeckner E, Schnitzler KG, Schnur R, Strassman K, Weaver A, Yoshikawa C, Zeng N (2006) Climate-carbon cycle feedback analysis: Results from the C4MIP model intercomparison. Journal of Climate 19(14):3337–3353

Friedlingstein P, Meinshausen M, Arora VK, Jones CD, Anav A, Liddicoat SK, Knutti R (2013) Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. Journal of Climate 27(2): 511–526

Held IM, Soden BJ (2006) Robust responses of the hydrological cycle to global warming. Journal of Climate 19(21):5686–5699

Hibler W (1979) A dynamic thermodynamic sea ice model. Journal of Physical Oceanography 9(4):815–846

Holden P, Edwards N (2010) Dimensionally reduced emulation of an AOGCM for application to integrated assessment modelling. Geophysical Research Letters 37(21): L21707

Holden P, Edwards N, Gerten D, Schaphoff S (2013a) A model-based constraint on CO_2 fertilisation. Biogeosciences 10(1):339–355

Holden P, Edwards N, Müller S, Oliver K, Death R, Ridgwell A (2013b) Controls on the spatial distribution of oceanic $\delta^{13}C_{DIC}$. Biogeosciences 10(3):1815–1833

Holden P, Edwards N, Garthwaite P, Fraedrich K, Lunkeit F, Kirk E, Labriet M, Kanudia A, Babonneau F (2014a) PLASIM-ENTSem v1. 0: a spatio-temporal emulator of future climate change for impacts assessment. Geoscientific Model Development 7(1):433–451

Holden PB, Edwards NR, Garthwaite PH, Wilkinson RD (2015) Emulation and interpretation of high-dimensional climate model outputs. Journal of Applied Statistics 42(9):1–18

Houghton R (2008) Carbon flux to the atmosphere from land-use changes 1850-2005. In TRENDS: A compendium of data on global change. carbon dioxide information analysis center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tenn., USA.

IPCC (2013) Summary for Policymakers. In: Stocker T, Qin D, Plattner G, Tignor M, Allen S, Boschung J, Nauels A, Xia Y, Bex V, Midgley P (eds) Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

IPCC (2013) Summary for Policymakers. In: Stocker T, Qin D, Plattner G, Tignor M, Allen S, Boschung J, Nauels A, Xia Y, Bex V, Midgley P (eds) Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

IPCC (2014) Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Edenhofer O, Pichs-Madruga R, Sokona Y, Farahani E, Kadner S, Seyboth K, Adler A, Baum I, Brunner S, Eickemeier P, Kriemann B, Savolainen J, Schlömer S, von Stechow C, Zwickel T and Minx J (eds.), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jarvis A, Leedal D, Hewitt C (2012) Climate-society feedbacks and the avoidance of dangerous climate change. Nature Climate Change 2(9):668–671

Jones RN (2000) Managing uncertainty in climate change projections–issues for impact assessment. Climatic Change 45(3-4):403–419

Joshi SR, Vielle M, Babonneau F, Edwards NR, Holden PB (2014) Physical and economic consequences of sea-level rise: A coupled GIS and CGE analysis under uncertainties. Environmental and Resource Economics. Submitted.

Knutti R, Masson D, Gettelman A (2013) Climate model genealogy: Generation CMIP5 and how we got there. Geophysical Research Letters 40(6):1194–1199

Labriet M, Joshi SR, Kanadia A, Edwards NR, Holden PB (2013)Worldwide impacts of climate change on energy for heating and cooling. Mitigation and Adaptation Strategies for Global Change doi:10.1007/s11027-013-9522-7

Luderer G, Bosetti V, Jakob M, Leimbach M, Steckel JC, Waisman H, Edenhofer O (2012) The economics of decarbonizing the energy systemâ€”results and insights from the recipe model intercomparison. Climatic Change 114(1):9–37

Meinshausen M, Meinshausen N, Hare W, Raper SC, Frieler K, Knutti R, Frame DJ, Allen MR (2009) Greenhouse-gas emission targets for limiting global warming to 2°C. Nature 458(7242):1158–1162

Meinshausen M, Raper S, Wigley T (2011a) Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6–Part 1: Model description and calibration. Atmospheric Chemistry and Physics 11(4):1417–1456

Meinshausen M, Smith SJ, Calvin K, Daniel JS, Kainuma M, Lamarque J, Matsumoto K, Montzka S, Raper S, Riahi K, et al. (2011b) The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. Climatic Change 109(1-2):213–241

Mercure JF (2012) FTT:Power : A global model of the power sector with induced technological change and natural resource depletion. Energy Policy 48(0):799– 811

Mercure JF, Pollitt H, Chewpreecha U, Salas P, Foley A, Holden P, Edwards N (2014) The dynamics of technology diffusion and the impacts of climate policy instruments in the decarbonisation of the global electricity sector. Energy Policy 73(0):686–700

Mercure JF, Pollitt H, Bassi A, Viñuales J, Edwards N (2015) Braving the tempest: Methodological foundations of policy-making in sustainability transitions. arXiv preprint arXiv:150607432

Moss RH, Edmonds JA, Hibbard KA, Manning MR, Rose SK, van Vuuren DP, Carter TR, Emori S, Kainuma M, Kram T, et al. (2010) The next generation of scenarios for climate change research and assessment. Nature 463(7282):747–756

Oberthür S, Roche Kelly C (2008) EU leadership in international climate policy: achievements and challenges. The International Spectator 43(3):35–50

O'Neill BC, Oppenheimer M (2004) Climate change impacts are sensitive to the concentration stabilization path. Proceedings of the National Academy of Sciences of the United States of America 101(47):16,411–16,416

Parry M, Lowe J, Hanson C (2009) Overshoot, adapt and recover. Nature 458(7242):1102–1103

R Development Core Team (2013) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org

Randalls S (2010) History of the 2°c climate target. Wiley Interdisciplinary Reviews: Climate Change 1(4):598–605

Ridgwell A, Hargreaves J, Edwards NR, Annan J, Lenton TM, Marsh R, Yool A, Watson A (2007) Marine geochemical data assimilation in an efficient Earth system model of global biogeochemical cycling. Biogeosciences 4(1):87–104

Ridgwell A, Hargreaves J (2007) Regulation of atmospheric CO_2 by deep-sea sediments in an Earth system model. Global Biogeochemical Cycles 21(2)

Semtner AJ (1976) A model for the thermodynamic growth of sea ice in numerical investigations of climate. Journal of Physical Oceanography 6(3):379–389

Schaeffer M, Gohar L, Kriegler E, Lowe J, Riahi K, van Vuuren D (2013) Mid-and long-term climate projections for fragmented and delayed-action scenarios. Technological Forecasting and Social Change

Shindell D, Kuylenstierna JC, Vignati E, van Dingenen R, Amann M, Klimont Z, Anenberg SC, Muller N, Janssens-Maenhout G, Raes F, et al. (2012) Simultaneously mitigating near-term climate change and improving human health and food security. Science 335(6065):183–189

Stone D, Allen MR, Selten F, Kliphuis M, Stott PA (2007) The detection and attribution of climate change using an ensemble of opportunity. Journal of Climate 20(3):504–516

Tebaldi C, Arblaster JM (2014) Pattern scaling: Its strengths and limitations, and an update on the latest model simulations. Climatic Change 122(3):459–471

Tebaldi C, Knutti R (2007) The use of the multi-model ensemble in probabilistic climate projections. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 365(1857):2053–2075

Van Vuuren DP, Edmonds J, Kainuma M, Riahi K, Thomson A, Hibbard K, Hurtt GC, Kram T, Krey V, Lamarque JF, et al. (2011) The representative concentration pathways: an overview. Climatic Change 109(1-2):5–31

Venables WN, Ripley BD (2002) Modern applied statistics with S. Springer

Weaver AJ, Eby M, Wiebe EC, Bitz CM, Duffy PB, Ewen TL, Fanning AF, Holland MM, MacFadyen A, Matthews HD, et al. (2001) The UVIC earth system climate model: Model description, climatology, and applications to past, present and future climates. Atmosphere-Ocean 39(4):361–428

Willeit M, Ganopolski A, Dalmonech D, Foley AM, Feulner G (2014) Time-scale and state dependence of the carbon-cycle feedback to climate. Climate Dynamics 42(7-8):1699–1713

Williamson M, Lenton T, Shepherd J, Edwards N (2006) An efficient numerical terrestrial scheme (ENTS) for Earth system modelling. Ecological Modelling 198(3):362–374

Zickfeld K, Eby M, Weaver AJ, Alexander K, Crespin E, Edwards NR, Eliseev AV, Feulner G, Fichefet T, Forest CE, et al. (2013) Long-term climate change commitment and reversibility: An EMIC intercomparison. Journal of Climate 26(16):5782–5809