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:

'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 C4MIP 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.

Response to Reviewer #2

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 CO2 concentrations, and the text has been amended to reflect this:
   'Each individual simulated CO2 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 CO2 emissions. I guess the latter, as I could not see a 90% reduction on Figure 3.**

    Figure 3 refers to total CO2 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-CO2 RCP is chosen, but it then says that RCP8.5 non-CO2 is used for all scenarios. Which is it?**

    RCP8.5 non-CO2 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-CO2 radiative forcing, we select the RCP that best matches the CO2 concentrations associated with the baseline scenario(RCP 8.5) and force GENIEem with the non-CO2radiative forcing associated with that RCP. The RCP 8.5 non-CO2 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-CO2 here means just methane and nitrous oxide? Is that correct? If so, what is the implication of just considering these non-CO2 gases rather than the wider mix including the short-lived pollutants?**

    Please refer to response to comment 3. CH4 and N2O 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-CO2 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 CO2 is to sum the forcings before calculating the equivalent CO2. There is a hint in the next sentence that the CO2 seen by the model is the sum of equivalent CO2's calculated individually for actual CO2 and non-CO2 forcings. Perhaps the difference is negligible, but it would be worth clarifying.**

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

14. **1294:5. "due to the effect of non-CO2 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 CO2, the combination of remaining CO2 forcing and non-CO2 forcing still have a warming effect. This has been clarified as follows:
    'While the mitigation policies explored generate reductions in CO2 emissions from the energy sector, due to the effect of non-CO2 radiative forcing on climate, combined with remaining CO2 emissions, CO2 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.