Dear editor,

Please consider this revised manuscript for publication in Biogeosciences. It is of interest to the broad global change community, and very timely in that it characterizes a land cover issue in CMIP5 and highlights a critical path forward for CMIP6 land cover and land use considerations. The central issue is that without land cover harmonization among models, the integrated assessment models and earth system models are all simulating effectively different terrestrial scenarios for a single prescribed Representative Concentration Pathway (RCP), especially for RCP4.5. These different terrestrial scenarios have different carbon cycle effects and local to global climate effects that make model intercomparison extremely difficult because as a result of these differences some models could be simulating a completely different scenario space than the one prescribed.

We have revised the manuscript in accordance with the reviewers suggestions. The main changes were to 1) frame our work in the context of the CMIP5 scenario-based process and clearly present why the inconsistencies need to be addressed and to 2) clarify the text through a) additional and reorganized methods and b) restructured discussion that refers more to the results. We have also added the requested figures. Our point-by-point responses to the reviewers are included in the supplemental 'response' file, as well as a manuscript with highlighted text denoting where changes have been made. We have included this highlighted manuscript in place of a list of all relevant changes because the review responses effectively list the relevant changes and also because it is easier to see where we modified the text than to cross-reference a descriptive list.

Sincerely,
Alan (on behalf of all co-authors)

Response to Anonymous Referee #1, bgd-11-C2671-2014

Title: "From land use to land cover: restoring the afforestation signal in a coupled integrated assessment - earth system model and the implications for CMIP5 RCP simulations"
by Di Vittorio et al.

We appreciate your thorough and thoughtful review and suggestions. We agree that there is a simple message, but the issue is much more complicated than the simple message implies. The complications contribute to the very existence of the reported inconsistencies and the effects on the global modeling. The following responses to your comments show how the manuscript will be improved.

Major comments

We disagree that this manuscript relies "quite a bit" on the in-review GMD paper by Bond-Lamberty et al. and an in-prep paper by Collins et al. We cite these papers to refer to additional technical details that do not need to be presented here in order to understand this paper on land coupling between ESMs and IAMs. In fact, the Bond-Lamberty et al. paper focuses on the climate feedback part of the loop that does not contribute to the reported inconsistencies, and the Collins et al. paper focuses on technical development of the model and its code. We can definitely remove these references without affecting this paper, but then we would be omitting two very relevant references. We can certainly clarify the relationship of these references to the current paper.

The use of emissions in our simulations can easily be clarified. As you note, we do specify that our simulations use emissions and the RCP4.5 scenario. The CMIP5 land use/cover data we present were used for both emissions and concentration driven CESM simulations, although we think only the concentration driven outputs are available from the CMIP5 archive. This does not matter for this paper because we only look at the land use/cover trajectory data from CMIP5. All other data are from emissions-driven simulations.

The effects of restoring afforestation on atmospheric CO_2 can also be easily clarified. The vegetation carbon and atmospheric CO_2 gain changes (19 Pg C, -8 ppm, respectively) are model outputs. So the 8ppm is the net reduction in atmospheric CO_2 gain with a fully coupled carbon cycle operating, by 2040. And this is actually a big deal because this is over only the first 25 years of 2/3 of prescribed afforestation. There are 60 more years until 2100, during which additional afforesting occurs and the previously afforested area continues to grow. We do explicitly state that the other numbers are linear extrapolations to make the point that the full afforestation over the entire century would likely have a very different atmospheric CO_2 concentration (~40 ppm difference). Unfortunately, our simulations cover only until 2040 because they were performed during a developmental phase. These simulations are very expensive to run, and we needed to reserve computational time for our final production simulations, which do run to 2100. The simulations presented in this paper do cover the most rapid period of afforestation from 2015-2020 and the subsequent 20 years.

We agree that regional biophysical effects of land use/cover change are very important, and in many cases more significant that global impacts on the carbon cycle, but the focus of this paper is on overall consistency of the land surface, which is required in order to adequately evaluate regional effects. We do discuss the regional impacts in the introduction, and can mention them in the discussion as well.

The land cover in CLM can be changed by only one component at a time: either the dynamic vegetation module or the land use change module. Here we use the land use change module and thus do not account for potential biogeographic vegetation shifts due to changing climate. While this is a shortcoming of the model, we are not concerned about this limitation because most current studies show that the biogeographical effects of climate change on vegetation distribution are small compared to the effects of land use change on vegetation distribution, both in recent history and in 21st century projections.

We do, however, discuss how non-crop vegetation changes when cropland and pasture change. The constraint of 'potential vegetation' is presented in section 2.3.1 (page 7161,line 17) but we should explain what it means (land cover as it would be today if no land use change had ever occurred). The algorithms for land cover change are presented in figures 3 and 4. We further discuss how this constraint (page 7167) limits afforestation in the OLDLUT and in CMIP5. We remove this constraint to increase afforestation, and if the conditions are not right where forest is added, then the forest should not grow well in CLM, which would have a negative feedback on afforestation. We further discuss this issue on page 7168 where we explain that the prescribed afforestation assumes that silvicultural inputs are available (water, fertilizer, etc.) while CLM does not include such inputs. So to meet the RCP4.5 scenario, afforestation needs to occur in CLM, but it might not produce the biomass that the integrated assessment model expected. This is one of the inconsistencies that we point out in this paper.

Minor comments

We can definitely clean up the abstract so that is presents a more clear message.

page 7155 lines 5-9: The mention of C4MIP may not be necessary, and we can remove it, but it is not "totally unrelevant." It draws a relationship between uncertainty in atmospheric variables and uncertainty in carbon uptake due to land use/cover change.

Yes, there are several "land" terms throughout the paper. We will make every effort to consolidate, clarify, and explain our "land" terms. We do define "land use harmonization" on page 7155, and use it consistently throughout the paper.

page 7157 lines 6-7: Yes, this does refer to land carbon uptake, and goes along with the next sentence and corresponding citation.

page 7157 lines 15-17: Yes, we are sure that the radiative forcing targets do not include

the direct effects of land use/cover change. See http://tntcat.iiasa.ac.at/RcpDb/dsd?Action=htmlpage&page=welcome# for this description:
"The RCPs are named according to their 2100 radiative forcing level as reported by the individual modeling teams. The radiative forcing estimates are based on the forcing of greenhouse gases and other forcing agents - but does not include direct impacts of land use (albedo) or the forcing of mineral dust."

The document you refer to appears to show calculations for many radiative forcing components, but not all of them (e.g. land use) were included in the CMIP5 RCP targets.

page 7157, lines 21-26: We disagree that this sentence is irrelevant. The radiative effects of GHGs and some aerosols are included in the RCP targets, so these forcings are included in the shared socioeconomic pathways that try to meet the RCP targets. The biogeophysical forcing effects of land use/cover, however, are not included in the target calculations. So while the atmospheric constituents change to meet the target, there is no biogeophysical forcing constraint on changes in land use/cover, which changes the total forcing from the target (see the Jones et al., 2013a reference). The only land constraint is on how much emissions are released from land use/cover change.

page 7158 line 6: We can clarify that the time varying vegetation productivity in CESM is used by GCAM at 5-year intervals.

page 7158: "lost afforestation signal": We can certainly provide more context here, or even where we introduce the RCPs on page 7156. Rcp4.5 is indeed an afforestation scenario. IAM land projection is driven primarily by human needs and economics, with some assumptions about vegetation productivity. The IAMs use a relatively simple global climate model to determine the effects of emissions at an aggregate global level, and generally do not include the effects of globally aggregated climate on their systems. To our knowledge, no global IAM uses a dynamic vegetation model to estimate biogeography.

page 7158 lines 24-25: We can take out this reference to the second stage.

page 7159 line 17: Yes, this isn't entirely clear. GCAM projects a single year of land use/cover distribution, once every five years.

page 7159: "ingesting": We can replace 'ingesting' with another word. In this case we can use 'using.'

page 7160: Yes, thank you for the suggestion, "land use run" needs to be changed.

page 7160: "GCAM initial conditions": We will clarify that the initial GCAM state is initialized to real world statistics. This state includes production amounts, costs, prices, land areas, etc.

page 7161 lines 19-21: GLM's harvest comes from 5 categories within the main categories of 'primary' and 'secondary' land. However, CLM harvest is from forest only. So the GLM harvest area is normalized by the total area available for harvest (primary + secondary), and then this fraction of harvestable area is used as the fraction of forest area harvested in CLM.

page 7161-7162 lines 22-5: This paragraph is quite dense, especially sentences 3 and 5. We will clarify this. Basically, climate effects on vegetation in CLM are used by GCAM to update land use/cover projections at 5-year intervals.

page 7162 line 8: We will put this into context, as mentioned in response to your previous comment. CESM is supposed to simulate the land use of the RCP4.5. This includes the afforestation of RCP4.5. CESM with land use change does not use a dynamic vegetation module, and even if it did, CESM should still simulate the scenario-induced changes in forest area.

page 7162 lines 13-15: Yes, the spatial allocation of cropland and pasture. GLM maintains its own map of potential forest land. New ag land preferentially replaces forest, and when ag land is lost, it is removed preferentially from area that is considered potential forest land.

page 7162 lines 23-26: The more explicit explanation follows in steps a-c on page 7163. We can rephrase this sentence to be more descriptive or to refer to steps a-c.

page 7164 section 2.3.4 title: This should explicitly refer to "land use harmonization," which is specifically introduced and defined on page 7155.

page 7164 section 3.1 title: We will make this more specific to refer to land cover area inconsistencies. However, the global land area is not exactly the same in each model, which is another inconsistency in the overall coupling.

page 7164 lines 17-18: We use "RCP4.5" in this way to distinguish these CMIP5 GCAM outputs from the GCAM outputs in our iESM simulations, which also simulate the RCP4.5 scenario. We will consider replacing "RCP4.5" with a different label here.

page 7165 line 21: We will clarify this as suggested (area covered by herbaceous PFTs).

page 7165: We will introduce figs 6 and 7 more clearly as the changes and absolute values, respectively, of the same results.

page 7165 lines 22-24: This needs to be clarified. The meaning is that the cropland area in CLM is more representative with NEWLUT than with OLDLUT. And the "normalization" here is a bug fix that makes this improved representation. It is literally a normalization of GLM cropland area to a CLM reference area.

page 7166 line 1: We mean discrepancies between scenario-prescribed land use/cover and the corresponding simulated PFT areas.

page 7166 line 9: Yes, we mean that the ESMs need to simulate the energy/climate scenarios as generated by IAMs.

page 7166 lines 13-15: This needs clarification. The sharing is between the source of land cover info and the ESMs. There are two sources relevant to this discussion: historical data and the IAMs.

page 7166 line 1 to 7167 line 4: We will make these details clear in the methods section.

page 7167 line 27: The difference plot is a cleaner and easier way to see the difference in NEE between the two simulations. We certainly can provide a plot of both simulations. It is unclear whether the reviewer is concerned about the term "significant," of which we do not have a statistical test for here, or if the reviewer prefers to see the separate plots. We might be able to calculate a t-test comparison between the simulations for each of two time periods to find out if they are statistically significantly different: 2005-2020 and 2020-2040. But even though the NEE during the two time periods might not be statistically different based on this test, land carbon uptake does increase with NEWLUT and has considerable effects on vegetation carbon and atmospheric co2 concentration, as reported on page 7168.

page 7168 lines 6-7: We can include the atmospheric co2 plot.

page 7168 line 9: Based on our current simulations, we are seeing very linear responses in both vegetation carbon and atmospheric co2 between 2005 and ~2070. Also, we are extrapolating the difference in gain, which means that any nonlinearities introduced into both simulations by climate or fossil fuel emissions should be somewhat accounted for. So here we use linear extrapolation as a simple estimate, and to mitigate the effects of rapid change starting in 2015, we start our extrapolation at 2005. ~70 of total prescribed afforestation occurs by 2040, but it does not start until 2015. Forest expansion in CLM reduces forest leaf area index to accommodate the new forest, which initially reduces carbon assimilation on a per area basis. As the new forests age they gain leaf area index and carbon assimilation capacity, up to a point dictated by environmental conditions. Throughout the century both forces are acting to maintain forest carbon uptake: new forest area and increasing forest leaf area index. So this linear extrapolation gives a reasonable estimate.

page 7169 lines 13-14: We mean that the ESM land use/cover distribution must match the scenario-prescribed land use/cover distribution to ensure that the ESM is actually simulating the prescribed scenario.

page 7169 lines 25-26: We will rephrase this statement to indicate that the CESM simulation does not use appropriate changes in PFT area.

page 7170 lines 9-10: We do not say that all additional carbon went into vegetation. We do correctly state, as quoted, "..afforestation has a significant impact on iESM's global carbon cycle through increased vegetation carbon and decreased atmospheric CO2 concentration." It turns out that the difference in soil carbon gain between the two simulations is only about 1.5 PgC from 2005-2040, with the NEWLUT gaining soil carbon at a slightly lower rate than the OLDLUT. This decrease in soil carbon gain is small compared to the 19 PgC increase in vegetation carbon gain. The atmospheric co2 concentration is calculated from all fluxes, and the primary change in land carbon is in additional vegetation carbon.

page 7170 line 23: Yes, the RCP4.5 scenario for CMIP5 was simulated by GCAM.

page 7172 line 9: Actually, we mean a spatially explicit land area data set. Currently, each model has its own estimate of global land area and where that land area is located (e.g. different islands may be absent/present in different land area data sets). And technically, global land area is not constant, although it is for the purposes of these simulations.

page 7172 point 4: This needs clarification. Gross transitions are all losses and gains in area between two points in time. Net transitions are the sum of gross transitions between two points in time.

Response to Anonymous Referee #2, bgd-11-C3782-2014

Title: "From land use to land cover: restoring the afforestation signal in a coupled integrated assessment - earth system model and the implications for CMIP5 RCP simulations"
by Di Vittorio et al.

We appreciate your critical and helpful review and suggestions.
The following responses to your comments show how the manuscript will be improved.

General Comments

We will make it more clear that we argue for a more consistent and complete land coupling between IAMs and ESMs for robust scenario-based simulations of global carbon and climate. One of our main points is that ESMs need to simulate the scenario-prescribed land use and land cover in order to robustly estimate the impact on carbon and climate of anthropogenic emissions and land use. Land use and land cover are interdependent and need to be treated as such when used as a scenario condition for earth simulations. Furthermore, land use and land cover sometimes refer to the same thing. For example, forests, which are at the heart of this paper, can be a land use for sequestering carbon, yet are generally treated as land cover.

Our discussion does use the results to explain the problems arising from the CMIP5 land coupling design, but some reorganization and more thorough discussion of why these problems need to be fixed will improve the manuscript. We discuss the considerable impacts on the global carbon cycle on page 7168 as one of the reasons for improving the coupling, and intend to include additional figures showing these effects on vegetation carbon and atmospheric co2.

Summary of the paper

We actually did investigate the reasons for the discrepancies in forest area, and determined that it results from mismatches in model structure, assumptions, and definitions among all 3 models, such that not all appropriate information was passed between the models. We discuss this on pages 7166-7167, with references to three of our figures. The effects of veg-climate interactions and forest management between GCAM and CESM, as discussed on page 7168, are superposed on the incomplete sharing of information.

We actually suggest integrating land cover change with land use change, not replacing land use with land cover. While these two concepts are uniquely defined, they are interdependent; land use influences land cover and vice versa to generate the observed spatial distribution of vegetation.

More discussion needed

We suggest incorporating land cover and land use information, not replacing a land use scenario with a land cover scenario. So we could discuss the pros and cons of land use only vs. land use and land cover.
We agree that we need to state our perspective on the purpose of RCP simulations. It would open up the discussion from one based just on the obvious impacts on the fidelity of the simulated carbon cycle to the prescribed scenario, which we present, to one that addresses the meaning and utility of scenario development and simulation.

In answer to your two example questions: in order to have a robust multi-model intercomparison of responses to atmospheric composition and land use/cover, the multiple models need to simulate the same basic earth (i.e. atmosphere and land changes), as prescribed by a given socio-economic scenario, which would also provide the most accurate projection of global change.

Again, we advocate adding land cover information, not replacing land use information. Whether this would reduce or increase multi-model spread is not known, but it should improve the uncertainty range because it would constrain the sampling space for a given land use/cover change scenario to a more realistic range. For example, on page 7167 we point out that another CMIP5 model increased forest area by 11%, when the prescribed increase was 24% from 2005-2100 (see Davies-Barnard et al. 2014). Additionally, another CMIP5 model (see http://www.biogeosciences-discuss.net/11/5443/2014/bgd-11-5443-2014.pdf) started ~24 M km^2 of forest in 2005 and increased this by 35% by 2100 (which still wasn't total area of prescribed afforestation), while CESM started with ~39 M km^2 of forest area and increased this by only 6%. The RCP4.5 scenario, as simulated by GCAM, started with about 41 M km^2 of forest. The differences among these models for this single scenario are too large to simply be covering a realistic sampling space; there is a lot of room for error reduction here.

Clarity needed

We will clarify manuscript with the aid of your specific comments, and also the comments by the other referee.

Specific Comments

page 7156 line 14: We argue that it will improve earth system simulations by making all components of the earth system, including the human components that drive the earth scenarios, more consistent with each other.

page 7158 line 24: Yes, we will clarify in the text that GCAM also makes use of heterotrophic respiration.

section 2.3.1 paragraph 2: Yes, this paragraph can be moved to section 2.1. We will probably keep it for completeness, and clarify it. We analyzed CESM outputs of NPP, HR, and vegetation and soil carbon densities to develop this particular feedback method. The carbon density values were too sensitive to changes in vegetation area, thereby masking the climate feedback that we wanted to implement. NPP and HR were more robust climate feedback proxies. Also, the the below ground factor is not the inverse of the HR ratio, this is incorrect; it is an average of the above ground NPP factor and the opposite fractional change of the HR ratio ((NPPratio + (1-(HRratio-1))/2)). The effect of soil carbon increases are accounted for by the increase in NPP. The GCAM carbon densities are used to determine how much carbon to value in a particular place. These scaling factors are also used by GCAM to adjust crop productivity. Both the carbon value and crop productivity are used to make land use projections in GCAM.

section 2.3.1: We will describe the relationship between pasture and grass and shrub in CLM. There is some discussion of this on page 7167, but this needs to be described in the methods.

section 3.1: We will clarify this section. Using CESM in place of CLM would reduce acronym usage. It was difficult to clearly distinguish the CMIP5 GCAM RCP4.5 scenario from the GCAM scenario used in the iESM (which is also RCP4.5). If we just use 'GCAM' for both of them the reader will not know to which we are referring. We will work to make this clear.

page 7166 line 1: We argue that this is a gap. One of our main points is that the ESMs need to simulate the scenario-prescribed land use and cover in order to estimate the impact on carbon and climate of anthropogenic emissions and land use. The range of carbon cycle and climate responses are going to come from differences in the ESMs' biogeochemistry and physics modules (see page 7169 lines 15-22). The range of land cover responses is robust only if 1) all the ESMs start with the same land use and land cover areas and spatial distributions (which they do not), 2) correctly simulate the prescribed land use changes (which they do not, e.g. afforestation land use is not passed to them), and 3) then use different assumptions about how land use changes affect land cover distributions (which they do). The bottom line here is that even condition 2 is not satisfied and so the range of land use scenarios in ESMs is unconstrained, rather than being given by the four RCPs, as designed. It is then quite difficult to compare the effects of land use on cover/carbon/climate when each ESM implements different land use under the same RCP scenario. More generally, the interdependence of land use and land cover requires more consistency among models and more complete information to robustly estimate the possible range of land cover responses to land use change, and how these responses impact carbon and climate. We will, as mentioned above, discuss the goals of CMIP5 and explain why more consistency and information is needed to meet these goals.

page 7166 line 10: We can clarify sentences to make the following more clear: The rest of page 7166 through line 7 of page 7167 focuses on the main source of the forest area inconsistency: the land use harmonization, its lack of land cover info, and its relationship to CESM. Page 7167 focuses on the role of the CESM land use translator in this inconsistency. The ESM simulates given areas of plant functional types, a type of land cover, and we discuss how this relates to pasture as a land use, via the land use translator, also on page 7167. The inconsistency does not arise from errors in the IAM, but on page 7168 we do discuss how certain land use assumptions in the IAM generate afforestation that might not be consistent with land use/management/cover assumptions in CESM (see figure 8). This study presents an experiment in using the land use translator to allow CESM to use the prescribed land use. We found that this is not enough, and that more consistency and information shared across the models is necessary for accurate simulations of scenarios. The culmination of the gained insights is presented in the list of suggestions for improving IAM - ESM land coupling on pages 7171-7172.

page 7168 lines 1-2: This needs some clarification, or maybe reasonable is not the right word to use here. The point is that for RCP4.5 we need CESM to have more forest area, so finding an upper limit, given the current structures of the translator and the model, and the available information, is a desirable constraint. Going beyond this constraint to add forest area is completely arbitrary and unwarranted.

page 7168 paragraph 1: Yes, this would be interesting. Comparing with Jones et al. 2013b figure 2 we see that the iESM vegetation uptake by 2040 due to more afforestation would noticeable change its veg carbon trajectory. More importantly, we see from this figure that RCP4.5 has a large spread of slopes in vegetation carbon change, which might be partially due to different levels of afforestation in the different models. Another large spread of slopes appears to be for RCP8.5, which has net deforestation and might also be affected by the lack of land cover info sharing. The slopes for the other two RCPs appear to be more consistent across models. This same paper plots the atmospheric co2 concentrations of the four IAMs that generated the RCPs (prescribed concentrations), and at 2040 the iESM with improved afforestation brings the atmospheric co2 concentration down by an amount comparable to the difference between the prescribed RCP4.5 and RCP6.0 concentrations. This being said, the emissions driven CMIP5 CESM RCP4.5 simulation has a co2 concentration that is on the order of 50ppm higher than the prescribed concentration at 2040

page 7168 line 17: The potential vegetation constraint is a main reason why CESM did not simulate the prescribed afforestation. And by incomplete we do mean that it is smaller than the GCAM forest area increase.
We need to explain CESM's land cover in more detail in the methods. CESM with land use change cannot use the dynamic vegetation module that would change vegetation area based on climate. Only the land use change changes vegetation area in our simulations. Regionally, the land use initiated changes generally swamp the biogeographical climate-induced changes in vegetation area. Changes in CESM forest area track GCAM forest area changes annually. The main time-lag differences would be related to the biogeochemistry differences between the models, i.e. age and growth rate vs carbon amount and leaf area index. This contributes to carbon cycle differences between GCAM and CESM, which do not focus on here (we discuss this briefly on page 7169 lines 9-22).

page 7169 lines 21-22: Changes in land cover due to changes in climate are relatively small compared to changes in land cover due to land use change. At coarse resolution and regional to global levels, changes due to land use dominate. At fine resolution in local areas, however, climate induced shifts in vegetation can be very important for a

variety of reasons, including biodiversity. Ideally, the climate-induced shifts would be included with the land use shifts, but there needs to be a mechanism for separating the effects of each, so that proper multi-model intercomparisons can be made. Models with the same configurations need to be compared (e.g. use vs use, use+climate vs use +climate, climate vs climate), otherwise the addition of climate induced shifts might confound results more than they would clarify them. Currently, the state of modeling appears to still be struggling with the dominant effects of land use.
With respect to uncertainty in land cover area in general, it is of interest, and should generally be introduced at the scenario level for proper multi-model intercomparisons. This would make the best use of available state-of-the-art current day data sets model projections and their uncertainties, which could be better characterized as input to the global models. Part of the reasoning here is that land cover and land use are interdependent, and thus should be addressed jointly. The climate-indued shifts in vegetation would most likely reside within the ESMs, however, but again, this is an additional process that needs a consistent land base to work from. With the iESM we are looking at one form of this climate-induced vegetation uncertainty on land use by adding climate impacts to the land use projections through productivity changes rather than spatial range changes. The land use model in GCAM then determines spatial distributions of land use and cover, albeit without a direct biogeographic comment. Overall, this is a complex and iterative process that still needs to be researched. Furthermore, uncertainty in land cover will always be introduced by the ESMs because they each implement the land surface differently. Some use plant functional types, some use land cover designations, and their categories of land cover are not identical (this is also the case for historical land use/cover data sets). So given the exact same input land cover distribution, each model will have a slightly different representation of the land surface based on its translation to its native land system. This structural model uncertainty is important, and yet it can be constrained to a more realistic range by clear definitions and accurate mappings between land use/cover systems.

page 7169 lines 24-27: Yes, the hypothesis that we test is whether we can match GCAM's forest area by modifying the land use translator. And the result is that we cannot, which means that more information is needed to meet the prescribed land use. And so if the GCAM afforestation land use area is passed through to CESM, then CESM will be able to simulate the correct afforestation area, like you say. We explain both in the methods and the discussion why the GCAM afforestation does not get passed through. The mechanism to pass it through does not exist. Only cropland, pasture, secondary, and primary land areas and transitions (and harvests, which we do not address here) are passed through. The results obviously show that this is not enough to do the job of providing the prescribed land use, even after modifying the land use translator. So we can say that a "lack of specific land cover type information being shared among GCAM, GLM, and CESM in the iESM as the primary cause of CESM simulating very little afforestation and effectively no change in herbaceous vegetation."

page 7171 lines 10-13: Yes, each ESM would need a scenario-specific method to meet the prescribed land use/cover. This is essentially the same as what was already done, but with the added constraint of using the land cover outputs of the four IAMs as additional constraints. As no land cover information was passed between IAMs and

ESMs, each ESM used its own set of transitions between land cover types as it matched cropland and pasture area. Some models were able to use the primary and secondary areas and even the transitions between these four categories, but in the end each ESM decided which land cover to add or remove in response to changes in land use. So the land use area trajectory (cropland, pasture, secondary, primary), would not change, but how these land use changes affect land cover (and land uses such as afforestation) would change to better match the prescribed land use/cover scenarios.

The captions for figures 3 and 4 are very similar because of the descriptive explanations of certain processes within the flow charts, both of which depict the land use translator. But figure 3 does state it is for the OLDLUT, and figure 4 for NEWLUT.

From land use to land cover: Restoring the afforestation signal in a coupled integrated assessment - earth system model and the implications for CMIP5 RCP simulations

Alan V. Di Vittorio[1,*], Louise P. Chini[2], Ben Bond-Lamberty[3], Jiafu Mao[4], Xiaoying Shi[4], John Truesdale[5], Anthony Craig[5], Kate Calvin[3], Andrew Jones[1], William D. Collins[1], Jae Edmonds[3], George C Hurtt[2], Peter Thornton[4], Allison Thomson[3]

[1]Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[2]University of Maryland, College Park, MD, USA
[3]Joint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, MD, USA
[4]Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN, USA
[5]Independent contractor with Lawrence Berkeley National Laboratory, Berkeley, CA, USA

[*]Lawrence Berkeley National Laboratory
Earth Sciences Division
One Cyclotron Road, Mail Stop 74R316C
Berkeley, CA 94720-8268
avdivittorio@lbl.gov

Abstract

Climate projections depend on scenarios of fossil fuel emissions and land use change, and the IPCC AR5 parallel process assumes consistent climate scenarios across Integrated Assessment and Earth System Models (IAMs and ESMs). The CMIP5 project used a novel "land use harmonization" based on the Global Land use Model (GLM) to provide ESMs with consistent 1500-2100 land use trajectories generated by historical data and four IAMs. A direct coupling of the Global Change Assessment Model (GCAM), GLM, and the Community ESM (CESM) has allowed us to characterize and partially address a major gap in the CMIP5 land coupling design: the lack of a corresponding land cover harmonization. For RCP4.5, CESM global afforestation is only 22% of GCAM's 2005 to 2100 afforestation. Likewise, only 17% of GCAM's 2040 afforestation, and zero pasture loss, were transmitted to CESM within the directly coupled model. This is a problem because GCAM relied on afforestation to achieve RCP4.5 climate stabilization. GLM modifications and sharing forest area between GCAM and GLM within the directly coupled model did not increase CESM afforestation. Modifying the land use translator in addition to GLM, however, enabled CESM to include 66% of GCAM's afforestation in 2040, and 94% of GCAM's pasture loss as grassland and shrubland losses. This additional afforestation increases CESM vegetation carbon gain by 19 PgC and decreases atmospheric CO_2 gain by 8 ppmv from 2005 to 2040, which demonstrates that CESM without additional afforestation simulates a different RCP4.5 scenario than prescribed by GCAM. Similar land cover inconsistencies exist in other CMIP5 model results, primarily because land cover information is not shared between models. Further work to harmonize land cover among models will be required to increase fidelity between IAM scenarios and ESM simulations and realize the full potential of scenario-based earth system simulations.

1. Introduction

Land use plays a major role in determining terrestrial-atmosphere mass and energy exchange (Adegoke et al., 2007; Raddatz, 2007), which in turn influences local to global climate (Brovkin et al., 2013; Jones et al., 2013a; Pitman et al., 2009). Despite much recent progress, we still have a limited understanding of how historical land use has affected, and continues to affect, climate (Brovkin et al., 2013; Jones et al., 2013a; Pitman et al., 2009) and carbon (Anav et al., 2013; Arora and Boer, 2010; Houghton, 2010; Houghton et al., 2012; Hurtt et al., 2006; Jain et al., 2013; Jain and Yang, 2005; Jones et al., 2013b; Smith and Rothwell, 2013), and high uncertainty as to how land use might evolve in the future (Hurtt et al., 2011; van Vuuren et al., 2011a; Wise et al., 2009). Part of the uncertainty in future land use trajectories is due to inherent unpredictability of human actions, and part to the high diversity of potential climate mitigation and adaptation scenarios. Several energy and land strategies have been proposed to mitigate climate change (Rose et al., 2012;Smith et al., 2013a), and while these strategies have similar overall goals, some strategies will likely compete for land and other resources if implemented simultaneously. For example, afforestation and bioenergy production both aim to reduce atmospheric CO_2 concentrations, but both activities require land area, and both strategies would impact crop production and markets through effects on crop area (Reilly et al., 2012).

Reflecting this limited understanding of land use effects on climate and carbon, Global Climate Models (GCMs), and also next generation Earth System Models (ESMs) that include fully coupled atmosphere-land-ocean carbon cycles, implement a wide range of land use/cover approaches with varying degrees of detail and limited inclusion of managed ecosystems and land use practices (Brovkin et al., 2013; Pitman et al., 2009). The Land Use and Climate, IDentification of robust impacts (LUCID) activity employed seven GCMs to determine whether land use change has significant regional climate impacts and farther-reaching teleconnections due to biophysical changes in land surface. The results for 1972-2002 revealed significant but inconsistent changes in temperature, precipitation, and latent heat in some areas where land use change had occurred. The authors concluded that the model disagreement was due mainly to differences in land use and land cover change implementations and corresponding land cover distributions, with contributions from methodological differences in crop phenology, albedo, and evapotranspiration (Pitman et al., 2009). The environmental factors addressed by LUCID are also key factors for determining carbon uptake by vegetation, and thus it is not surprising that the Coupled Climate-Carbon Cycle Model Intercomparison Project (C^4MIP) activity generated ESM projections that range from the land being a carbon source to a large carbon sink by 2100 (Friedlingstein et al., 2006).

To advance the scientific understanding of the effects of land use change on climate, phase 5 of the Coupled Model Intercomparison Project (CMIP5) (Taylor et al., 2012) applied a novel “land use harmonization” approach to produce the required land use change information for all participating GCMs and ESMs. The Global Land use Model (GLM) was used for this land use harmonization to generate the first set of continuous, spatially gridded land use change scenarios for the years 1500-2100 (Hurtt et al., 2011). GLM computes land use states and transitions annually at half-degree, fractional spatial resolution, including secondary land age, area, and biomass, and the spatial patterns of shifting cultivation and wood harvesting (Hurtt et al., 2006). Land use products from GLM have successfully been used as inputs to both regional and global dynamic land models (Baidya Roy et al., 2003; Hurtt et al., 2002; Shevliakova et al., 2009) and fully coupled ESMs (Jones et al., 2011; Shevliakova et al., 2013). The land use harmonization process ensures a continuous transition from the historical reconstructions to the future projections made by Integrated Assessment Models (IAMs).

The land use harmonization methodology was designed to satisfy the demands of a broad range of models and to provide a consistent set of land use inputs for GCMs and ESMs. The historical period of the land use harmonization (1500-2005) was based on version 3.1 of the Historical Database of the Environment (HYDE; Klein Goldewijk et al., 2011) and Food and Agriculture Organization (FAO) wood harvest data. For the future period (2005-2100), the land use harmonization process utilized land use data from the four Representative Concentration Pathways (RCPs), each provided by a different IAM. The RCP scenarios were designed to each meet a different radiative forcing target (2.6, 4.5, 6.0, and 8.5 W m^{-2}), and due to differences among the IAMs these scenarios spanned a range of approaches in all sectors, including land use, for meeting the targets (van Vuuren et al., 2011a). As a result, forest cover change varied widely from deforestation to afforestation across the scenarios. Once the land use data were passed through the land use harmonization, each GCM/ESM utilized a unique subset of the harmonized outputs, based on model capabilities, and applied it to a unique set of land use and land cover types (e.g. Lawrence et al., 2012). Although this process was largely successful in enabling the first spatially explicit land use driven climate change experiments, it introduced considerable uncertainty into the climate response for a given RCP in part because of model-specific translation requirements between harmonized land use outputs and GCM/ESM simulated land cover. This uncertainty due to inconsistent land cover distributions among models precluded robust intercomparison of land-atmosphere processes (e.g., carbon uptake, evapotranspiration) because differences among models were dominated by the differences among simulated land cover distributions (Brovkin et al., 2013). As land use and land cover are interdependent, a more detailed specification of the relationship between land use and land cover may reduce uncertainty in earth system simulations such that experiments can focus on land-atmosphere process uncertainty rather than be confounded by inconsistent land use/cover distributions.

Recent analyses of CMIP5 results using prescribed CO_2 concentrations have also showed the land ranging from a carbon source to a sink in 2100 for a given scenario (Brovkin et al., 2013; Jones et al., 2013b). The LUCID activity was repeated for five CMIP5 ESMs and the results demonstrated that large inter-model spreads of key regional land surface variables (temperature, precipitation, albedo, latent heat, and available energy) were still due mainly to differences in land use and land cover change implementations and corresponding land cover distributions. Inter-model spreads of CO_2 emissions, however, were attributed mainly to differences in land carbon cycle process parameterizations. As a result, different land cover distributions among the models gave significantly different regional changes in climate associated with land use change, but with insignificant effects on global mean temperature. Furthermore, the range of net cumulative land use change emissions from 2006 to 2100 for RCP8.5 was 34 to 205 PgC, with the high estimate likely due to the combination of relatively high levels of land carbon and the inclusion of all land use transitions rather than just net land use change (Brovkin et al., 2013). Additionally, not all of the models used the GLM wood harvest data, further contributing to the spread of model results. For comparison, estimates of net cumulative carbon emissions during 1700-2000 (1850-2000) range from 138-250 PgC (110-210 PgC) (Table 3 in Smith and Rothwell, 2013). The differences in land use and land cover implementations are also a main factor in the large spread of 21st century land carbon uptake and of compatible fossil fuel emissions allowable for a given RCP. In fact, the inter-model spreads in land carbon uptake for individual scenarios are greater than the inter-scenario spreads for individual models (Jones et al., 2013b). It is apparent that further work is needed to resolve inconsistencies among land use and land cover approaches to reduce climate uncertainty, especially for regional impact assessment.

Additional sources of climate uncertainty related to land use are the RCP radiative forcing targets, which include only emissions of GreenHouse Gases (GHGs) and some aerosols and reactive gases (van Vuuren et al., 2011a). These targets do not include radiative forcing from albedo change or other direct climate effects associated with land use change. In a recent modeling experiment, two different carbon tax policies with dramatically different land use scenarios met the same radiative forcing target (4.5 W m^{-2}) in the IAM used for RCP4.5 but had significantly different radiative forcing in an ESM (difference of 1 W m^{-2}) due to albedo differences between the land use scenarios (Jones et al., 2013a). Likewise, the Shared Socioeconomic Pathways (SSPs) for mitigation, adaptation, and impact studies in the Intergovernmental Panel on Climate Change (IPCC) fifth Assessment Report (AR5) are likely to produce different land use scenarios that meet the same RCP target, but have different radiative forcing in the ESMs due to the direct effects of land use and land cover change on climate. However, one of the goals of the RCP process was to provide a set of radiative forcing targets for ESMs that remains consistent with respect to the diversity of SSPs associated with each RCP target (Moss, et al., 2010). As a result of the wide range of land use and land cover related uncertainties in climate projections, an increased emphasis on land use and land cover dynamics is a high priority for CMIP6 (Meehl et al., 2014).

A more consistent and complete land use and land cover coupling between IAMs and ESMs will facilitate more accurate projections of global change scenarios and more robust multi-model intercomparisons of climate and carbon cycle interactions with anthropogenic drivers such as fossil fuel emissions and land use change. These expected outcomes are in line with a primary goal of a scenario-based approach, such as the RCPs, which is “to better understand uncertainties in order to reach decisions that are robust under a wide range of possible futures” (Moss et al., 2010; p. 747). The RCPs were designed to better understand uncertainties in global climate projections by providing distinct scenarios of atmospheric radiative forcing and land use change. Intra-scenario comparison of ESM simulations offers insights to uncertainties in ESM processes, while inter-scenario comparison of ESM simulations offers insights to uncertainties due to a range of possible futures. However, the efficacy of this approach depends on the fidelity of the ESM simulations to the RCP scenarios. Without this fidelity, intra-scenario comparison is not possible, because the ESMs are not simulating the same scenario, and inter-scenario comparison might include futures outside the prescribed range of possibility.

The IAMs projected a complete terrestrial surface (along with ice, rock, and urban) for each given scenario because land use and land cover are interdependent. For example, carbon stocks in various ecosystems might be valued under a carbon price policy, so land cover would need to be determined along with land use. Or a land policy might restrict certain land cover conversions. Within the CMIP5 coupling process, however, GCMs and ESMs determine their own land cover while remaining consistent with the land use harmonization data, thus potentially reducing the fidelity of the full climate simulations to the RCP scenarios. This was a practical design that obviated the redesign of GCM/ESM land use and land cover implementations, but also precluded analysis of the climate impacts of different land cover responses to land use change because such analysis is robust only within a single model where everything but land cover response remains consistent. Another challenge posed by the interdependence of land use and land cover is the implementation of geographic shifts in land cover due to bioclimatic changes. While these shifts are often implemented within ESMs, such shifts are a second-order effect that is superposed upon land use change and might be better implemented as a feedback from ESMs to IAMs to inform land use and land cover projection. Incorporating both land use and land cover into the coupling between IAMs and ESMs is a fundamental step toward realizing the full potential of the scenario-based RCP process.

Our approach to addressing inconsistencies between IAMs and ESMs is to integrate an IAM and an ESM into the first fully coupled model that directly simulates human-environment feedbacks. The resulting integrated ESM (iESM) includes climate feedbacks on vegetation productivity and ecosystem carbon from the Community ESM (CESM) to the Global Change Assessment Model (GCAM) to facilitate land use projection at five-year intervals. The iESM uses GLM as in the CMIP5 land use harmonization, along with the CESM Land Use Translator (LUT) that converts land use harmonization outputs to CESM land cover and wood harvest area. Our initial iESM simulations showed that time varying factors based on CESM simulated Net Primary Production (NPP) and Heterotrophic Respiration (HR) were successfully used by GCAM for land use projection. However, these simulations also demonstrated that the large RCP4.5 afforestation signal was not being passed through from GCAM to CESM. GCAM simulated afforestation as a carbon-sequestering strategy to help meet the RCP4.5 target, but this additional forest area was not included in the land use harmonization. As a result, most of this forest area was not included in CESM simulations, both for CMIP5 and in an early version of iESM.

Here we test the feasibility of restoring the lost afforestation signal by using the iESM as a test bed to explore alternative coupling strategies. We focus on modifications to the CESM LUT because initial modifications to GLM did not restore CESM afforestation. One advantage of focusing on a post-land use harmonization approach is that it could be applied to other ESMs independently without changing the land use harmonization product. Section 2 includes model description and experimental design, Section 3 presents results and demonstrates that this problem exists in CMIP5, and Section 4 discusses the limitations of our current approach and the implications for the CMIP5 archive with respect to land use and climate. We conclude with suggestions for improving IAM to ESM land coupling for future model inter-comparisons.

2. Methods

2.1. iESM Description

The iESM integrates GCAM, GLM, and CESM to evaluate the effects of human-environment feedbacks on the earth system (Figure 1). We have completed the first coupling stage that allows GCAM to project land use distribution in five-year increments based on the previous five years of CESM vegetation productivity. Here we give an overview of how the three main components interact. A more detailed description of iESM development will be presented in a forthcoming paper (Collins et al., in prep).

GCAM v3.0 ((Calvin et al., 2011); henceforth referred to as GCAM) is a tightly coupled IAM of human and biogeophysical processes associated with climate change. GCAM's human system components simulate global economic activity within energy, agriculture, and forest product markets with respect to 14 geopolitical regions. A previous version of GCAM projected land use and land cover distributions for each of the 14 geopolitical regions (Wise et al., 2009) and was used to generate the CMIP5 RCP4.5 scenario (Thomson et al., 2011). Currently, GCAM incorporates a range of improvements to the Agriculture and Land Use (AgLU) module, including the capacity to operate on 151 geographical land units to generate a more detailed and accurate spatial distribution of land use. There are three land cover types that remain constant over time (urban, tundra, and rock/ice/desert) and 24 land use and land cover types available for redistribution, including 12 food and feed crops, five bioenergy crops, and seven managed and unmanaged ecosystems (Kyle et al., 2011; Wise and Calvin, 2011). The "geographical land units" are defined by intersecting 18 global agro-ecological zones (Lee et al., 2005) with the 14 geopolitical regions. In the iESM, GCAM projects land use and land cover distributions within each of these land units at five-year intervals. These distributions are based on profit shares calculated from agricultural costs, prices, yields, and the application of a carbon price to vegetation and soil carbon densities.

In a second and intermediate step, GLM uses GCAM's cropland, pasture, and forest areas (and wood carbon harvest) to compute all annual, fractional land use states and transitions. As part of this process it disaggregates GCAM's geographical land unit data to a half-degree global grid by computing spatial patterns and also ensures consistency with the historical land use reconstructions (Hurtt et al., 2011; Hurtt et al., 2006). GLM has been slightly modified from its CMIP5 implementation to better facilitate forest area change matching with GCAM (Section 2.3.2). This modification enables GLM to use forest area output from GCAM that was not incorporated into the CMIP5 land use harmonization. Nonetheless, iESM still follows the CMIP5 implementation for CESM in using these GLM land use harmonization outputs: cropland, pasture, primary, and secondary land area, as well as wood harvest areas on primary and secondary forested and non-forested land.

CESM (Bitz et al., 2011; Gent et al., 2011) has fully coupled atmosphere, ocean, land, and sea ice components. Within CESM, the Community Land Model v4.0 (CLM; Lawrence et al., 2011) receives the selected GLM outputs via a translator that converts these outputs to 16

CLM Plant Functional Types (PFTs; eight forest, three grass, three shrub, one bare soil, and one crop) (Lawrence et al., 2012). The CLM dynamic vegetation module, which estimates bioclimate-driven geographical shifts in CLM PFTs, cannot run at the same time as the land use change module presented here; only one of these modules can change CLM PFT areas per simulation. While the iESM does not directly estimate bioclimatic shifts in land cover, the NPP and HR feedbacks to GCAM do incorporate bioclimatic effects on ecosystems into GCAM's land use cover projections. The version of iESM used in this study was based on CESM v1.0beta9, which is a pre-release version of the model used for the CMIP5 simulations.

The iESM climate feedbacks on vegetation and carbon were implemented by passing annual climate scaling factors from CESM to GCAM based on NPP and HR. These factors were used to scale GCAM crop yields and vegetation and soil carbon densities every five years. To calculate the scaling factors, the per-pixel, PFT-specific CESM 5-year annual average NPP and HR values for a given GCAM time step were divided by base-period average annual values (1990-2004). These NPP and HR ratios were then filtered to exclude outliers based on a median absolute deviation method, and finally aggregated to GCAM's geographical land units and land use and land cover types (for details see Bond-Lamberty et al., in review). Crop yields and vegetation carbon densities for GCAM's next land use projection were scaled by the NPP ratio, while soil carbon densities were scaled by a combination of the NPP and HR ratios $((NPP_{ratio} + (1 - (HR_{ratio} - 1))) / 2)$.

2.2. Simulations

Our iESM simulations cover 2005 to 2040 with fully coupled CESM components and prescribed RCP4.5 emissions and carbon price path. These simulations use the land use change module, a dynamic ocean (Smith et al., 2013b), Community Atmosphere Model v4 physics (Gent et al., 2011), carbon-nitrogen biogeochemistry (Thornton et al., 2007), and active land-atmosphere-ocean carbon dynamics, at approximately 1° resolution (0.9375°x1.25°). The iESM initial conditions are the culmination of a CESM spinup run followed by a CESM 1850-2005 transient historical run with land use change. GCAM initial conditions are calibrated to 2005 wood harvest, land use area, and energy and agriculture costs and production, as reported by individual countries and processed and archived by international organizations (e.g. FAO, International Energy Agency). The GCAM RCP4.5 scenario was described fully by Thomson et al. (2011).

We performed two fully integrated simulations to compare two iESM cases: 1) original CESM land use translator (OLDLUT) and 2) modified CESM land use translator (NEWLUT) (Table 1). In fact, OLDLUT was our initial fully integrated simulation with iESM and, as reported below, it revealed inconsistencies within iESM that needed to be addressed prior to scientific experimentation. OLDLUT also showed that the updated GLM did not increase CESM afforestation with respect to a previous simulation performed by manually passing data between the respective iESM models. The NEWLUT case was used to test our hypothesis that the lost afforestation signal could be recovered by modifying only the CESM component of iESM. These fully integrated runs included climate feedbacks on vegetation productivity and ecosystem carbon in GCAM's land use projections, which occurred at five-year intervals. Analysis of the effects of introducing these feedbacks on land use, carbon, and climate will be presented in a forthcoming paper (Thornton et al, in prep).

2.3. Land use coupling

2.3.1. OLDLUT land use coupling within iESM

The OLDLUT iESM land use coupling followed the CMIP5 land use harmonization algorithm (Figure 2), but with a slightly modified version of GLM (see Section 2.3.2). The coupling was designed to match GCAM and CESM changes in absolute cropland and pasture area. For CMIP5, GLM received only crop and pasture areas from GCAM, but for the iESM GLM also receives forest area from GCAM to better facilitate forest area change matching (see Section 2.3.2). GLM also receives wood products demand from GCAM (in tons of carbon), which is spatially distributed to determine the extent of harvested area in each of five wood harvest types (primary forest harvest, primary non-forest harvest, secondary mature forest harvest, secondary immature forest harvest, and secondary non-forest harvest). The OLDLUT (Figure 3) uses only the cropland and pasture area outputs from GLM to update CESM PFT areas in conjunction with maps of potential vegetation (the vegetion most likely to be present if no land use change had occurred; Ramankutty and Foley, 1999). Non-crop PFT area reductions are made in proportion to their respective existing grid-cell fractions, while additions are made in proportion to their respective potential vegetation grid cell fractions. The OLDLUT does not use the primary and secondary land area information for updating PFT areas because CESM does not keep track of these land use designations. The OLDLUT does, however, use the primary and secondary land area to calculate the harvested fraction of GLM harvestable area (sum of the five wood harvest type areas divided by the total area of primary and secondary land). Wood is harvested from only forest in CESM, and so the GLM harvested fraction is applied to forest area to determine the harvested area in CESM (Lawrence et al., 2012).

The OLDLUT makes specific assumptions about pasture area change because CESM does not keep track of pasture area (Figure 3). Changes in GLM cropland result directly in CESM changes in crop PFT area, but changes in pasture area are constrained by forest PFT area and reflected in changes in grass and shrub PFT area. More specifically, pasture addition is limited to replacement of existing forest PFT area with grass PFT area, and pasture removal is limited to the replacement of grass and shrub PFT area by potential forest PFT area. This means that grass and shrub PFT area changes associated with pasture area change can be only as large as the available existing or potential forest area.

2.3.2. Modifying the GLM spatial distribution algorithm

For the iESM, GLM was modified to better facilitate forest area change matching with GCAM in an effort to increase the forest area simulated by CESM. These modifications included operating on GCAM's 151 geographical land units (rather than the 14 regions used for CMIP5) in addition to using GCAM's forest area output, which was not previously shared between the models. For CMIP5, GLM applied the cropland and pasture area changes to the 2005 half-degree map of cropland and pasture while preserving the total cropland and pasture area changes within GCAM regions. Spatial allocation of cropland and pasture areas to the half-degree grids was done with a preference for expanding agricultural area onto non-forested land and reducing agricultural area where GLM would expect a forest to grow, while also preserving 2005 spatial patterns of land use by allocating new cropland and pasture near to existing agricultural areas (Hurtt et al., 2011).

The new GLM algorithm uses GCAM forest area from each geographical land unit at each time step and attempts to preserve the forest area changes within each geographical land unit in addition to preserving the cropland and pasture area changes. GLM has previously defined "forest" as natural vegetation that is growing on land where the potential biomass density, based on an internal potential vegetation growth model, is greater than 2 kgC m^{-2}. Using this definition the potential forestland within GLM is fixed and, as a result, the GLM algorithm cannot grow forest outside of this forestland. In the new algorithm, GLM matches GCAM forest area changes by moving cropland and pasture around within each geographical land unit to "expose" enough potential forestland for regrowth to meet the GCAM forest area changes (see the following steps a-c). In addition, to meet GCAM's land requirements for afforestation, GLM uses a different definition of "forest" (potential biomass density greater than 1 kgC m^{-2}, rather than 2 kgC m^{-2}) than the definition used elsewhere in the GLM code (e.g. for computing the spatial pattern of wood harvesting). The new GLM algorithm operates in three main steps:

a) Decreases in cropland and pasture occur first on the highest potential biomass land and increases in cropland and pasture occur first on the lowest potential biomass land.

b) If the forest area change within a geographical land unit is not met, a redistribution of cropland and pasture within that geographical land unit occurs such that, when possible, existing cropland and pasture is moved from high biomass density land to low biomass density land.

c) If the forest area change within a geographical land unit is still not met, the algorithm attempts to allocate any "unmet" forest area change within another land unit (or across multiple land units) within the same region, using a similar method to (b) above.

2.3.3. Modifying the CESM land use translation algorithm

To test our hypothesis that the lost afforestation signal could be recovered solely by the ESM component, we focused on modifying the LUT (NEWLUT; Figure 4) to capture GCAM afforestation via changes in agricultural land. This approach is more expedient than redesigning the coupling code and LUT to receive forest area changes directly from GLM because such redesign would logically require implementation of a single, consistent land surface and carbon cycle among all iESM components. Specifically, the NEWLUT adds tree PFTs when cropland and pasture are removed. Furthermore, the NEWLUT preferentially removes tree PFTs when cropland and pasture are added. Forest area information is still not shared between GLM and the NEWLUT (other than forest harvest). The NEWLUT also includes proper grid cell fraction matching between GLM and CESM, which primarily affects crop, grass, and shrub PFTs.

2.3.3. CMIP5 RCP4.5 land use and land cover distributions among GCAM, GLM, and CESM

The OLDLUT iESM land use coupling was also used in CMIP5, albeit with 14 regions rather than 151 geographical land units and without the GLM modifications and climate feedbacks described above, and so we explored the extent to which the afforestation signal was lost in the CMIP5 simulations. We compared the RCP4.5 pre-land use harmonization forest and pasture area outputs from GCAM with the GLM land use harmonization values and also with the corresponding PFT area inputs for the CESM1.0-BGC simulations submitted to the CMIP5 archive. CESM1.0-BGC served as the base code for iESM and thus contains the same versions of the model components.

3. Results

3.1. CMIP5 RCP4.5 land use and land cover area inconsistencies

The GCAM afforestation signal was dramatically decreased in the CESM simulations, and the total area covered by CESM herbaceous (grass and shrub) PFTs increased while GCAM pasture decreased (Figure 5). CESM forest area increased by 23% of the 4.82 million km^2 of afforestation between 2005 and 2020, and by 22% of the 10.98 million km^2 of afforestation, by 2100. GLM captured 64% and 56% of the afforestation in 2020 and 2100, respectively. GCAM and GLM pasture decreased by 4.69 million km^2 from 2005 to 2100 while CESM herbaceous PFTs increased by 1.11 million km^2 over the same period. The changes in global cropland area were faithfully transmitted (CESM decreases were only 7% less than GCAM decreases), but absolute CESM cropland area was approximately 1.5 million km^2 less than GCAM cropland area throughout the simulation (data not shown). Changes in GLM pasture and cropland areas were essentially identical to GCAM changes, and GLM absolute area values were slightly higher and lower, respectively, than GCAM pasture and cropland areas (cropland data not shown).

3.2. Restored afforestation in iESM

The OLDLUT simulation revealed that only changes in crop area were being faithfully transmitted from GCAM to CESM (Figure 6; changes in global area). In contrast, CESM forest area increased by only 17% of GCAM's 5.40 million km^2 of afforestation between 2015 and 2020, and by only 17% of the 7.73 million km^2 of afforestation between 2015 and 2040. Changes in GLM forest area, on the other hand, reflected changes in GCAM forest area quite well (Figure 6), but at the cost of dramatically overestimating absolute forest area within GLM due to a low biomass threshold for defining forest (Figure 7; absolute values of global area). Within GLM, the new algorithm captured 93% of afforestation between 2015 and 2020 and 84% between 2015 and 2040, as compared to the original GLM algorithm that captured only 14% and 20% over the respective periods in a previous simulation performed by manually passing data between the respective iESM models (data not shown). Changes in GCAM pasture were not reflected by changes in CESM herbaceous PFTs, but were faithfully output by GLM (Figure 6).

The NEWLUT simulation shows improved forest and cropland area changes in CESM with a corresponding change in CESM herbaceous PFT area. The main improvement is that CESM forest area increases by 64% of GCAM's 2015-2020 afforestation and by 66% of the 7.71 million km^2 of afforestation from 2015-2040 (Figure 6). This additional forest area in NEWLUT reduces total area covered by CESM herbaceous PFTs by 94% of the 4.36 km^2 of GCAM pasture loss by 2040. Figure 8 shows the spatial tradeoff between forest and herbaceous PFTs that achieves this level of afforestation, and Figure 9 demonstrates a sustained increase in average annual land carbon uptake after 2020 due to additional afforestation. In comparison to OLDLUT, the NEWLUT increase in land carbon uptake results in a 19 PgC increase in vegetation carbon gain and an 8 ppmv decrease in atmospheric CO_2 gain between 2005 to 2040 (Figure 10). NEWLUT also improves the CESM absolute cropland area (Figure 7) through proper matching of GLM and CESM grid cell fractions. The effect of this proper matching is apparent in the cropland and pasture area changes from 2005 to 2006 (Figures 6 and 7). GLM NEWLUT outputs follow the GCAM NEWLUT outputs with relationships between GLM and GCAM similar to those for OLDLUT (data not shown).

4. Discussion

The iESM and CMIP5 land cover area discrepancies (Figures 5-7) result from a gap in the original CMIP5 land coupling design that allows inconsistent forest area and land cover type definitions across models (Figure 2), along with different underlying carbon cycles. The land use harmonization was, however, ambitious and largely successful in developing consistent land use definitions and data without requiring extensive redevelopment of land use and land cover components of all participant models (Hurtt et al., 2011). As our study attests, such redevelopment is challenging and model-specific, but might be required for ESMs to adequately simulate the IAM-prescribed anthropogenic drivers and their corresponding effects on carbon and climate. Thus, while this is a specific case, the lost iESM afforestation signal is instructive of the shortcomings of the CMIP5 design and the restoration of this signal offers insights into improving land use and land cover coupling for model inter-comparisons.

A primary challenge for improving the CMIP5 land coupling is to increase the amount of specific land cover information being shared between IAM (and historical) scenarios and ESMs. For CMIP5, the land use harmonization was designed to harmonize land use data between models, and as such GLM did not receive forest area or any other land cover information from any of the IAMs (Masui et al., 2011; Riahi et al., 2011; Thomson et al., 2011; van Vuuren et al., 2011b). Thus, at the first coupling step, scenario-prescribed land cover associated with any IAM policy that valued carbon within unmanaged ecosystems (e.g., grassland, wetland, forest) was lost. While GLM does, however, keep track internally of forested and non-forested land (according to its own definition of forest, which likely differs from those within IAMs and ESMs), the output land use harmonization product includes only cropland, pasture, primary, and secondary land areas and transitions, and the age and biomass density of secondary land (and harvest areas, carbon amounts, and transitions, which we do not address here). As each ESM characterizes the land surface by its own suite of vegetation and management types (Brovkin et al., 2013), additional land use and land cover information could be lost in the second coupling step between GLM and the ESMs. For example, some ESMs were able to use the primary, secondary, and transition information, but they might have been applying this information to different land covers than those used by GLM, thus introducing a second shift away from the original IAM scenario. Our specific case demonstrates an even greater inconsistency due to the use of only cropland and pasture information. GCAM has 19 crop types (the CMIP5 version had 10) and seven managed and unmanaged land cover types while CESM has 16 PFTs, only one of which is a crop type. The LUT algorithm uses only the GLM cropland and pasture area information to adjust PFTs because CLM does not keep track of primary versus secondary land. The resulting spatial pattern of non-crop PFTs is determined by the existing PFT distribution and CESM's internal representation of potential vegetation cover (Lawrence et al., 2012; Ramankutty and Foley, 1999). An additional source of error that we did not investigate here is the relationship between individual PFTs and land cover types that may comprise several PFTs (e.g. forest land may consist of 60% trees and 40% grass).

Due to the lack of a prescribed land cover input associated with the land use input, forest area changes in CESM (and iESM) are effectively residual changes that are only indirectly linked to GCAM forest area through changes in cropland and pasture areas. The LUT calculates cropland area changes first and pasture area changes second (Figures 3 and 4). In CMIP5 CESM simulations, cropland area changes cause non-crop PFTs to be added or removed in proportion to their potential or existing grid-cell fractions, respectively. Pasture is more complicated because it is not tracked as such: pasture is not a single PFT and its changes are represented as changes in herbaceous and tree PFTs. Specifically, tree PFTs are removed when pasture is added, and non-crop PFTs are added in proportion to their potential vegetation grid-cell fractions when pasture is removed (Lawrence et al., 2012). This residual PFT determination, combined with independent and unique forest definitions across GCAM, GLM, and CESM, causes the bulk of prescribed afforestation to not appear in the CESM land surface. As a direct consequence, CESM grass area (and shrub area to a lesser extent) increases while GCAM pasture decreases dramatically (Figure 5). CESM has this same limitation for all four RCP scenarios, and the other CMIP5 ESMs implement similar inconsistencies to varying degrees due to the lack of specific vegetation types in the land coupling between IAMs and ESMs. For example, Davies-Barnard et al. (2014) recently reported that the HadGEM2-ES RCP4.5 forest area increased 11% from 2005-2100, while the GCAM forest area increased by 24%. Additionally, the GCAM 2005 forest area was 41.1 Mkm^2, the GLM 2005 forest area was 39.9 km^2, but the MPI-ESM 2005 forest area was about 24 M km^2. As a result, the 35% increase in MPI-ESM RCP4.5 forest area by 2100 (Wilkenskjeld et al., in review) was still only 77% of GCAM's afforestation. It is apparent from these inconsistencies that interdependent land use and land cover need to be faithfully transmitted from IAMs to ESMs to robustly simulate the effects of prescribed scenarios on the earth system.

Even partial restoration of the lost afforestation signal in iESM demonstrates the potentially dramatic effect on global carbon and climate of using IAM land cover and land use information in ESMs. As soon as 25 years after the initial increase in forest area, and with only 64% of GCAM's afforestation area, the NEWLUT has a significant impact on global carbon balance (Figure 9). The assumption that forest exclusively replaces abandoned cropland and pasture in GCAM's land use projection (Figures 6-8) sets the upper limit for CESM because there is no other information to constrain forest area, and may be applicable only to the RCP4.5 scenario. Although this limits NEWLUT to including only two-thirds of the total afforestation, adding more forest area to CESM would be arbitrary without additional land cover information. Nonetheless, the increased afforestation in NEWLUT results in an increase in net land carbon uptake over the OLDLUT case due to a sustained increase in average annual land carbon uptake after 2020 (Figure 9). As a result, the NEWLUT simulation increases vegetation carbon gain by 19 PgC and decreases atmospheric CO_2 gain by 7.7 ppmv from 2005 to 2040 in comparison to OLDLUT (Figure 10). The NEWLUT simulation also decreases soil carbon gain by about 1.5 PgC over this period (data not shown).

Simple linear extrapolation of the iESM vegetation carbon gain and atmospheric CO_2 gain from 2005 to 2100 increases these changes to approximately 52 PgC and 21 ppmv, and extending CESM forest area to match GCAM total afforestation could potentially increase these changes to 88 PgC and 36 ppmv in 2100. These are rough estimates that use 2005 as a starting point to reduce the high slope associated with the initial increase from 2015-2020, and also assume that additional forest area continues to gain carbon for 60-80 years after it is established. Regardless of the absolute accuracy of these extrapolations, the potential gain in vegetation carbon alone for CESM with full afforestation is on the order of estimates of net cumulative land use change emissions during 1850-2000, which range from 110-210 PgC (Table 3 in Smith and Rothwell, 2013). For comparison, the range of CMIP5 vegetation carbon stock gains for RCP4.5 is about 50 to 300 PgC from 2005 to 2100, with most gains being less than 150 PgC and relatively linear (Figure 2 in Jones et al., 2013b). An increase in gain of 88 PgC would dramatically shift CESM vegetation carbon dynamics in relation to the other ESMs. The corresponding 36 ppmv decrease in atmospheric CO_2 is nearly one-third of the difference between the prescribed 2100 concentrations of the RCP4.5 (~540 ppmv) and RCP2.6 (~420 ppmv) scenarios (Figure 1 in Jones et al., 2013b). More importantly for CESM's ability to robustly simulate the effects of the RCP scenarios on the earth system, the prognostic CESM atmospheric CO_2 concentration in 2100 for RCP4.5 is 610 ppmv (Keppel-Aleks et al., 2013), and a decrease from 610 to 574 ppmv has an approximate decrease in radiative forcing of 0.33 W m^{-2}, which is non-trivial with respect to the 4.5 W m^{-2} target. While these carbon cycle changes in the CESM component of iESM may have a significant effect on climate, it is important to note that the carbon cycle effects of afforestation in CESM are not identical to those in GCAM or GLM because these three models have different biogeochemistry and vegetation models. These differences in carbon cycles, however, do not obviate the need for making both land cover and land use consistent between IAMs and ESMs in order to best match the prescribed radiative forcing scenario.

Different implementations of land cover and land use among IAMs and ESMs also reduce the fidelity between RCP scenarios and their associated effects on the earth system. Figure 8 shows that most of the additional forest area in NEWLUT occurs on grassland and shrubland, and that these lands generally coincide with areas of limited potential forest. The OLDLUT could not add forest area where no potential forest area exists, and the rate of forest carbon accumulation is constrained by environmental conditions. GLM also limits forest area and growth based on potential forest and environmental conditions, but with a different growth model and map of potential forest area than used by CESM. On the other hand, GCAM afforestation is a strategy to expand forest area for carbon sequestration, and assumes that it is cost effective to use agricultural inputs (e.g., water, fertilizer) to achieve the expected forest growth. This disagreement among the three models hampers communication of forest area changes and contributes to the differences in forest area among the models, both in CMIP5 (Figure 5) and in the iESM (Figures 6 and 7). Nonetheless, sharing forest area between GCAM and GLM does improve the fidelity between GCAM and GLM's forest area changes (Figures 5 and 6). GLM and CESM do not simulate agricultural inputs for forests, but the NEWLUT can simulate most, but not all, of the prescribed afforestation (Figures 6 and 7) by adding forest area based on GCAM's cropland and pasture changes, rather than on potential forest area. The additional forest might not grow as well in CESM as in GCAM, but the CESM forest productivity is fed back to GCAM for subsequent land use projections, so environmental restrictions on forest growth will influence future land use and land cover. This feedback does not, however, fully compensate for the lack of bioclimatic or agricultural input availability constraints on GCAM's land use projection, which might contribute to an overly optimistic afforestation projection. More generally, this feedback mechanism opens a path for more robustly simulating interdependent land use and land cover through incorporation of potential, bioclimate-driven geographic shifts in land cover. ESMs could estimate bioclimatic drivers or geographic shifts for given land use/cover scenarios, and then feed this information back to the IAMS for incorporation into land use/cover projection. Implementing such a feedback for scenario-based simulations would consolidate land use/cover determination into internally consistent modules within the IAMs, thereby increasing fidelity between the scenario-prescribed land surface and the one used by the ESMs.

We have focused on understanding the effects of mismatched land cover areas on global simulations, rather than on mismatched carbon cycles, because the spatial distribution of land cover and land use is a scenario-determined boundary condition for ecosystem-specific processes such as biogeochemical dynamics. For global simulations this boundary condition is generally provided by historical data and IAMs, and, as we have shown, a mismatch in this boundary condition causes CESM to simulate non-scenario effects on carbon and climate (due to a non-scenario land surface), rather than the scenario-driven effects of the land surface prescribed for meeting the RCP4.5 target. Mismatched carbon cycles among IAMs and ESMs, on the other hand, along with differences in atmospheric radiation code, will preclude exact matches in radiative forcing for a given RCP scenario, but should not cause significant deviations among models in the carbon and climate effects of a given scenario. While we plan to completely reconcile land use and land cover inconsistencies within the iESM by implementing a single carbon cycle with consistent land surface characterization among the components, it is not desirable, nor feasible, for all IAMs and ESMs to have the same biogeochemistry and vegetation growth components. For example, a diversity of terrestrial models can help characterize uncertainty in global simulations. This uncertainty, however, is most useful if these models simulate the same spatial distribution of land cover and land use change. Therefore, iESM redevelopment that ensures land use and land cover consistency between GCAM and CESM could provide a template for improving the fidelity between IAM scenarios and ESM simulations in the next CMIP. In fact, land cover information is currently planned to be included in the CMIP6 land coupling, along with a more extensive land use model intercomparison project (Meehl et al., 2014).

5. Conclusion

We have identified the lack of specific land cover type information being shared among GCAM, GLM, and CESM in the iESM as the primary cause of CESM having very little afforestation and effectively no change in herbaceous PFT area in contrast to GCAM's large RCP4.5 afforestation and corresponding pasture reduction. Initial efforts to fix this problem through GLM modifications and the sharing of forest area between GCAM and GLM improved only the fidelity of forest area changes between GCAM and GLM. We then focused on modifying the algorithm that translates GLM land use harmonization outputs to CESM PFTs. While these land use translator modifications have been successful at capturing two-thirds of GCAM's RCP4.5 afforestation signal and corresponding reductions in herbaceous PFT area, they are not sufficient to completely overcome the limitations imposed by not passing specific land cover types from GCAM through to CESM. These modifications are also specific to the GCAM RCP4.5 scenario, and might need to be altered for the other RCP scenarios. Furthermore, we have not addressed the lack of constraints on GCAM forest area expansion, nor mismatches between land cover and PFT definitions. Nonetheless, this partial restoration of afforestation has a significant impact on iESM's global carbon cycle through increased vegetation carbon and decreased atmospheric CO_2 concentration.

The iESM framework follows the CMIP5 land coupling design, and as such we have characterized a major gap in this design that precludes accurate translation of projected IAM land surface scenarios to ESMs by focusing only on land use such as cropland and pasture (albeit successfully), and not including specific land cover types such as forest, grassland, and shrubland. The relationship between land use and land cover is handled uniquely by individual ESMs, which means that the effects of scenario mismatch will be model-specific and more relevant for some RCPs than others. The resulting land cover discrepancies are likely most pronounced for the large RCP4.5 afforestation signal, which was greatly reduced in the CMIP5 CESM and HadGEM2-ES (see Davies-Barnard et al., 2014) simulations, but could also arise for other large land cover changes such as the extensive deforestation of RCP8.5. As total land area is conservative, errors in the distribution of one land cover are complemented by errors in the distributions of other land covers. In GCAM's RCP4.5 scenario, pasture decreases over the 21st century, but the CMIP5 CESM runs have increasing grass and shrub areas over the same period. It is very important that the land use and land cover changes (which determine land use change emissions and the total capacity for vegetation carbon assimilation) match between the IAMs and ESMs because the CMIP5 experimental design is predicated on the fidelity between IAM scenarios and ESM simulations such that they have similar, specific radiative forcings for a given scenario, including CO_2 emissions from land use change (Moss et al., 2010). Furthermore, future radiative climate targets are likely to include the biogeophysical forcings of land use change because it has been shown that the modeled climate system is sensitive to changes in these forcings due to the spatial distribution of land use and land cover change (Brovkin et al., 2013; Jones et al., 2013a; Pitman et al., 2009), making it imperative that IAM and ESM land use and land cover distributions match as closely as possible. Maintaining the diversity of global biogeochemical and vegetation models also calls for GCMs and ESMs to match historical and projected land cover and land use distributions as closely as possible, so as to isolate carbon cycle contributions to uncertainty from contributions due to differences in land use and land cover. Fortunately, our results indicate that it might be possible to adjust land cover in other CMIP5 models to better match RCP4.5 afforestation and the corresponding climate scenario, while still using the standard land use harmonization data.

We conclude that the land coupling between IAMs and ESMs for future model intercomparisons needs to ensure greater consistency in land cover and land use among the models in order to realize the full potential of scenario-based earth system simulations. In short, the models need to agree on the actual land area and the annual spatial distribution of major (non-) vegetation land covers and land uses. In other words, the ESMs need to simulate the same basic land surface as prescribed by the IAM-generated RCP scenarios. To achieve the required consistency, we suggest that the next CMIP land coupling design provides land cover and land use information, and a standard mapping between land cover and plant functional types. Fortunately, this is an emerging priority for the CMIP6 Land Use Model Intercomparison Project (LUMIP, http://www.wcrp-climate.org/index.php/modelling-wgcm-mip-catalogue/modelling-wgcm-mips/318-modelling-wgcm-catalogue-lumip , http://www.wcrp-climate.org/wgcm/WGCM17/LUMIP_proposal_v4.pdf). The following gridded data with fractional shares within grid cells are specifically recommended:

1) Annual land cover states with complete, contiguous spatial coverage within grid cells. Land cover needs to include at least the basic categories of cropland, grassland, shrubland, woodland, forest, and other (bare/sparse, ice, urban, water). This will allow consistency in major (non-) vegetation types for model intercomparison (with the "other" category having fixed area). The "other" categories could also be separated out for models that can use them, and in preparation for changing their areas also.
2) Annual land use states including primary and secondary land, wood harvest, and pasture (cropland should coincide with the land cover state). These uses should be provided with respect to the land cover categories. Wood harvest and pasture should include both area and amount of biomass/carbon harvested or removed by grazing.
3) A standard present-day land area data set to be used by all models. Land area includes all land cover and land use categories as described above.

4) Annual land use and land cover transitions. Land use transitions need to be accompanied by corresponding land cover transitions with complete, contiguous spatial coverage within grid cells. Net land use/cover transitions, which should be used for model intercomparison, are annual changes in individual land use and cover states, and may include additional detail about sources of wood harvest and grazed biomass. Gross land use/cover transitions are the transitions among particular land use/covers occurring within a particular year. These transitions sum to the net land use/cover transitions, and should also be provided to characterize shifting cultivation and other gross land conversions. While gross land use/cover transitions are very important and make a significant difference in the carbon cycle, until more models are able to make use of gross transitions they should not be included in model intercomparisons.

Acknowledgements

We are extremely grateful to Peter Lawrence (National Center for Atmospheric Research) for providing the original CESM land use translation code. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-AC02-05CH11231 as part of the Integrated Assessment Research and Earth System Modeling Programs. This project used resources of the National Energy Research Scientific Computing Center (NERSC), which is a DOE Office of Science User Facility. The CESM project is supported by the National Science Foundation and the Office of Science (Biological and Environmental Research) of the U.S. Department of Energy. We also gratefully acknowledge the support of the National Aeronautics and Space Administration.

References

Adegoke, J. O., Pielke Sr, R., and Carleton, A. M.: Observational and modeling studies of the impacts of agriculture-related land use change on planetary boundary layer processes in the central u.S, Agricultural and Forest Meteorology, 142, 203-215, http://dx.doi.org/10.1016/j.agrformet.2006.07.013, 2007.

Anav, A., Friedlingstein, P., Kidston, M., Bopp, L., Ciais, P., Cox, P., Jones, C., Jung, M., Myneni, R., and Zhu, Z.: Evaluating the land and ocean components of the global carbon cycle in the cmip5 earth system models, J. Clim., 26, 6801-6843, 10.1175/jcli-d-12-00417.1, 2013.

Arora, V. K., and Boer, G. J.: Uncertainties in the 20th century carbon budget associated with land use change, Global Change Biol., 16, 3327-3348, 10.1111/j.1365-2486.2010.02202.x, 2010.

Baidya Roy, S., Hurtt, G. C., Weaver, C. P., and Pacala, S. W.: Impact of historical land cover change on the july climate of the united states, Journal of Geophysical Research: Atmospheres, 108, 4793, 10.1029/2003jd003565, 2003.

Bitz, C. M., Shell, K. M., Gent, P. R., Bailey, D. A., Danabasoglu, G., Armour, K. C., Holland, M. M., and Kiehl, J. T.: Climate sensitivity of the community climate system model, version 4, J. Clim., 25, 3053-3070, 10.1175/jcli-d-11-00290.1, 2011.

Bond-Lamberty, B., Calvin, K., Jones, A. D., Mao, J., Patel, P., Shi, X., Thomson, A., Thornton, P., and Zhou, Y.: Coupling earth system and integrated assessment models: the problem of steady state, Geoscientific Model Development Discuss, 7, 1499-1524, doi: 10.5194/gmdd-7-1499-2014, 2014.

Brovkin, V., Boysen, L., Arora, V. K., Boisier, J. P., Cadule, P., Chini, L., Claussen, M., Friedlingstein, P., Gayler, V., van den Hurk, B. J. J. M., Hurtt, G. C., Jones, C. D., Kato, E., de Noblet-Ducoudré, N., Pacifico, F., Pongratz, J., and Weiss, M.: Effect of anthropogenic land-use and land-cover changes on climate and land carbon storage in cmip5 projections for the twenty-first century, J. Clim., 26, 6859-6881, 10.1175/jcli-d-12-00623.1, 2013.

Calvin K., Clarke, L. E., Edmonds, J. A., Eom, J., Hejazi, M. I., Kim, S. H., Kyle, G. P., Link, R. P., Luckow, P., Patel, P. L., Smith, S. J., and Wise, M. A.: GCAM Wiki Documentation, PNNL-20809, Pacific Northwest National Laboratory, Richland, WA, http://wiki.umd.edu/gcam/, 2011.

Davies-Barnard, T., Valdes, P. J., Singarayer, J. S., and Jones, C. D.: Climatic impacts of land-use change due to crop yield increases and a universal carbon tax from a scenario model, J. Clim., 27, 1413-1424, 10.1175/jcli-d-13-00154.1, 2014.

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, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K.-G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C., and Zeng, N.: Climate-carbon cycle feedback analysis: Results from the c4mip model intercomparison, J. Clim., 19, 3337-3353, doi:10.1175/JCLI3800.1, 2006.

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., Vertenstein, M., Worley, P. H., Yang, Z.-L., and Zhang, M.: The community climate system model version 4, J. Clim., 24, 4973-4991, 10.1175/2011jcli4083.1, 2011.

Houghton, R. A.: How well do we know the flux of co 2 from land-use change?, Tellus B, 62(5), 337-351, doi: 10.1111/j.1600-0889.2010.00473.x, 2010.

Houghton, R. A., House, J. I., Pongratz, J., van der Werf, G. R., DeFries, R. S., Hansen, M. C., Le Quere, C., and Ramankutty, N.: Carbon emissions from land use and land-cover change, Biogeosciences, 9, 5125-5142, 10.5194/bg-9-5125-2012, 2012.

Hurtt, G. C., Pacala, S. W., Moorcroft, P. R., Caspersen, J., Shevliakova, E., Houghton, R. A., and Moore, B.: Projecting the future of the u.S. Carbon sink, Proceedings of the National Academy of Sciences, 99, 1389-1394, 10.1073/pnas.012249999, 2002.

Hurtt, G. C., Frolking, S., Fearon, M. G., Moore, B., Shevliakova, E., Malyshev, S., Pacala, S. W., and Houghton, R. A.: The underpinnings of land-use history: Three centuries of global gridded land-use transitions, wood-harvest activity, and resulting secondary lands, Global Change Biol., 12, 1208-1229, 10.1111/j.1365-2486.2006.01150.x, 2006.

Hurtt, G. C., Chini, L. P., Frolking, S., Betts, R. A., Feddema, J., Fischer, G., Fisk, J. P., Hibbard, K., Houghton, R. A., Janetos, A., Jones, C. D., Kindermann, G., Kinoshita, T., Klein Goldewijk, K., Riahi, K., Shevliakova, E., Smith, S., Stehfest, E., Thomson, A., Thornton, P., Vuuren, D. P., and Wang, Y. P.: Harmonization of land-use scenarios for the period 1500-2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands, Clim. Change, 109, 117-161, 10.1007/s10584-011-0153-2, 2011.

Jain, A. K., and Yang, X.: Modeling the effects of two different land cover change data sets on the carbon stocks of plants and soils in concert with co2 and climate change, Global Biogeochemical Cycles, 19, GB2015, 10.1029/2004gb002349, 2005.

Jain, A. K., Meiyappan, P., Song, Y., and House, J. I.: Co2 emissions from land-use change affected more by nitrogen cycle, than by the choice of land-cover data, Global Change Biol., 19, 2893-2906, 10.1111/gcb.12207, 2013.

Jones, A. D., Collins, W. D., Edmonds, J., Torn, M. S., Janetos, A., Calvin, K. V., Thomson, A., Chini, L. P., Mao, J., Shi, X., Thornton, P., Hurtt, G. C., and Wise, M.: Greenhouse gas policy influences climate via direct effects of land-use change, J. Clim., 26, 3657-3670, 10.1175/jcli-d-12-00377.1, 2013a.

Jones, C. D., Hughes, J. K., Bellouin, N., Hardiman, S. C., Jones, G. S., Knight, J., Liddicoat, S., O'Connor, F. M., Andres, R. J., Bell, C., Boo, K.-O., Bozzo, A., Butchart, N., Cadule, P., Corbin, K. D., Doutriaux-Boucher, M., Friedlingstein, P., Gornall, J., Gray, L., Halloran, P. R., Hurtt, G., Ingram, W. J., Lamarque, J.-F., Law, R. M., Meinshausen, M., Osprey, S., Palin, E. J., Parsons Chini, L., Raddatz, T., Sanderson, M. G., Sellar, A. A., Schurer, A., Valdes, P., Wood, N., Woodward, S., Yoshioka, M., and Zerroukat, M.: The HadGEM2-ES implementation of CMIP5 centennial simulations, Geosci. Model Dev., 4, 543-570, doi:10.5194/gmd-4-543-2011, 2011.

Jones, C., Robertson, E., Arora, V., Friedlingstein, P., Shevliakova, E., Bopp, L., Brovkin, V., Hajima, T., Kato, E., Kawamiya, M., Liddicoat, S., Lindsay, K., Reick, C. H., Roelandt, C., Segschneider, J., and Tjiputra, J.: Twenty-first-century compatible co2 emissions and airborne fraction simulated by cmip5 earth system models under four representative concentration pathways, J. Clim., 26, 4398-4413, 10.1175/jcli-d-12-00554.1, 2013b.

Keppel-Aleks, G., Randerson, J. T., Lindsay, K., Stephens, B. B., Moore, J. K., Doney, S. C., Thornton, P. E., Mahowald, N. M., Hoffman, F. M., Sweeney, C., Tans, P. P., Wennberg, P. O., and Wofsy, S. C.: Atmospheric Carbon Dioxide Variability in the Community Earth System Model: Evaluation and Transient Dynamics during the Twentieth and

Twenty-First Centuries, J. Clim., 26, 4447-4475, 10.1175/jcli-d-12-00589.1, 2013.
Klein Goldewijk, K., Beusen, A., van Drecht, G., and de Vos, M.: The hyde 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years, Global Ecol. Biogeogr., 20, 73-86, 10.1111/j.1466-8238.2010.00587.x, 2011.
Kyle, G. P., Luckow, P., Calvin, K., Emanuel, W., Nathan, M., and Zhou, Y.: Gcam 3.0 agriculture and land use: Data sources and methods, Pacific Northwest National Laboratory, 58, 2011.
Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang, Z.-L., Levis, S., Sakaguchi, K., Bonan, G. B., and Slater, A. G.: Parameterization improvements and functional and structural advances in version 4 of the community land model, Journal of Advances in Modeling Earth Systems, 3, M03001, 10.1029/2011ms000045, 2011.
Lawrence, P. J., Feddema, J. J., Bonan, G. B., Meehl, G. A., O'Neill, B. C., Oleson, K. W., Levis, S., Lawrence, D. M., Kluzek, E., Lindsay, K., and Thornton, P. E.: Simulating the biogeochemical and biogeophysical impacts of transient land cover change and wood harvest in the community climate system model (ccsm4) from 1850 to 2100, J. Clim., 25, 3071-3095, 10.1175/jcli-d-11-00256.1, 2012.
Lee, H.-L., Hertel, T. W., Sohngen, B., and Ramankutty, N.: Towards an integrated land use database for assessing the potential for greenhouse gas mitigation, Purdue University, 83, 2005.
Masui, T., Matsumoto, K., Hijioka, Y., Kinoshita, T., Nozawa, T., Ishiwatari, S., Kato, E., Shukla, P. R., Yamagata, Y., and Kainuma, M.: An emission pathway for stabilization at 6 Wm^2 radiative forcing, Clim. Change, 109, 59-76, 10.1007/s10584-011-0150-5, 2011.
Meehl, G. A., Moss, R., Taylor, K. E., Eyring, V., Stouffer, R. J., Bony, S., and Stevens, B.: Climate model intercomparisons: Preparing for the next phase, EOS Trans. Am. Geophys. Union, 95, 77-84, 2014.
Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram, T., Meehl, G. A., Mitchell, J. F. B., Nakicenovic, N., Riahi, K., Smith, S. J., Stouffer, R. J., Thomson, A. M., Weyant, J. P., and Wilbanks, T. J.: The next generation of scenarios for climate change research and assessment, Nature, 463, 747-756, doi: 10.1038/nature08823, 2010.
Pitman, A. J., de Noblet-Ducoudré, N., Cruz, F. T., Davin, E. L., Bonan, G. B., Brovkin, V., Claussen, M., Delire, C., Ganzeveld, L., Gayler, V., van den Hurk, B. J. J. M., Lawrence, P. J., van der Molen, M. K., Müller, C., Reick, C. H., Seneviratne, S. I., Strengers, B. J., and Voldoire, A.: Uncertainties in climate responses to past land cover change: First results from the lucid intercomparison study, Geophys. Res. Lett., 36, L14814, 10.1029/2009gl039076, 2009.
Raddatz, R. L.: Evidence for the influence of agriculture on weather and climate through the transformation and management of vegetation: Illustrated by examples from the canadian prairies, Agricultural and Forest Meteorology, 142, 186-202, http://dx.doi.org/10.1016/j.agrformet.2006.08.022, 2007.
Ramankutty, N., and Foley, J. A.: Estimating historical changes in global land cover: Croplands from 1700 to 1992, Global Biogeochem. Cycles, 13, 997-1027, 10.1029/1999gb900046, 1999.

Reilly, J., Melillo, J., Cai, Y., Kicklighter, D., Gurgel, A., Paltsev, S., Cronin, T., Sokolov, A., and Schlosser, A.: Using land to mitigate climate change: Hitting the target, recognizing the trade-offs, Environ. Sci. Technol., 46, 5672-5679, 10.1021/es2034729, 2012.

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., and Rafaj, P.: Rcp 8.5--a scenario of comparatively high greenhouse gas emissions, Clim. Change, 109, 33-57, 10.1007/s10584-011-0149-y, 2011.

Rose, S. K., Ahammad, H., Eickhout, B., Fisher, B., Kurosawa, A., Rao, S., Riahi, K., and van Vuuren, D. P.: Land-based mitigation in climate stabilization, Energy Economics, 34, 365-380, 10.1016/j.eneco.2011.06.004, 2012.

Shevliakova, E., Pacala, S. W., Malyshev, S., Hurtt, G. C., Milly, P. C. D., Caspersen, J. P., Sentman, L. T., Fisk, J. P., Wirth, C., and Crevoisier, C.: Carbon cycling under 300 years of land use change: Importance of the secondary vegetation sink, Global Biogeochemical Cycles, 23, GB2022, 10.1029/2007gb003176, 2009.

Shevliakova, E., Stouffer, R. J., Malyshev, S., Krasting, J. P., Hurtt, G. C., and Pacala, S. W.: Historical warming reduced due to enhance land carbon uptake, PNAS, 110, 16730-16735, 10.1073/pnas.1314047110, 2013.

Smith, P., Haberl, H., Popp, A., Erb, K.-h., Lauk, C., Harper, R., Tubiello, F. N., de Siqueira Pinto, A., Jafari, M., Sohi, S., Masera, O., Böttcher, H., Berndes, G., Bustamante, M., Ahammad, H., Clark, H., Dong, H., Elsiddig, E. A., Mbow, C., Ravindranath, N. H., Rice, C. W., Robledo Abad, C., Romanovskaya, A., Sperling, F., Herrero, M., House, J. I., and Rose, S.: How much land-based greenhouse gas mitigation can be achieved without compromising food security and environmental goals?, Global Change Biol., 19, 2285-2302, 10.1111/gcb.12160, 2013a.

Smith, R., Jones, P., Briegleb, B., Bryan, F., Danabasoglu, G., Dennis, J., Dukowicz, J., Eden, C., Fox-Kemper, B., Gent, P., Hecht, M., Jayne, S., Jochum, M., Large, W., Lindsay, K., Maltrud, M., Norton, N., Peacock, S., Vertenstein, M., and Yeager, S.: The parallel ocean program (pop) reference manual, Los Alamos National Laboratory, 183, 2013b.

Smith, S. J., and Rothwell, A.: Carbon density and anthropogenic land-use influences on net land-use change emissions, Biogeosciences, 10, 6323-6337, 10.5194/bg-10-6323-2013, 2013.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of cmip5 and the experiment design, Bulletin of the American Meteorological Society, 93, 485-498, 10.1175/bams-d-11-00094.1, 2012.

Thomson, A., Calvin, K., Smith, S., Kyle, G. P., Volke, A., Patel, P., Delgado-Arias, S., Bond-Lamberty, B., Wise, M., Clarke, L., and Edmonds, J.: Rcp4.5: A pathway for stabilization of radiative forcing by 2100, Clim. Change, 109, 77-94, 10.1007/s10584-011-0151-4, 2011.

Thornton, P. E., Lamarque, J.-F., Rosenbloom, N. A., and Mahowald, N. M.: Influence of carbon-nitrogen cycle coupling on land model response to CO2 fertilization and climate variability, Global Biogeochemical Cycles, 21, GB4018, 10.1029/2006gb002868, 2007.

van Vuuren, D., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S., and Rose, S.: The representative concentration pathways: An overview, Clim. Change, 109, 5-31, 10.1007/s10584-011-0148-z, 2011a.

van Vuuren, D., Stehfest, E., Elzen, M. J., Kram, T., Vliet, J., Deetman, S., Isaac, M., Klein Goldewijk, K., Hof, A., Mendoza Beltran, A., Oostenrijk, R., and Ruijven, B.: Rcp2.6:

Exploring the possibility to keep global mean temperature increase below 2°C, Clim. Change, 109, 95-116, 10.1007/s10584-011-0152-3, 2011b.

Wilkenskjeld, S., Kloster, S., Pongratz, J., Raddatz, T., and Reick, C.: Comparing the influence of net and gross anthropogenic land use and land cover changes on the carbon cycle in the MPI-ESM, Biogeosciences Discuss., 11, 5443-5469, 10.5194/bgd-11-5443-2014, 2014.

Wise, M., Calvin, K., Thomson, A., Clarke, L., Bond-Lamberty, B., Sands, R., Smith, S. J., Janetos, A., and Edmonds, J.: Implications of limiting CO2 concentrations for land use and energy, Science, 324, 1183-1186, 10.1126/science.1168475, 2009.

Wise, M., and Calvin, K.: Gcam 3.0 agriculture and land use: Technical description of modeling approach, Pacific Northwest Naitonal Laboratory, 51, 2011.

Figure captions

Figure 1. Status of iESM implementation as of Spring 2014. The light blue arrows show information flow from GCAM to CESM. The light green arrows show information flow from CESM to GCAM. The dashed gray outline, including the crossed out arrows, represents the CMIP5 land coupling. The solid green outline, minus the arrow crossed out by green and including the 100-year emissions arrow, depicts the current iESM implementation. The dashed blue outline, minus both crossed out arrows, indicates ongoing development. The dashed red line, minus the crossed out arrows, includes the next stage of development. GCAM: Global Change Assessment Model. GLM: Global Land use Model. CESM: Community Earth System Model.

Figure 2. General integrated Earth System Model (iESM) land use coupling algorithm. Forest area is not passed from the Global Change Assessment Model (GCAM) to the Global Land use Model (GLM) in the CMIP5 land use coupling, but it is passed in the iESM simulations used in this study. NPP: Net Primary Productivity. HR: Heterotrophic Respiration. PFT: Plant Functional Type.

Figure 3. OLD Land Use Translator (OLDLUT) algorithm for dynamic Plant Functional Type (PFT) coverage. When cropland and pasture decrease, non-crop PFTs are added in proportion to potential vegetation fractions. When cropland and pasture increase, non-crop PFTs are removed in proportion to reference year fractions.

Figure 4. NEW Land Use Translator (NEWLUT) algorithm for dynamic Plant Functional Type (PFT) coverage. When cropland and pasture decrease, tree PFTs are added in proportion to

potential vegetation fractions. When cropland and pasture increase, tree PFTs are removed first, then other non-crop PFTs, in proportion to reference year fractions.

Figure 5. Projected global forest, pasture, grass, and shrub areas for the CMIP5 4.5 W m^{-2} Representative Concentration Pathway (RCP4.5), in million km^2. CESM: Community Earth System Model. GLM: Global Land use Model. PFT: Plant Functional Type.

Figure 6. Integrated Earth System Model (iESM) land use and forest area changes with respect to 2015. The GLM-NEWLUT forest and pasture data are nearly identical to the GLM-OLDLUT data and are not shown for clarity. Similarly, the GLM-NEWLUT cropland data are nearly identical to the GCAM-NEWLUT data. CESM: Community Earth System Model. GCAM: Global Change Assessment Model. GLM: Global Land use Model.

Figure 7. Integrated Earth System Model (iESM) land use and forest area. The GLM-NEWLUT forest and pasture data are nearly identical to the GLM-OLDLUT data and are not shown for clarity. Similarly, the GLM-NEWLUT cropland data track the GCAM-NEWLUT data, but with the same offset as for the GLM-OLDLUT data. CESM: Community Earth System Model. GCAM: Global Change Assessment Model. GLM: Global Land use Model.

Figure 8. Spatial distributions of iESM increased forest Plant Functional Types (PFTs), decreased grass and shrub PFTs, and potential forest PFTs, as percentages of land area within each grid cell. a) Difference in 2040 forest PFT area (NEWLUT - OLDLUT). b) Difference in 2040 grass plus shrub PFT area (NEWLUT - OLDLUT). c) Potential forest PFT area.

Figure 9. Net Ecosystem Exchange (NEE) comparison between iESM simulations. a) NEE for each simulation. b) NEE difference (NEWLUT minus OLDLUT). These data show more land

carbon uptake (negative NEE), associated with the additional trees, for the NEWLUT simulation during the afforestation period (2015 forward).

Figure 10. Comparison between iESM simulations of a-b) vegetation carbon and c-d) atmospheric CO_2 concentration. Differences are NEWLUT minus OLDLUT. Due to additional forest area, the NEWLUT simulation significantly increases vegetation carbon gain and decreases atmospheric CO_2 gain over the OLDLUT simulation.

Table 1. Two integrated Earth System Model (iESM) simulations performed for this study.

	OLDLUT	NEWLUT
Modified Land Use Translator	N	Y
Vegetation productivity feedbacks	Y	Y
Updated Global Land use Model	Y	Y

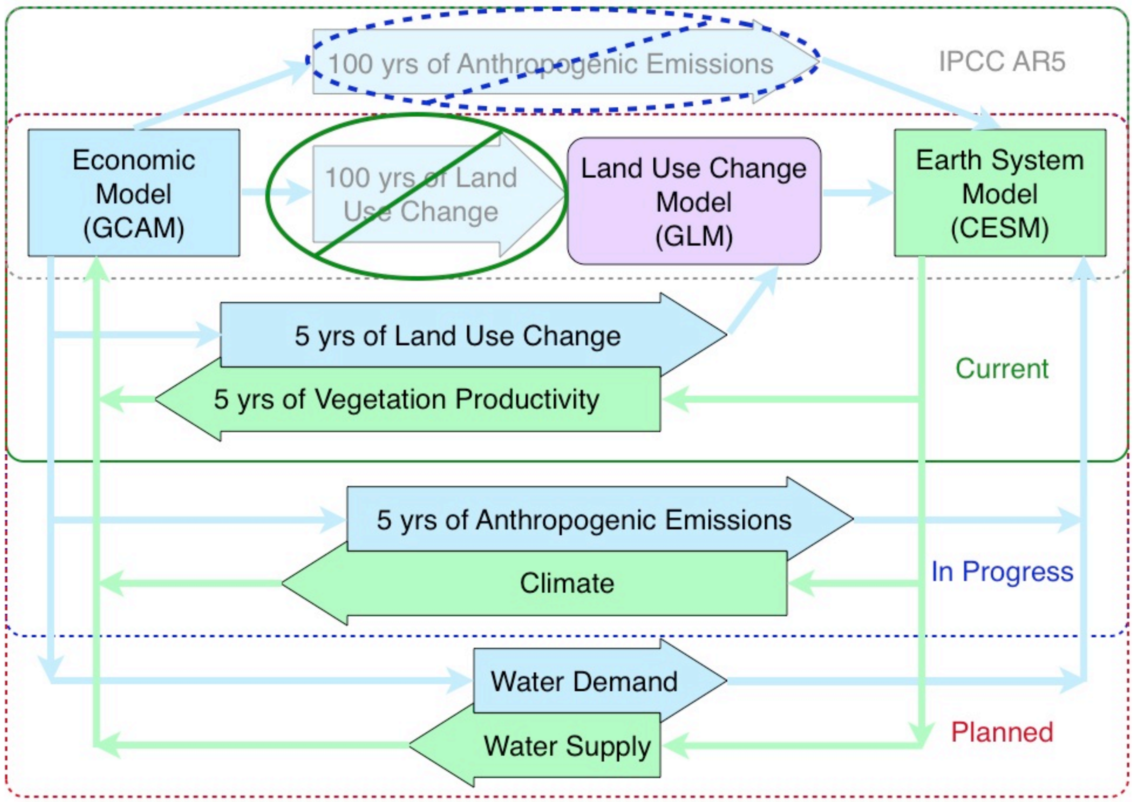


Figure 1. Status of iESM implementation as of Spring 2014. The light blue arrows show information flow from GCAM to CESM. The light green arrows show information flow from CESM to GCAM. The dashed gray outline, including the crossed out arrows, represents the CMIP5 land coupling. The solid green outline, minus the arrow crossed out by green and including the 100-year emissions arrow, depicts the current iESM implementation. The dashed blue outline, minus both crossed out arrows, indicates ongoing development. The dashed red line, minus the crossed out arrows, includes the next stage of development. GCAM: Global Change Assessment Model. GLM: Global Land use Model. CESM: Community Earth System Model.

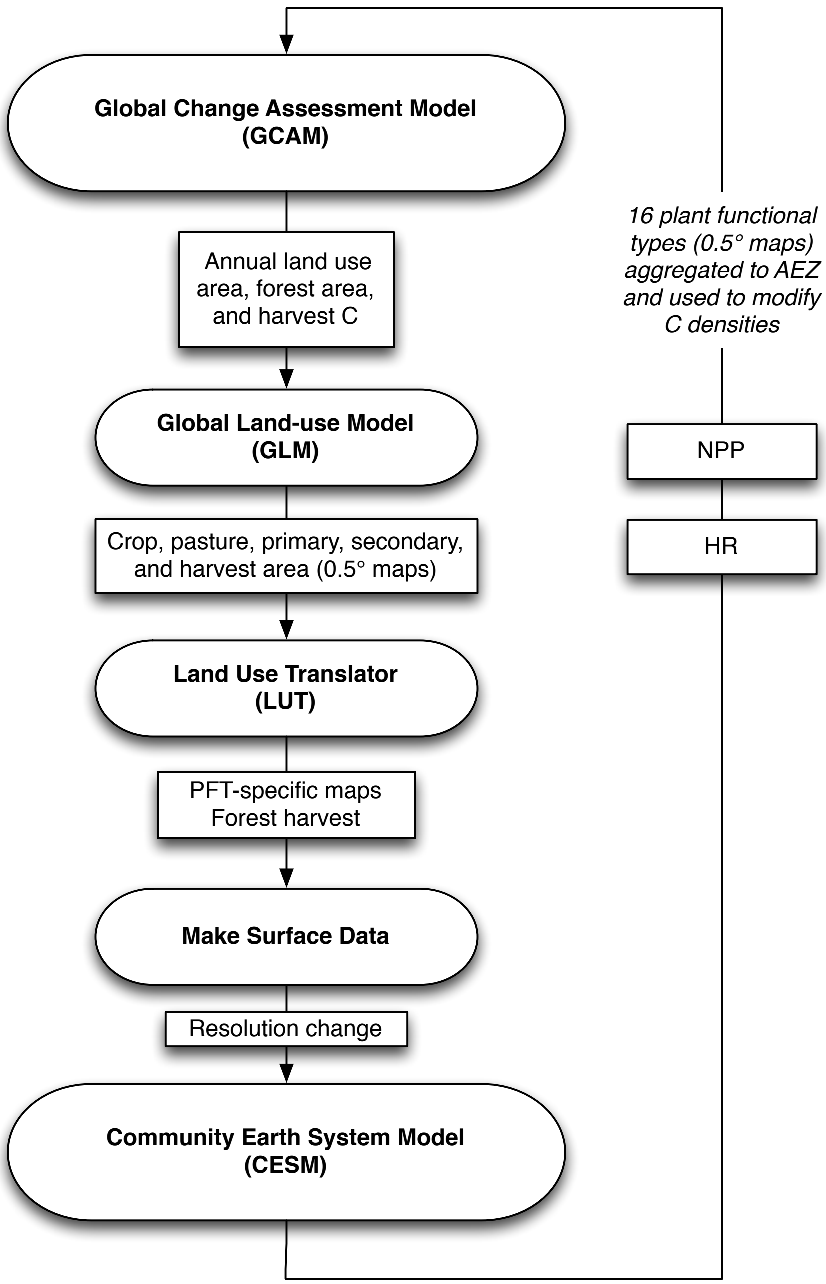


Figure 2. General integrated Earth System Model (iESM) land use coupling algorithm. Forest area is not passed from the Global Change Assessment Model (GCAM) to the Global Land use Model (GLM) in the CMIP5 land use coupling, but it is passed in the iESM simulations used in this study. NPP: Net Primary Productivity. HR: Heterotrophic Respiration. PFT: Plant Functional Type.

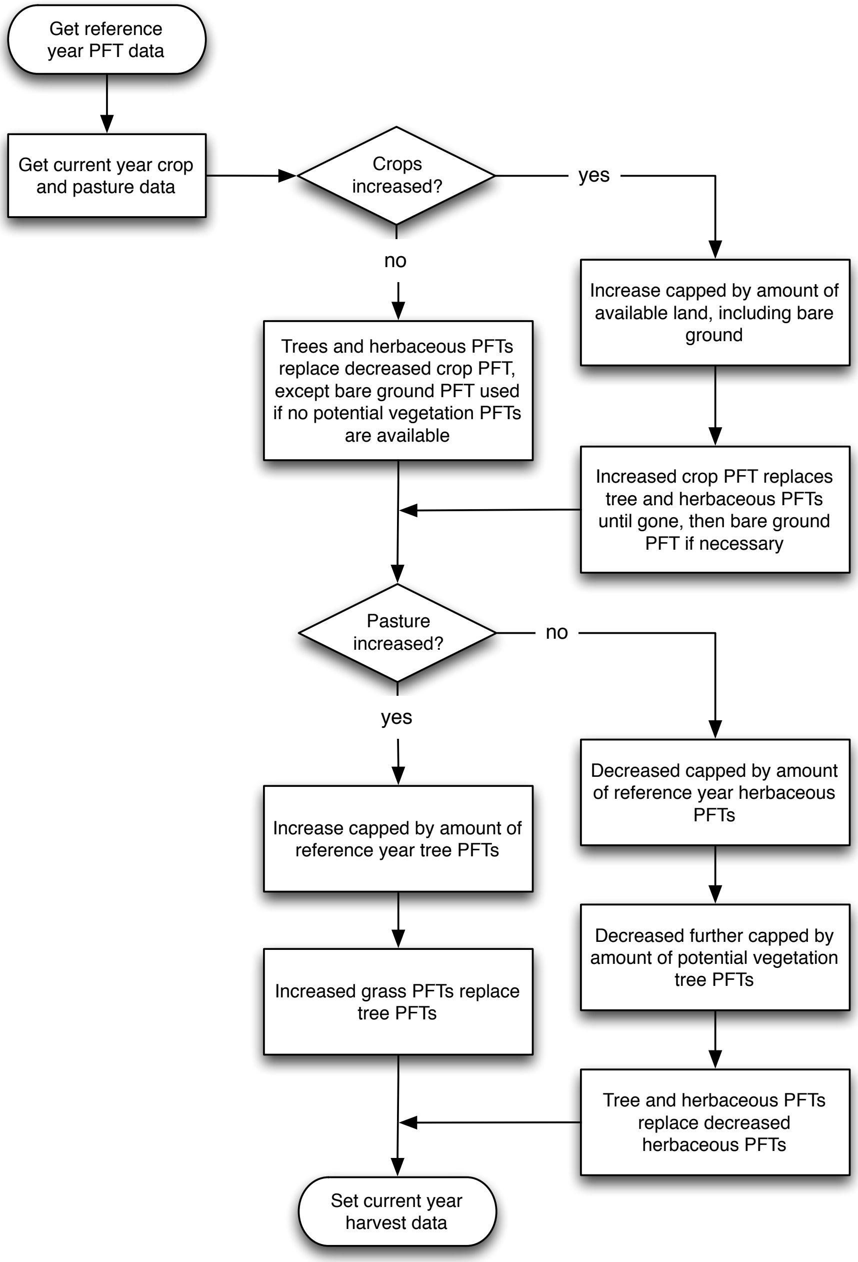


Figure 3. OLD Land Use Translator (OLDLUT) algorithm for dynamic Plant Functional Type (PFT) coverage. When cropland and pasture decrease, non-crop PFTs are added in proportion to potential vegetation fractions. When cropland and pasture increase, non-crop PFTs are removed in proportion to reference year fractions.

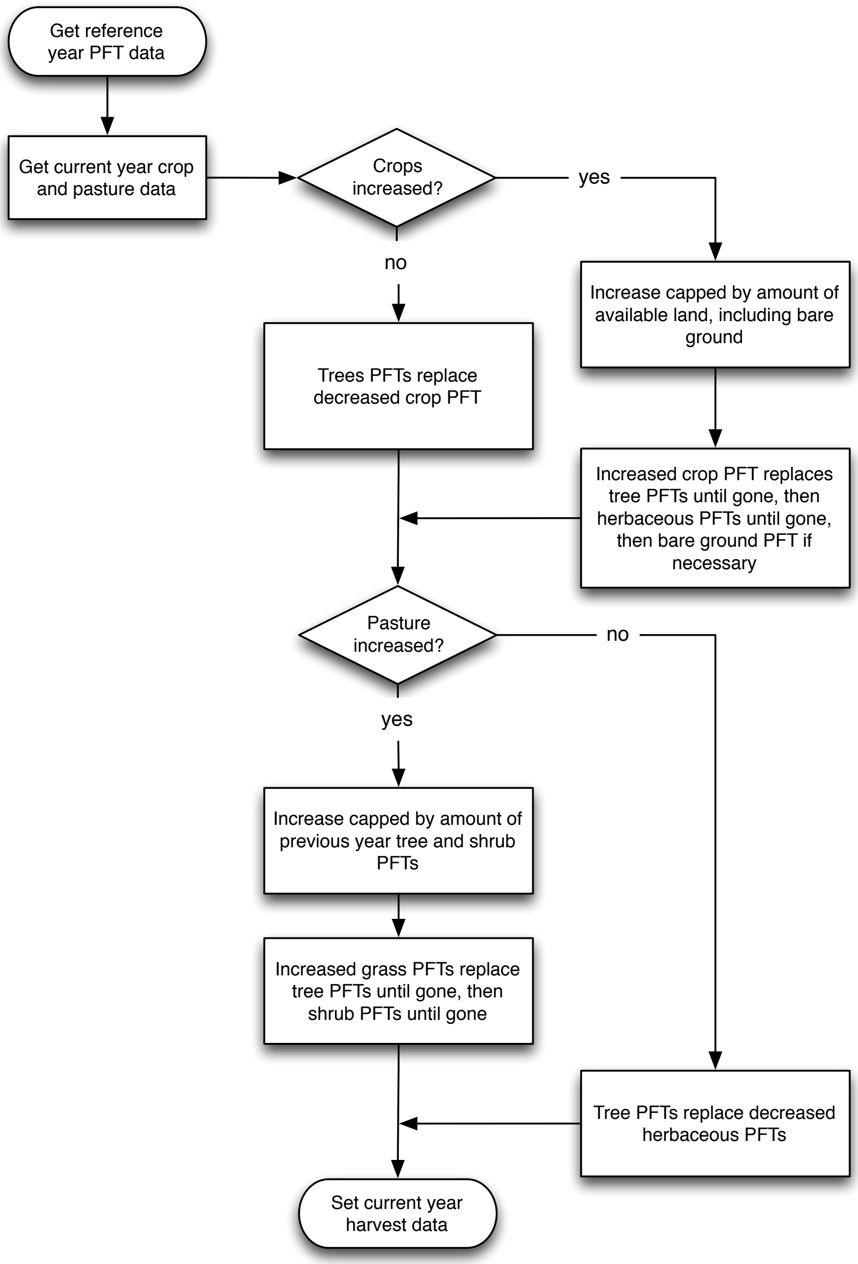


Figure 4. NEW Land Use Translator (NEWLUT) algorithm for dynamic Plant Functional Type (PFT) coverage. When cropland and pasture decrease, tree PFTs are added in proportion to potential vegetation fractions. When cropland and pasture increase, tree PFTs are removed first, then other non-crop PFTs, in proportion to reference year fractions.

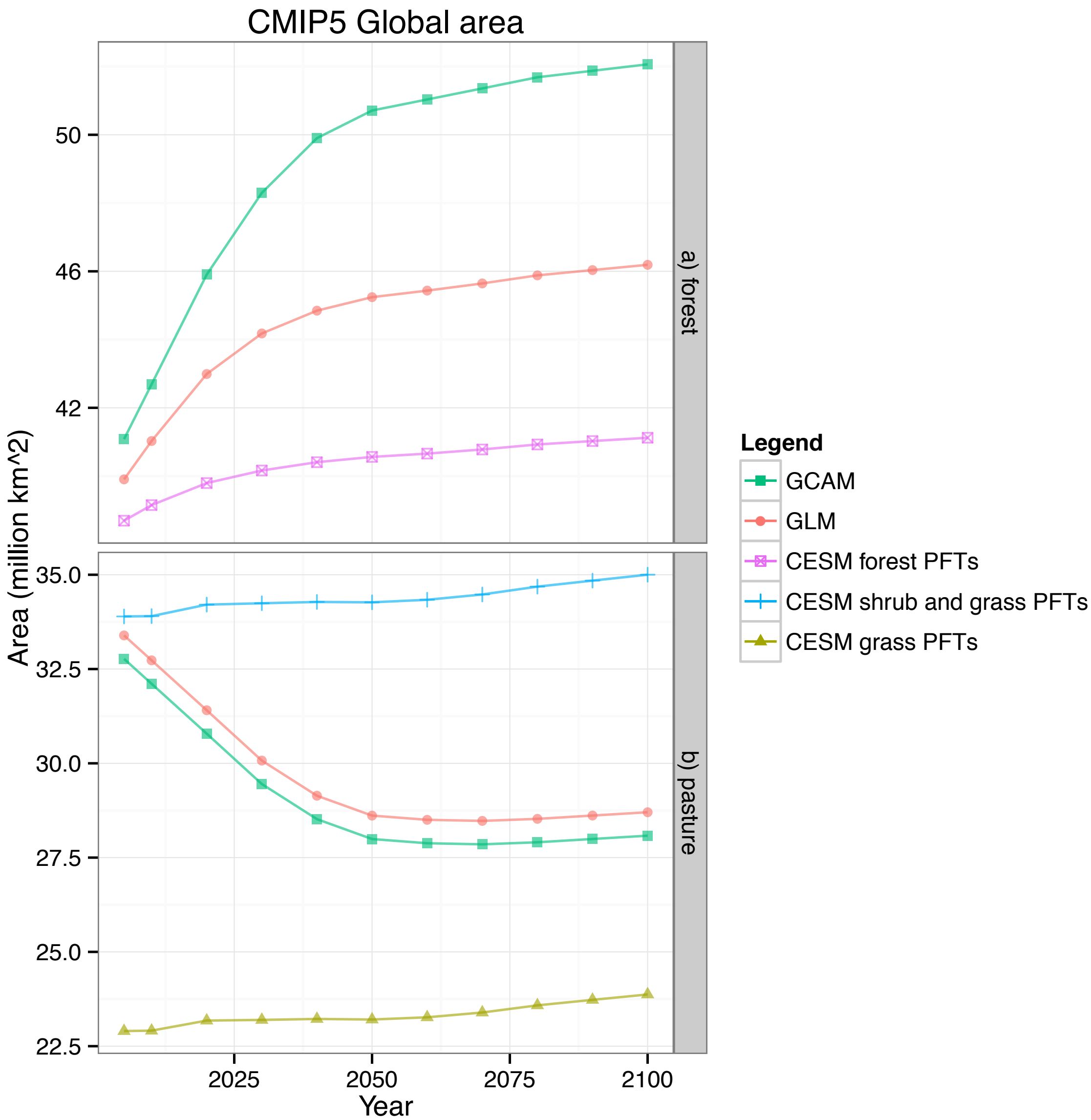


Figure 5. Projected global forest, pasture, grass, and shrub areas for the CMIP5 4.5 W m^{-2} Representative Concentration Pathway (RCP4.5), in million km^2. CESM: Community Earth System Model. GLM: Global Land use Model. PFT: Plant Functional Type.

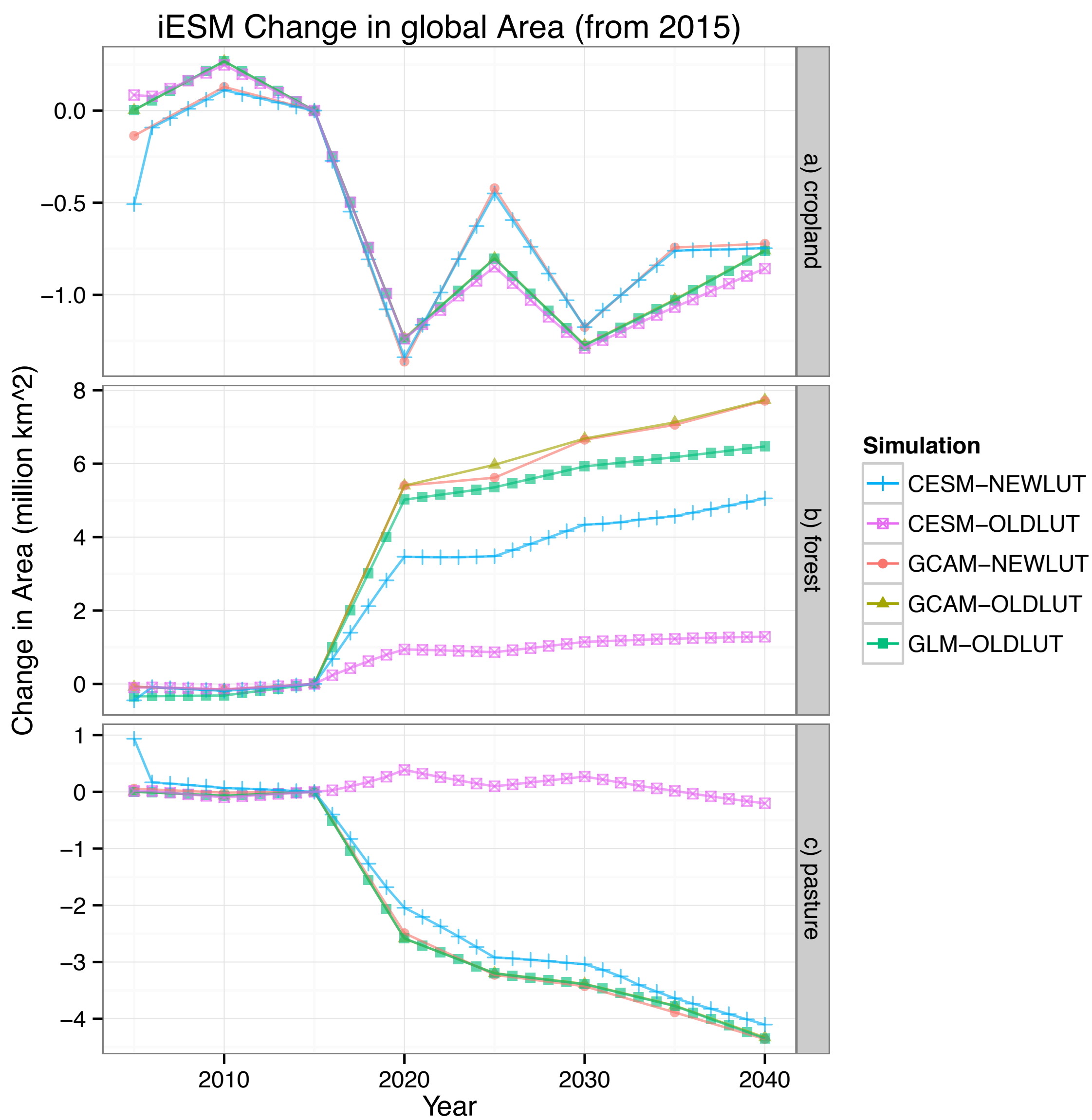


Figure 6. Integrated Earth System Model (iESM) land use and forest area changes with respect to 2015. The GLM-NEWLUT forest and pasture data are nearly identical to the GLM-OLDLUT data and are not shown for clarity. Similarly, the GLM-NEWLUT cropland data are nearly identical to the GCAM-NEWLUT data. CESM: Community Earth System Model. GCAM: Global Change Assessment Model. GLM: Global Land use Model.

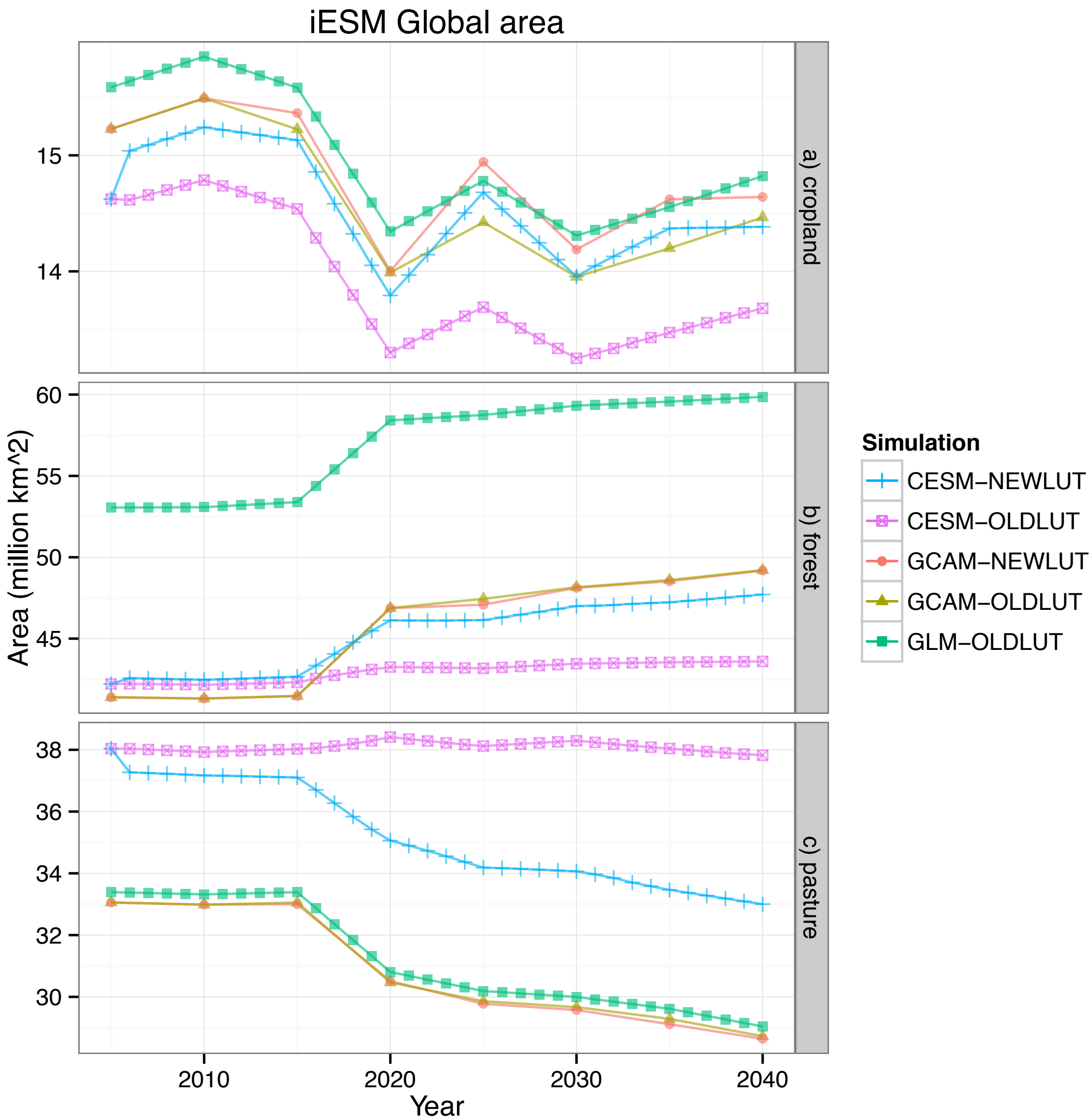
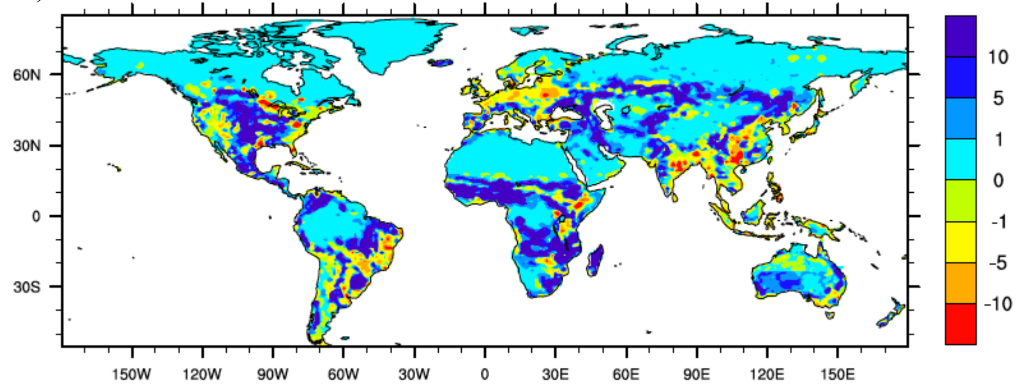
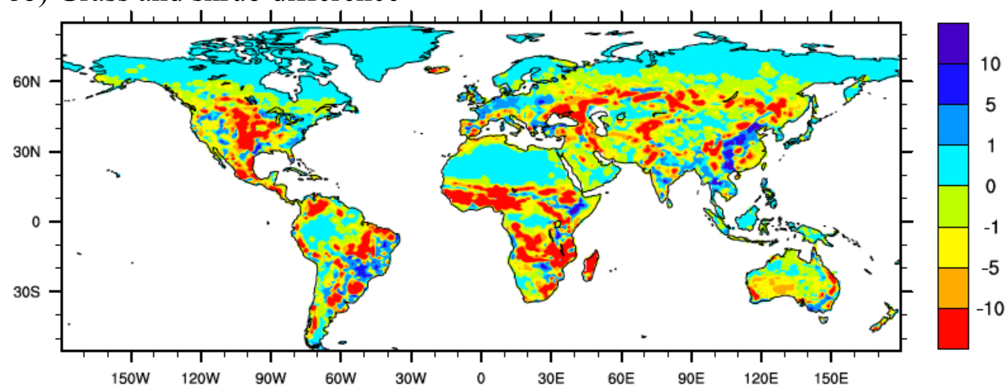


Figure 7. Integrated Earth System Model (iESM) land use and forest area. The GLM-NEWLUT forest and pasture data are nearly identical to the GLM-OLDLUT data and are not shown for clarity. Similarly, the GLM-NEWLUT cropland data track the GCAM-NEWLUT data, but with the same offset as for the GLM-OLDLUT data. CESM: Community Earth System Model. GCAM: Global Change Assessment Model. GLM: Global Land use Model.

8a) Forest difference



8b) Grass and shrub difference



8c) Potential forest

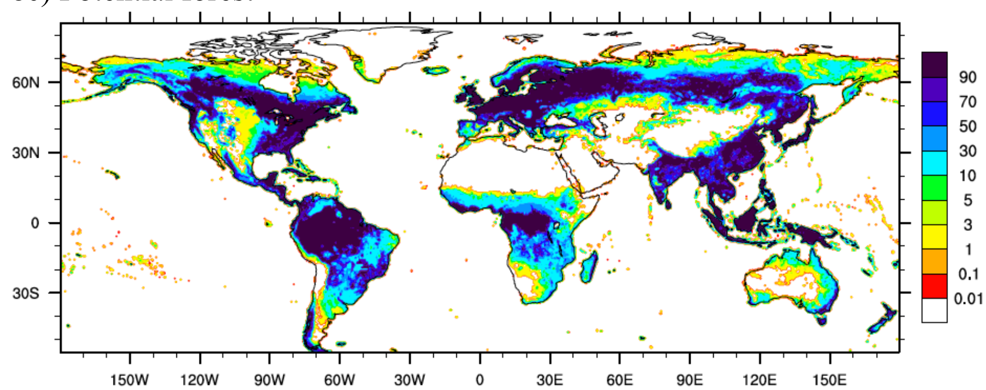


Figure 8. Spatial distributions of iESM increased forest Plant Functional Types (PFTs), decreased grass and shrub PFTs, and potential forest PFTs, as percentages of land area within each grid cell. a) Difference in 2040 forest PFT area (NEWLUT - OLDLUT). b) Difference in 2040 grass plus shrub PFT area (NEWLUT - OLDLUT). c) Potential forest PFT area.

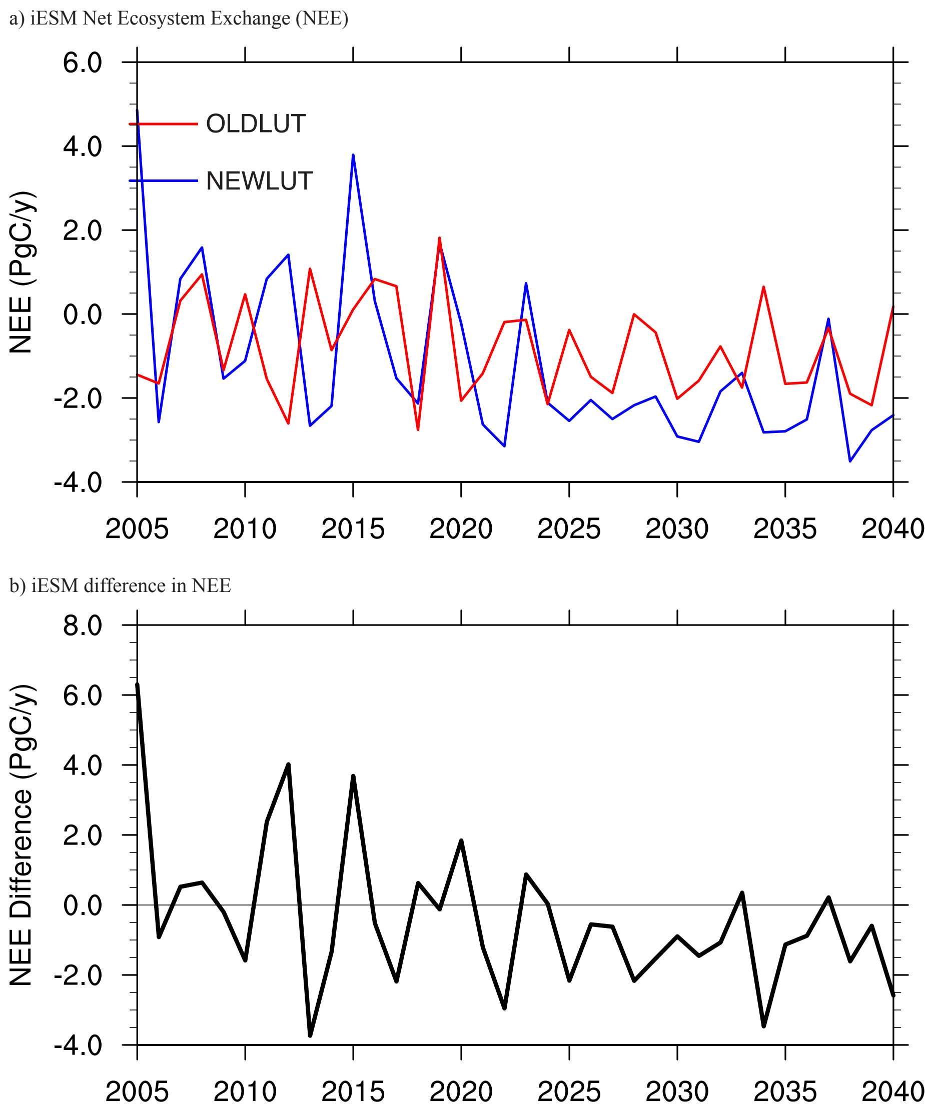


Figure 9. Net Ecosystem Exchange (NEE) comparison between iESM simulations. a) NEE for each simulation. b) NEE difference (NEWLUT minus OLDLUT). These data show more land carbon uptake (negative NEE), associated with the additional trees, for the NEWLUT simulation during the afforestation period (2015 forward).

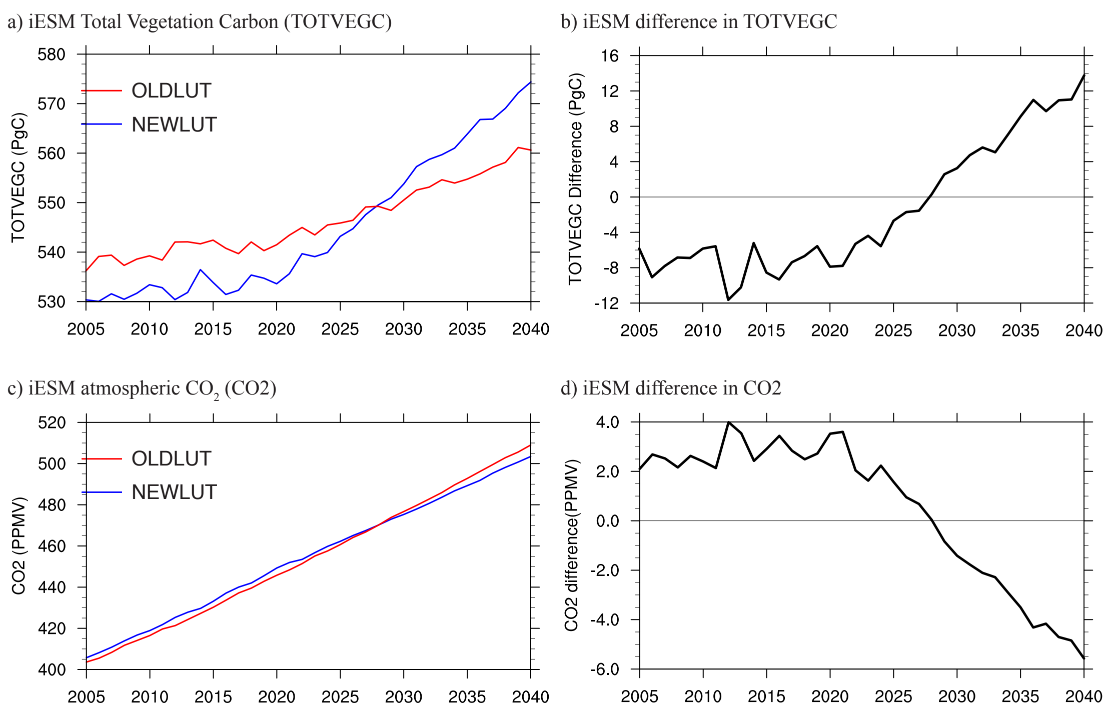


Figure 10. Comparison between iESM simulations of a-b) vegetation carbon and c-d) atmospheric CO_2 concentration. Differences are NEWLUT minus OLDLUT. Due to additional forest area, the NEWLUT simulation significantly increases vegetation carbon gain and decreases atmospheric CO_2 gain over the OLDLUT simulation.