Overview of revisions

We greatly appreciate the constructive review by Julia Pongratz (below referred to as "reviewer 1") and the anonymous reviewer (below referred to as "reviewer 2"). We hope that added and modified text has served to improve our manuscript. In summary:

- Model simulations have not been repeated and all results remain unchanged since the first submission.
- We conducted additional simulations to investigate and illustrate differences in indirect fluxes, caused by environmental change (Δf) on natural and agricultural land, and how indirect fluxes (Δf^{FF} and Δf^{LUC}) combine linearly (new Fig. 1). This shows that non-linearities are negligible during the historical period. A treatment of non-linearities has been included in our formalism and we show which non-linearity terms remain when deriving *e*RSS and *e*LFB from simulations (Eqs. 14 and 15 in the new manuscript).
- We included an additional figure to illustrate the spatial distribution (map) of the C sinks and sources triggered by environmental change ($\Delta f_{\mathrm{nat}}^{\mathrm{FF+LUC}}$ and $\Delta f_{\mathrm{agr}}^{\mathrm{LUC}}$) and of the non-linearity effect mentioned above ($\Delta f_{\mathrm{nat}}^{\mathrm{FF}} + \Delta f_{\mathrm{nat}}^{\mathrm{LUC}} - \Delta f_{\mathrm{nat}}^{\mathrm{FF+LUC}}$).
- We have re-structured some contents of the text, trying to address both reviewers' comments. Specifically, we now have a separate section for an overview of the D1, D3, and E2 methods, including a formalistic description of their setup (Eq. 2, 3, 4). We then introduce the formalism of flux components rigorously as suggested by reviewer 2 (contents previously in Appendix 1), and use this to identify the conceptual difference between D3 and E2 methods. This rearrangement provides more and more concise information, avoids the appendices and limits additional text.
- In view of the ambiguity of the choice of methods, we dropped the strict recommendations, e.g. the sentence "In summary, we recommend not to rely on results from method D3 or E2 in the context of the global (or regionalized) carbon budget, but to apply method D1 (under preindustrial conditions)."
- While trying to add information and analysis where required by the reviewers, we tried to maintain the conciseness of the paper to qualify it as an ESD Short Communication.

In this document, quoted reviewer comments are indented and in blue font. New and/or modified text is in green font.

Response to Reviewer 1

> This study quantifies the net land use flux in a coupled Earth system model setup and the corresponding offline setup of the vegetation model. Based on these quantifications the authors describe the difference in flux components contained in each, with the explanation focussing on the terms "replaced sources and sinks" and "land use feedback". They conclude with recommendations of using the "D1" approach, i.e. quantifying the net land use flux under constant pre-industrial environmental conditions, for global carbon budget estimates and for model intercomparisons.

The manuscript contains some new information that is worth publishing, and ESD is the obvious outlet for it given that two other studies on the topic of flux components contained in different net land use flux definitions have also been published in this journal (Gasser and Ciais, 2013 and Pongratz et al, 2014). I find it interesting to see at least a subset of the at least 9 different methods identified in previous studies be quantified consistently in the same model framework and with state-of-the-art land use datasets; this does not give the complete picture, but it is a good start for a comprehensive study at a later point. However, my feeling is that the way the manuscript is framed currently it will contribute to the confusion in the climate-carbon cycle community about the net land use flux rather than reduce it. While I think that little change is needed to the modeling itself and the analysis of the results to resolve this issue, the discussion and interpretation needs substantial shortening (very justifiable for a "short communication" manuscript type of ESD), clarification, and re-structuring.

We would like to thank Julia Pongratz for her in-depth review of our manuscript. The points mentioned above are addressed in our responses to specific comments below.

General comments

I see the main value of the paper in two points:

(1) To give a quantification of some of the effects discussed in earlier studies consistently in the same modeling framework. Note, however, that these quantifications have partly been performed in consistent model setups by Strassmann et al 2008 (SM08) and Arora and Boer, GCB, 2010 ("Uncertainties in the 20th century carbon budget associated with land use change"). Here, more explanation is warranted in the introduction about the novelty of the current estimates. [...]

The present study was motivated by the recent publications of Pongratz et al. (2014) (PG14) and Gasser and Ciais (2013) (GC13), published in Earth System Dynamics. Our aim was to combine the merits of the two studies: (i) providing an overview and formalism of different methods to quantify *e*LUC and (ii) providing consistent quantifications with a single model, whereby differences in *e*LUC estimates are solely attributable to conceptual differences in applied methods.

Furthermore, we extend the scope of Pongratz et al. (2014) and Gasser and Ciais (2013) by providing an analysis up to year 2100 following a business-as-usual, high-fossil-fuel-emission and low-land-conversion scenario (RCP8.5). These trajectories have important implications as illustrated by the substantial and rising contribution of component fluxes in the 21st century. This underlines the necessity of clarifying definitions of *e*LUC quantifications.

To clarify these aspects and to emphasize the difference to the study by SM08 and Arora and Boer (2010), we added and modified text in the introduction. New and modified text in Section 1 reads:

Here, we apply a single model, use a simple formalistic description of *e*LUC flux components inspired by GC13 and SM08, and follow the classification of PG14 to distinguish different methods of *e*LUC quantification. We quantify these differences for the historical period and a future business-as-usual scenario (RCP8.5). In contrast to earlier studies (Strassmann et al., 2008; Arora and Boer, 2010), we designed model setups to limit differences in *e*LUC to merely conceptual ones by using climate and CO_2 outputs from the coupled simulations to drive offline simulations, instead of using observational data for the latter. We will demonstrate that such definition differences imply inconsistencies of estimated land use emissions on the order of 20% on the global scale and may increase to 30% under a future business-as-usual scenario. This is directly relevant for territorial C balance accountings and national greenhouse gas balances under the Kyoto Protocol and thus inherently carries a political relevance.

We elucidate the implications of the choice of definition for the residual terrestrial C sink and global C budget accountings and discuss how *e*LUC quantifications may most appropriately be defined in studies that rely on multiple methodological approaches. In such cases, we propose, following Houghton (2013), to resort to the "least common denominator", following the bookkeeping approach (method D1 in PG14), where LUC emissions are defined without accounting for any indirect effects on terrestrial C storage caused by transient changes in CO_2 or climate.

We limited our analysis to differences between D1, D3, and E2 as these are widely used methods, applied by process-based models (DGVMs) and bookkeeping models. We write at the beginning of Sect. 2 (slightly modified text):

We focus our analysis on the discrepancy between *e*LUC derived from bookkeeping and offline vegetation models (D1 and D3 methods) and coupled ESMs (E2 method). Results the D3 method feature prominently in model intercomparison studies (McGuire et al., 2001; Sitch et al., 2008), the Global Carbon Project (Le Quéré et al., 2015) and the IPCC (Ciais et al., 2013), and are often presented along with and compared against D1-type estimates. Yet, a consistent separation of commonly identified component fluxes can only be achieved by ESMs (see below).

> [...] A quantitative comparison is needed to these studies as well as to the summary of the size of effects such as land use feedbacks by Pongratz et al 2014 (PG14).

We now provide a quantitative comparison of the difference between estimates following the D3 and E2 methods and relate our estimates of *e*RSS and *e*LFB to previous quantifications. As discussed now in the revised manuscript, *e*LFB can consistently be separated from the simulations presented in Pongratz et al. (2009) and Stocker et al. (2011). These studies did not include the FF forcing and thus the difference between their fully coupled simulations (E1) and their uncoupled simulations (D1) is equal to our definition of *e*LFB. We calculated this difference to 74 GtC in Stocker et al. (2011) and 64 GtC in Pongratz et al. (2009) (their Table 2, sum of 800-1850 and 1850-2000, "primary emissions" vs. "net emissions"). This corresponds to a 32% and 40% reduction relative to $e\mathrm{LUC}_0$ due to *e*LFB. The results by Arora and Boer (2010) do not provide consistent quantitative insight into *e*LFB. Pongratz et al. (2014) used their data from a D3 and E2 simulation in their Figure 3. This does not correspond to *e*LFB. In the same Figure, D1 vs. E2 is used from the SM08 study. This neither corresponds to *e*LFB. Quantitative comparison to other *e*RSS quantifications are not straight-forward either. Our estimate is close to the one by SM08 (see added text below) and on the same order as summarised in PG14 (their table 10). Differences between D3 and D1 are not identical to *e*RSS as discussed in Section 3 in our manuscript, which inhibits the use of other studies. To provide quantitative comparison wherever possible, we added/modified text in Section 5:

SM08 applied observational CO_2 and climate in simulations used for D3. They found slightly higher differences of D3 vs. E2 (30% higher in their D3). Arora and Boer (2010) report a difference of ~100% for a case where they only used CO_2 concentrations from their interactive $F_{\mathrm{LUC}}^{\mathrm{FF+LUC}}$ to force their $F_0^{\mathrm{FF+LUC}}$ simulation. A stronger effect in this case appears plausible as the replaced sinks/sources flux due to climate and CO_2 effects are generally opposing (Strassmann et al., 2008).

and

During the historical period (1850 – 2005 AD), *e*RSS cumulatively adds 6% to primary emissions, similar as in SM08 (5%), while *e*LFB reduces them by 17%, similar as in SM08 (18%) but less than in Pongratz et al. (2009) and Stocker et al. (2011) (30-40%).

> (2) To make it easier for the community to understand all the complications with net land use flux definitions discussed by SM08, Gasser and Ciais 2013, and PG14 by giving a concrete example of model setup. This contribution in eliminating confusion in the scientific community is substantially hampered, however, by not being clear about what is old and what is new. [...]

We added references to clarify this wherever possible. The model setups used in D1, D3, and E2 were not explicitly described in our initial manuscript version. They appeared in Section 2, given by subscripts and superscripts to F and f. We now provide information of model setup by Equations 2, 3, and 4 to describe the difference between D1, D3 and E2 early on. This also motivated a re-structuring of Section 2 (Formalism) to avoid double-mentioning of the same equations.
We also added text in the introduction to emphasize that our description of different methods is essentially a repetition of PG14 but necessary to lay out our arguments further on:

We start by revisiting the classification of PG14 for a subset of *e*LUC quantification methods identified in their study. We focus our analysis [...]

> [...] The discussion in Sec. 5 is all right and interesting, but the key thoughts therein are identical with what Pg14 and Houghton, GCB, 2013 (opinion piece on "Keeping management effects separate from environmental effects in terrestrial carbon accounting"; not cited, but should be!) presented as conclusions. Presenting it here again in so much detail and without pointing to earlier discussions suggests to the reader that this is yet another effect and this is certainly contra-productive to eliminating confusion.

The reference to Houghton (2013) has been added and we now treat his arguments in our discussion in Sect. 6.1 (previously 5.1). However, our discussion in Sect. 6 (previously 5) addresses points that (to our knowledge) have not mentioned explicitly in previous publications. For example:

- When replaced sinks/sources (*e*RSS) are ascribed to the *e*LUC, then the residual C sink represents a hypothetical where no natural vegetation has been converted. This virtual flux confounds comparisons to up-scaled local-to-regional scale observation-based information.
- Including the *e*LFB term into *e*LUC quantifications is inconsistent with the term for fossil-fuel CO_2 emissions in the global C budget.

> The manuscript contains "recommendations" for a specific choice of method, which, as explained below, I find subjective. It is up to the editor to accept such statements and treat the paper partly as an opinion piece, but this is not my understanding of ESD's scope.

We changed this so that the choice of method for global C budget accountings doesn't appear prescriptive. Modified text in the abstract reads:

Therefore, we argue that synthesis studies should resort to the "least common denominator" of different methods, following the bookkeeping approach [...]

And reformulated text in Section 5 points out that it may be "misleading" rather than saying "we recommend...":

The D3 method ascribes replaced sinks/sources (*e*RSS) to *e*LUC. This implies that the residual terrestrial sink represents a flux occurring in a hypothetical state before land conversion. This may be misleading in view of the actual reduction of land C sinks due to the reduction of natural vegetation. [...]

Furthermore, we dropped the "recommendation sentence" at the end of Section 5.1 ("In summary, we recommend not to rely on results from method D3 or E2 in the context of the global (or regionalized) carbon budget, but to apply method D1 (under preindustrial conditions).")
According to our understanding of an ESD Short Communication, a perspective from a somewhat subjective angle is in line with the scope this manuscript type.

Specific Comments
Title:
The title in its current form is the conclusion from Pg14. The title needs to include what is new, along the lines of "Quantification of net land use flux components in coupled vs offline climate-vegetation models".

We changed the title to "Quantifying differences in land use emission estimates implied by definition discrepancies".

Abstract:
l. 1 (and elsewhere, e.g. Introduction l. 21-22): "anthropogenic land use and land use change": Why not stick with the common "land use and land cover change (LULCC)"? Land use implies the human component, and it does not hurt to spell out the land cover change part.

The abbreviation to be used to describe what we are talking about here is far from being standardized within the community. LULUC, LULUCF, LULCC, LUC, ALCC all have been used to describe essentially the same thing. 'LUC' captures the single underlying cause (land use change) that affects land cover as well as soil properties. Anthropogenic land cover change is in this sense the result of land use change. Furthermore, we used 'LUC' in our previous publication (Stocker et al., 2014), GC13 used 'LUC' as well, and it is practical for being short (relevant for many equations where 'LUC' appears in superscripts and subscripts).

l. 7 (and elsewhere where "secondary" is used): "primary and secondary components": I am flattered that the authors use the term "primary emissions" that we introduced some years ago, and understand that "secondary" is the logical counterpart. Still, the more common term are "indirect effects" of LULCC, see e.g. Houghton, GCB, 2013. I suggest to stick with "indirect", but if not at least the relation between indirect and secondary needs to be stated.

We used the term "secondary" as it appears intuitive and is the logical counterpart as reviewer 1 states. To be clear about what this refers to, we use the term consistently throughout the manuscript. E.g., slightly re-formulated text in the abstract:

We establish how coupled Earth System Models may be applied to separate **secondary** component fluxes of *e*LUC [...]

And in Section 1:

SM08 distinguished between primary emissions that capture the direct effects of land conversion, and secondary effects arising from the interaction of land conversion and environmental change (CO_2 and climate). SM08 further separated the secondary fluxes into the *land use feedback flux* and the *replaced sinks/sources flux*. We term these *e*LFB and *e*RSS, respectively, and provide definitions in Sect. 3 and quantifications in Sect. 5.

And in Section 3:

[...] studies generally agree that total C fluxes induced by LUC can be split into primary emissions, $e\text{LUC}_0$, that capture the direct effects of land conversion, and secondary effects arising from the interaction of land conversion and environmental change (CO_2, climate).

> l. 9 (and elsewhere): "up to 20% when quantified from ESM versus offline vegetation models": Arora and Boer 2010 estimate the same difference, of D3 vs E2, to be more on the order of 100% (i.e., estimates for D3 are twice as high as E2). The manuscript would benefit very much from a more quantitative comparison against published estimates, both the individual studies and the summary by Pg14. For this specific line, a "in our modeling framework" has to be added at least.

As mentioned above, we have added quantitative comparisons to these previous studies. We note here that no previous publication applied a setup to exclude differences of ESM and offline vegetation model quantifications arising from using observational climate and CO_2 in the offline steup vs. interactively simulated climate and CO_2 in the coupled setup. To be more specific about this point, we reformulated the respective passage in the abstract:

Here, we review differences of *e*LUC quantification methods and apply an Earth System Model (ESM) of Intermediate Complexity to quantify them. We find that the magnitude of effects due to merely conceptual differences between ESM and offline vegetation model-based quantifications is ~20% for today. Under a future business-as-usual scenario, differences tend to increase further due to slowing land conversion rates and an increasing impact of altered environmental conditions on land-atmosphere fluxes.

> l. 15 ff: "Therefore, we argue that synthesis studies and global carbon budget accountings should resort to the "least common denominator" of different methods, following the bookkeeping approach where only primary land use emissions are quantified under the assumption of constant environmental boundary conditions.": See comment on "opinion piece" above. The choice of definition depends first and foremost on the application. For example, for global C budget estimates I argue that the prime aim is to keep the C budget closed and not attribute processes we know of to the residual (error) terms, thus the choice of eLUC definition needs to be made in agreement with the model setup for the residual terrestrial sink.

We agree that the choice of definition depends on the application. This is expressed in Section 5 where we distinguish between "Carbon budget accounting" and "LUC in the Earth System". In the context of the global C budget, it may also be argued that process, that are not directly due to LUC and that we have relatively little understanding of (CO_2 fertilisation and climate impacts on C cycle) should be ascribed to the residual term (Houghton, 2013). We refer to this point in Section 5.1, but we now try to refrain from being prescriptive in the choice of methods, while still expressing our (to some extent subjective) view. See also response above.

Introduction:
l. 58 ff: The study by McGuire et al accounted for climate/CO2 changes due to both LULCC and fossil-fuel emissions, not just "LULCC-induced environmental change".

We don't state that McGuire et al., accounted for "just LULCC-induced environmental change". We added "observed" to clarify this (this implies that it's the result of all anthropogenic impacts, including fossil fuel emissions):

The first such study using a set of process-based vegetation models with prescribed, transiently varying climate and CO_2 from observed historical data was presented by McGuire et al. (2001).

l. 83-84: "They state that the discrepancy between methods stems from the inclusion of the land use feedback on actual natural land." Should rather read "...stems predominantly. . .". As can be seen from Pg14 Fig. 2, E2 and D3 also differ in terms of fluxes on potential natural vegetation and synergy fluxes.

Done, added "predominantely".

l. 90-98: It sounds as if the present study would quantify all possible methods, while indeed you focus on D1, D3, and E2. Please clarify this (and justify why you deem consistent quantification of these the most important).

As mentioned above, we limited our analysis to differences between D1, D3, and E2 as these are widely used methods, applied by process-based models (DGVMs) and bookkeeping models. Focussing on these three methods is illustrative for investigating differences implied by how environmental effects on land-atmosphere fluxes are accounted for and serves to support our conclusions within the limits of a ESD Short Communication. We try to point this out by added/reformulated text at the beginning of Section 2 (see also response above).

l. 115: "total CO2 emissions from land use change": misleading, better is "total C fluxes induced by LULCC" as also sink terms (e.g. the feedback) are included.

Done.

l. 120 ff and eq. 3: The replaced sources/sinks term as defined in this study indeed just refers to changes in environmental conditions by causes other than LULCC, but as discussed in Pg14 this term (under this or the other names, loss of additional sink capacity or land use amplifier effect) may also refer in published studies to LULCC-induced changes, or both. This needs to be clarified. In the definition used in eq.3 it also comprises the effect of environmental changes on instantaneous emissions and legacy fluxes (the delta-f-I and delta-f-L fluxes in Pg14), which is worth clarifying because this is not obvious and because it makes comparison with PG14 easier. Similarly, state that the last term in eq. 4 includes delta-l-I and delta-l-L.

Regarding the definition of eRSS: In PG14, the definition of this flux ("Loss of Additional Sink Capacity") is given as $\delta(E_m - E_p)$ (their Eq. 8). It is not entirely clear, if their δ is a δ_l, δ_f, or $\delta_{l,f}$ (environmental change caused by LUC, FF, or both). Translated into our formalism, this is $\Delta A(\Delta f_{\mathrm{agr}} - \Delta f_{\mathrm{nat}})$. If this is to be interpreted as $\Delta A(\Delta f_{\mathrm{agr}}^{\mathrm{FF+LUC}} - \Delta f_{\mathrm{nat}}^{\mathrm{FF+LUC}})$, then their definition is also not identical to the one applied in SM08 ($e\mathrm{RSS}_{\mathrm{SM08}} = \Delta A\,(\Delta f_{\mathrm{agr}}^{\mathrm{FF+LUC}} - \Delta f_{\mathrm{nat}}^{\mathrm{FF}})$ and the one used here. However, these discrepancies are comparatively small, as indicated in new Figure 1 (dotted lines). SM08 provide a more specific definition of the eRSS flux. We now provide an explicit discussion of these differences. Modified/extended text regarding eRSS reads:

In spite of the variety of terminologies presented in the published literature, studies generally agree that total C fluxes induced by LUC can be split into primary emissions, $e\mathrm{LUC}_0$, that capture the direct effects of land conversion, and secondary effects arising from the interaction of land conversion and environmental change (CO_2, climate). However, the exact delineation between secondary emissions eLFB and eRSS differs (Strassmann et al., 2008; Pongratz et al., 2009, 2014). Here, we chose a definition so that eRSS arises due to environmental changes (e.g. CO_2, climate, N-deposition, ozone, air pollution, etc.) that are not caused by LUC, whereas eLFB is due to environmental changes driven by LUC. According to Eq. 8 and for a reference state without land under use, eRSS can be interpreted as the difference in sources/sinks between land under potential natural vegetation ($\Delta f_{\mathrm{nat}}^{\mathrm{FF}}$) and agricultural land ($\Delta f_{\mathrm{agr}}^{\mathrm{FF}}$) and scales with the area of land converted ΔA. The LUC-feedback flux eLFB (Eq. 9) describes the flux arising as a consequence of LUC-induced environmental changes (e.g., CO_2, climate change). eLFB occurs on non-converted (natural) and converted (agricultural) land, with different sink strength ($\Delta f_{\mathrm{nat}}^{\mathrm{LUC}}$ and $\Delta f_{\mathrm{agr}}^{\mathrm{LUC}}$). To sum up, eRSS arises from secondary effects of fossil fuel emissions (and N deposition, etc.), whereas eLFB is driven only by LUC. This is reflected by the fact that only superscript 'LUC' occurs in the definition of eLFB, whereas only 'FF' occurs in the definition of eRSS. The definitions of eRSS, and hence of eLFB differ slightly between publications (Strassmann et al., 2008; Pongratz et al., 2014). SM08 defined eLFB so that this flux only occurs on remaining natural land. Specifically, the term ($\Delta A\,\Delta f_{\mathrm{agr}}^{\mathrm{LUC}}$) appears in eLFB here, while it is ascribed to eRSS in SM08. However, this flux component is relatively small (see Fig. 1). As indicated by PG14, eRSS may also be defined as $e\mathrm{RSS} = \Delta A\left(\Delta f_{\mathrm{agr}}^{\mathrm{FF+LUC}} - \Delta f_{\mathrm{nat}}^{\mathrm{FF+LUC}}\right)$, implying that $e\mathrm{LFB} = A_0 \Delta f_{\mathrm{nat}}^{\mathrm{LUC}}$. Our choice of eRSS and eLFB has the advantage that it follows an intuitive separation between underlying environmental drivers (FF vs. LUC) and that eLFB can identically be separated in coupled ESM-type simulations where the FF forcings are excluded. This corresponds to the E1 definition in PG14, with $e\mathrm{LUC}_{\mathrm{E1}} = F_{\mathrm{LUC}}^{\mathrm{LUC}} - F_0^0 = e\mathrm{LUC}_0 + e\mathrm{LFB}$, and was applied by Pongratz et al. (2009) and Stocker et al. (2011).

Regarding where legacy vs. instantaneous fluxes are subsumed, we now write in Section 2.1 (description of D1):

$e\mathrm{LUC}_{\mathrm{D1}}$ is equivalent to primary emissions (see Section 3) and captures instantaneous CO_2 emissions occurring during deforestation and C uptake during regrowth, as well as delayed (legacy) emissions from wood product decay and the gradual readjustment of soil and litter C stocks to altered input levels and turnover times.

And in Section 3 we write

f^0 are direct emissions in response to land conversion under constant environmental conditions and comprise instantaneous and legacy fluxes as identified by I_u and L_u in PG14; $\Delta f_{\mathrm{agr}}^{\mathrm{FF+LUC}}$ is its modification due to environmental change (δI and δL in PG14). Note that on long time scales, the cumulative flux of ($\Delta f_{\mathrm{agr}}^{\mathrm{FF+LUC}}$) is independent of the magnitude of f^0.

> l. 123: It would help many readers if you clarify that the "natural" in "natural land" refers to the vegetation types, not to the land area (which is agricultural). The term "potential natural vegetation" Pg14 adopted from the geography community was meant to do exactly this. In this context it won't harm to explicitly state that eLFB affects both natural and agricultural vegetation (l. 122).

We accounted for this suggestion. This text passage is now slightly re-worded anyway. It reads:

According to Eq. 8 and for a reference state without land under use, eRSS can be interpreted as the difference in sources/sinks between land under potential natural vegetation ($\Delta f_{\mathrm{nat}}^{\mathrm{FF}}$) and agricultural land ($\Delta f_{\mathrm{agr}}^{\mathrm{FF}}$) and scales with the area of land converted ΔA.

> l. 128: What do you mean by "direct"?

This text passage is now modified as well. Corresponding modified text reads:

f^0 are direct emissions in response to land conversion under constant environmental conditions and comprise instantaneous and legacy fluxes as identified by I_u and L_u in PG14 [...]

> l. 146 and elsewhere: You equate the bookkeeping approach to D1, which is what SM08 did but not what PG14 did, so best is to be explicit how you use the term "bookkeeping" in your study.

We did not use the term bookkeeping very consistently before. We now re-worded corresponding text in Section 2.1 (introducing D1), and explictly resolve the point raised by the reviewer here:

Environmental boundary conditions thus implicitly represent fixed conditions under which the observations are taken, i.e. climate, CO_2, and N-deposition levels of recent decades. Process-based vegetation models can be run in a conceptually corresponding setup ("bookkeeping method" in SM08 and thereafter) by holding environmental boundary conditions constant. While bookkeeping models are designed to derive LUC-related C emissions from a single simulation, process-based models commonly take the difference in net land-to-atmosphere carbon flux (F) between a simulation with and one without LUC:

> l. 145: "abandonment of agriculture": Which model would not include this term in eLUC?

True. Changed.

> l. 157: I believe the OSCAR model by Gasser and Ciais uses NPP, not C stocks, as input.

Because of this, their model is not strictly-speaking a bookkeeping model. After all, they account for a modification of NPP and heterotrophic respiration by climate, CO_2, and N-deposition. Thus they are mimicking (emulating) a process-based model, while keeping track of the fate of C in different cohorts (which makes it a bookkeeping model in some sense). We decided to stick with the formulation we have.

Sec. 2.2: Most of this section is rephrasing from Pg14. As stated above, this is confusing. I suggest to shorten the implications parts of this section and refer to the corresponding equations in SM08 and Pg14.

We re-arranged contents in the text. However, some rephrasing of PG14 cannot be ommited as we need to introduce the methods in order to investigate them later on. These contents are now moved to Section 2 and clear reference is given to PG14 at the beginning of this section:

We start by revisiting the classification of PG14 for a subset of *e*LUC quantification methods identified in their study.

We also state in Section 2.3 (introducing E2) that the systematic differences between methods has been noted before:

As noted also in earlier publications (Strassmann et al., 2008; Arora and Boer, 2010; Pongratz et al., 2014), here, in contrast to the D3 method, environmental conditions in the LUC and non-LUC simulation differ. [...]

Finally, the following point has never been stated clearly and explicitly before:

In other words, while primary emissions $e\text{LUC}_0$ can be consistently derived from offline DGVMs by simply holding environmental conditions constant, the secondary fluxes derived from such studies are neither equal to *e*RSS, nor *e*LFB, nor the sum of the two. In other words, *e*RSS and *e*LFB cannot be separated as shown here using offline vegetation models.

The conceptual analysis we provide allows us to concisely describe the difference between E2 and D3 quantifications. Particularly after the publication by GC13, readers may have been left with the impression that their indirect (or secondary) fluxes ΔELUC and ΔLSNK are the same as *e*RSS and *e*LFB. But they aren't.

Eq. 8 ff: State explicitly that A0 refers to potential and actual natural vegetation. Further, the statement "Unlike suggested by PG14, the formal treatment presented here reveals that the difference is related to the land use feedback for the reference distribution of natural land and not the actual distribution." is wrong. If you subtract the fluxes of E2 and D3 in Fig. 2 of Pg14 one can easily see that exactly the fluxes due to LUC-induced changes in environmental conditions on land of potential natural and actual natural vegetation remain. Eq. 13b and 15c in Pg14 show the same. Pg14 even show that there is an additional synergy term by which the two methods further differ. The formalization used by Pg14 may differ to the one presented in this study, but I don't see any disadvantage thereof and cannot see any effect presented in the present study that has not been properly described in SM08 and Pg14 already. The key difference is that in the present study you refer to one very specific modeling framework, which allows you to eliminate much of the ambiguity that the land use feedbacks and the replaced sources sinks/loss of additional sink capacity/amplifier effect contain across earlier studies using different frameworks (discussed in Pg14 Sec. 4.1, 4.2 and 4.5). If you shorten this section, reference earlier work and clarify that you look at these two terms under a very well defined modeling framework then this section would not be confusing and set the basis well for the following analysis.

In Section 3 ("Defining flux components") we now state

A_0 is the initial (reference) area.

With regards to the choice of the reference state, we now provide additional clarification, e.g.:

F_0^0 is the land-atmosphere flux in the reference state, which may either be forced with the land use distribution at the beginning of the transient simulation (year 1700 here, see Section 4) or zero anthropogenic land use.

And

Note that the reference state may also include agricultural land. But under commonly used model setups, the extent of agricultural land in the reference state is small in comparison to the area under natural vegetation.

The decomposition of $(\Delta f^{\mathrm{FF+LUC}})$ into (Δf^{FF}) and $(\Delta f^{\mathrm{LUC}})$ indeed allows us to cancel terms and arrive at short and simple expressions, e.g., for the difference between $e\mathrm{LUC}_{\mathrm{D3}}$ and $e\mathrm{LUC}_{\mathrm{E2}}$ or expressing fluxes as a sum of component fluxes (Table 2) or proving that component fluxes can be isolated by a sum/difference of total fluxes in different simulations (Eq. 13 and 14). We consider this to be a clarifying insight, not expressed as such in SM08 and PG14. The underlying linearity assumption is now assessed by new Fig. 1.
Regarding our contention "Unlike suggested by PG14, the formal treatment presented here reveals that the difference is related to the land use feedback for the reference distribution of natural land and not the actual distribution." This statement was based on the sentence in PG14 that reads "A larger discrepancy between methods, however, stems from the inclusion of the land use feedback *on actual natural land* $(\delta_l E_n)$,". As we now understand, we misunderstood this. Indeed, subtracting 13b and 15c in PG14 and using $\delta_{l,f} = \delta_l + \delta_f + \sigma_{l,f}$ yields a difference between D3 and E2 of $(\delta_l + \sigma_{l,f})\ (E_n + E_p)$. Translated into our formulation, this is equivalent to $(\Delta f_{\mathrm{nat}}^{\mathrm{LUC}} + \delta)A_0$, the same as we identified as the difference. Therefore, we modified the text describing this difference, and emphasize that the difference is not equal to eLFB as defined here.

Ignoring the non-linearity term δ_{nat}, the discrepancy can thus be interpreted as a flux, triggered by environmental changes caused by LUC, but occurring on land not converted since the reference period $(f_{\mathrm{nat}}^{\mathrm{LUC}})$. Note that this is not identical to eLFB as defined here. The same theoretical result can be found when applying the formalism of PG14 and their definition of flux components in $e\mathrm{LUC}_{\mathrm{E2}}$ and $e\mathrm{LUC}_{\mathrm{D3}}$, with the difference turning out to $(\delta_l + \sigma_{l,f})(E_n + E_p)$.

And added text at the beginning of Section 3.1:

SM08, PG14, and GC13 establish a formalism to describe and discuss the different definitions of total eLUC and its component fluxes. Here, we synthesize these previous frameworks to a minimal description that allows us to identify the different flux components contained in eLUC provided by the offline DGVM setups (D3 method), coupled ESM model setups (E2 method), and the bookkeeping aproach (D1 method). We then show that $e\mathrm{LUC}_{\mathrm{E2}} = e\mathrm{LUC}_0 + e\mathrm{RSS} + e\mathrm{LFB}$ plus nonlinearities. We propose a definition for the delineation between component fluxes that follows a separation along underlying drivers of environmental changes, and that allows a consistent identification of component fluxes in coupled model setups with and without the FF forcing. The formalism presented below sets the basis for the analysis and discussion in subsequent sections.

Following GC13, [...]

However, we consider the formalistic description of flux components and differences between methods necessary here to concisely describe what our modelling analysis shows. Furthermore, the distilled formalism presented in our manuscript serves to emphasize three key points that are, in our reading, not clearly expressed in P14 and GC13, and therefore worth pointing out:

- $e\mathrm{LUC}_{\mathrm{E2}} = e\mathrm{LUC}_0 + e\mathrm{RSS} + e\mathrm{LFB}$ (plus non-linearities δ specified in the manuscript). Note that delineation between eLFB and eRSS chosen by SM08 differs slightly from ours, as now pointed out in this Section (3.1)).
- The secondary flux terms quantified from D3 are not identical to eRSS or eLFB, nor to their sum: $e\mathrm{LUC}_{\mathrm{D3}} - e\mathrm{LUC}_0 \neq e\mathrm{RSS} + e\mathrm{LFB}$
- The difference between E2 and D3 quantifications is $\Delta f_{\mathrm{nat}}^{\mathrm{LUC}} A_0$ (plus non-linearities δ specified in the manuscript)

Methods:
l. 204 ff: Pre-industrial environmental conditions for the year 1700 are a pretty wild mix of 1700-CO2 and early-20th-century climate. The model shows substantial sensitivity of C stocks to environmental conditions (shown in l. 328), which I would expect to stem not only from difference in atmospheric CO2, but also in climate.

Using a recycled early-1900-CRU climatology for the pre-1900 period is standard practice and is applied identically in TRENDY simulations for the Global Carbon Project. To explain our choice of an early start of the transient LUC simulations, we now state:

We focus on results after 1800 but chose an early start of the transient simulation (1700) in order to minimise effects of the initial equilibrium assumption for LUC-related fluxes.

l. 210: Since the community knows and appreciates your earlier studies at other resolutions, could you quantify what the change in resolution means for eLUC for easier comparison to your other studies?

We added a reference to the Results section here, stating:

This has negligible effects (see Section 5).

Indeed, the difference is negligible: 165 GtC (Stocker et al., 2014) vs. 164 GtC (here). We added text in the Results section:

Stocker et al. (2014) applied the same model at a $1^\circ \times 1^\circ$ resolution following the D3 and D1 methods to quantify "total" and "primary" LUC emissions. Results at the finer resolution (165 GtC for "total GNT" in their Table 3) are virtually identical to the present estimate.

I cannot find the description of the setup for the bookkeeping approach.

References to relevant equations are given (Eqs. 2, 3, and 4).

l. 223: "As outlined in Appendix I": It is not obvious where eRSS is derived in the Appendix.

We re-structured these contents. Contents from appendices are included in the main text and tables, including new Table 2 (Flux decomposition for model setups described in Table 1). We modified text to describe how we get to Eqs.12 and 13:

The model is run in a set of simulations that allows us to disentangle flux components eRSS and eLFB. Using the description of decomposed fluxes given in Table 2 and the definition of eRSS in Eq. 7, the replaced sinks/sources flux component can be derived from simulations described in Table 1 as [...]

And analogously for eLFB.

> l. 226: Here is another intuitive way to understand eRSS: $((F_{\mathrm{LUC}}^{\mathrm{FF+LUC}} - F_0^{\mathrm{FF+LUC}}) - (F_{\mathrm{LUC}}^{\mathrm{LUC}} - F_0^{\mathrm{LUC}}))$, which is eq. 9 assuming linearity $F_0^{\mathrm{FF+LUC}} = F_0^{\mathrm{FF}} + F_0^{\mathrm{LUC}}$. This is the difference between with- and without- LULCC simulations under total vs only LULCC-induced environmental changes.

We thank the reviewer for this hint. We included this point:

Alternatively, eRSS can also be derived as $(F_{\mathrm{LUC}}^{\mathrm{FF+LUC}} - F_0^{\mathrm{FF+LUC}}) - (F_{\mathrm{LUC}}^{\mathrm{LUC}} - F_0^{\mathrm{LUC}})$, wich is formally identical to Eq. 15, assuming linearity.

> Results: l. 236: "~23%": Compare to earlier work, e.g. Arora and Boer 2010, and discuss model-dependence of the estimates.

We added text:

SM08 applied observational CO_2 and climate in simulations used for D3. They found slightly higher differences of D3 vs. E2 (30% higher in their D3). Arora and Boer (2010) report a difference of ~100% for a case where they only used CO_2 concentrations from their interactive $F_{\mathrm{LUC}}^{\mathrm{FF+LUC}}$ to force their $F_0^{\mathrm{FF+LUC}}$ simulation. A stronger effect in this case appears plausible as the replaced sinks/sources flux due to climate and CO_2 effects are generally opposing (Strassmann et al., 2008).

Unfortunately, Arora and Boer (2010) provide no numbers for their simulation given by the purple line in their Fig. 14 ("LUC emissions obtained by driving CTEM offline with specified CO_2 and reanalysis data").

> l. 243: I am confused: Why do you cite CRU climate here when you prescribe ESM climate to your simulations? (And does the 364 ppm refer to your ESM results or to observations?)

ESM climate in Bern3D-LPX is simulated using a pattern scaling approach. Ths means that temperature, precipitation, and cloud cover are the sum of a baseline-climatology (recycled CRU 1901-1931) plus an anomaly patterns scaled by the global mean temperature anomaly relative to a reference state (preindustrial). Primary emissions are simulated without accounting for the scaled anomaly component; only the baseline climatology is applied, thus the model may also be run offline. To clarify this, we added text:

During the transient simulation (1700-2100), climate is simulated by adding an anomaly pattern, scaled by global mean temperature change relative to 1700, to the continuously recycled CRU climatology (temperature, precipitation, cloud cover).

> l. 263: "cumulative total emissions": Clearer to speak of "cumulative net LULCC/LUC flux". What do the percentages refer to, i.e. 6 and 21% relative to which method?

Flux components $e\mathrm{LUC}_0$, eRSS, and eLFB add up to $e\mathrm{LUC}_{\mathrm{E2}}$. Relative contributions are therefore w.r.t. $e\mathrm{LUC}_{\mathrm{E2}}$. To clarify this, we modified text at the beginning of this paragraph:

Figure 3 illustrates the different flux components of total emissions from LUC following the E2 method [...]

> l. 264: "total (eLUCE2) and primary emissions": Why "total"? Might "net" be the better term again? Add "(eLUC0)" to "primary emissions".

Added "(eLUC$_0$)". "Total" appears intuitive in the sense that it is the sum of eLUC$_0$, eRSS, and eLFB. "Net" is just as ambiguous because it may refer to a flux after accounting for regrowth, or after accounting for eLFB, or be used in the context of gross land use transitions vs. net land use transitions, as used, e.g., in Stocker et al. (2014).

> **Discussion**:
> See comments above. Much of the discussion in Sec. 5.1 is not new over previous studies. In particular here, Houghton, GCB, 2013 argues the same way as you do in l. 308 -310; l. 313 f (implications for the residual terrestrial sink) has been discussed in Pg14.

As pointed out above, we wanted to highlight additional points not expressed explicitly in either of these earlier publications (see also other responses).

> l. 318: "The inclusion of secondary LUC fluxes [...] in eLUC and in turn in estimates of the implied residual sink is misleading when comparing to observational data": This does not seem to be a valid argument since attributing C fluxes to the net LULCC flux as compared to the residual terrestrial sink is not doable on the large scale based on observations and always requires modeling.

We modified the discussion on this point. New text reads

The D3 method ascribes replaced sinks/sources (eRSS) to eLUC. This implies that the residual terrestrial sink represents a flux occurring in a hypothetical state before land conversion. This may be misleading in view of the actual reduction of land C sinks due to the reduction of natural vegetation. This reduction of the residual sink due to the replacement of natural by agricultural vegetation is only captured when basing its quantification on D1-type eLUC estimates.

> l. 334: "quantifying eLUCD1 under preindustrial conditions is a viable and pragmatic solution": I agree that using 1700-climate is a pragmatic solution, but what is the evidence that it is a viable one? Given the long history of LULCC, the choice of any time period is a subjective one; e.g. mid-Holocene conditions or Little Ice Age temperatures would give different results but could equally be justified. Wouldn't the only objective decision be to use environmental conditions from the time land use change first emerged? Similarly the definition of "present-day" is subjective; LULCC emission quantifications will likely be performed 10 years from now as well, when the atmospheric CO2 concentration will have substantially risen.

The argument for using eLUC$_{D1}$ under preindustrial conditions being viable is based on the analysis given in the preceeding sentences: "the case where C stocks are responding to transient changes in CO_2 and climate (F_{LUC}^{FF+LUC} – the closest analogue to what observational data represent) is farther from its equilibrium to be attained under present-day conditions than its equilibrium under preindustrial conditions."

> l. 345 ff: “In summary, we recommend not to rely on results from method D3 or E2 in the context of the global (or regionalized) carbon budget, but to apply method D1 (un- der preindustrial conditions).” I agree with recommending D1 for deriving net LULCC flux estimates from multi-model studies when the aim is to narrow down the model spread, because it excludes some of the highly model-dependent processes such as CO2-fertilization strength. But as mentioned above for global C budget estimates the method for eLUC depends on the method chosen for the residual terrestrial sink, if an independent method is used for the latter, to close the C budget. If the present-day residual terrestrial sink is derived as residual from coupled simulations and eLUC es- timates, then fluxes induced by environmental changes since the preindustrial era, in particular higher emissions due to higher standing biomass stocks, are attributed to the residual terrestrial sink if eLUC is derived from D1 simulations. This is so counterintuitive that I would recommend method D3 instead (it also disagrees with the authors' argumentation that D1 represents observable processes: observable changes in C stocks upon deforestation include effects of higher standing biomass). My point is: Such a recommendation is subjective and unless the editor decides that this is an opinion piece I would refrain from any subjective recommendations. I recommend that the authors restrict the text to the discussion on the advantages of method D1 in terms of ease of setup in common MIPs and narrowing down of the model spread, which are undisputable and relevant points. Then the limitations of D1 as e.g. above should shortly be mentioned, but for longer discussions of the purpose of different methods the reader can be referred elsewhere.

Observable C stocks are observable under present-day conditions. When using this information for estimates of *e*LUC, one arrives at a D1-PD-type quantification. We completely agree with the reviewer that in year 1700, “observable” meant preindustrial conditions, and one would arrive, using the same argument, at a D1-PI quantification. The point we are trying to make here, is that using transiently varying conditions is not a realistic compromise between D1-PI and D1-PD, as it includes secondary fluxes that drastically affect the trend of emissions (D1-PI and D1-PD have the same trend!) and imply a conceptual definition of the residual sink that may be misleading (see response above).
Note however that we dropped the sentence “In summary, we recommend not to rely on results from method D3 or E2 in the context of the global (or regionalized) carbon budget, but to apply method D1 (under preindustrial conditions).”

> Sec. 5.2: I am not sure what the purpose of this section is. To state that C fluxes are just one of many effects of LULCC on the Earth system? But then the effects in paragraph 1 seem a bit random. What about biogeophysical effects, or other greenhouse-gas-related management effects such as wetland management effects on CH4? The discussion of paragraph 2 on eRSS and eLFB is repetitive from earlier sections. See comment before (“opinion piece”) on the recommendation of l. 370 f.

The purpose of Section 5.2 is described at the beginning of Section 5: “Below, we outline two different perspectives on what “emissions from LUC” may represent and discuss their methodological implications.” Biogeophysical effects are mentioned (“Vegetation cover change affects the local surface energy and water balances.”). In response to the reviewer's comment, we modified/added text:

Vegetation cover change also affects the local surface energy and water balances (biogeophysical effects) and emissions of other greenhouse gases. Deforestation by purposely set fires is associated with emissions of a range of radiatively active compounds (e.g., CH_4, CO, NO_x), wetland management may have strong effects on CH_4 emissions, and the application of mineral fertiliser and manure, [...]

l. 372 f: "We argue that offline-vegetation model setups are not capable of separating eRSS and eLFB as defined here.": No need to argue, this is a fact.

Thank you.

Conclusions:
l. 377: See comments above on "recommendations". Again, who defines what "preindustrial" means?

Please note that here, we make this recommendation "in order to guarantee comparability and continuity". A preindustrial control simulation is a well-defined CMIP standard and the term is widely used as a generic concept. An exact definition for what "pre-industrial" means in our simulations is provided in the Sect. 4 (Methods) of the present manuscript. In our view, using this term in the conclusions does not necessitate an exact specification.

Appendix:
The equations are all right and good, but I would structure them in a more helpful way: Show with the equations on the one hand which simulations need to be subtracted from each other to isolate individual flux components (basically the information of Tab. 2), and on the other hand how each of these components would be expressed in the area-flux notation. For example, it is not immediately clear that l. 117 refers to ePS, l. 420 to eLUC0, or l. 425 to eLUC0 + eLFB. Derivation of eRSS seems to be missing.

In new Table 2, we now provide the information of how each of the total land-atmosphere CO_2 exchange flux would be expressed in the area-flux notation and how it can be written as a sum of the flux components we have identified (*e*PS, *e*RSS, *e*LFB, and *e*LUC_0). Equations for which simulations have to be subtracted in order to isolate flux components are given by Eq.12 and Eq.13 in the main text of Section 4. Contents previously in appendices are thus integrated into the main text now, also in response to the suggestion by reviewer 2 to provide a clear mathematical description of the formalism.

l. 412: Add again the assumption about linearly adding fluxes.

This is now expressed slightly differently and the non-linearity assumption is clearly expressed and validated by Fig. 1. New/modified text in Section 3 reads:

Figure 1 reveals that global fluxes due to FF and due to LUC forcing alone combine in an almost perfectly linear fashion to the flux induced by the combined effect of FF and LUC up to present and discernible deviations (δ) emerge only in a future scenario of continuously rising CO_2 and changing climate and contribute ~10–20% by 2100 in RCP8.5.

l. 418-419: "the land does not "see" any changes in climate and CO2 (no fossil fuel emissions)": Confusing. How about "A run with LUC but with prescribed environmental conditions unaltered by LULCC and fossil-fuel burning"?

This is now expressed differently. Table 2 now provides this information with the model setups concisely described by F_i^k.

Figure and Tables:
Fig. 2 looks very similar in content and numbers to Fig. 8 by SM08. Given that also much of the discussion in the manuscript is rephrasing SM08 and Pg14 I feel the current form of the manuscript is really overselling the novelty of this study.

We hope that our responses and edits convince the reviewer about the merits of our study.

Typos etc:
l. 8: Introduce abbreviation "ESM" here (used later).

Done.

l. 13: Explain abbreviation "C" here. It seems a bit arbitrary at which occasions the authors use "C" and at which "carbon" throughout the manuscript.

Done.

l. 99: quantification should be singular.

Respective text was changed.

l. 110: ST08 should be SM08; "of" missing after "definitions".

Respective text was changed.

l. 113 and elsewhere: Be consistent: book-keeping or bookkeeping? (I believe the latter)

Done.

l. 121: "which" should be "that".

Both are correct. Still, we changed it.

l. 137: typo in deposition

Done.

l. 149: "is" missing

Respective text was changed.

l. 160: remove "for"

Respective text was changed.

l. 372: Remove hyphen in "offline-vegetation model setups" (the setup is offline, not the vegetation).

Done.

l. 391: typo in following

Respective text was changed.

Response to Reviewer 2

The submitted manuscript by Stocker and Joos investigates differences in anthropogenic land use and land-cover change (LULCC) emissions (eLUC) arising from different methodologies in the literature and presents a case with stand-alone DGVM and a coupled model. The study is a step in the right direction and will help to resolve some existing confusion in the LU literature, but it still needs a sharper focus and clarity. It attempts to cover both a methodological discussion (i.e. different definitions) and an analysis of differences in eLUC estimates in stand-alone and fully coupled models. The current conceptual scope of the former is too idealized and it does not help to understand in depth the latter. I would recommend to expand the analysis of the simulations and to downscale the discussion of the flux definitions to only those relevant to that analysis.

We would like to thank the anonymous reviewer for the time spent on reviewing our manuscript and the very helpful comments. We took several measures to account for the above mentioned points:

- We have expanded the analysis of our model results. Specifically, we demonstrate the validity of the crucial linearity assumption that underlies much of our formalism ($\Delta f^{\mathrm{FF+LUC}} = \Delta f^{\mathrm{FF}} + \Delta f^{\mathrm{LUC}}$, see new figure 1). Furthermore, we included the new Fig. 4 to show maps of the C sink/source on natural and agricultural land and of the non-linearity. In the response document below we also provide maps showing the flux components ($e\mathrm{LUC}_{\mathrm{E2}}$, $e\mathrm{LUC}_0$, eRSS, eLFB).
- We rearranged the text to include the appendix describing the concepts in the main text to better help the reader understand the concepts applied and support the interpretation of our results.
- We restrict the discussion of model setups (D1, D3 and E2) and flux components ($e\mathrm{LUC}_0$, eRSS, eLFB) to those quantitatively analysed in the results section using our model.
- Table 1 and new table 2 provide a description of model setups and flux components. This is a minimum number of model setups, necessary to disentangle these differences and component fluxes treated in our paper.
- The level of complexity chosen here is very comparable to the one in PG14 (no time, no space, only distinction between natural and managed land). We added discussion on limitations of our formalism (e.g., choice of the reference state) and on interactions of interannual climate variability and eLUC derived from differencing modelled land-atmosphere fluxes.

  My first criticism is that the manuscript needs a cleaner presentation of the mathematical formalisms relevant to analysis of the experiments (may be in an appendix):

  - It does not present mathematical equations used to produce figures 1-3 and table 3, just conceptual definitions. How such definitions are used for the cases with heterogeneous and time-varying LULCC? Can they be as easily linearized?

To clarify this, we now provide additional information in the figure captions. For example added text in caption of Fig. 2:

Time series are calculated following Eqs. 2-4, where F is the global total land-atmosphere CO_2 flux in the respective simulation.

Added text in caption of Fig. 2 reads:

Time series are calculated following Eqs. 2, 4, 13, and 14.

Added text in caption of Tab. 2 reads:

Note that fluxes F generally refer to global totals for a given point in time t. Thus, for example $F_0^{\mathrm{FF}}(t) = \int_{x,y} A_0(x,y)\,\Delta f_{\mathrm{nat}}^{\mathrm{FF}}(x,y,t)\,dx\,dy$. For simplicity, we have dropped the time and space dimensions.

Table 1 and new table 2 together now provide the necessary information of how the model is set up in each simulation, and how total fluxes in each of these setups can be decomposed. The linearity assumption is now explicitly assessed (see new Figure 1).

- It also would help to state from the beginning if the formalisms refers to cumulative or net fluxes. It appears that figures show the net fluxes but the methods section states that the equations 5, 6, and 7 compute cumulative CO2 emissions from land use change as a difference in terrestrial C storages.

This was described wrongly in our manuscript ("Cumulative CO_2 emissions from land use change are calculated as the difference in terrestrial C storage [...]"). This is inconsistent with the formalism and calculations we present and has probably caused confusion here. We modified respective text in Section 4:

LUC-related CO_2 emissions are calculated as the difference in the land-atmosphere CO_2 exchange flux between the simulation with and without LUC using Eq. 2 for the bookkeeping, 4 for the coupled, and Eq. 3 for the offline setup.

All equations are valid irrespective of whether they describe annual fluxes or cumulative fluxes. Generally, we describe F to the be land-atmosphere CO_2 exchange **flux**, i.e. not cumulative. Also Figures show annual fluxes. We added text in Section 2.1 (Introduction of D1) reads:

In general, F refers to a global annual flux, but equations provided here are valid also for cumulative fluxes and smaller spatial domains.

Wherever we refer to cumulative fluxes, this is clearly expressed (Captions of Figures and Table 3, Section 5).

- Furthermore, it would be useful to include a list of all mathematical terms and what experimental setups they represent. There are a number of Fs with different sub- and super-scripts and it's hard to follow the equations without having all notations in one place.

For a better overview, model setups and their component fluxes are now all given in Table 1 and new Table 2. Mathematical terms used to describe model setups, areas, and flux components are described in the text and in captions of Table 1 and 2. E.g., in Section 3, where we introduce the formalism to describe component fluxes, we write:

F denotes again a carbon flux (e.g. in GtC yr^{-1}), f a carbon flux per unit area, and Δ a change with respect to the reference period/start of the simulation. Superscripts '0', 'LUC', and 'FF' refer to the driver of environmental conditions: no forcing, emissions from LUC, and fossil fuel plus other non-LUC forcings. Subscript 'agr' refer to converted land and subscripts 'nat' to land that has not changed its status over the course of the simulation. ΔA is the total area that has been converted, e.g., from natural to agricultural, up to the point in time of interest. A_0 is the initial (reference) area. $\Delta f_{\text{nat}}^{\text{FF+LUC}}$ is the change of the area-specific flux occurring on unconverted land due to environmental impacts caused by the combination of FF and LUC.

- The methods used also make a critical assumption that the environmental effects from LUC and FF combine linearly. I think the validity of this assumption needs to be demonstrated and discussed, both for local and global scales.

We added such analysis, now provided in new Figure 1. In section 3, where we use this assumption to decompose fluxes, we write:

Figure 1 reveals that global fluxes due to FF and due to LUC forcing alone combine in an almost perfectly linear fashion to the flux induced by the combined effect of FF and LUC up to present and discernible deviations (δ) emerge only in a future scenario of continuously rising CO_2 and changing climate and contribute ~10–20% by 2100 in RCP8.5.

Second, manuscript does not discuss implication of unforced climate variability for the eLUC in the coupled and stand-alone simulations. I don't think the SM08 and GC13 approach takes care of natural climate variability; it would be good to include that aspect into consideration as well.

We thank the reviewer for this suggestion. Added text in the revised manuscript addresses the issue of unforced climate variability and *e*LUC. In our formalism, the land-atmosphere CO_2 exchange flux due to unforced climate variability is F_0^0. We have added explanations to clarify this issue. In the introduction, we now write:

Internal, unforced climate variability may affect the quantification of *e*LUC as climate variability affects the land-atmosphere carbon flux F. Ideally, the model setup should be such that internal, unforced variability evolves identically in both simulations. Then the land-atmosphere fluxes from land not affected by LUC and caused by internal variability would cancel when evaluating Eq. 2. In practice, this may be difficult to achieve for some state-of-the-art Earth System Models as LUC affects heat and water fluxes and thus climate. A potential solution is to run the land module offline in both simulations or to force the land module in the simulation with LUC by using climate output from the reference simulation without LUC.

And in Sect. 2.3 addressing *e*LUC derived from coupled models, we now write:

Unforced climate variability will evolve differently in the two ESM simulations as the applied forcing is different. The component in $F_{\text{LUC}}^{\text{FF+LUC}}$ and F_0^{FF} arising from differences in internal variability will be attributed to $e\text{LUC}_{\text{E2}}$ according to Eq. 4. This misattribution could be significant in particular when considering small regions and short time scales. Ensemble simulations would be required to quantify the impact of internal climate variability on $e\text{LUC}_{\text{E2}}$. Alternatively, averaging over a large spatial domain and temporal smoothing tends to moderate the influence of unforced variability on $e\text{LUC}_{\text{E2}}$.

> Furthermore, I am not sure if it's actually possible for many current DGVMs and ESMs to compute the difference between sources on agricultural and natural lands (i.e delta fs) in the same experiment, because most models cannot separately compute physical and biogeochemical soils under agricultural and natural lands. Perhaps the authors could provide figures illustrating how delta fs compare to one another in their model, which would be fairly novel illustrations.

The model applied here, as well as other DGVMs, rely on a gridcell-tiling to separately simulate C dynamics on natural land, croplands, and pastures, affected by land conversion and environmental conditions. This is described in more details in SM08. Some models, including the one applied here, include separate gridcell tiles for primary and secondary (abandoned agricultural) land (Stocker et al., 2014), and some even distinguish between cohorts of agricultural land (GC13) or secondary land (Shevliakova et al., 2009) of different age (time after abandonment).
We appreciate the reviewer's suggestion to include results for how Δf_{nat} and Δf_{agr} compare. We now included new Figure 4 that provides this information and added text:

Secondary emissions are determined by the magnitude of C sinks and sources induced by environmental change, occurring differently on disturbed (agricultural) and undisturbed (natural) land. Fig. 4 reveals that the C sink capacity on natural land under rising CO_2 and a changing climate (year 2100, RCP8.5) is greatest in semi-arid regions of the Tropics and Subtropics and along the boreal treeline. In contrast, agricultural land at low latitudes acts as a net C source under environmental change and a net sink at high latitudes. The difference between the sink strength on natural and agricultural land is related to the *e*RSS component flux and reveals that the Tropics are the most efficient potential C sinks. Interestingly, at high latitudes, agricultural vegetation is an even more efficient C sink than natural vegetation. Fig. 4 also provides information about the spatial distribution of non-linearities from the combination of the FF and LUC forcings, corresponding to the differences between the red and the black curves in Fig. 1 in year 2100. The sum of individual effects is greater than their combination in almost all vegetated areas, but most pronounced along the transition zone between forest and open woodland. Opposite effects are simulated in individual gridcells and are likeley related to the threshold-behavior of the dominant vegetation type.

> Third, if models compute spatial fluxes why does analysis focuses only on global totals and ignores spatial details? It will be useful to go beyond global net flux trajectories, such as in figures 1 and 2, and show maps of LULCC effects. Unlike the global effect of CO2 on climate, the effect of LULCC on carbon is not global but local, and is highly heterogeneous and time varying. If the Bern model is able to compute delta f values separately for agricultural and natural lands in their simulations, they can actually clarify how changes in the C fluxes on different kinds of lands (at the core of the used formalism) relate to differences in total fluxes. Furthermore,.

Again, we appreciate the reviewer's suggestion to provide more details on spatial information. As mentioned now in the manuscript, secondary emissions (*e*RSS and *e*LFB) are determined by the magnitude of C sinks and sources induced by environmental change, occurring differently on disturbed (agricultural) and undisturbed (natural) land. This information is provided by new Figure 4. Below, we show cumulative component fluxes across space (Figures 1 and 2). However, we chose not

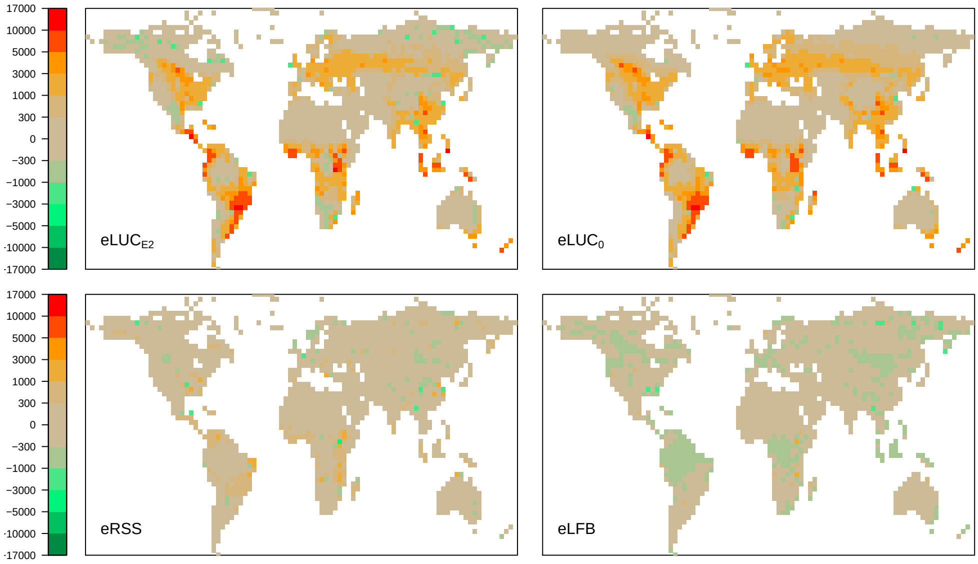

Figure 1: Cumulative component fluxes for the historical period (1850-2004). eLUC$_0$ = eLUC$_{\mathrm{D1}}$.

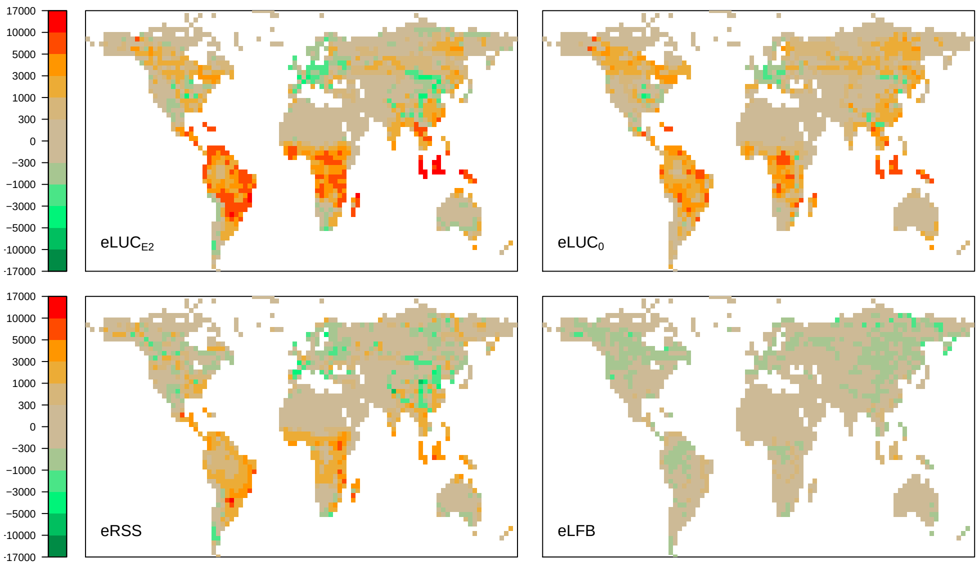

Figure 2: Cumulative component fluxes for the future period (2005-2100). eLUC$_0$ = eLUC$_{\mathrm{D1}}$.

to include this figure in the manuscript to keep the presentation of results to a minimum.

> Fourth, the used definition of the bookkeeping flux as a difference between two experiments is incorrect. The original bookkeeping approach of Houghton 83 and all subsequent Houghton's estimates compute LULCC emissions (i.e. eLUCD1 in the manuscript) only for the lands affected by LULCC in the same simulation (there is no F_0^0), not as a difference between fluxes in two experiments as presented in equation 5. The difference equation 5 was introduced in stand-alone models and EMICs studies.

This is absolutely true. We did not clearly distinguish between actual bookkeeping models and process-based models following a "bookkeeping method". Added/modified text in Section 2.1, where we describe the D1 method, reads:

Process-based vegetation models can be run in a conceptually corresponding setup ("bookkeeping method" in SM08 and thereafter) by holding environmental boundary conditions constant. While bookkeeping models are designed to derive LUC-related C emissions from a single simulation, process-based models commonly take the difference in net land-to-atmosphere carbon flux (F) between a simulation with and one without LUC:

> I personally believe that the differencing approach, even if it's the most widely used in the literature, is not a good strategy for characterizing emissions from lands affected by LULCC. The difference in total land fluxes under the method in equation 1, 6 and 7 is caused by LULCC but it does not represent emissions from lands affected by LULCC, it's a different metric. Perhaps the authors can clarify this in the results sections. Most models have to invoke differencing approach because of their technical limitations: stand-alone land models and ESMs do not keep track of belowground soil BGC pools separately on lands affected by LULCC and natural lands; as a result, cannot compute soil respiration on natural and agricultural lands. A few models do (e.g. JSBACH and MPI-ESM, GFDL's LM3 and ESMs, as well as some EMICs). In ESMs the fluxes on natural land are not the same in F^0_{LUC} and F^0_0 because of climate variability and a biophysical feedback on climate (the two are of about the same magnitude).

We agree with the reviewer in that the differencing approach implies that effects of LUC-related changes of climate variability on the land-atmosphere CO_2 fluxes are ascribed to eLUC. As noted above, we have added text discussing this aspect ("Internal unforced climate variability ..." in Sect. 2.1 and "Unforced climate variability will evolve differently ..." in Sect. 2.3). In the model used here, climate variability is not internally simulated but prescribed from the observational data (31-year baseline climatology, see Sect. 4). The differencing approach thus largely cancels this effect and what is ascribed to eLUC is only the LUC-related modification of this flux. This is also valid for other models that use prescribed climate (and climate variability). However, we note that "Emissions from lands affected by LULCC" are by definition not the same thing as "emissions attributable to LULCC". The differencing approach allows a separation of the latter by comparing a world with and a world without LUC, rigorously achieved by the E2 method, using coupled ESMs.

> While it's beyond the scope of this manuscript, generally it would be much more productive to analyze how LULCCs affect stored carbon in vegetation and soils where such LULCC are taken place – not just differences in fluxes between a simulation X and simulation Y.

Clearly, analysing LUC effects on C pools (before and after conversion) , e.g., by comparing simulations and observations has a high priority for future research. Such benchmarking activities are under way and will provide essential information to constrain models and quantify uncertainty. Here, we are restricted in space (ESD Short Communication), and we have to limit the analysis of results.

References

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

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., A., C., DeFries, R., J., G., Heimann, M., Jones, C., Le Quéré, C., Myneni, R., S., P., and Thornton, P.: Carbon and Other Biogeochemical Cycles, in: Climate Change 2013: The Physical Science Basis. Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., pp. 571–658, Cambridge University Press, 2013.

Gasser, T. and Ciais, P.: A theoretical framework for the net land-to-atmosphere CO_2 flux and its implications in the definition of "emissions from land-use change", Earth System Dynamics, 4, 171–186, doi:10.5194/esd-4-171-2013, URL http://www.earth-syst-dynam.net/4/171/2013/, 2013.

Houghton, R. A.: Keeping management effects separate from environmental effects in terrestrial carbon accounting, Glob. Change Biol., 19, 2609–2612, 2013.

Le Quéré, C., Moriarty, R., Andrew, R. M., Peters, G. P., Ciais, P., Friedlingstein, P., Jones, S. D., Sitch, S., Tans, P., Arneth, A., Boden, T. A., Bopp, L., Bozec, Y., Canadell, J. G., Chini, L. P., Chevallier, F., Cosca, C. E., Harris, I., Hoppema, M., Houghton, R. A., House, J. I., Jain, A. K., Johannessen, T., Kato, E., Keeling, R. F., Kitidis, V., Klein Goldewijk, K., Koven, C., Landa, C. S., Landschützer, P., Lenton, A., Lima, I. D., Marland, G., Mathis, J. T., Metzl, N., Nojiri, Y., Olsen, A., Ono, T., Peng, S., Peters, W., Pfeil, B., Poulter, B., Raupach, M. R., Regnier, P., Rödenbeck, C., Saito, S., Salisbury, J. E., Schuster, U., Schwinger, J., Séférian, R., Segschneider, J., Steinhoff, T., Stocker, B. D., Sutton, A. J., Takahashi, T., Tilbrook, B., van der Werf, G. R., Viovy, N., Wang, Y.-P., Wanninkhof, R., Wiltshire, A., and Zeng, N.: Global carbon budget 2014, Earth System Science Data, 7, 47–85, doi:10.5194/essd-7-47-2015, URL http://www.earth-syst-sci-data.net/7/47/2015/, 2015.

McGuire, A. D., Sitch, S., Clein, J. S., Dargaville, R., Esser, G., Foley, J., Heimann, M., Joos, F., Kaplan, J., Kicklighter, D. W., Meier, R. A., Melillo, J. M., Moore III, B., Prentice, I. C., Ramankutty, N., Reichenau, T., Schloss, A., Tian, H., Williams, L. J., and Wittenberg, U.: Carbon balance of the terrestrial biosphere in the twentieth century: Analyses of CO_2, climate and land use effects with four process-based ecosystem models, Global Biogeochem. Cycles, 15, 183–206, 2001.

Pongratz, J., Reick, C. H., Raddatz, T., and Claussen, M.: Effects of anthropogenic land cover change on the carbon cycle of the last millennium, Global Biogeochemical Cycles, 23, GB4001+, doi:10.1029/2009GB003488, 2009.

Pongratz, J., Reick, C. H., Houghton, R. A., and House, J. I.: Terminology as a key uncertainty in net land use and land cover change carbon flux estimates, Earth System Dynamics, 5, 177–195, doi:10.5194/esd-5-177-2014, URL http://www.earth-syst-dynam.net/5/177/2014/, 2014.

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 Biogeochem. Cycles, 23, GB2022, doi:10.1029/2007GB003176, 2009.

Sitch, S., Huntingford, C., Gedney, N., Levy, P. E., Lomas, M., Piao, S. L., Betts, R., Ciais, P., Cox, P., Friedlingstein, P., Jones, C. D., Prentice, I. C., and Woodward, F. I.: Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs), Global Change Biology, 14, 2015–2039(25), 2008.

Stocker, B. D., Feissli, F., Strassmann, K., Spahni, R., and Joos, F.: Past and future carbon fluxes from land use change, shifting cultivation and wood harvest, Tellus B, 66, doi:10.3402/tellusb.v66.23188, URL `http://www.tellusb.net/index.php/tellusb/article/view/23188`, 2014.

Stocker, B. D., Strassmann, K., and Joos, F.: Sensitivity of Holocene atmospheric CO2 and the modern carbon budget to early human land use: analyses with a process-based model, Biogeosciences, 8, 69–88, doi:{10.5194/bg-8-69-2011}, 2011.

Strassmann, K. M., Joos, F., and Fischer, G.: Simulating effects of land use changes on carbon fluxes: past contributions to atmospheric CO_2 increases and future commitments due to losses of terrestrial sink capacity, Tellus B, 60, 583–603, doi:10.1111/j.1600-0889.2008.00340.x, 2008.