Point by point replies II

Point by point replies - Jonathan Sanderman (2. round)

Preface: The authors have thoughtfully addressed all reviewer comments and have done a thorough job in revising and improving the manuscript. With fresh eyes, I found the manuscript to be extremely clear and insightful, and I believe it will have a big impact on our field. I have only a few comments and suggestions for further discussion of your findings:

Answer to preface:

Dear Dr. Sanderman,

thank you for reviewing our paper again and adding very helpful points. We believe especially the hint to swidden agricultural practices is very important, even so we could not integrate this into the study, but are happy to take it up for future research. We here respond point-by-point to the your comments. Please note that we only included the updated (and not the outdated) version of the respective statements and paragraphs. We refer to the marked-up manuscript version for showing the changes in more detail.

1. L365 - The IPCC2006 values for the tropics were highly criticized as being way too high. The 2019 refinements incorporated >10x as much data. The fact that your results are showing large losses in tropics is problematic and worth some discussion (see additional comment below). Note IPCC factors are derived under the assumption that there is a linear change between steady states over 20 years. Are Fscf factors consistent with this assumption? If not, this might not be a fair comparison.

Answer to 1: We added the following clarification in the method and results part:

LN 188 (in the method part): "Note that IPCC factors are derived under the assumption that there is a linear change between steady states over 20 years."

LN 394 (in the results part): "Note that IPCC Tier 1 factors are derived under the assumption that there is a linear change between steady states over 20 years, whereas our aggregated factors just reflect the relative change compared to a given potential natural vegetation reference stock without specifically tracking the age of the cropland."

LN 398 (in the results part): "[For the tropical regions the IPCC factors increased notably from the guidelines in 2006 (Lasco et al., 2006) to the update in 2019 (Ogle et al., 2019)] due to the inclusion of more data points."

We reframed the discussion (also including the point on swidden agriculture practices) - see answer to point 8.

2. L378 – This section can go for more than one sentence. Can you describe the results? Goodness of fit for natural veg and cropland (add R2 values to Fig 7)? Individually the results look fairly poor.

Answer to 2.:

We shifted the figure into the appendix and added additional information within the manuscript. It now reads:

LN 414: "In Fig. A3 we correlate our SOC results for natural vegetation and cropland in 2010 with literature values from point measurements (for data base see appendix of Sanderman et al., 2017). The goodness of the fit is very low with an R^2 of 0.13. Individually, the correlations are even lower with a R^2 of 0.09 for cropland

and 0.08 for areas of natural vegetation. This points to the fact that differences between land-use type SOC stocks could be better matched than the spatial pattern of the rather small point measurement data base. Due to the low number of small-scale measurements, statistical properties of the point data variability are not derived and thus, could not be used to improve the point-to-grid-cell comparison (see Rammig et al., 2018)."

3. Fig 6 - There is a big difference between soilgrids and soilgrids2.0 - in v2.0, they explicitly excluded litter (i.e. organic) horizons on top of dominantly mineral soils. I think this makes your cropland results more comparable to v2.0 but for global stocks, the original soilgrids product may be more reasonable particularly for northern soils where thick litter layers dominate SOC stocks.

Answer to 3.:

We add clarification on the difference between the two SoilGrids products and its comparability. We rewrote the paragraph to:

LN 408: "SoilGrids (Hengl et al., 2017) especially stands out with its high estimate, since they include the litter horizon on top of the soil that might dominate especially polar and boreal soils. SoilGrids 2.0 (Poggio et al., 2021) however, excludes litter C and thus marks the lower end particularly for northern regions. For the same reason it is also more comparable to our results, which do not account for litter C as well."

4. L383 - It seems hardly surprising to find good correlation between model results since both models have similar structure albeit with different levels of complexity

Answer to 4.:

We agree with the reviewer that this comparison is not very central to evaluate model performance and therefore moved the figure to the appendix. Additionally, we added the following sentence to the results section to clarify the added value of this comparison from our perspective:

LN 422: "As the Tier 2 modeling approach (Ogle et al., 2019) is not specifically parameterized for natural vegetation, it is important to evaluate its suitability to produce reasonable results in that domain as well at least comparable to other modeling approaches."

5. L407 – Thinking about structural uncertainly, how would you reconcile your findings with that of the CMIP6-LUMP (Ito et al 2020 ERL - https://iopscience.iop.org/article/10.1088/1748-9326/abc912) which found strong increases in SOC? I think there are strong CO2 feedbacks in those models.

Answer to 5.: The SOC debt — as the difference of a natural state to a actual state — only indirectly include changes of SOC with climatic or atmospheric conditions. As our analysis is focused on time trends of the SOC debt and not the SOC stocks itself, most of the effect of CO2 feedbacks cancel out.

However, we added some results on total SOC stock dynamic to the results part and also pick up the above study together with the added numbers within the discussion section. It now reads:

LN 326 (in the results part): "Note that the SOC stock itself — without comparing it to a PNV state — increases during the period between 1975 and 2010 from 705 GtC to 712 GtC, which corresponds to an overall SOC stock increase of 0.2 $\mathrm{GtC\,yr^{-1}}$."

LN 507 (in the discussion part): "Additionally, we find our 0.2 $\mathrm{GtC\,yr^{-1}}$ SOC stock change within the period between 1975 and 2010 for the first 30 cm of the soil profile at the upper end of estimates comparing it to estimates of Ito et al. (2020) of $0.18 \pm 0.41\,\mathrm{GtC\,yr^{-1}}$ within the period between 1850 and 2014 for the whole soil profile. This might be not surprising as the CO_2 effect is most likely stronger and land-use change effects weaker within our later and shorter modeling period compared to a mean value in the period between 1850 and 2014. The large uncertainty within estimates of SOC stock and its changes (Ito et al., 2020) again points to the large structural uncertainty within SOC modeling."

LN 516 (in the discussion part): "Moreover, the temporal dynamic of the SOC debt and stock change factors is not (or only to a small degree) altered by climatic or atmospheric effects on SOC stocks, as they cancel out by taking the difference (for the SOC debt) and ratio (for the stock change factors) of cropland SOC and SOC under hypothetical PNV conditions."

6. L425 - I disagree that input uncertainty is the largest source of error. Structural uncertainty is likely much larger – look at inter-model comparison studies such as CMIP5 (Todd-Brown publications) and others that actually include novel model structures (i.e. microbially-explicit model classes)... structural uncertainty leads to differences in +/- 100s of GtC. Wieder et al (https://doi.org/10.1111/gcb.13979) also demonstrated this with a testbed of constant input data.

Answer to 6.:

We agree with the reviewer that the effect of structural uncertainty were underestimated. It seems hard to make a judgement on the most important uncertainties. We reframed the sentence. It now reads:

LN 465: "This is why, in our view, a large part of the uncertainty with respect to the impacts of SOC management — next to the parametric and structural uncertainty of the soil model — is included in the management data itself."

7. L455 - again, model to model comparison doesn't mean much if the model structure is similar. The comparison to PNV point data (with the caveat that a large grid cell and a point measurement are hard to reconcile) does not look nearly as good.

Answer to 7.:

We reframed the sentence slightly and added two sentences on the point-data validation. It now reads:

LN 497: "The comparison of simulated PNV data with LPJmL4 shows the model's capability in reproducing PNV SOC stocks (Fig. A5). Concurrently, the point-data comparison (see Fig. A4) shows low correlation for PNV, however also for cropland sites. This might also point to the fact that SOC stocks can vary at field and local scale considerably and thus a very high number of point data is needed to derive statistical properties that could improve the point-to-grid-cell comparison (see Rammig et al., 2018)."

8. L473 – This is not quite accurate. Erosion moves a lot of SOC into aquatic reservoirs, not just to somewhere else on the land, and the 'dynamic' replacement (i.e., high net sink strength on eroding surfaces) of that carbon is what is leading to the offset in losses.

Answer to 8.: We added a references supporting the claim made by the reviewer and weakened the sentence in order to show the still controversial nature of the erosion sink paradox. It now reads:

LN 531: "In our model, erosion might however only affect the spatial pattern but not the aggregate SOC pool. As pointed out by Doetterl et al. (2016), the final fate of leached or eroded carbon is uncertain and might even offset LUC emissions (Wang et al., 2017). Concurrently, other studies have claimed erosion moves SOC into aquatic reservoirs (Zhang et al., 2020) and thus changing total global terrestrial SOC stock."

9. L487 - Another thought on the particularly high losses in the central african basin region – there is a lot of swidden agriculture practiced here - several years of cropping followed by 10 or more years of fallow to allow soil fertility (and presumably SOC) to rebuild. Is this type of cropping pattern represented in your management data?

Answer to 9.: Thank you very much for this comment. We think this is a very valid point and added the missing carbon from fallow land (as part of the swidden agriculture practices) to the discussion. We rephrased a large part of the paragraph on stock change factors. It now reads:

LN 543: "Yet, even when comparing our estimates to the low-input stock change factors of the IPCC, our SOC loss is roughly twice as large as the revised 2019 IPCC default values (2019), while it shows good agreement with the older default values from IPCC (2006). However, the revised estimates of the IPCC included much more and more recent data points, calling for a closer look on causes for the large deviations between our results and the refined Tier 1 factors. On the one hand, our approach does not account for unharvested carbon inputs from weeds or biomass cover on short-term fallows. Shifting agriculture with fallow periods might be prominent in tropical regions. While long-term fallow land (>4 years) is excluded from FAOSTAT as cropland, short-term fallow is not. Thus, our carbon inputs for these areas might be underestimated, leading to too low stock change factors. On the other hand, older studies by Don et al. (2011) estimated SOC losses for tropical soils of around 25% on average corresponding to a stock change factor of 0.75, but also reported a wide range of measured SOC changes from -80% to +58%. Fujisaki et al. (2015) however found much lower loss rates of around 9%, attributing the difference to the different time period lengths since the conversion to cropland. As our results do not specifically account for cropland age and most of the cropland is older than 20 years (as assumed for the default IPCC Tier 1 stock change factors) our stock change factors have to be lower by definition following the steady-state assumption that cropland will continue to approach a new equilibrium."

10. L536 – It is not surprising that tillage didn't have much of an impact. In line with observations, Century model does not have a strong tillage feedback. Dangal et al 2022 (JAMES - https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2021MS002622) showed shifting to no-till had pretty minimal effects of SOC when running DayCent across parts of the US.

Answer to 10.:

Thanks for that suggestion. We added this citation. It now reads:

LN 609: "In line with this, Dangal et al. (2022) finds that no-tillage has only minor impacts on SOC dynamics across parts of the US."

Point by point replies - second reviewer

0. This paper is a convincing step forward for global cropland modelling of soil carbon, based on land use and agricultural management data that are dynamic in time going back to 1975, and a model framework that is reduced in complexity in order to allow global modelling. The paper is rather long but of great value for the soil carbon and modelling community. The authors clearly state the limitations of the approach and involve several model evaluation steps that are of great help to see the limitations but also the robustness of the results. The paper is written well. At some points it could be improved in order to make it better understandable and easier to read. In particular a clear differentiation between land use change effects and agricultural management effects within croplands would be important. A topic that is not fully explored is the effect of climate change vs. anthropogenic direct impact on global soil carbon dynamics (but this maybe topic for a next paper).

Answer to 0.: Dear reviewer, thank you for the thorough and helpful review of our manuscript. We agree with the reviewer that more detailed analysis of the magnitude of management effects compared to land-use change effects is needed and added two additional global maps on attribution of SOC debt change (see Answer to 14.). Here, we respond point-by-point to the reviewers' comments. Please note that sometimes due to the amount of changes introduced within this update, we only refer to the section within the new manuscript rather than including the entire changed paragraph within these replies.

Detailed remarks:

1. l. 1-2: The first sentences of the abstract are too general and has been repeated too often – please remove.

Answer to 1.:

We shortened the respective sentences to:

LN 1: "Soil organic carbon (SOC), one of the largest terrestrial carbon (C) stocks on Earth, has been depleted by anthropogenic land-cover change and agricultural management."

2. l. 17: The carbon pool in the lithosphere (including fossil carbon) is much larger than the soil C pool.

Answer to 2.: We removed the statement so that the sentence now reads:

LN 16: "Soil organic carbon (SOC), the amount of organic carbon stored in the Earth's soil, exceeds the carbon in the atmospheric and vegetation pools multiple times (Batjes, 1996)."

3. l. 41: N-deposition does not play a role in agricultural systems that are fertilized or in areas with no N-deposition. Also, CO2 fertilisation is of minor importance compared to agricultural management including breeding.

Answer to 3.:

We agree that these aspects may not be of central importance for SOC dynamics on agricultural land and have removed the claim. Now it reads:

LN 38: "BKMs are designed to estimate LUC related emissions and have largely improved in estimating additional emissions from wood harvest and shifting cultivation. However, state of the art models do not consider impacts of varying agricultural management (Frielingstein et al., 2020; Houghton et al., 2012_carbon_2012; Hansis et al., 2015; Bastos et al., 2021)."

4. L, 46: What are stylized scenarios? And in l. 54: stylized future management? Please explain or rephrase.

Answer to 4.:

We added some more detail to what we mean with stylized. It now reads:

LN 42: "Pugh et al. (2015) explicitly consider agricultural management in the form of tillage, irrigation and biomass extraction at harvest, but worked with uniform scenario assumptions on management rather than with historical management data."

LN 47: "In global-scale carbon cycle assessments, management systems are typically represented as spatially explicit patterns that are static in time (e.g. for growing seasons in Portmann et al., 2010; multiple cropping systems in Waha et al., 2020; irrigation systems in Jägermeyr et al., 2015) or as stylized (in the sense of uniform management assumptions) scenarios (e.g. Pugh et al., 2015; Lutz et al., 2019)."

5. l. 56: Please explain why no net sink is possible? In the next paragraph you write about soil C sequestration. The reader might ask why if it is not possible.

Answer to 5.:

In the above mentioned sentence, we refer to the results of one specific study (Herzfeld et al. 2021), which reports that none of their studied changes in management (if applied globally) could counteract the legacy flux caused by the initial land conversion. We have revised the sentence as follows to avoid any confusion about the general possibility of net carbon sinks. It now reads:

LN 52: "Within their stylized future management scenarios under future climate change they find that none of the management aspects considered (residue management, no-tillage) can create a positive SOC stock change on current cropland areas that counteracts the still continuing legacy flux from the initial land-use change."

6. l. 106: How did you treat grasslands that are no pastures (mowed meadows e.g.)?

Answer to 6.:

We changed pastures to grasslands and added more detailed explanations here. It now reads:

LN 102: "We distinguish two land-use types: cropland and uncropped land under potential natural vegetation as representative for all other land-use types including forestry and grasslands (referred to as natural vegetation in the following). Forage crops are included within cropland, whereas pastures (including mowed meadows (perennials) and rangelands) are assigned to natural vegetation. Carbon flows connected to livestock are only considered in this study when they originate from cropland feed sources, while the manure originating from pasture biomass is disregarded, implicitly assuming that this manure is excreted or applied to pastures."

7. Fig 1: I am not sure if this figure is required to understand the paper. However, please check the values for SOC stocks at the y-axis. They are too low for 0-30 cm SOC stocks for most regions of the world.

Answer to 7.: We changed the unit of the values to 0 - 100 t/ha to represent a more realistic SOC stock. Another reviewer found this figure to be of value help to understand, why we do not have to distinguish between newly converted and existing cropland, so that we are convinced that it should remain in the manuscript.

8. l. 169: The assumption that agricultural management from 1510 to 1965 was the same (the same as in 1965) is an assumption with high uncertainty. You might need to explain how important or not important this assumption is for your results.

Answer to 8.: Indeed, this is an important point. There are no data and any simple assumption is obviously wrong. We acknowledge this shortcoming in the methods section and discuss the initialization uncertainty (see answer to point 18). We added:

LN 171: "We acknowledge that this introduces a bias as agricultural management has changed prior to 1965, but this approach follows other studies on effects of land-use change and management (e.g. Schaphoff et al., 2018; Herzfeld et al., 2021) and is limited by data availability on harvest statistics (and other management effects)."

9. Chapter 2.2: This consistency check with the IPCC Tier 1 approach is very useful and good.

Answer to 9.: Thanks. We improved readability of the results section on the IPCC Tier 1 comparison as well (see answer to 15.).

10. The start of the results section (l. 304) with two references from the same authors may give the impression that results are published in other papers. Please indicate here that you refer to supplementary studies supporting the actual study with data and code.

Answer to 10.: We clarified that both citations refer to data and script repositories, so it now reads:

LN 318: "Detailed results for the spatially explicit global SOC budget including intermediate results on input data as well as SOC stock results for all scenario runs can be found in the data repository from Karstens (2020a). In the following, the most important results (see Karstens (2020b) for post-processing scripts) are summarized."

11 Fig 2b) The choice of the colours is suboptimal since the different green cannot be distinguished by eye. Please use more contrasting and different colours. The Fig 2b-d give the impression that most land on earth is cropland even though only around 10% of land surface is cropland.

Answer to 11.: We agree with the reviewer and changed the color scale to a more distinguishable palette. We already excluded all cells from the global maps with less than 1000 ha cropland. Maps that additionally to the SOC information (stock, diff or change factor) include information on the cropland extent (e.g. via saturation of the color) are in our opinion more complex to understand. Thus, we decided against this and stated in legend and caption that this maps are exclusively for cropland soils.

12 l. 339: This sentence is not clear to me. It would be better to provide sums of the global input to cropland soils and not to the agricultural system (with undefined system borders). It would be also useful to include terms such as net primary production (NPP) here (also in Fig 4). Also, the term human appropriated fraction of NPP (HANPP) maybe useful also to compare with in the discussion. Derived from the data you stated that 2463 Mt C entered global croplands each year (fig 4). The fraction of manure (16%) is unexpected high since manure can only be transferred to croplands if it is collected (mostly in stables), which is not common in many parts of the world. How much of this manure is feedstock grazing on croplands?

Answer to 12.: We agree with the reviewer that "agricultural system" is rather undefined and changed it for all occurrences to "cropland system" as our study focus is on croplands only.

On the second point on HANPP, we like kindly have to decline the reviewers suggestion to include this into this study. As our study focus only on the cropped NPP/HANPP not covering grazed system and the whole forestry sector, a comparison is beyond the scope of this study.

At the last point, we like to emphasis that our results came with large uncertainties especially for the manure recycling shares, as they are derived indirectly from feed consumption. This might be the cause for systematic errors that might led to high shares of manure C input. We point to this already within our discussion part (LN 467). Concurrently, even so advanced animal waste management systems (AWMS) are not very common in most parts of the world, a large share of the livestock production is however concentrated in intensive livestock systems often featuring advanced AWMS. Please also see our answer to your point 6, where we added more details to feed biomass from cropland and pastures and LN 102 of the method section 2.3.3, where we outline our assumption on stubble grazing. Moreover, as mentioned in beginning of our result section (LN 318) intermediate results, like the 18 MtC stubble grazed manure, can be derived from the data upload (hopefully quite easily).

13. The fraction of above ground biomass C input to the soil (55%) looks to me is quite high, see e.g. Bolinder et al. providing lower fractions. The root fraction should also include rhizodeposits, which might be not considered in your study.

We agree with the reviewer that other studies have shown lower rates of residue inputs. We discuss this already in LN 478 in the discussion part by referring to Keel et al. (2017), who did a comparison of residue allometric equation (including studies from Bolinder et al.). They report rhizodeposists to be included in the residue estimates and also find that below-ground residue shares from the IPCC methodology are often higher than in other approaches (also included in our discussion LN 482).

14 l. 306: This chapter gives a general overview on agricultural land use and management effect on SOC. The impact of land use changes from natural vegetation to croplands on SOC is well known and thus the maps shown here are in many regions of the world in line with the global cropland maps. The new aspect of this study is the agricultural management within croplands. It would be interesting to show maps for agricultural management effects, e.g. a standard or a worst-case scenario vs. the real data scenario. This is part of the next subchapter (3.3.). However, there it is not showed spatially explicit (with maps).

Answer to 14.: We agree with the reviewer and added spatial explicit maps on the split up of land-use change and agricultural management effects (see Fig. 3). We also added the following to the result part Sect. 3.3:

LN 370: "Using the constManagement results that only include land-use change (LUC) related changes of the SOC debt between 1975 and 2010, we are able to subtract the LUC effect from the overall SOC debt change within the histManagement results. The remaining effect can be attributed to the changing agricultural management (MAN) as other drivers such as climatic effects have been already canceled out by taking the difference to a PNV reference state when calculating ΔSOC. The increasing SOC debt on global cropland are primarily caused by LUC (red areas in Fig. 3(b2)). Deteriorated management also contributed to increasing SOC debt in parts of Sub-saharan Africa and Central Asia. In contrast, agricultural management has led to an decrease in SOC debt in the USA, Europe, as well as in parts of China and India (blue areas in Fig. 3(b3)), which is not visible in the total ΔSOC change as LUC was happening at the same time."

15. Tab. 3 is not easy to read. For the IPCC values you display low medium and high values. For your modelled data, only an average (I guess medium) value are shown. Would it be possible to also provide low and high values derived from your models for each climate region? It might be also helpful to convert this table into a figure. More important, the effect of agricultural management is not visible in this data set since the IPCC values are developed for land use changes. Again, it would be important to distinguish between land use change and land management effects.

Answer to 15.:

We agree with the reviewer and replace the table with a graphical representation including a range of values (5th, 50th to 95th percentile for each climate region) instead of a medium value in a table. The Tier 1 approach of the IPCC guidelines also tracks management decisions e.g. on tillage, but more importantly on input regimes. We included these input regime differentiation into the comparison figure as well and reframed the sentence in the result part. It now reads:

LN 398: "Stock change factors for temperate climate zones of this study are lower than the default values of the IPCC. For the tropical regions the IPCC factors increased notably from the guidelines in 2006 (Lasco et al., 2006) to the update in 2019 (Ogle et al., 2019) due to the inclusion of more data points. Our results span over a broad range due to the different ages of the cropland but also due to different agricultural management practices within climate regions."

Please also note that we added more detail to the discussion of the stock change factors from LN 540 on.

16. Fig. 6 looks a bit strange with SOC on the x-axis instead of the y-axis. I would also like to see total stocks for croplands and for all other land use types separately.

Answer to 16.:

Thanks for the suggestion. We changed x- any y-axis. We agree with the reviewer that land-use type specific comparison would be very useful. Unfortunately, to our knowledge, non of the validation data sources used for this comparison (maybe with the exception of LPJmL) reports land-use type specific carbon stocks. Thus, we can not present such a comparison within this study.

17. Chapter 3.4.3: There are no results reported in this chapter for the point data comparison. Moreover, it is questionable if point based data are useful for model validation since SOC stocks can vary at field and local scale considerable and thus a very high number of point data (I would recommend >5000 points globally) of high quality (including bulk density measurements) are required for such an exercise. In the cited database of Sandermann et al. includes less than 300 sites. I think the comparison with the SoilGrid2.0 data is sufficient.

Answer to 17.:

We shifted the figure into the appendix and added additional information within the manuscript. We did not dropped the section as a whole since it was requested by another reviewer. It now reads:

LN 414: "In Fig. A3 we correlate our SOC results for natural vegetation and cropland in 2010 with literature values from point measurements (for data base see appendix of Sanderman et al., 2017). The goodness of the fit is very low with an R^2 of 0.13. Individually the correlations are even lower with a R^2 of 0.09 for cropland and 0.08 for areas of natural vegetation. This point to the fact that differences between land-use type SOC stocks could be better matched than the spatial pattern of the rather small point measurement data base. Due to the low number of small-scale measurements, statistical properties of the point data variability are not derived and thus, could not be used to improve the point-to-grid-cell comparison (see Rammig et al., 2018)."

18. l. 399: There might be a direct link between underestimated SOC and overestimated SOC stock increases and vice versa since models are sensitive with their modeled SOC trend to the initial SOC stock. Thus, it needs to be carefully checked and discussed whether underestimation or overestimation of SOC stocks are the reason for the predicted SOC trends. For example, the strong increase in SOC in arid regions might also be a results of an underestimation of initial SOC stocks.

Answer to 18.:

We agree with the reviewer that SOC initialization is crucial to SOC stocks, SOC debt and its changes over time, since the size of the legacy fluxes will be affecting these values strongly. We add our sensitivity analysis on the initialization choice (see Sect. 3.3), which was included in the first version of our manuscript, back including an updated scenario definition (see Sect. 2.4) and a figure in the Appendix (see Fig. A2). It gives the reader an impression of the uncertainty towards the initialization choice and shows that the impact of the dynamic agricultural management is robust to the initialization of SOC stocks. We also included the following paragraph within the discussion section:

LN 519: “The initialization of SOC stocks, however, is important for the size of the SOC debt and its change over time, since the presence and size of legacy fluxes affect these values strongly (see Fig. A2). According to our sensitivity analysis the SOC debt varies between 33.3 GtC and 50.7 GtC depending on the initialization choice, with our best guess at 39.6 GtC. Concurrently, our results indicate that the impact of the dynamic agricultural management is robust to the initialization of SOC stocks.”

19. l. 500: Its not only tillage that can affect subsoil SOC but all land use and land management options can affect SOC in the subsoil below about 30 cm depth. Powslon et al 2014 found not significant tillage effect below 25 cm depth. This might be due to the low sample size of 43 sites. However, there is no evidence that tillage effects subsoil SOC stronger than other agricultural management or land use change.

Answer to 19.:

The think the reviewer is right in pointing out that subsoil effects not only from tillage are neglected. However, other management practices considered in this study might be more equally effecting top- and subsoil, whereas tillage practices by definition are the only management option that relocates carbon vertically (as explained in the paragraph following the reviewers comment). We rewrote the paragraph also adding a reference that the effect to the subsoil and the overall C sequestration effect is still debated. It now reads:

LN 565: “Second, changes to the subsoil are neglected, which is most important for tillage effects. Other management practices might be more equally effecting top- and subsoil as they do not directly the relocates carbon vertically. As Powlson et al. (2014) have shown, the subsoil can be make a large difference in evaluating total SOC losses or gains for no-tillage systems. No-tillage effects may seem larger than they actually are if only topsoil is considered. SOC transfers to deeper soil layers under tillage might enhance subsoil SOC compared to no-till practices. However, the effect of no-till to the subsoil as well as its overall importance as a SCS measure is still debated (Ogle et al., 2019).”

20. l. 528: Please provide more detailed data either here or in the results section how increasing yields affected SOC stocks in the past.

We reframed the paragraph to more clearly on the direction and cause of the yield effect on SOC stocks. It now reads:

LN 596: “Their study moreover concludes that yield gains (by 18% in their simulations) do not lead to a substantial decline in SOC debt (less than 1% SOC increase). Historical yield increases, however, are estimated to be well above 50% (Pellegrini and Fernández, 2018; Ray et al., 2012; Rudel et al., 2009) and often lead to an increase in below- and above-ground residue biomass inputs to the soil. While we find substantially larger SOC increase in response to productivity gains than the 1% reported by Pugh et al. (2015), this is not sufficient to compensate SOC losses from global cropland expansion of around 11% between 1974 and 2010.”

21. l. 544. It would be more helpful to relate and compare the estimated SOC loss with deforestation in Gt/a to the estimated total land use change induced emissions of about 2.5 Gt/a.

Answer to 21.: Unfortunately our model does not vegetation biomass, and thus we are only able to capture soil deforestation effects. However, comparing SOC related emissions to over all Land use change induced emissions seems to be a valuable addition. We added to the discussion:

LN 581: “Comparing our SOC loss rate (the change of SOC debt) of $0.14\,\mathrm{GtC\,yr^{-1}}$ to estimates on land-use change induced emissions of $2.0 \pm 1.0\,\mathrm{GtC\,yr^{-1}}$ (sum of ‘Bookkeeping LULCC emissions’ and ‘Loss of additional sink capacity’ for the years 2009–2018 in Gasser et al. (2020)), we find SOC emissions of the first 30 cm of the soil profile to be a minor contributor to overall land-use change induced emissions.”

22. l. 549: This sentence (“one fifth of total annual C sequestration by crops is lost through soils (0.8 GtC per year”) is not clear. Please be very careful with the term “C sequestration” throughout the manuscript. C sequestration is the removal and long-term storage of CO2 from the atmosphere. I guess you are referring to the plant photosynthesis flux. The figure 4

is rather complex and may need further description in order to make it understandable also here in your discussion.

Answer to 22.:

Thank you for the suggestion to improve readability and clarity of that paragraph. We believe major confusion came from the un-updated numbers that were not in line with the figure presented in that version of the manuscript. We updated the numbers and reframed the sentence. It now reads:

LN 621: "As shown in Fig. 4, the annual SOC respiration (1.3 GtC per year) is slightly above one quarter of total annual net C uptake by crops (4.6 GtC per year). C compounds have to be respired by soil organisms to maintain basic soil functions and regulate the nutrient cycle, which often leaves limited options to decrease C losses via SOC respiration (Janzen, 2006). However, similar C losses occur at the end of the food supply chain (1.2 GtC per year), at the soil surface (1.5 GtC per year), and smaller but still considerable during residue burning (0.2 GtC per year) and within animal waste management systems (0.2 GtC per year)."

We also revised all occurrences of "sequestration" and changed to "enhancement", "accumulation" or other terms were it was needed.