Reply to the reviewers

Below, we have copied the relevant sections of the two reviews (copied sections are given in black, italic font). We have addressed the reviewers' comments separately, and have given our reply below each copied 'reviewer' section (reply is given in blue font).

Anonymous referee #1

I found this a generally well written and interesting manuscript on a timely topic.

Thank you for this positive assessment

However I feel that the authors have to sharpen their arguments and they should remove several formal weaknesses (especially regarding mathematical notation and use of units), which make it difficult understanding the text.

The manuscript deals with the influence of accumulating erosion on yield and in turn on carbon sequestration. However, it ignores basic agronomic knowledge and agricultural concepts and thus has (presently) limited real-world relevance. From an agronomic point of view it is very clear that soil truncation has NO influence on yield in contrast to the basic assumption by the authors. My strong statement is easily proven because highest yields are possible without any soil (for extreme examples see the conceptual studies for a future Mars mission). What changes is not yield but the effort-yield relationship. The effort may increase to maintain yield. Some erosion effects can be changed with little or even no effort (e.g. nutrient losses in over-fertilized landscapes); other may require more effort (e.g., irrigation to compensate losses in water holding capacity). The authors may wish to argue that their relation holds true for a given and constant effort. Such behaviour may be found in controlled plot experiments but it is agronomically invalid because it would require that a farmer stops making decisions while in fact he has to decide and adjust his management every day. There is no other explanation why farmers accept soil losses that are above what soil scientist regard tolerable than that they regard the increase in effort to maintain yields smaller than the efforts needed to lower erosion. Note that usually it is assumed that erosion decreases productivity. This is something different than yield and switching from productivity to yield is not a trivial modification and would call for a discussion of its implications.

We welcome your comment and clarify that our study is based on observed relations between biomass productivity and soil erosion. We would like to emphasize that the data used to construct the functional relationships are not derived from manipulation experiments but from the comparative analysis of eroding soils and their stable non-eroding counterparts that have received the same external inputs. They thus represent typical farming management and we therefore argue that they are representative and have some real-world relevance.

We agree that farmers in high-input systems will take measures to compensate the loss in crop yields. Heavy mechanization and intensive practices have contributed to increase yields and cope with most of the otherwise expected decline. However, as several studies pointed out (see Fenton et al. 2005, Reyniers et al, 2006, Kosmas et al., 2001), these measures may not be sufficient in low to medium input production systems and may not fully compensate the decline in biomass productivity. Also, changes in management practices can take time, and there might be a time lag between the increase in external inputs and the decrease in productivity, resulting in declining crop yields.

Finally, even in high-input system, it has been shown that the spatial patterns of biomass production are related to topography-driven erosion processes (e.g. Reyniers et al 2006).

The presented data do not result from an assumption about the relationship between soil erosion and biomass productivity but from observed cases in eroding landscapes under controlled amendment. We propose to discuss this explicitly in the revised manuscript.

I wonder why EPIC was not used. Doesn't this do essentially the same job but allows a better control of agronomic practices and all other parameters that influence yields (which all are completely ignored in the manuscript). EPIC would allow deriving yields from productivity. This also leads to the next influence that the authors do not consider: some causes of productivity decline by soil truncation are difficult to remove while this is easy for others. For instance the authors expect the largest effect on SOC decline from a loss of nutrients due to erosion (although this is pure speculation). Such a loss of nutrients would be easy and cheap to replace in many countries. Reversibility of productivity again points to the importance of the effort-yield relationship.

The main goal of the paper was to assess the potential impact of soil erosion on crop productivity and yield, assuming no changes in external inputs. Although we agree that it would be interesting to include the effects of agricultural management practices in the model, this is beyond the current scope of our study as we do not have the necessary input data to constrain the spatio-temporal evolution of external inputs. We will revise the literature review on mechanistic SOC models, and further clarify the scope of our paper in the introduction.

In our study, we used the relationship between soil truncation (as a result of soil erosion) and relative yields published by Bakker et al. (2004) in a process-based SOC dynamics model. The simple model structure allows us to keep the number of input parameters in balance with the available data input.

The basic relation between soil depth and yield is given in figure 1. This figure suggests that the study used data but this is misleading. In fact only one conceptual relation was used although the authors suggest that this relation can be separated into three different cases. My main critique regarding this figure is twofold:

(i) It ignores a fourth rather common case, namely that productivity first increases with increasing soil truncation (often up to a truncation of 20 cm to 40 cm) and then starts to decrease. This behaviour can be found in many loessial landscapes and the effect is so strong that at least in former times without subsidies farmers paid higher prices for land where the clay depleted AE horizon had been lost and the better structured Bt horizon improved the properties of the Ap.

Our Figure 1 is based on data that were published in the review paper by Bakker et al. (2004) on "The crop productivity-erosion relationship: an analysis based on experimental work". This publication compiled data from 24 experimental studies, and they analyzed the effect of soil truncation on yield by comparing the yield to a reference yield. Following their review, we used a subset of these data that exemplify the relationship between soil truncation and yield. In the dataset published by Bakker et al (2004), there is one case study (Olson et al 1999) where an increase in yield was observed as a result of soil truncation. This change was reported to be maximum 1.1 times the reference yield, and was observed after 25cm of soil truncation:

The Olsen et al (1999) study shows a decline in relative yield to 0.9 for the first 7.5 cm of soil truncation, followed by a slight increase to 1.1 relative yield. Furthermore, in the Belgian loess belt, Reynier et al (2006) studied the effect of soil truncation on yield and found that, even with soil amendment, yields were lower on the slopes than on the plateau as a result of soil truncation and removal of topsoil. Fenton et al. (2005), and Gregorich et al (1998) came to similar conclusions for sites in the US, and Dusar et al. (2011) and Kosmas et al. (2001) for the Mediterranean Region. The latter two studies indicated that, even with input of fertilizers, yields were decreasing because of soil erosion.

We agree with the reviewer that yield increases are possible for specific cases but the available data suggests that it may not be representative for a more generally applicable soil erosion – productivity relationship.

(ii) The interpretation of these three conceptual cases is brave. The authors explain a steep decrease in yield at little truncation by nutrient limitation. This is quite opposite to text book knowledge of plant nutrition. Since the early times of Mitscherlich we know from the law of diminishing returns that a reduction in nutrient availability has little effect when starting at high availability. For the topic of the manuscript it is completely irrelevant whether the one curve is caused by nutrient loss and the other curve is caused by loss of water holding capacity. These interpretations, which are repeatedly treated in the manuscript like truth although any proof is missing, should entirely be removed.

Based on the experimental data published in the meta-analysis by Bakker et al (2004), we have identified one mathematical expression that allows to express the change in relative yield as a result of soil truncation. The use of a simple mathematical expression facilitates its integration in the SOC dynamics model. Bakker et al. (2004) state that, following the literature, the three main regressors explaining yield losses due to soil truncation are water-availability, nutrient depletion and physical hindrance.

We agree with the reviewer that the interpretation of the three functional forms is not always straightforward. As this is not the main point of the paper, we will rework this section and remove the interpretations.

I would suggest that the authors strictly follow the rule of notation in mathematics. E.g.: sometimes they ignore the multiplication sign and AB means A × B, in other cases AB means one variable; sometimes variable are in italics, sometimes not; sometimes even mathematical signs are in italics (it should be dt). Units are similarly ambiguous (e.g., the unit coulomb is reported but not meant). I suggest following the “Guide for the Use of the International System of Units (SI)” (https://physics.nist.gov/cuu/pdf/sp811.pdf).

We apologize, and will make the necessary corrections to the annotations.

The data that were used to calibrate the model come a bit out of the blue. “we used data from ten study sites” but I am not sure whether the five references distributed within this paragraph were the origin of the data. Without clear reference there is no information about their reliability and the boundary conditions under which they were carried out.

Our apologies for this confusion. In fact, the data that were used to calibrate the model were presented in Van Oost et al. (2007), and the characteristics of each site are resumed in the supplementary material of Van Oost et al (2007). To avoid redundancy, we have referred to Van Oost et al. (2007). We will rework the text, and clarify the source of the datasets.

Some assumptions inherent in the model and some equations seem doubtful and would need better justification or modification:

(i) The model treats organic manure and plant residues identical (eqn 2). I wonder whether this is true because digestibility of fresh plant material is around 75%. Hence only 25% is left after the passage of the digestive tract and it is likely to assume that he remaining 25% are more resistant to further degradation than the initial material. Furthermore, in solid manure often stabilization processes take place that do not occur with plant residues on the field.

Manure and residues have a different humification coefficient values (respectively 0.3 and 0.125) which, in effect, leaves different amounts of C entering the first layer of the soil profile. These values result from parameter calibrations presented in the original ICBM model paper by Andren and Katterer (1997).

(ii) Surprisingly, the humification factor then distinguishes between manure and crop residues although this is not possible at this stage anymore because eqn2 has already mixed manure, crop residues and other young carbon into one young pool.

At each time step, the humification values are calculated based on the input from crop and manure at the considered time step, then the values of the C pools are computed. For the sake of clarity, the order in which the equations are presented in the text differs from the order of the calculations in the model. This probably caused the confusion and we will clarify this in the revised manuscript.

(iii) The model considers only temperature as climate and edaphic (!?) factor (which temperature is not said), while usually soil moisture is the most dominant influence on SOC stabilisation (see Jenny 1941).

We argue that soil temperature is important for the C mineralization rate and the evolution of the SOC stocks. The ICBM model takes the moisture into account in the factor “r” (climatic factor) (Andren and Katterer, 1997)

(iv) The model does not consider any preferential loss of SOC or clay by erosion. The results may thus only be valid for tillage erosion.

We agree with the comment that we did not include selective erosion. However, based on several studies (e.g.Wang et al., 2010, 4 years monitoring), the observed enrichment is relatively small (1.3) which indicates that most of the erosion occurs under aggregated form, at least in loessical landscape. We will develop this issue in the discussion.

(v) Only roots incorporate SOC into subsoil. Bioturbation, leaching and other processes are omitted.

It is correct that our model does not take these processes into account. We argue that at a timescale of 60 to 200 years, SOC dynamics are largely dominated by soil redistribution processes and that bioturbation and leaching, although important processes on long timescales, account for a minor part of SOC fluxes and dynamics in this context (Doetterl et al., 2016, Minasny et al., 2015, Kirkels et al., 2014). We will identify this shortcoming in a revised version of the manuscript.

(vi) It is not clear, what follows in the model below 100 cm depth. I had the impression hard rock (i.e. the model does not shift the entire soil profile downward, when topsoil is lost). In this case, the model would be far too simple because hydrology then becomes tricky. Lateral water movement could not be ignored anymore when large parts of the soil were removed. Modelling would be easier and the results likely more realistic if soft rock would follow below.

We consider the following boundary condition: the soil properties (SOC, clay) observed at 1 m depth are representative for the soil/soft rock below 1 m. In the model implementation, the values of SOC and clay at the 100^{th} layer are assumed to represent the soil characteristics below 1 m. Assuming a constant bulk density of the soil, soil characteristics are advected upward in response to soil erosion, proportionally to the amount of removed topsoil. As a response to soil erosion, the soil properties of the 100^{th} layer are also continuously advected upward. In the physical hindrance case, the amount of coarse fragments is actually given by the low absolute clay content (in volume). We will describe this boundary condition more clearly in the revised manuscript.

(vii) Eqn (9) seems to be wrong because all carbon that leaves the young pool is delivered to the atmosphere although large part of this carbon (see humification factor) enters the old pool.

Equation 9 is just the difference between the input of carbon entering the soil and the mineralized carbon leaving the young and old pools. Eq 9 is based on Fig. 1 from Andren and Katterer (1997) who developed the ICBM model. We checked Eq 9 with the original ICBM formulation (Andren and Katterer, 1997) and it is correctly presented in our manuscript.

(viii) A value of 0.55 or 0.6 seems to be more appropriate for alpha than 0.7. This could have considerable influence on the results.

The authors decided to use RRMSE for optimization (eqn (10)). Why? Isn't this a bad decision because it puts larger weight on layers with low SOC content although those layers are rather unimportant and relative measuring error is larger there? The authors also seem to have forgotten that they used RRMSE because they frequently report units of RRMSE (e.g. in Fig. 2) although this parameter cannot have a unit.

We argue that the shape of the SOC profile is as important as topsoil SOC content for our study. As Kirkels et al. (2014) pointed out, SOC stock and lateral fluxes follow a two-phase evolution in which the very high rate of loss in the first decades is followed by a period of lower loss rates. The evolution is similar for the vertical C fluxes as the C uptake increases fast at the beginning of the erosion period while the rate of increase is slowing down over time. This temporal evolution is due to the lower SOC content in the subsoil. Hence, the shape of the SOC profile determines the intensity and evolution of both lateral and vertical C exchanges. As these fluxes are a key part of our analysis, the RRMSE was kept so that parametrization of SOC profile would ensure a good representation of (i) observed SOC profile and (ii) an accurate representation of the impact of soil erosion on C fluxes.

Is a model error of 93% or even 121% acceptable (see Table 1b)? I would not be satisfied. Table 2a+b: How can the contribution of all parameters sum up to more than 100%?

The model error is indeed high for the cumulative vertical C fluxes, in contrast to the SOC stock loss' prediction. We point out in the discussion that this discrepancy as well as the high error is mainly due to the yield effect. We hypothesize that a site-specific yield effect would allow to decrease the error, but these data were not available. Finally, the long timescales considered should be taken into account when analyzing the model errors: the model predictions are in the correct order of magnitude and the relative differences between the sites are well represented.

In FAST analysis, the sum of contributions can be more than 1 when two (or more) variables are correlated. In our case, erosion rate and yield response to soil truncation are correlated.

Table 2b: How can erosion rate have an influence on the result although erosion rate was set constant?

A FAST analysis can show small positive contributions for constant parameters when (i) the number of runs is too small and (ii) due to mathematical dispersion. In our case, both cases are plausible. This also applies to negative contributions. We will discuss this in the revised manuscript.

The Results chapter does not differ in style and content from the preceding chapters, which were assigned to Material and Methods. Most results are in fact reported in the preceding chapters. The manuscript requires better structuring

We will better separate materials and methods from results.

Fig. 4: Units of the left panel? Shouldn't be a time unit in the right panel? What do the black lines denote?

The left panel represents the relative SOC loss $1 - \frac{\text{SOC (final)}}{\text{SOC (initial)}}$. It is thus without dimension.

Fig. 5+6 are in poor quality. Use the same font size as in fig. 4

This may be due to the compression applied when submitting the manuscript. We will provide figures with better quality.

Fig. 7: the information about the treatments is repeated three times (twice in the figure and once in the caption). What do the boxes and whiskers show (there is no convention on this)?

The boxes represent the interquartile range and whiskers represent the 95 % quantiles of the distribution. We will correct the information of Figure 7 and add a description of the meaning of the boxes and whiskers.

I didn't like the Discussion. What I missed at the very beginning is a paragraph about the assumptions and simplifications of the model and which influence they can have on the results (a little bit on this can be found at the very end but this is not stringent enough). Be more critical regarding your work. This would increase its value. At the moment it is of little value for me because I do not know under which conditions the results would apply and under which conditions nothing could be said. Studies are cited which seem to be in agreement with your results but this does not mean much. It only becomes meaningful if we know your assumptions and simplifications because then we also know that these assumptions and simplifications would not be important for the other study.

On the other hand there are parts in the discussion that could be written even without the preceding results (e.g. the last paragraph of chapter 4.1). They could be deleted in order not to increase the length of the discussion. Also all speculations about hindrance or nutrients should be deleted. They are all unsubstantiated and misleading.

We agree with this comment and will add a paragraph on model simplifications and assumptions in the revised manuscript. We will also remove some of the less relevant parts of the discussion and modify the discussion as suggested by the reviewer.

Details:

In general, the use of blanks is strange. After semicolon the authors do not like blanks. Also periods are often omitted (e.g. in i.e.)

We will correct the typo and punctuations issues.

Figure 2 only allows for yield reduction. Yield increase would also be possible (as in the already mentioned case of alfisols or in the case where an acidified topsoil is lost; there may be more cases).

See discussion above. However, based on your comment, we will add this effect in our simulations and will evaluate the response of the model.

I wonder why the authors used different orientation of Table 1a and 1b. The same orientation in both parts would be possible. I suggest using the same orientation as in Table 1b also in Table 1a because this is the standard orientation (variables in columns, cases in rows). Table 1 b shows the vertical C balance. In all other cases this is called vertical C flux (at least I assume that this is the same). Be consistent.

We will change the orientation of Table 1.

Fig. 3: Aren't the red profiles calibrated profiles (the word "simulated" would then be misleading). I thought the manuscript was about arable soils but apparently these soils do not have a plough horizon. Is the manuscript about grassland or woodland soils?

We agree that these are calibrated profiles. The study is only for arable lands, and we did not take tillage into account, only water erosion.

The manuscript frequently reports 1000 parameters. Fortunately the model has less. I guess the authors mean 1000 parameter sets.

We apologize for this, it is correct that we generated 1000 sets of parameters values.

There are many more technical details (e.g. inconsistent tenses, omitted periods and blanks, inconsistent formatting of references) but given that large changes are necessary it does not make sense reporting these details.

We will carefully revise the manuscript, and pay attention to the formatting of text, references and tables.