Dear Referees and Editor:

Thank you very much for your time and comments. Your suggestions are appreciated and helpful to improve the manuscript. Below are our replies to the individual questions. All the relevant changes are also listed out at the end of the file.

Referee 1

(1) As the authors already acknowledge, the data provided is very limited data for such large statements, nevertheless, they devote a large part of their discussion to discuss the global implications of their findings. Thus, I would suggest the authors to strongly reconsider the focus of their discussion and put their efforts into explaining the mechanisms behind their observations rather than drawing global conclusions with large uncertainties (not even quantified).

Answer: We understand your concerns in terms of extrapolating data from a single soil type under certain simulated rainfalls to global scales. As stated in our manuscript, the potentially significant deposition of eroded SOC within the terrestrial system inferred from this study aims at illustrating the risk of overestimating the erosion-induced CO_2 sink strength, rather than quantitatively determining the significance of such biased estimation. Similar concerns have also been raised by several other referees on the way the risk was calculated in this paper. However, Referee 9 also acknowledges that the estimation in this study is put in a proper perspective. Therefore, we would prefer to still draw attention to the comparison in the paper, but without too specific numbers to justify the uncertainty. In order to accurately deliver our statement, the relevant section will be changed in the revised manuscript:

“A 15.5% SOC enrichment of sediment re-deposited in the terrestrial system corresponds to the proportion of eroded SOC estimated to be deposited in permanent sinks (e.g., 0.12 Pg of SOC eroded per year by van Oost et al. 2007). While the effects of aggregation on SOC redistribution and subsequent fate cannot be assessed based on one experiment, most sediment is transported in form of aggregates (Walling, 1988; Walling and Webb, 1990). Ignoring the effect of aggregation on erosion and redistribution of SOC, therefore, bears the risk of overestimating the erosion-induced carbon sink effect. As a consequence, the behavior of aggregated sediment requires a reconsideration of existing approaches. Further study of different soil types, their aggregation and aggregate breakdown while moving through landscapes of varying topography during rainfall events of different intensity, frequency and duration, is required to assess the relevance of aggregation for SOC movement and fate identified in this study.”

(2) In addition, the authors report that there were no differences in the particle size distribution of soil and sediments, contrary to what has been observed in other laboratory studies and field experiments. Given the fact that particles in the flume are moved by interrill erosion, my guess would be that the flume might not be long enough for redeposition of large particles to occur during the transport phase and, thus, a selective transport of fine particles further on. On this behalf, were there signs of sediment deposition along the flume (and not only at the bottom collection point?) How would having a larger flume might have changed your results? Sediment arriving at colluvial sites in agricultural landscapes might often have traveled very large distances. Adding some explanation on how the methodological constraints which make an extrapolation of the results undesirable (p.e. slope gradient of 15%, leveling of the

surface,: : :) might have affected the results would also contribute to strengthen the manuscript.

Answer: The 1.5 m long flume with a 15% gradient was particularly chosen to ensure sufficient runoff to initiate non-selective erosion on this particular silt loam (see section 2.1). Besides, the soil surface was levelled to exclude large roughness and depressions that might inhibit the movement of aggregates. Relevant text as following will be added into the revised manuscript: “During the simulated rainfall events, the sediments were seen to move continuously with runoff, and no evident selective deposition was observed on the soil surface.”

Using a longer flume might result in two reactions: 1) If the flume was extended with a flatter pathway, then the slowed down runoff would in theory result in even more significant preferential deposition of eroded sediments. 2) If extended with the same or even steeper gradient, then a longer slope would probably accelerate the speed of the runoff, thus improve the transport capacity of the runoff. In this case, the sediments transport would be more likely to be transported in form of aggregates rather than mineral particles, as a result of the combined effects of re-entrainment and preferential deposition (Beuselinck et al., 1999c). This would generally increase the transport distances of eroded sediment and the associated SOC.

In slope scale, previous research has pointed out that sediment delivery ratios are up to 90% smaller than soil erosion rates, even in catchments with soils of fine texture where all soil particles should move as suspended load (Walling, 1983; Beuselinck et al., 1999b, 1999c; Parsons et al., 2006). This demonstrates that most of the eroded sediments are re-deposited during transport processes (Beuselinck et al., 2000). There could be two possible explanations: 1) sediment is not eroded and transported as mineral particles, but in form of aggregates (Beuselinck et al., 1999c). Aggregates do not move that far as individual mineral particles, due to the accelerated settling velocity of aggregates by the greater masses and larger sizes. 2) Runoff is not always continuous, but of certain transport capacity. Preferential deposition occurs along the transport pathway, once sediment fractions are out of the transport capacity of runoff. These re-deposited fractions would then likely to be subjected to repeated erosion processes (Starr et al., 2000; Jacinthe et al., 2002; Lal et al., 2004; Lal and Pimentel, 2008).

In addition, we assume that the aggregate size distribution during prolonged transport processes would not change significantly. The proportional composition of the six EQS classes in each sediment collection interval did not significantly differ over rainfall time (ANOVA, single factor, $P > 0.05$, n=18). Experiments from another study (Xiao et al., in preparation) also show that increasing raindrop impact to aggregates, within a certain extent, does not reduce aggregate size distribution much more.

(3) I would also suggest the authors to look at the effect of sediment re-aggregation in future experiments. In relation to the respiration measurements, it would be interesting to have some insight into the quality of the SOC within each aggregate size class. Could smaller size particles respire less due to the fact that they contain older or more ‘recalcitrant’ SOC?

Answer: We agree with your comments. This study serves merely to identify the potential error introduced by the effects of aggregation on SOC redistribution, rather than quantitatively determining the significance of such an error. In the future, more experiments with soils of different aggregation and various SOC contents need to be carried out to examine the aggregation effects on the silty loam studied here to a wider range of soils. Long-term monitoring is also required to determine the mineralization potential of different SOC fractions. Further research should also focus on the effects of preferential deposition of eroded aggregates, and the fate of SOC in these aggregates, whilst in-transit towards downslope during multiple rainfall events. Effects of varying rainfall characteristics as well as a range of crust and moisture conditions of soil surface as well as soil management (e.g., Wang et al., 2008; Hu et al., 2013a) onto SOC transport should also be investigated.

(4) Can you show the standard deviation of the original soil as well in Figures 3, 4,5? It is difficult to tell otherwise if the observed differences are relevant or not.

Answer: The standard deviation of the original soil will be added in the revised manuscript. Please see the new figures at the end of this file (or the supplement pdf file):

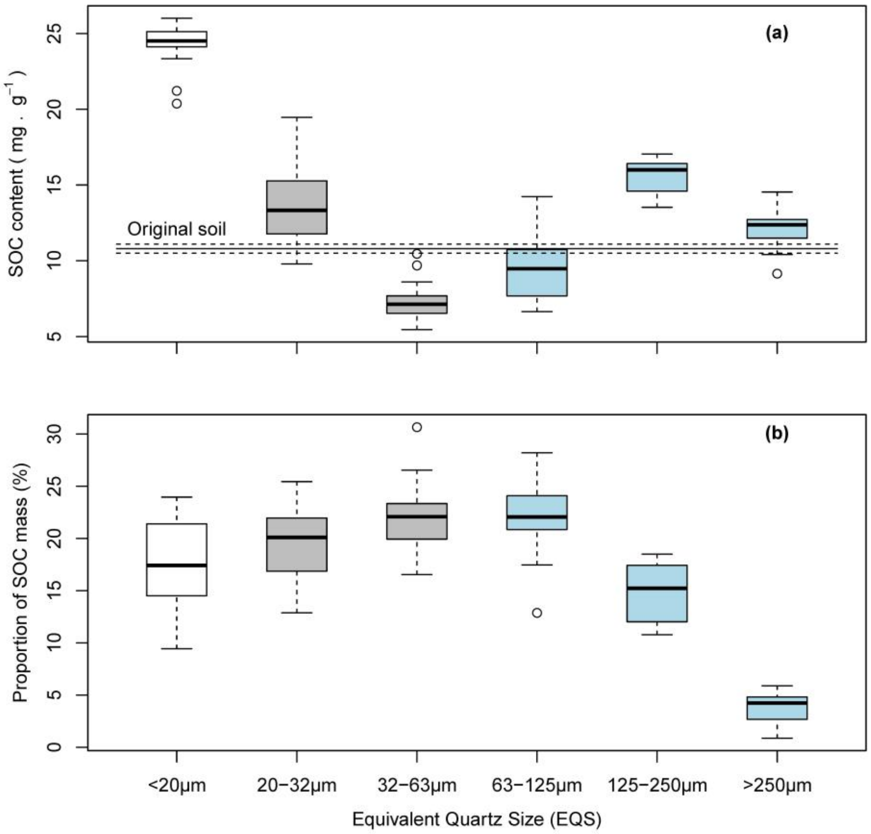

Figure 3 The distribution of soil organic carbon (SOC) (a), and soil organic carbon (SOC) mass (b) in different Equivalent Quartz Size (EQS) classes of the sediment. The bold and dashed lines in (a) denote the average and standard deviation of soil organic carbon (SOC) of the original soil. Colors of the boxes correspond to the likely fate of each fraction after erosion. See section 4.1 and Figure 6 for definitions and explanation of the three manners of likely fate. Bars in the boxes represent median values. Whiskers indicate the lowest datum within 1.5 interquartile range of the lower quartile, and the highest datum within 1.5 interquartile range of the upper quartile (*n* = 18).

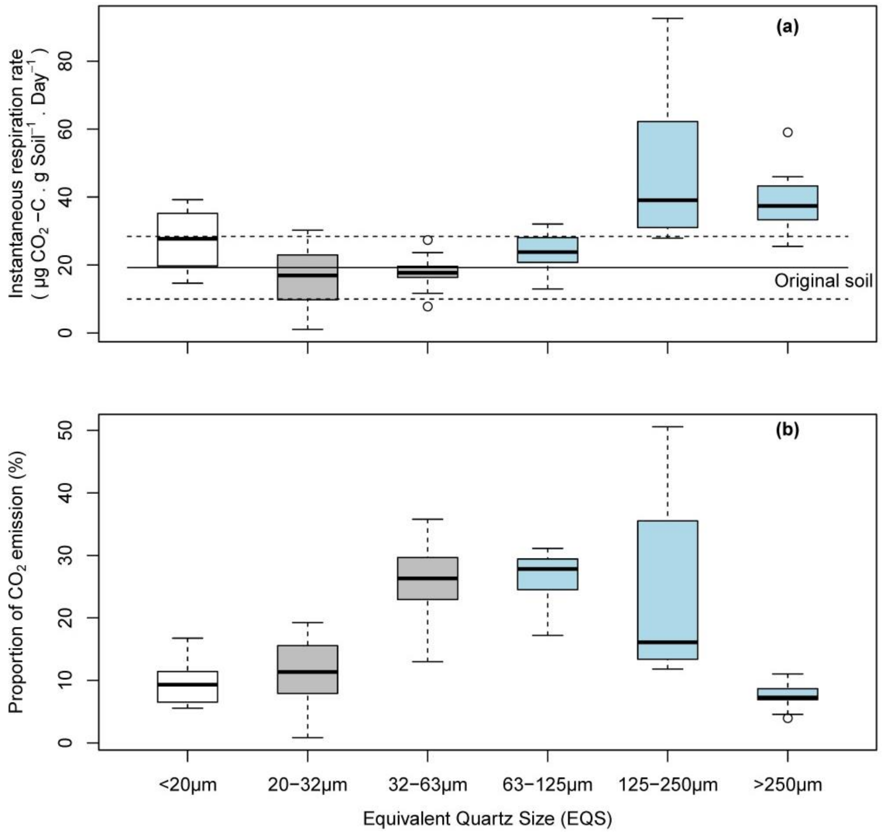

Figure 4 The distribution of instantaneous respiration rate (a); and potential CO_2 emission (b) in different Equivalent Quartz Size (EQS) classes of the sediment. The bold and dashed lines in (a) denote the average and standard deviation of instantaneous respiration rate of the original soil. Colors of the boxes correspond to the likely fate of each fraction after erosion. See section 4.1 and Figure 6 for definitions and explanation of the three manners of likely fate. Bars in boxes represent median values. Whiskers indicate the lowest datum within 1.5 interquartile range of the lower quartile, and the highest datum within 1.5 interquartile range of the upper quartile (*n* = 18).

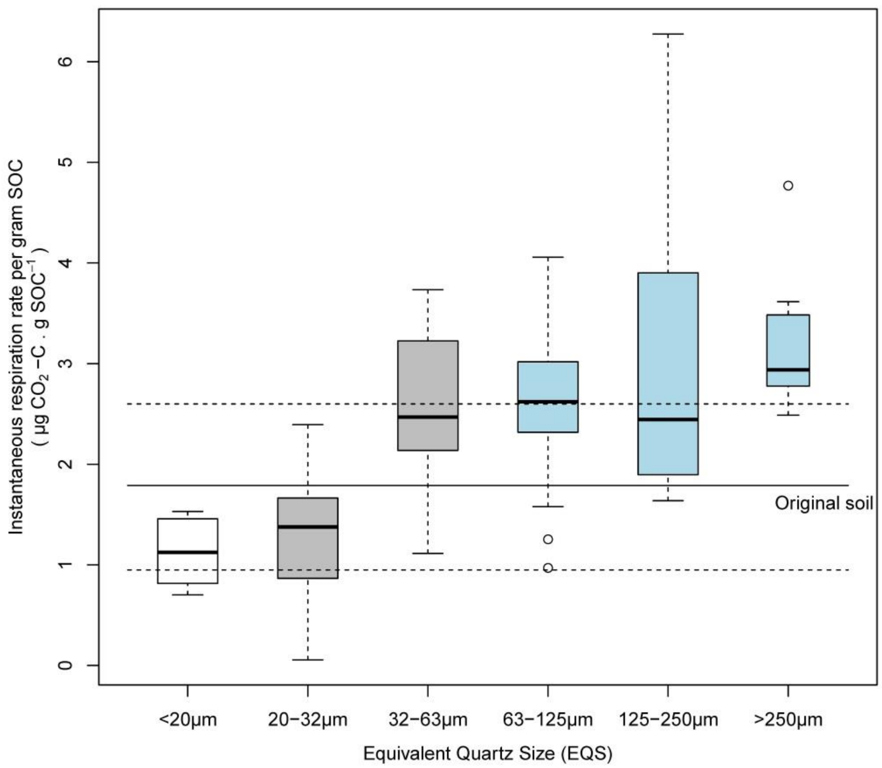

Figure 5 The distribution of instantaneous respiration rate per gram of soil organic carbon (SOC) in different Equivalent Quartz Size (EQS) classes of the sediment. The bold and dashed lines denote the average and standard deviation of instantaneous respiration rate per gram SOC of the original soil. Colors of the boxes correspond to the likely fate of each fraction after erosion. See section 4.1 and Figure 6 for definitions and explanation of the three manners of likely fate. Bars in boxes represent median values. Whiskers indicate the lowest datum within 1.5 interquartile range of the lower quartile, and the highest datum within 1.5 interquartile range of the upper quartile (*n* = 18).

(5) In relation to figure 4, if most CO_2 comes from the fine fractions (due to their total mass) part of this is at the same time related to the loess soil you are using, which does not reflect global soil diversity (in conflict with your extrapolation to the global scale). At the same time, where is the burial effect taken into account there? SOC mineralization is physically constrained by burial. How could this change your results?

Answer: According to Figure 4 and Figure 7, most CO_2 emissions (53%) came from the coarse fractions of EQS > 63 µm. This is, on one hand, because of the greater total mass of these coarse fractions (light blue boxes in Figure 2); on the other hand, due to their greater susceptibility to mineralization (light blue boxes in Figure 5). Such observations from a single soil type cannot represent the global soil diversity. However, as a first study aimed at quantifying the potential effects of aggregation onto the movement and fate of eroded SOC, these findings just illustrate the need for further investigation on various soil types.

An enrichment of SOC in terrestrial deposition suggests a reduction of net lateral SOC transfer from eroding sites to depositional sites than estimated in previous reports. This further implies a reduction in all the long-term deposition sites further downslope than the colluvial deposition, and thus a reduced effect of deep burial. Although the significance of such reducing effects remains speculative, the net effects of erosion as a CO_2 sink/source requires an investigation of aggregation on current slope-scale carbon balances. Further investigations on other relevant factors (see answer above) are required.

(6) The authors question in a couple of occasions the often reported association of SOC with mineral particles. They show that while 61% of the sediment fractions were in EQS of 32-125 m, containing 65% of SOC. However, the difference between 61% and 65% does not seem enough to support the affirmation that SOC is not associated to mineral particles.

Answer: We did not state that SOC is not associated to mineral particles. We stress that sediment fractions and the associated SOC are moved and transported in form of aggregates, rather than individual mineral particles. The reference to this particular result, "while 61% of the sediment fractions were in EQS of 32-125 µm, containing 65% of SOC", is not intended to compare the differences between 61 and 65%, but to illustrate the contrast of aggregate specific SOC distribution (Figure 2, 3) and the mineral particle specific SOC distribution shown in Table 1. According to the mineral particle distribution, 62% of the particles were < 32 µm carrying about 81% of the total SOC.

(7) Section 4.1 is weak and not supported by any references other than a previous experiment from the same authors. I suggest to either integrate this as part of the results or provide additional discussion and contrast with results from other studies.

Answer: This study is the first using settling tube to fractionate eroded sediment and investigate the transport distance of eroded SOC. So, there are barely other studies to refer to. In order to strengthen our argument, comparison with other fractionation methods using wet sieving, although not strictly comparable, can be added. For instance, the distinct SOC distribution across aggregate size classes is consistent with the field investigation by Polyakov and Lal (2008), where the coarse aggregates (1–0.5 mm) fractionated by wet-sieving, contained up to 4.5 times more SOC than the finest fraction (< 0.05 mm).

(8) In page 8832 lines 18-19 explain these "diverse impacts".

Answer: The qualities and stabilizing mechanisms of SOC in the soil matrix are varying with aggregate conditions (Six et al., 2002). For instance, physically-stabilized SOC within macro- and micro-aggregates is protected by forming physical barriers between microbes and enzymes and their substrates, and thus very susceptible to mineralization after aggregates break-up. Chemically-stabilized SOC results from the chemical or physicochemical binding between SOC and soil minerals (i.e. clay and silt particles). Such stabilization is also likely to be disturbed by aggregates break-up. Biochemically-stabilized SOC is resulted from the inherent or acquired biochemical resistance to decomposition. Aggregates break-up might also affect the biodegradability of the SOC or the exposure to hydrolyzation. Therefore, erosion, either breaking aggregates from soil matrix or out of an aggregate itself, may have distinct impacts on mineralization of eroded SOC.

(9) In relation to the incubation, did you take into account the effect of re-wetting on the CO2 initial pulse?

Answer: The fractionated sediments were re-wetted on previous day before conducting the respiration measurements. In this way, the initial CO_2 pulses of rewetted soils should be largely excluded (Orchard and Cook, 1983). Even if there were any CO_2 pulses induced by rewetting, this exactly mimics the natural processes, where dry sediments deposited from previous rainfall events, experience a second time of erosion and transport processes. To determine the quality of eroded SOC, it requires monitoring long-term mineralization potential.

Referee 4

(1) Results observed from one soil cannot represent all other soil types

Answer: We agree with your comments. This study serves merely to identify the potential error introduced by the effects of aggregation on SOC redistribution, rather than quantitatively determining the significance of such an error. In the future, more experiments with soils of different aggregation and various SOC contents need to be carried out to examine the aggregation effects on the silty loam studied here to a wider range of soils. Long-term monitoring is also required to determine the mineralization potential of different SOC fractions. Further research should also focus on the effects of preferential deposition of eroded aggregates, and the fate of SOC in these aggregates, whilst in-transit towards downslope during multiple rainfall events. Effects of varying rainfall characteristics as well as a range of crust and moisture conditions of soil surface as well as soil management (e.g., Wang et al., 2008; Hu et al., 2013a) onto SOC transport should also be investigated.

(2) Transport process & aggregate breakdown

Answer: We agree with your concerns that further breakdown of aggregates into fine particles during transport process can potentially increase the transport distance of eroded SOC and thus increase the likelihood of eroded SOC to be transferred into rivers. However, in slope scale, previous research has pointed out that sediment delivery ratios are up to 90% smaller than soil erosion rates, even in catchments with soils of fine texture where all soil particles should move as suspended load (Walling, 1983; Beuselinck et al., 1999b, 1999c; Parsons et al., 2006). This demonstrates that most of the eroded sediments are re-deposited during transport processes (Beuselinck et al., 2000). There could be two possible explanations: 1) sediment is not eroded and transported as mineral particles, but in form of aggregates (Beuselinck et al., 1999c). Aggregates do not move that far as individual mineral particles, due to the accelerated settling velocity of aggregates by the greater masses and larger sizes. 2) Runoff is not always continuous, but of certain transport capacity. Preferential deposition occurs along the transport pathway, once sediment fractions are out of the transport capacity of runoff. These re-deposited fractions would then likely to be subjected to repeated erosion processes (Starr et al., 2000; Jacinthe et al., 2002; Lal et al., 2004; Lal and Pimentel, 2008).

In addition, we assume that the aggregate size distribution during prolonged transport processes would not change significantly. The proportional composition of the six EQS classes in each sediment collection interval did not significantly differ over rainfall time (ANOVA, single factor, $P > 0.05$, n=18). Experiments from another study (Xiao et al., in preparation) also show that increasing raindrop impact to aggregates, within a certain extent, does not reduce aggregate size distribution much more.

(3) Rainfall intensity, duration & aggregate size

Answer: As discussed in previous question, this study is merely the first step to investigate the entire erosion-transport-deposition process. Further research should, therefore, focus on the effects of preferential deposition of eroded aggregates, and the fate of SOC in these aggregates, whilst in-transit towards downslope after multiple rainfall events. Effects of varying rainfall characteristics as well as a range of crust and moisture conditions of soil surface as well as soil management (e.g., Wang et al., 2008; Hu et al., 2013a) onto SOC transport should also be investigated in the future.

(4) Erosion size (interrill or rill erosion) & sediment load

Answer: We agree that the splash or interrill erosion is more likely to selectively erode soil fractions and thus form various sediment compositions. But regardless of selective splash and interrill erosion, or non-selective rill erosion, sediment fractions are all likely to experience preferential deposition. Therefore, the SOC redistribution by either selective or non-selective erosion, is strongly depending on the transport distance of eroded aggregates.

(5) Calibration and efficiency of ultrasound dispersion

Answer: It is true that the calibration and efficiency of ultrasound dispersion are controversial (Beuselinck et al., 1999a; Kaiser et al., 2012). But, the application of ultrasound dispersion in this study aims at the comparison of size distributions between aggregated fractions and non-aggregated fractions. Although the ultrasound energy used in Hu et al. (2013b) was not enough to thoroughly disperse the original soil into real mineral particles (Kaiser et al., 2012), such extent of dispersion was notable enough to demonstrate the potential under-estimation of applying mineral particle size distribution to predict the settling velocity of eroded SOC. In addition, complete dispersion, if so difficult to thoroughly carry out, is probably not feasible to apply in any other erosion models, either.

(6) Aggregate specific SOC distribution vs. mineral SOC distribution

Answer: We agree with you that, in the silty loam used in this study, SOC may not play an important role in forming and stabilizing coarse aggregates, which would be immediately re-deposited. But, according to our results (Table 1, Figure 2 and Figure 3), the aggregation effects were very pronounced in forming medium size fractions, such as EQS of 32 to 125 µm.

(7) Particle size classification

Answer: We agree with you that the EROSION 3D model and LiSEM model are very powerful in representing many particle classes. But for aggregated soils, settling velocities are affected by the actual size, irregular shape, porosity of soil fractions and their incorporation with SOC of light-density (Kinnell and McLachlan, 1988; Loch, 2001; Hu et al., 2013b). Hence, the mineral particle size classes, no matter how efficiently applicable in erosion models, are not the decisive factor to determine the actual settling behavior or movement of aggregates. In addition, the upper limits of the mineral particle size classes used in current erosion models are often smaller than the sizes of coarse EQS applied in this study. For instance, the largest class in van Oost et al. (2004) is only 90 µm, whereas up to 250 µm for rill erosion in Morgan et al. (1998). Such limits may also skew the estimation on settling velocity of eroded sediment.

(8) Separation of clay

Answer: It is definitely possible to separate clay from other fractions by settling velocity. However, in this study, a long settling tube is required to sufficiently fractionate coarse aggregates that have fast settling velocities. If using the 1.8 m long settling tube as in this study, it will cost a clay particle (size of 2 µm) about 140 hours to settle from the top to the bottom. Such a long time is simply not practical for a laboratory experiment. In addition, fine suspended fractions are considered as one group “exported” out of the terrestrial system as suspended sediment. Hence, with the current settling tube (length of 1.8 m), any fractions finer than 20 µm (settling time longer than 1.5 h) were not further fractionated to save time.

Referee 5

(1) The soil preparation in the flume may represent in a best case one field situation, while many others can be expected in the field (depending e.g. on soil moisture, tillage operations, temp., earth worm activity ...). Hence, a transfer of results to similar soil seemed to be impossible.

Answer: We understand your concerns in terms of extrapolating data from a single soil type under certain simulated rainfalls to global scales. As stated in our manuscript, the potentially significant deposition of eroded SOC within the terrestrial system inferred from this study aims at illustrating the risk of overestimating the erosion-induced CO_2 sink strength, rather than quantitatively determining the significance of such biased estimation. Similar concerns have also been raised by several other referees on the way the risk was calculated in this paper. However, Referee 9 also acknowledges that the estimation in this study is put in a proper perspective. Therefore, we would prefer to still draw attention to the comparison in the paper, but without too specific numbers to justify the uncertainty. In order to accurately deliver our statement, the relevant section will be changed in the revised manuscript:

A 15.5% SOC enrichment of sediment re-deposited in the terrestrial system corresponds to the proportion of eroded SOC estimated to be deposited in permanent sinks (e.g., 0.12 Pg of SOC eroded per year by van Oost et al. 2007). While the effects of aggregation on SOC redistribution and subsequent fate cannot be assessed based on one experiment, most sediment is transported in form of aggregates (Walling, 1988; Walling and Webb, 1990). Ignoring the effect of aggregation on erosion and redistribution of SOC, therefore, bears the risk of overestimating the erosion-induced carbon sink effect. As a consequence, the behavior of aggregated sediment requires a reconsideration of existing approaches. Further study of different soil types, their aggregation and aggregate breakdown while moving through landscapes of varying topography during rainfall events of different intensity, frequency and duration, is required to assess the relevance of aggregation for SOC movement and fate identified in this study.

(2) The transport mechanisms in the flume only partly address all processes involved when transporting soil from erosional sites into water bodies (e.g. aggregate breakdown during prolonged transport with water).

Answer: We agree with your concerns that further breakdown of aggregates into fine particles during transport process can potentially increase the transport distance of eroded SOC and thus increase the likelihood of eroded SOC to be transferred into rivers. However, previous research has pointed out that sediment delivery ratios are up to 90% smaller than soil erosion rates even in catchments with soils of fine texture where all soil particles should move as suspended load (Walling, 1983; Beuselinck et al., 1999b, 1999c; Parsons et al., 2006). This demonstrates that most of the eroded sediments are re-deposited during transport processes (Beuselinck et al., 2000). There could be two possible explanations: 1) sediment is not eroded and transported by mineral particles, but in form of aggregates (Beuselinck et al., 1999c). Aggregates do not move that far as individual mineral particles, due to the accelerated settling velocity of aggregates by the greater masses and larger sizes. 2) Runoff is not always continuous, but of certain transport capacity. So, preferential deposition occurs along the transport pathway, once aggregates are out of the transport capacity of runoff.

In addition, we assume that the aggregate size distribution during prolonged transport processes would not significantly change. The proportional composition of the six EQS classes in each sediment collection interval did not significantly differ over rainfall time (ANOVA, single factor, $P > 0.05$, n=18). Experiments from other colleague (Xiao et al., in preparation) also show that increasing raindrop impact to aggregates, within a certain extent, does not reduce aggregate size distribution much more. In future study, more simulation as well as field experiments will be carried out to investigate the effects of various transport processes (such as slope length, slope gradients, field barriers) onto the mechanism of aggregate breakdown and aggregate specific SOC distribution.

Referee 6

(1) Please add to the Abstract: which exactly classes of aggregates have been obtained; quantity information; reduce the introduction part.

Answer: The relevant part in the Abstract will be changed to: "Both the eroded sediments and undisturbed soils were fractionated into six different size classes using a settling tube apparatus according to their settling velocities: > 250, 125 to 250, 63 to 125, 32 to 63, 20 to 32, and < 20 µm." The current "Introduction" is necessary to clarify all the relevant background information in the most concise manner, and therefore cannot be reduced any shorter.

(2) L90 Please clarify the depth of the A horizon.

Answer: L1 on page 8833 in revised manuscript will read as: "A-horizon material (top 20 cm) was sampled from a gentle shoulder slope (< 5%) in March 2012."

(3) L128 Remove "but" in the start of the sentence.

Answer: "But" is removed.

(4) L191 How long the incubation was done?

Answer: The fractionated sediments were re-wetted on the previous day, approx. 20 hours before conducting the respiration measurements, to exclude the initial CO2 pulses possibly caused by the rewetting effects (Orchard and Cook, 1983). In this study, the respiration measurements were only carried out for once, mainly to determine the instantaneous respiration rates. This can effectively mimic the natural processes, where dry sediments deposited from previous rainfall events experience a second time of erosion and transport processes. However, instantaneous respiration can only partly reflect the long-term respiration potential of the eroded SOC. Further investigation on the quality of eroded SOC is required to determine the fate of eroded SOC.

(5) L345-350 Please, split the sentence.

Answer: The sentences will be changed in revised as: "As a consequence, the preferentially deposited SOC could potentially generate a further error in the carbon source-sink balance. Such error would be particularly significant, when repeated erosion and deposition processes along hill-slopes cause further disintegration of large aggregates (Kuhn et al., 2003; van Hemelryck et al., 2010). This would thereby result in additional SOC exposure and mineralization (Jacinthe et al., 2002; Six et al., 2002)."

(6) Table 1. Term "concentration" is usually used for the solutions, for the solid substances term content has to be used. The dimension mg g-1 soil is not so typical, better to use g kg-1 soil. "General SOC" what does this mean? I suggest to leave SOC here, and in case of SOC in aggregates write in the left column SOC (g kg-1 fraction). Please, present the standard errors by the normal way (_value).

Answer: The text body and Table 1 in Method will be modified accordingly to:

"A-horizon material (top 20 cm) was sampled from a gentle shoulder slope (< 5%) in March 2012. Previous research on the same silty loam showed that aggregation increased the settling velocity of original soil fractions, particularly the medium sized fractions, in comparison with that expected based on the texture of the original soil (Hu et al., 2013b). The mineral particle specific SOC distribution, average SOC content (LECO RC 612 at 550°C), and aggregate stability of original soil (method adapted from Nimmo and Perkins, 2002), are shown in Table 1. The mineral particle size distribution was fractionated by wet-sieving, after dispersed by ultrasound using a Sonifier Model 250 from Branson, USA. The energy dissipated in the water/soil suspension was 60 $J \cdot ml^{-1}$ (i.e. Energy = output power 70 W × time 85 s / suspension volume 100 ml). The SOC mass proportions across mineral particle size classes were calculated only from average values of individual weight and SOC content. Although the ultrasound energy used in Hu et al. (2013b) was not enough to thoroughly disperse the original soil into real mineral particles (Kaiser et al., 2012), such extent of dispersion was notable enough to demonstrate the potential under-estimation of applying mineral particle size distribution to predict the settling velocity of eroded SOC. Hence, it is speculated that similar increasing effects would also occur to sediment fractions, and thus make the silty loam suitable to investigate the potential effects of aggregation of original soil on the transport distances of differently sized sediment fractions."

Table 1 Mineral particle size distribution, soil organic carbon (SOC) distribution across mineral particle size, average SOC of bulk soil, and the percentage of stable aggregates greater than 250 µm in the silty loam used in this study.

	Mineral particle size (µm)					SOC of original soil (g kg^{-1})	Aggregates greater than 250 µm (%)
	< 32	32-63	63-125	125-250	> 250		
Weight (%)	62.0 $_{\pm 0.3}$	29.1 $_{\pm 0.4}$	6.6 $_{\pm 0.3}$	1.2 $_{\pm 0.1}$	1.1 $_{\pm 0.1}$	10.8 $_{\pm 0.4}$	67.2 $_{\pm 6.9}$
SOC (g kg^{-1})	13.7 $_{\pm 0.7}$	3.0 $_{\pm 0.3}$	8.9 $_{\pm 2.6}$	21.9 $_{\pm 0.8}$ ^{a}	26.4 $_{\pm 1.3}$ ^{a}		
SOC mass proportion (%)	80.8	8.3	5.6	2.5	2.8		

NOTE: a) might be over-estimated due to the mixture of minute amount of residue or straw, which was previously incorporated into the aggregates but then released by dispersion and blended with coarse particles.
Lower case numbers indicate the range of minimum and maximum values (*n* = 3).

(7) The same is for table 3. In the marks under the table please remove the information about methods, this information is for materials and methods section.

Answer: Table 3 will be changed in revised manuscript to:

Table 3 Summary of the erosional responses of Möhlin soil over 180 min of rainfall time. Subscripted numbers indicate the minimum and maximum range of the parameters (*n* = 3).

Steady state (after 120 min)			Total runoff (kg)	Runoff coefficient (%)	Total sediment yield (g)
Runoff rate ($mm \cdot h^{-1}$)	Sediment discharge rate ($g \cdot m^{-2} \cdot h^{-1}$)	Sediment concentration ($g \cdot L^{-1}$)			
18.0 $_{\pm 0.9}$	168.7 $_{\pm 14.4}$	9.4 $_{\pm 0.1}$	40.7 $_{\pm 3.1}$	20.6 $_{\pm 1.6}$	475.8 $_{\pm 74.6}$

Referee 7

(1) There are several grammatical errors and inconsistencies. For the purposes of further scientific discussion, I won't itemize this, but please carefully edit the paper.

Answer: Thank you for your notice. The manuscript has been carefully edited, and you will find the minor corrections in the revised manuscript.

(2) The use of only the first 10 minutes of run-off: Why run out to 30 minutes to ensure full breakdown, if you aren't going to study the most difficult-to-destroy (ie, most stable, potentially most C) aggregates: What was the point of separately pooling all the fractions captured in that last 20 minutes while considering a detailed fractionation of the first 10 minutes? Is slaking not a function of time? And wouldn't slaked aggregates release differently reactive C?

Answer: We understand your concerns on the temporal variation of aggregate breakdown over rainfall time. The runoff and sediment suspensions were collected every 30 min throughout the simulated rainfall events. Within each of the 30 min intervals, only samples from the very first 10 min were used to carry out settling fractionation, limited by the volume of the injection device. In total, there were six sediment collection intervals over the 3 hours rainfall events, and a settling fractionation test was carried for each of the six sediment collection intervals. Therefore, any potential breakdown of the aggregates during the simulated rainfall was not ignored, but examined through the consecutive sampling.

In addition, the possible changes of slaking effects, as concerned by the referee, did not differ significantly over rainfall time (ANOVA, single factor, $P > 0.05$, $n = 18$). Therefore, the variations of the proportional composition of eroded sediments were on purpose simplified to accentuate the differences across EQS size classes.

(3) You reported that 95% of the sediments settled during your 1-hour pre-treatment. But this was not a random sampling of 95% of the particulates. The remaining 5% are a fraction sharing the common trait of low density/settling velocity; it is entirely possible and consistent with your own hypotheses that this 5% features uniquely in the C accounting you are attempting to resolve. This suspended fraction appears to be retained and analyzed, but it was added back with the finest settled fraction. Why?

Answer: The collection beakers used in pre-treatment had height of 20 cm, which, in theory, were short settling tubes as compared to the 1.8 high settling tube apparatus. Hence, the pre-treatment of settling in the collection beakers can be considered following the same theory of settling velocity. Thus, this pre-treatment was designed to save the time for the fine fractions from settling over long time in the 1.8 m settling tube. The supernatant and remaining suspended sediment after 1 h of settling in the collection beakers corresponded to EQS < 8 μm. Fine suspended fractions are considered as one group "export" out of the terrestrial system. Hence, the fine fractions of EQS < 8 μm from the pre-treatment were added back to the fractions of EQS < 20 μm suspended in the 1.8 high settling tube, for the sake of terrestrial or aquatic perspective. Where settling tubes of different lengths are applicable, sediment fractions of distinct settling velocities should be separately treated.

(4) One flaw in your arguments (to me), is that the C associated with the different EQS fractions is not all the same in terms of decomposition risk. As you make the link from C

stocks to fluxes, instantaneous respiration from a subfraction is not likely representative of the respiration observed at the deposition site. Some discussion of the C forms and stabilization mechanisms is needed to put the predictions of gas fluxes in a more relevant context.

Answer: We agree with your concerns on the quality and the mineralization susceptibility of SOC stored in various sediment fractions. The instantaneous respiration only partly reflects the mineralization potential of the eroded SOC. The scope of this study will be clarified in the revised manuscript. Future investigations, such as determining the SOC quality using isotope, and monitoring the long-term respiration potential of different SOC fractions, are required to quantify the net effects of erosion on global carbon cycling.

(5) How can you use the classic water-stable aggregate profile for this soil to tie this experiment to broader experiments in the literature? The WSA data you provide in soil characterization isn't really used. The calculations of percent change against various denominators are loosely discussed.

Answer: The aggregate stability of the original soil showed that 67% of the soil fractions were stable aggregates > 250 µm (Table 1), whilst the settling velocities of the eroded sediments indicate that only 4% of the fractions were of EQS > 250 µm (Figure 2). Such remarkable differences between the two fractionation methods are partly induced by the bias of using Stokes' Law to estimate the settling velocity of coarse aggregates. But more importantly, the differences clearly show that aggregates experienced sufficient breakdown during the simulated erosion events. This, furthermore, suggests the relevance of using the settling velocities of actual aggregated sediment to estimate the likely transport distances of eroded sediments rather than texture or any arbitrary stress induced when measuring aggregate stability.

(6) I'd like to see the calculations presented more crisply. And how did you make the leap from "approximately 41% of the eroded SOC from the silt loam used in this study would be re-deposited along eroding hill slopes" to "Our data show that 41% of the eroded SOC from a silty loam was incorporated into aggregates of EQS > 63_m"? I missed that connection.

Answer: According to the conceptual model developed by Starr et al. (2000) (Figure 6), the six EQS classes can be further grouped into three separate groups, each with a different likely fate: EQS < 20 µm would be likely remain suspended in runoff and hence, transferred to rivers, and all EQS > 63 µm would be re-deposited along eroding hill-slopes (Table 2). The intermediate EQS of 20 to 32 µm and 32 to 63 µm can have either fate, depending on localised flow hydraulics. Figure 3b shows that approximately 41% of the eroded SOC from the silt loam used in this study was stored in aggregates of EQS > 63 µm. Therefore, according to the conceptual model described above, approximately 41% of the eroded SOC from the silt loam used in this study would be re-deposited along eroding hill slopes (Figure 7b).The clarity the figures presented in our paper will be improved in the revised manuscript.

Referee 8

Comment: I do not think it is needed to provide some big numbers that are probably irrelevant due to their uncertainty and from which it is hard to learn something. However, this comment does not call for removing the discussion around weaknesses existing in erosion models. It is an important matter per se and this does not need a very uncertain global estimate to be justified.

Answer: We highly appreciate your understanding. We also understand your concerns in terms of extrapolating data from a single soil type under certain simulated rainfalls to global scales. As stated in our manuscript, the potentially significant deposition of eroded SOC within the terrestrial system inferred from this study aims at illustrating the risk of overestimating the erosion-induced CO_2 sink strength, rather than quantitatively determining the significance of such biased estimation. Similar concerns have also been raised by several other referees on the way the risk was calculated in this paper. However, Referee 9 also acknowledges that the estimation in this study is put in a proper perspective. Therefore, we would prefer to still draw attention to the comparison in the paper, but without too specific numbers to justify the uncertainty. In order to accurately deliver our statement, the relevant section will be changed in the revised manuscript:

A 15.5% SOC enrichment of sediment re-deposited in the terrestrial system corresponds to the proportion of eroded SOC estimated to be deposited in permanent sinks (e.g., 0.12 Pg of SOC eroded per year by van Oost et al. 2007). While the effects of aggregation on SOC redistribution and subsequent fate cannot be assessed based on one experiment, most sediment is transported in form of aggregates (Walling, 1988; Walling and Webb, 1990). Ignoring the effect of aggregation on erosion and redistribution of SOC, therefore, bears the risk of overestimating the erosion-induced carbon sink effect. As a consequence, the behavior of aggregated sediment requires a reconsideration of existing approaches. Further study of different soil types, their aggregation and aggregate breakdown while moving through landscapes of varying topography during rainfall events of different intensity, frequency and duration, is required to assess the relevance of aggregation for SOC movement and fate identified in this study.

Referee 9

(1) One of the few ways I think this paper could be improved would be by including a slightly more involved explanation of the Equivalent Quartz Sizes approach and the settling tube apparatus. The authors appropriately cite their earlier work introducing these techniques and I understand their (or the journal's) reluctance to include extraneous text. However, it is difficult to understand what the authors did without reading the Hu et al 2013b paper.

Answer: We agree with you that a brief description of the settling tube apparatus is necessary to elaborate the experimental rationale. Relevant parts in the Method will be changed as following:

"A 1.8m settling tube (Fig. 1b) was used to fractionate the eroded sediment fractions according to their respective settling velocities. The settling tube apparatus consists of four components (Figure 1a): the settling tube, through which the soil sample settles; the injection device, by which the soil sample is introduced into the tube; the turntable, within which the fractionated subsamples are collected; and the control panel, which allows an operator to control the rotational speed and resting/moving intervals of the turntable. Details about the settling tube apparatus were described in Hu et al. (2013b)."

"Six particle size classes, based on the concept of Equivalent Quartz Size (EQS) described in Hu et al. (2013b), were selected according to their likely transport distances after erosion (Table 2). EQS represents the diameter of a nominal spherical quartz particle that would fall with the same velocity as the aggregated particle for which fall velocity is measured (Loch, 2001)."

(2) I was somewhat skeptical of the extrapolations the authors made from their findings from one experiment to the global impact of erosion and deposition. However, they were careful to note the shortcomings of such projections. For example, they note in the Methods section that although they used a suitable soil for the purposes of the study, more soils need to be investigated. In addition, they discuss that the respiration rates of the soil fractions were estimated from very brief observations, but the authors note that these are instantaneous respiration rates that don't represent long-term decomposition processes. It might be seen as a weakness, but the fact that they are even considering how aggregate protection plays into the carbon balance of erosion is a strength of the paper.

Answer: Thank you very much for thorough reading this manuscript and your generous appreciation.

(3) Section 3 L15–20: Can the authors provide evidence that SOC in the various classes didn't differ over time? E.g. means and variation for the time points or a P value if this was explicitly tested?

Answer: Both the weight distribution and EQS specific SOC content did not differ significantly among samples over rainfall time (ANOVA, single factor, $P > 0.05$, $n = 18$)

(4) Technical concerns Page 8839 L6: typo in "eroded"

Answer: Typo will be corrected in the revised manuscript.

List of all the relevant changes made in the manuscript:

L20 to 23 now reads: Both the eroded sediments and undisturbed soils were fractionated into six different size classes using a settling tube apparatus according to their settling velocities: > 250, 125 to 250, 63 to 125, 32 to 63, 20 to 32 and < 20 µm.

L69 to 73 now reads: In addition, the upper limits of the mineral particle size classes used in current erosion models are often smaller than the actual aggregate sizes. For instance, the largest class in van Oost et al. (2004) is only 90 µm, whereas up to 250 µm for rill erosion in Morgan et al. (1998). Such limits may also skew the estimation on settling velocity of eroded sediment.

L84 to 94 now reads: For instance, physically-stabilized SOC within macro- and micro-aggregates is protected by forming physical barriers between microbes and enzymes and their substrates, and thus very susceptible to mineralization after aggregates break-up (Six et al., 2002). Chemically-stabilized SOC results from the chemical or physicochemical binding between SOC and soil minerals (i.e., clay and silt particles). Such stabilization is also likely to be disturbed by aggregates break-up, as often occurs during erosion and transport (Starr et al., 2000; Lal and Pimentel, 2008; van Hemelryck et al., 2010). Biochemically-stabilized SOC is resulted from the inherent or acquired biochemical resistance to decomposition. Aggregates break-up might also affect the biodegradability of the SOC or the exposure to hydrolyzation.

L107 to 126 now reads: A-horizon material (top 20 cm) was sampled from a gentle shoulder slope (< 5%) in March 2012. Previous research on the same silty loam showed that aggregation increased the settling velocity of original soil fractions, particularly the medium sized fractions, in comparison with that expected based on the texture of the original soil (Hu et al., 2013b). The mineral particle specific SOC distribution, average SOC content (LECO RC 612 at 550°C), and aggregate stability of original soil (method adapted from Nimmo and Perkins, 2002), are shown in Table 1. The mineral particle size distribution was fractionated by wet-sieving, after dispersion by ultrasound using a Sonifier Model 250 from Branson, USA. The energy dissipated in the water/soil suspension was 60 J·ml^{-1} (i.e. Energy = output power 70 W × time 85 s / suspension volume 100 ml). The SOC mass proportions across mineral particle size classes were calculated only from average values of individual weight and SOC content. Although the ultrasound energy used in Hu et al. (2013b) was not enough to thoroughly disperse the original soil into real mineral particles (Kaiser et al., 2012), such extent of dispersion was notable enough to demonstrate the potential under-estimation of applying mineral particle size distribution to predict the settling velocity of eroded SOC. Hence, it is speculated that similar increasing effects would also occur to sediment fractions, and thus make the silty loam suitable to investigate the potential effects of aggregation of original soil on the transport distances of differently sized sediment fractions.

L180 to 186 now reads: The settling tube apparatus consists of four components (Figure 1a): the settling tube, through which the soil sample settles; the injection device, by which the soil sample is introduced into the tube; the turntable, within which the fractionated subsamples are collected; and the control panel, which allows an operator to control the rotational speed and resting/moving intervals of the turntable. Details about the settling tube apparatus were described in Hu et al. (2013b).

L189 to 191 now reads: In total, there were six sediment collection intervals over the 3 h rainfall events, and a settling fractionation test was carried for each of the six sediment collection intervals.

L199 to 201 now reads: EQS represents the diameter of a nominal spherical quartz particle that would fall with the same velocity as the aggregated particle for which fall velocity is measured (Loch, 2001).

L232 to 237 now reads: The re-wetting was done on the previous day before the respiration measurements. This way, the initial CO_2 pulses of rewetted soils should be largely excluded (Orchard and Cook, 1983). Even if there were any CO_2 pulses induced by rewetting, this exactly mimics the natural processes, where dry sediments deposited from previous rainfall events, experience a second time of erosion and transport processes.

L259 to 261 now reads: During the simulated rainfall events, the sediments were seen to move continuously with runoff, and no evident selective deposition was observed on the soil surface.

L264 to 267 now reads: Preliminary data analysis had shown that while the absolute sediment mass increased during the simulated rainfall events, the proportional composition of six EQS classes in each sediment collection interval did not differ significantly between samples (ANOVA, single factor, $P > 0.05$, $n = 18$).

L316 to 320 now reads: The distinct SOC distribution across aggregate size classes also agrees with the field investigation by Polyakov and Lal (2008), where the coarse aggregates (1 to 0.5 mm) fractionated by wet-sieving, contained up to 4.5 times more SOC than the finest fraction (< 0.05 mm).

L356 to 369 now reads: A 15.5% SOC enrichment of sediment re-deposited in the terrestrial system corresponds to the proportion of eroded SOC estimated to be deposited in permanent sinks (e.g., 0.12 Pg of SOC eroded per year by van Oost et al. 2007). While the effects of aggregation on SOC redistribution and subsequent fate cannot be assessed based on one experiment, most sediment is transported in form of aggregates (Walling, 1988; Walling and Webb, 1990). Ignoring the effect of aggregation on erosion and redistribution of SOC, therefore, bears the risk of overestimating the erosion-induced carbon sink effect. As a consequence, the behavior of aggregated sediment requires a reconsideration of existing approaches of sediment behavior in erosion models. Further study of different soil types, their aggregation and aggregate breakdown while moving through landscapes of varying topography during rainfall events of different intensity, frequency and duration, is required to assess the relevance of aggregation for SOC movement and fate identified in this study.

L380 to L385 now reads: As a consequence, the preferentially deposited SOC could potentially generate a further error in the carbon source-sink balance. Such error would be particularly significant, when repeated erosion and deposition processes along hillslopes cause further disintegration of large aggregates (Kuhn et al., 2003; van Hemelryck et al., 2010). This would thereby result in additional SOC exposure and mineralization (Jacinthe et al., 2002; Six et al., 2002).

L418 to 421 now reads: More simulations as well as field experiments are also needed to examine the effects of various transport processes (such as slope length, slope gradients, field barriers) onto the mechanism of aggregate breakdown and aggregate specific SOC distribution.

Tables: As suggested by the reviewers, the notes which were previously below Table 1 and 4 are now rephrased and added to the Method from L107 to 126.

Figures: As requested by the reviewers, the standard deviation of the original soil was added Figure 3, 4, and 5.