Response to Reviewer 2

We would like to thank the anonymous reviewer for his or her constructive comments. In this response we provide an answer to all the comments and then indicate the changes that will be applied in the revised manuscript.

Comment 1: *First of all the paper lacks clear aims (or research questions). In Line 68 to 83 the authors give an overview of the contents of the paper, but I think the entire paper would improve substantially if clear aims would be given here. For example, (i) introduce a coupled soil erosion and C turnover model with an LSM model which is applicable on regional scales. (ii) Rigidly test the model for the Rhine Catchment against other modelling results and regionally available data. (iii) Analyze the sensitivity/uncertainty of the model results due to weak input data and a priori model assumptions. (regarding (iii) see comments below.*

Answer: We will clarify the aims of our study in the Introduction section of the revised manuscript. See potential changes below.

Changes to manuscript:

L76-92: To address these knowledge gaps, we present a parsimonious process-based Carbon Erosion Dynamics Model (CE-DYNAM), which integrates sediment dynamics resulting from water erosion with the SOC dynamics at the regional scale. The SOC dynamics are calculated consistently with drivers of land use change, CO_2 and climate change by a process-based land surface model, with a simplified reconstruction of the last century increase of crop productivity. This modelling approach consists of a global sediment budget model coupled to the SOC removal, input, and decomposition processes diagnosed from the ORCHIDEE global land surface model (LSM) in an offline setting (Naipal et al., 2018). The main aim of our study is to quantify the horizontal transport of sediment and C along the continuum of hillslopes, floodplains and rivers, and at the same time analyze its impacts on the land-atmosphere C exchange. We calibrate and validate the new model with regional observations and high-resolution modelling results of the Rhine catchment. It should be noted here that the structure of CE-DYNAM is designed in a way that the model can be adapted easily to other large catchments and finally run globally. Finally, we also discuss the model uncertainties and the sensitivity of the model to changes in key parameters and assumptions made. In the next sections we give a detailed overview of CE-DYNAM model structure, the coupling of erosion, deposition and transport with the coarse-resolution SOC dynamics of ORCHIDEE, model application and validation for the Rhine catchment, and its potentials and limitations.

L83-88 goes to the discussion section where we also discuss the potential of CE-DYNAM: to combine erosion, transport, and re-deposition of soil material, for which small scale differences in topography are of utter importance, with a state-of-the-art representation of large-scale SOC dynamics driven by land use, climate, and atmospheric CO_2 as simulated by the ORCHIDEE LSM. The flexible structure of CE-DYNAM makes the model adaptable to the SOC dynamics of any other LSM. By coupling soil erosion with the C dynamics of LSMs it is possible to study the main processes behind the linkages of soil erosion and the global C cycle.

Comment 2: *Taking the temporal and spatial scale into account which should be later on analyzed with the model I think the authors found a good balance between model complexity and simplicity. However, the model is full of a priori assumptions, which will fundamentally affect the modelling results, so I personally do not think any model results can be interpreted without some estimates of at least the sensitivity of the model against these assumptions. The most important assumptions which could be tested easily are: C input via plants especially crops depending on erosion status, C enrichment during erosion and depletion during deposition, reduced C turnover in alluvial soils due to wetter conditions, etc. Overall, it is one of the major shortcomings of the paper that the modeling results in section 3.2 are presented single values (e.g. for 159 Tg C for C removal by erosion) and also conclusions based on this single model results are presented. I strongly suggest performing a a sensitivity analyses (including as far as possible effects of a priory assumptions) and giving results with a reasonable range. I am fully aware that it would be hardly possible to do a full uncertainty analysis and even an sensitivity analysis might be quite ambitious given the catchment size and the complexity of the involved* *models. However, it is not enough just stating in the discussion some important processes are not taken into account.*

Answer: We agree that an uncertainty analysis is important for a regional modelling study such as ours. Therefore, we performed additional simulations with a minimum and maximum soil erosion scenario, based on the uncertainty ranges in the rainfall erosivity and land cover factors of the Adjusted RUSLE model. Chapter 3 of the revised manuscript will be modified to include the new uncertainty results. We will also modify figure 9 to include the uncertainty ranges in the C budget components.

Regarding the sensitivity analysis of the model we tested the assumption of C enrichment during erosion as suggested by the reviewer. Here, we performed two additional simulations with an enrichment factor of two adapted from the study of Lugato et al. (2018): S1_EF (erosion only) and S2_EF (erosion with deposition and transport). We also tested the rate of C transport between floodplains by letting the basin average sediment residence time to vary between a 50% lower and 50% higher value compared to the default. For this purpose we did another two additional simulations (S2_Tmin and S2_Tmax). However, we abstained from testing the model performance to a changed C turnover in alluvial soils as a result of wetter conditions. Previous studies show that there are still large uncertainties related to the turnover of C in depositional environments, and more specifically of alluvial soils, as they represent complex soil profiles with a wide range in physical, chemical and biological parameters that affect the C turnover in interaction with climatic variables such as soil moisture. For example, the studies of Doetterl et al. (2018) and Rasmussen et al. (2018) show that the C turnover of alluvial soils is determined by C stabilization affected by the availability of minerals (such as Iron, Aluminium) and nutrients, mediated by soil microbes and by the formation of peat deposits on river banks. Yet, old alluvial soils can be far from water-saturated, in which case the C turnover would not be substantially decreased as a result of additional oxygen limitation. In our study we also include floodplains that do not get flooded regularly. Therefore, it is not clear if these alluvial soils are in general 'wetter' than the colluvial soils and would therefore have a significantly different C turnover. Also, our model does not include a good representation of groundwater dynamics and a soil moisture function for alluvial soils. After performing an extensive literature study on C turnover in alluvial soils we could not find a way to easily but realistically modify the C turnover of alluvial soils, for example by using a simple turnover reduction factor derived from observations. We will introduce a new section 4.3 in chapter 4, where we will discuss the results from the sensitivity simulations, see changes to manuscript below.

Changes to manuscript:

Section 4.3: Sensitivity analysis

The carbon enrichment factor (EF) represents a higher C concentration in eroded soil compared to the original soil, due to the selectivity of erosion (Bertol et al., 2017). By increasing the EF from 1 to 2, we assume a strong enrichment of C during erosion and find that this enrichment results in a gross C erosion flux that is 1.5 times larger compared to the flux without enrichment. This leads also to a larger dynamic replacement of C on eroding sites in combination with a larger burial in depositional sites, which is in accordance with the study of Lugato et al. (2018). The resulting C sink from the enrichment simulation is 1.5 times larger than the sink under default conditions. The cumulative POC flux during enrichment is somewhat lower (Table 5).

To show the potential effects of a different sediment residence time on the C dynamics, we performed a sensitivity study where we changed the basin average sediment residence time to be 50% higher or 50% lower but keeping the maximum sediment residence time at 1500 years. By changing the average sediment residence time and keeping the maximum fixed, it will be the grid cells with the lowest residence times that will undergo the largest changes in residence time and consequently in the floodplain SOC storage and export. The higher the residence time, the longer the deposited soil C will reside in the floodplains, where it can either be respired or buried in deeper soil layers. Therefore, we find that the effects of the sediment residence time on the SOC dynamics are non-linear. Under default conditions we find the highest SOC storage. A 50% higher average sediment residence time leads to the lowest total SOC storage, with a decrease of 20% compared to default conditions. Here, the erosional C sink is reduced by 19% (Table 5). This could be explained by a higher C decomposition flux for floodplains due to the long residence time of C in deposition areas. Especially, in mountainous regions where the soil erosion flux is large and removes a large part of the labile C, a higher sediment residence time will lead to higher C decomposition emissions in floodplains. The turnover seems to dominate over the C burial in deeper layers and export. A 50% lower average sediment residence time also leads to a decrease (of 9%) in the total SOC storage and a decrease of 8% in the erosional C sink compared to default conditions. Also here, the largest changes are found in the mountainous regions where a low sediment residence time leads to a large export of C, which is then deposited in lower lying, more extensive floodplains. Thus, increasing or decreasing the residence time leads to a smaller total SOC storage, resulting from different spatial distributions of this SOC storage (this will be shown in a Supplementary figure). The POC flux under the low sediment residence time scenario is substantially higher than under default conditions (Table 5).

	Ce	dSOC_tot	C sink/source	POC
Default	181	152	163	0.138
enrichment	269	230	240	0.137
τmin	181	139	150	0.198
τmax	181	121	132	0.117

Table 5: Sensitivity analysis. The impacts of enrichment and changes to the sediment residence time on the cumulative gross C erosion (Ce), the cumulative change in the total SOC stock ($dSOC_{tot}$), the net C sink and the cumulative particulate organic C export flux (POC) of the Rhine catchment. Units: Tg C

L297, Eq 15: $k_E = \frac{f*(\frac{E}{365})}{BD*dz} * EF$

Where EF is the enrichment factor, set to 1 by default.

Section 2.11: We also performed 4 additional sensitivity simulations and 4 additional uncertainty simulations. Simulation S1_EF and S2_EF are performed to test the model assumption of a C enrichment during erosion. Here, we changed the enrichment factor EF to two, based on the study of Lugato et al. (2018). Simulations S2_Tmin and S2_Tmax are performed to test the rate of C transport between floodplains. Here we modified the average sediment residence time for the Rhine catchment to a minimum of 60 years (50 % lower than the current value), and to a maximum of 128 years (50% higher than the current value), respectively. However, we kept the maximum sediment residence time at 1500 years.

For the uncertainty analysis we performed simulations S1_min and S2_min with a minimum soil erosion scenario, and S1_max and S2_max with a maximum soil erosion scenario. These soil erosion scenarios are based on the uncertainty ranges in the rainfall erosivity and land cover factors of the erosion model. See the supplementary material for information on how these uncertainty ranges are derived. All the model simulations are summarized in table 2.

Default simulations	Soil erosion	Tau	Enrichment
S0	0		
S1	3.94	94	1
S2	3.94	94	1
Uncertainty simulations			
S1_min	1.52	94	1
S2_min	1.52	94	1
S1_max	5.95	94	1
S2_max	5.95	94	1
Sensitivity simulations			
S2_Tmin	3.94	60	1
S2_Tmax	4.94	128	1
S1_EF	5.94	94	2
S2_EF	6.94	94	2

Table 2: Model simulations, with changes to the basin average gross soil erosion rate (t ha^{-1} y^{-1}), the basin average sediment residence time (years), and the enrichment factor.

L471-475: We find an average annual soil erosion rate of 4.66+- 2.22 t ha^{-1} $year^{-1}$ over the period 1850-2005, which is about two times larger than the average erosion rate simulated for the last millennium (Naipal et al., 2016) and about four times larger than the average erosion rate of the Holocene (Hoffmann et al., 2013). This soil erosion flux mobilized around 181+-117 Tg of C over the period 1850-2005, of which 35+-9 % is deposited in colluvial reservoirs, 65+-9 % is deposited in alluvial reservoirs and 0.08 % is exported out of the catchment.

L495-498: We find that the cumulative C erosion removal flux of 181+-68 Tg of C leads to a cumulative net C sink of 163+- 31 Tg C for the whole Rhine region (Fig 8E). This is about 1.3 – 1.9 % of the cumulative NPP and about 35-51% of the cumulative land C sink of the Rhine without erosion.

Comment 3: *From my understanding of the paper and accounting for the scope of the journal testing such new model against data is essential. The authors did try doing so but here a lot of improvement is easily possible: (i) Include a section under methods explain which data are used to test the model and also explain in some detail how this is done. For example, the comparison with other models as given in Fig. 3 and 5 is not clear, as the following information is missing: (a) Were the data from the more high resolution models aggregated to the raster cell size of CE-DYNAM to do a raster-by-raster comparison? (b) If a CE-DYNAM raster cell consist of erosional and depositional sites, which are not resolved in the raster cell, how to compare with gross erosion of a high resolution model (e.g. Panagos et al. 2015) which might have different proportions of erosional and depositional raster cells in this large 8 x 8 km 2 raster cell. (c) It is not clear at all what is compared as all model results from literature do not focus on the time span from 1850 to 2005. These details are essential for the reader to understand your model validation. (ii) From the figures I have some doubts that the different models fit very well (why not giving statistical goodness-of-fit-parameters?). So the question is how good the other models are (please see e.g. the scientific debate regarding the Panagos et al. (2015) map. So, at least in the discussion this model to model comparison needs to be stressed. (iii) Generally, the erosion (partly deposition) validation of CE-DYNAM is mostly done against other models also using USLE technology (USLE factors might be even derived from same data sources), so an extended discussion if this is meaning full is needed.*

Answer: To better clarify the model validation and comparison against data and other models, we will include an additional section 2.12 in the revised manuscript where we discuss the data used, and how the validation is done in more detail. In this new section we will also discuss reasons why we do this model to model comparison, where we will give more background information on the various models. Finally, we provide a statistical goodness-of-fit summary by comparing the total soil erosion rates of sub-basins of the Rhine . See changes to manuscript below.

Changes to manuscript:

Section 2.12: We performed a detailed model validation of the sediment and the C part of the model based on the following steps: (1) Validation of soil erosion rates using high-resolution model estimates for Europe from Panagos et al. (2015) and model estimates/observations from Cerdan et al. (2010), (2) validation of C erosion rates using high-resolution model estimates for Europe from Lugato et al. (2018), (3) validation of the spatial variability of C storage using observational results from Hoffmann et al. (2013), (4) validation of SOC stocks using observational data from a global soil database and a European land use survey.

For the validation of gross soil erosion rates we used the high-resolution model estimates from the study of Panagos et al. (2015), who applied the RUSLE2015 model at a 100m resolution at European scale for the year 2010. The RUSLE2015 is derived from the original RUSLE model with some modifications to the model parameters L, C and P. The erosion module of CE-DYNAM is also based on a modified version of the RUSLE (Adj.RUSLE) which , however, lacks the L and P factors. It calculates the potential soil erosion rate under the assumption of no erosion control scenarios, in contrast to RUSLE2015, which does represent erosion control practices. Adj.RUSLE also differs from RUSLE2015 in the use of more coarsely resolved input datasets (see table 1), for which the equations for the R and S factors have been modified. The extensive validation of the Adjusted RUSLE model in this study and previous studies (Naipal et al., 2015, 2016, 2018), shows that despite its coarse resolution, the methodology works for large spatial scales. In contrast, RUSLE2015 uses largely similar equations as in the original RUSLE model presented in Renard et al. (1997). Thus, even though both Adj.RUSLE and RUSLE2015 are derived from the same erosion model, the differences between the models are large, and would justify our model comparison.

We also used independent high-resolution erosion estimates from the study of Cerdan et al. (2010), available at a 1km resolution, which were based on an extensive database of measured erosion rates under natural rainfall in Europe. For the comparison we aggregated the high-resolution model results of both datasets to the resolution of CE-DYNAM.

For the validation of C erosion rates, we used the high-resolution model results from Lugato et al. (2018), where they coupled the RUSLE2015 erosion model to the Century biogeochemistry model. These model results were available at a resolution of 1km, where each gridcell was composed of an erosion and deposition fraction. The C erosion rate provided by Lugato et al. (2018) was multiplied with the erosion fraction of a 1km grid cell. Then, the C erosion rates were aggregated to the resolution of CE-DYNAM. In CE-DYNAM the C erosion rates from simulation S1 are multiplied with the hillslope area to get the total C erosion flux of a grid cell. As the study of Lugato et al. (2018) considers only agricultural areas, we considered only the crop fraction of a grid cell. Also, in this model comparison it should be noted that the SOC dynamics scheme of CE-DYNAM, which is derived from ORCHIDEE LSM, is based on the Century model. However, there are large differences between the Century model used by Lugato et al. (2018) and the C dynamics scheme of ORCHIDEE used in this study. For example, in the Century model the crop productivity is mediated by nitrogen availability, which is not the case in the ORCHIDEE version used for this study. The Century model also includes some management practices such as crop rotations, which are not represented in ORCHIDEE. The Century model runs at a much higher resolution and is calibrated for agricultural land, while ORCHIDEE also simulates forest, grasslands and bare soil. In this way, the final SOC stocks derived with CE-DYNAM are also a result of erosion from other land cover types and land use changes. This is an important feature for land use change, which is not included in the Century model. Furthermore, the ORCHIDEE LSM has been used in many global intercomparisons and extensively evaluated for C budgets (Mueller et al., 2019; Todd-Brown et al., 2013). Also an important advantage of ORCHIDEE is that it includes the last century change in crop production calibrated against data (Guenet et al., 2018).

For the validation of the spatial variability of the SOC stocks between hillslopes and floodplains we used the scaling relationships between basin area and SOC storage derived by Hoffmann et al. (2013) from C14 observational data of the Rhine catchment. Hoffmann et al. (2013) did an inventory of 41 hillslope and 36 floodplain sediment and SOC deposits related to soil erosion over the last 7500 years. They found that the sediment and SOC deposits were quantitatively related to the basin size according to certain scaling functions, where floodplain deposits increased in a non-linear way with basin size while the hillslope deposits showed a linear increase with basin size.

The study by Naipal et al. (2016) already applied these scaling relationships for the Rhine to validate the simulated sediment accumulation over the last 1000 years derived with the global sediment budget model. They find that the global sediment budget model is able to reproduce the scaling parameters, and after analyzing the dependence of the scaling behavior on the main parameters of the model, they argue that the scaling is an emergent feature of the model and mainly dependent on the underlying topography. This indicates that the scaling features of floodplain and hillslope sediment and C storage should also be applicable to the more recent time period, such as in our study. In our study we aim to evaluate the ability of CE-DYNAM to reproduce this scaling behavior for the SOC storage of the Rhine. For this purpose we selected the grid cells that contained the points of observation of the study of Hoffmann et al. (2013) and performed a regression of the basin area (defined as the upstream contributing area) and the SOC storage of that gridcell for floodplains and hillslopes separately. Comparing the individual sediment and SOC storages of each grid cell from Hoffmann et al. (2013) was not possible due to the difference in the time-period of the studies, where Hoffmann et al. (2013) focussed on the entire Holocene, while our study focussed only on the period from 1850 AD.

Section 3.1: We also compared the total gross soil erosion rates of 17 sub-basins of the non-Alpine region of the Rhine. We find that the spatial distribution of the gross soil erosion rates compare well to the findings of other studies. A summary of the statistical goodness of fit between the soil erosion rates is given in the following table.

	E Cerdan et al.	E Auerswald et al.	E RUSLE2015

r-squared	0.72	0.97	0.94
RMSE	0.68	1.98	0.92

Table 3: A goodness-of-fit summary of gross soil erosion rates (E) from CE-DYNAM and other studies. The RMSE is the root mean square error give in tons*1E-6

Specific comments:

Comment S1: *Line 68-83: see general comment.*

Answer: See our response to general comment 1

Comment S2: *Line 89: be more explicit regarding 'low number of parameters'*

Changes to manuscript: We will modify this part of the sentence to: "… 3) the low number of parameters compared to other carbon erosion models that operate at a high spatial resolution (Lugato et al., 2018; Billings et al., 2019), which allows running the model at large spatial scales…."

Comment S3: *Line 118 ff: I do not agree that not taking the L factor into account is a reasonable decision. I agree that it is somewhat difficult to estimate (for the German part of the Rhine catchment there are some estimates) but if you are interested in land use change it is an essential factor if you kick it out the entire basis of the USLE is set into question. (The P factor is simpler as it is set to 1 in most studies).*

Answer: We agree that leaving both the L and P factors out of the equation will induce some bias in the results, especially for agricultural land. In our next study we aim to make CE-DYNAM better applicable for agricultural land, where these factors play an important role. For this purpose we will focus on the development of new methods that can quantify the L and P factors reliably at the global scale, and will need to re-calibrate the erosion module of CE-DYNAM, the Adj.RUSLE. Our decision of leaving out the L and P factors from the erosion equation in our study is based on the global study of Doetterl et al. (2012), which showed that the S, R, C and K factors explain approximately 78% of the total erosion rates on cropland in the USA. This indicates that on cropland the L and P factors, which are related to agriculture and land management, contribute only for 22 % to the overall erosion rates. This percentage is comparable to the uncertainty range in the estimation of the S, R, C and K factors at the regional scale from coarse resolution data. Renard and Ferreira (1993) also mention that the soil loss estimates are less sensitive to slope length than to most other factors.

Furthermore, various studies argue that the estimation of the L factor for large areas is complicated and thus can induce significant uncertainty in soil erosion rates calculated based on coarse resolution data (Foster et al., RUSLE2 user guide; Kinnell, 2007). Especially, for natural landscapes, such as forest, the estimation of the L factor is not straightforward as these natural landscapes usually include steep slopes (Elliot, 2004). In order to stay consistent with the estimation of potential soil erosion for all land cover types, we remove the L factor from the equation. The Adj.RUSLE has been already successfully validated at the regional scale, without the L and P factors where the spatial variability of soil erosion rates compares well to other high resolution modeling studies and observational data and the absolute values fall within the uncertainty ranges of those validation data (Naipal et al., 2015; Naipal et al., 2016; Naipal et al., 2018; and this study). Finally, the aim of this study was to develop and validate a carbon erosion module for applications at the global scale, where the estimations of the L and P factors is even more limited. By showing that the erosion rates from the Adj.RUSLE and CE-DYNAM are within the uncertainty of other data and modelling studies, we can assume that it will be applicable for other large catchments in the temperate region.

Changes to manuscript: In section 4.2 'Model limitations' of the revised manuscript, we will address the lack of the L and P factors in more detail in the same way as described above.

Comment S4: *Line 130: The statement ". . .has been calibrated and validated for the Rhine catchment. . ." is confusing here? If calibration and validation was already done why doing it again? If the model has changed you need a new validation (but what about calibration? Are you using parameters in CE-DYNAM which were calibrated before it is necessary to indicate this in detail).*

Answer: In our study we use the model parameters of the global sediment budget model as defined and calibrated by Naipal et al. (2016), such as the sediment residence time or the floodplain deposition factor. We did not perform an additional calibration of the sediment dynamics part of CE-DYNAM, only a validation, because of the use of different input datasets.
Changes to manuscript: We will add a sentence in the revised manuscript specifying why we redid the validation.

Comment S5: *Line 113: Alluvial soils are indicated in German soil maps, so the statement is not correct for the largest part of the Rhine catchment.*
Changes to manuscript: We will modify the sentence to: 'It should be noted that global soil databases do not identify floodplain soil as a separate soil class, although national soil databases might. However, the aim of this study is to present a carbon erosion model that should be also applicable for other catchments and eventually, globally. Therefore, we followed a 2-step methodology to derive floodplains in the Rhine catchment using hydrological parameters and existing data on hillslopes and valleys.'

Comment S6: *Line 141 / Eq. 2: Generally I think it would be good being more precise with the equation. For example in case of Eq. 2, I would expect a reference to the different raster cells (Aft(i) = Lstream(i) x Wstream(i); whereas i is the raster cell.) as for other equations e.g. Eq. 4a it was not clear if this refers to the entire catchment is calculated for each raster cell.*
Changes to manuscript: We will modify the equations accordingly in the revised manuscript.

Comment S7: *Line 148: If alpha and b are constants it means that the upstream area necessary to result in a stream is always the same. I understand that in case of a large scale model simplifications are necessary but this assumptions is for sure not true for the Rhine catchment (see papers from hydrology of maps of the stream system (which by the way would be available for the entire Rhine catchment).*

Answer: We agree, that they might not be the same for the entire catchment, and variations exist. But these constants have been derived from 467 cross-sections of the Rhine catchment combining 1:25 000 geological maps and catchment area extracted from the SRTM 3arcsec digital elevation model (Hoffmann et al., 2007).

Comment S8: *Line 159: '. . . at 8 km resolution. . . " I guess this means 8 km x 8 km raster cells. Should be changed throughout the text (also with other resolution given).*
Answer: We will modify this accordingly in the revised manuscript

Comment S9: *Line 163: I do not think the the assumptions of reduced hydrological and geomorphological connectivity in arable landscapes (compared to forest) is correct. From the recent studies dealing with flash floods it is obvious that it is a main problem that this landscapes have a very high connectivity as so many ditches, drainages etc. were built over the last century to get rid of any surplus of water on arable land. So, your assumptions for the range of the parameter f in different landscapes must be underlined by reasonable data.*
Answer: This assumption is underlied by several studies (Hoffmann et al., 2013; de Moor and Verstraeten, 2008; Gumiere et al., 2011;Wang et al., 2014) on the effect of erosion on sediment yield, where is shown that man-made activities on agricultural landscapes result in a trapping of eroded soil in colluvial deposition sites, reducing the sediment transport from hillslopes to the floodplains. The model parameter *f* has been calibrated for the Rhine catchment before in Naipal et al. (2016), where this range is found to produce a ratio between hillslope and floodplain sediment storage that was comparable to observations. The studies of Wang et al. (2010; 2014) identify a range for the hillslope sediment delivery to be between 50 and 80 %, which is similar to the range in the (1-f) factor of our model.
Changes to manuscript: We will better underline the choice for the *f* parameter in the revised manuscript as described above.

Comment S10: *Line 187: Does a multiple flow algorithm makes sense in case of a resolution of 8 km*

Answer: The multiple flow algorithm is especially effective in hilly regions, and we expect it to work better than the single flow algorithm for these regions. In such steep landscapes the river courses are meandering a lot. So the downstream part of a river can easily cut through the boundary of two downstream lying adjacent 8 km x 8 km cells. With the coarse resolution of 8 km, a single direction algorithm would lead to an extreme straightening of the river network, and it would underestimate the number of cells which have a proportion of floodplain area. However, we agree that the coarse resolution grid size will affect the results of both algorithms.

Comment S11: *Line 230 ff: In general this is a reasonable assumption for the crop residues. However, studies in small catchment clearly indicate that residue management is a key factor of SOC, so this a priori assumption has potentially a huge effect on the produced results. So, its importance must be analyzed with the model!*

Answer: We agree that the harvest and crop residues left on the field would have a large effect on the SOC dynamics of agricultural landscapes. It should be noted that we implemented an increase in the harvest index during the period 1850-present-day based on the study of Hay (1995), which already partly accounts for crop residue management. To explicitly quantify the potential impacts of crop residue management we performed an additional sensitivity simulation where we assumed that all above-ground crop residues are harvested. After running CE-DYNAM to equilibrium we find that the total crop litter C stock is about 70% smaller under the extreme crop residue management scenario. This leads to a total SOC stock (under steady state conditions) that is 60% smaller under no erosion (S0), and 35% smaller under erosion (S2). The transient sensitivity simulations considering crop residue management are still running. However, the preliminary results confirm that soil management practices such as residue management have a substantial effect on the SOC dynamics. We will present these findings in section 4.3 of the revised manuscript, sensitivity analysis.

Changes to manuscript: Table 2 will be modified to include the new sensitivity simulations and the new section 4.3 will include the discussion on the results of the effects of crop residue management.

Section 2.11: Input data and model simulations

We performed 3 additional sensitivity simulations (S0_RM, S1_RM, S2_RM) where we test the model assumption on crop residue management. In these simulations we assume that all the above-ground crop residues are continuously harvested.

Section 4.3: Sensitivity analysis

Here, we investigate the effect of reduced C input into the soil by an crop residue harvesting. After performing sensitivity simulations where all the above-ground crop residues have been harvested, we find that under steady state conditions the total crop litter C stock is about 70% smaller compared to the default case. This leads to a total SOC stock that is 60% smaller under no erosion (S0), and 35% smaller under erosion (S2). Our findings confirm that soil management practices such as residue management have a substantial effect on the SOC dynamics.

Comment S12: *Line 238-240: The given equation are a fundamental problem with modelling the effect of soil erosion on SOC turnover. For example, using standard SOC pool residence times for all landscape positions is of tremendous importance for the entire C balance effect of erosion. So, again it would be very important to know how sensitive the results are against this assumptions. At least give some estimates / measurements at different landscape positions in the discussion and comment of the potential effect in modelling results.*

Answer: The SOC pool residence times at different landscape positions (hillslope, depositional sites etc), is interrelated with weathering, soil erosion and sediment transport processes (Berhe et al., 2008). To be able to have different residence times for each SOC pool as a function of the landscape position, soil erosion effects on for example the aggregation of soil particles and transport of minerals and nutrients has to be included. This can currently not be done in global land surface models. However, there are efforts to change the current SOC dynamics scheme of

LSMs by introducing measurable SOC pools (Abramoff et al., 2018). In this case it might be possible to calibrate the residence time of each pool based on the landscape position.

Changes to manuscript:

Section 4.2: Model limitations

The current SOC scheme of CE-DYNAM does not account for different residence times of SOC as a function of landscape position along a hillslope. The SOC decomposition rates can vary significantly along a hillslope due to changes in soil moisture, temperature, aggregation, and the transport of minerals and nutrients (Burke et al., 1995; Doetterl et al., 2016). Currently, these processes are not resolved in coarse resolution LSMs, contributing to the uncertainty in the large-scale linkage between soil erosion and SOC dynamics.

Comment S13: *Line 265: ". . . The next soil layer contains less C and therefore at the following time-step les C will be eroded under the same erosion rate. . ." If this would be always true one would expect a continuous decline in SOC in soils. However, assuming a long- term forest use on a slope you will found the soil in an equilibrium between new C input via plants and small amount of erosion. So, in this case the eroded material will have a more or less constant C content.*

Answer: We understand that this sentence might be unclear. What we mean is that the reduction of C erosion at each timestep due to less C being available for erosion, and the existence of a compensatory C sink due to the erosional removal of C, will ultimately lead to an equilibrium state. Which is also reached in our model.

Changes to manuscript: We add the following sentences: " The removal of C by erosion also triggers a compensatory C sink due to the reduction in SOC respiration on eroding land. This compensatory C sink and reduced C erosion over time will ultimately lead to an equilibrium state."

Comment S14: *Line 277: Calculating a daily erosion fraction is a reasonable approach. However, if taking the episodic nature of erosion and deposition into account the C balance will be different compared to a small continuous process (see literature). Might be also discussed.*

Answer: We will mention this aspect in the discussion chapter of the revised manuscript.

Comment S15: *Line 291 ff: The assumption that there is no C selectivity (enrichment in eroded material and depletion at erosional sites) is taken in many modelling approaches. However, if there would be no enrichment of fines in the sediments transported in river systems, one would find e.g. sand in suspended sediments of larger rivers. Which is e.g. in case of the lowland Rhine not the case. Discuss this in the context of the scale of your paper. Also important regarding the loss of C to the ocean.*

Answer: We agree and performed a sensitivity analysis of the model where we changed the C enrichment factor to 2 , hereby, partly accounting for the selectivity of erosion. We find that although the POC export to the ocean is not significantly affected, the SOC storage and resulting C sink is increased. See our detailed response to general comment 2.

Comment S16: *Line 341: Where do the data regarding afforestation during the last two decades come from. To my knowledge this is a process already started in the late 1959th (please give reference)*

Answer: We use the land use change data from the study of Peng et al. (2017), who bases their estimates of forest cover on Houghton (2003,2008) for the period 1850-1990 and satellite data for the recent decades. The historical national forest area used by Houghton et al. since 1850 are from national surveys and they are arguably the best data available, although uncertainty arises when downscaling historical forest area change on a grid, using (uncertain) gridded reconstruction of agricultural land (HYDE) and land use transitions rules (see details for the method used by Peng et al. 2017).

Comment S17: *Line 379: (see also general comments). I wonder why you did not use other more specific and potentially profound national data. E.g. for Germany there are several maps for potential erosion which are much more elaborated than the map of Panagos et al. (2015). Moreover, I wonder why you did not use the sediment delivery data of the Rhine which are freely* *available - I guess since the 1950th - which would be a good and reliable additional data set for validation.*

Answer: We did not use the sediment delivery data of the Rhine, because the comparison to our simulated coarse resolution model results will most likely be not entirely justified. In our model we do not take into account daily changes in precipitation and runoff and how that affects the erosion rates and sediment transport. Instead we use yearly totals. We also do not take into account dams and other man-made structures that would affect the river transport of sediment. Further, we focus only on the rill and interrill erosion and do not account for other soil erosion processes and flash floods that might have a larger effect on the sediment delivery. Finally, our model has not been developed to simulate the river transport of sediment and C, but instead is focused on the redistribution of soil on land and the resulting sediment and SOC storage. See also our response to the general comment of reviewer 1. For the future development of CE-DYNAM we aim to better represent the river transport processes of sediment and C.

Regarding the validation of soil erosion with observations, we used the database of Cerdan et al. (2010), which already includes German national estimates on soil erosion. We also performed a comparison of our agricultural soil erosion model estimates to the agricultural soil erosion map of Auerswald et al. (2009), available at 250 m resolution, in the revised manuscript (see figure below).

Changes to manuscript:

Section 2.11: For the validation of gross soil erosion rates we also used the German national estimates on agricultural water erosion from Auerswald et al. (2009). This national data was available at a resolution of 250 m and is based on standardized erosion measurements from 27 studies and 1067 plot years.

Section 3.1: Inclusion of an additional figure (fig 3d) to show the comparison of our simulated agricultural soil erosion rates to the German map on agricultural soil erosion (Auerswald et al., 2009)

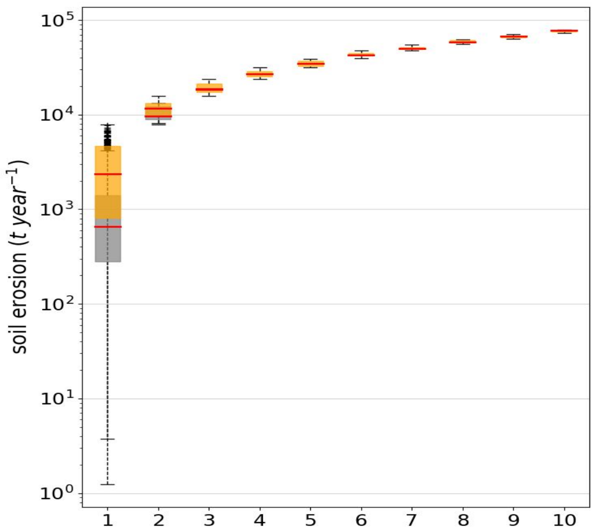

figure 3d: Quantile-whisker plot of simulated gross agricultural soil erosion rates (t/year) (grey whisker boxes), compared to (A) national estimated of agricultural soil erosion in Germany (Auerswald et al., 2009)

Comment S18: *Line 397-401: I suggest omitting these sentences and Fig. 4, because I do not see any additional value of this here. It is obvious from the model structure of all USLE based models (and all other erosion models) that an increase of erosivity and slope directly leads to an increase in erosion. Moreover, there is a coincidence in the catchment that highest erosivity and highest slopes occur at the same alpine area, but this is not any proof for the model. Hence, hence I think this is weakening your validation more that it would strengthen it. By the way: Erosivity and slope might explain 70% erosion if very different rainfall regimes and slopes (mountain areas and lowlands) are compared, but with a catchment like the Rhine (where except for the alpine part) the differences in slope and erosivity are relatively small soil cover (C factor) is getting much more important (erosion rates between grassland and arable land vary by a factor of 10-20).*

Answer: We agree and will omit figure 4 in the revised manuscript.

Comment S19: *Line 402: As I modeler I expect a goodness-of-fit parameter with this statement. (See also general comments regarding model to model comparison of different USLE implementations).*

Answer: We will include a goodness-of-fit summary related to the soil erosion result comparison, see our response to comment 3.

Comment S20: *Line 424 ff: The comparison with the data from Hoffman et al. (2013) underlines a deficit in all your comparisons. It is at no time clear what is compared exactly. Mean of 7500 years against 1850-2005?*

Answer: See the last paragraph in our response to general comment 3 and modifications to the manuscript.

Comment S21: *Line 438: Does the outflux fit to measured data? Would be easy to test even if this is not essential as only a very small amount will be delivered into the sea. (could be tested at several subcatchment, as data are available).*

Answer: See our response to general comment 1 of reviewer 1

Comment S22: *Line 451-452: This is a clear contradiction to your statement that differences in erosivity are very important for spatial differences in erosion.*

Answer: We agree and will remove this sentence in the revised manuscript.

Comment S23: *Line 453: The close link between C erosion and soil erosion is obvious from your modelling structure but not necessarily correct (C enrichment depending on event size?)*

Changes to manuscript: We will add the following sentence to this statement: "It should be noted that the correlation between soil and C erosion might be affected by processes not properly captured by the model such as the selectivity of erosion, which may also include the enrichment of C in eroded material."

Comment S24: *Line 434 ff: See also general comments*

Answer: We address this in our response to the general comments

Comment S25: *Line 446: I suggest not to over interpret the modeling results from the alpine area of the catchment as the modelling and the data are weakest there. (i) Increase in measured precipitation most uncertain; (ii) calculated R factor very uncertain in all USLE approaches; (iii) alpine USLE factors not very well underlined by data (compared to arable and grassland),*

Answer: We are aware of this bias and, therefore, presented model results without the Alpine region (see for example figures 8 and 9).

Changes to manuscript: We will add a sentence in the revised manuscript in the same section mentioning these biases for the Alpine region.

Comment S26: *Line 497ff. See comments regarding connectivity above. Moreover, even if the connectivity is high under forest (which I doubt), forest will produce not a lot of sediment and hence are not so important for building up alluvial soils at all.*

Answer: We agree that soil erosion in forests is minimal but forests also appear often in hilly landscapes that will contribute to the sediment production. When analyzing soil erosion rates over timescales longer than a few decades, extensive forest areas will contribute significantly to the overall removal of soil. Forests contain also a lot of SOC, and so minimal rates of soil erosion might be still significant for the SOC dynamics, also in depositional areas, as is demonstrated in the recent study of Billings et al. (2019).

Comment S27: *Line 493ff. I do not see from the results the CO2 fertilization plays an important role for an increase in dynamic replacement. I guess that the increase in yields due to changes in management are much more important (as reduced yields are not taken into account at erosional sites) as they boost dynamic replacement of eroded soils.*

Answer: We agree that increased yields due to management boost the dynamic replacement of eroded soil but the largest effect comes from the CO_2 fertilization due to increased atm. CO_2

concentrations. See below figure 1 representing the actual C replacement under land use change and climate change, and figure 2 representing the potential C replacement under a fixed climate (no change in temperature, precipitation, CO_2 atm. concentrations, but with a changing land use and management).

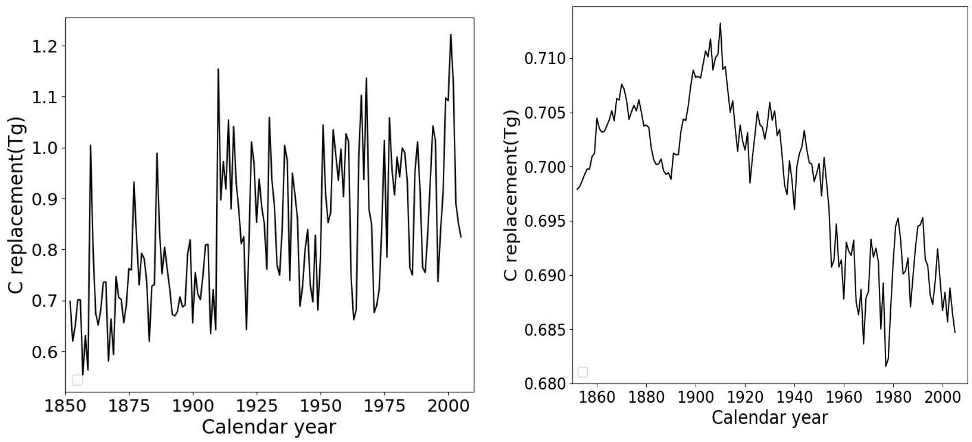

Fig 1 & 2: C replacement on eroding soils

Comment S28: *Line 501ff: It is obvious from first order kinetics that colluvial soils must have higher CO2 effluxes as they contain more C. So, this is not a very new finding.*

Answer: We agree that this might be an obvious finding, but this has not been quantified for such a large catchment before. With this finding we also indicate that the model reproduces process knowledge from field work.

Comment S29: *Line 506-507: Question: Is the modelled increase in respiration from floodplains resulting from an temperature increase or from an increase in depositional material, which would also result in an increase of respiration? Comment: Under real conditions the increase in respiration from floodplains is also a result of decreasing groundwater levels,*

Answer: In this case the increase in respiration of floodplains is mainly due to the additional deposition of material. In our model we have not a specific representation of ground water for floodplains.

Changes to manuscript: We will clarify this in the revised manuscript.

Comment S30: *Line 542-561: This is a nice collection of model deficits. However, for a modelling paper I would expect a bit more (see general comments).*

Answer: See our response to the general comments and resulting improvements made to the manuscript.

Comment S31: *Line 576-588: I think this conclusions are not fully supported by the results as the modelled C fluxes might be affected by a priori assumptions and model parameters which are not tested enough (see general comment regarding sensitivity analysis).*

Answer: We will add a 4th key finding in the conclusions related to the results of the sensitivity analysis.

Changes to manuscript: After performing a sensitivity analysis on key model parameters we find that the C enrichment by erosion, crop residue management and a different spatial variability of the residence time of floodplain sediment can substantially change the overall values of C fluxes and SOC storages. However, the main findings, such as soil erosion being a net C sink for the Rhine catchment, are not changed.

Comment S32: *Table 1: I guess the spatial resolution is always given in raster cells, e.g. 0.25 ◦ x 0.25 ◦*

Answer: We will change this in the revised manuscript

Comment S33: *Table 2: As the resolution of the data sets are different how to make sure that the comparison fits, e.g. the higher resolution data set might exclude the river network from the SOC calculation while the lower resolution data set might include this areas into the SOC stock calculation. Give somewhat more details.*

Answer: We give more details on the resolution and comparison in the new section 2.12 of the revised manuscript. See our response to comment 3.

Comment S34: *Fig. 3 and 5: What are the 10 classes given on the X-Axis?*

Answer: The x-axis represents bins or evenly spaced ranges between the minimum and maximum erosion or deposition rates. We will change the figure caption to include this information in the revised manuscript.

Comment S35: *Fig. 4. Omit this figure (see comment above)*

Answer: Will be removed

Comment S36: *Fig. 5c/d: What does it mean if CE-DYNAM has erosion rates which are up to a factor smaller than the Lugato model and C deposition rates which are more or less the same? Is it a result of different areas affected by erosion and deposition? Should be explained / discussed.*

Answer: This was a mistake as we used different bins on the x-axis for the different datasets. after changing the ranges for the bins we find that the simulated rates and those of Lugato et al. are similar. Thank you for pointing this out.

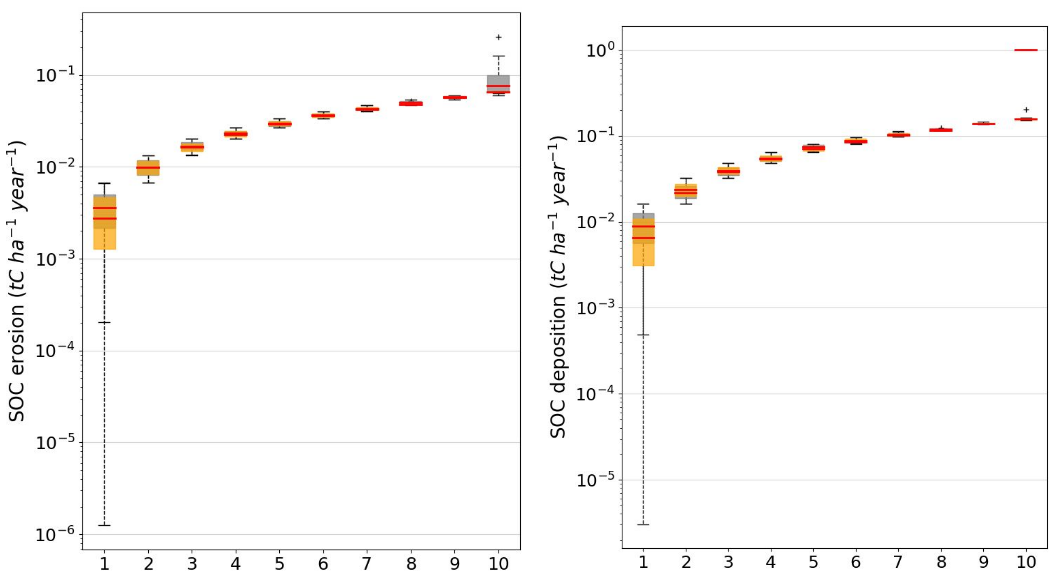

Figure 5c & d

Comment S37: *Fig. 7. Just a comment of an handicapped. About 4-8% of the male population are to a certain extend color blind (especially red/green is problematic), so if you do not what to* *lose these proportion of your readers you should adapted your color in your figures. There are color blind friendly color ranges available in most software packages. If the dashed lines range between min and max the outliers cannot be above or below the lines. So, I guess the lines represent something else.*

Answer: We apologize for this oversight and will use different colors for the figures that are better adapted for color-blind readers. The dashed lines do not represent the min and max but the outer extremes. The outliers are defined as values that are larger than the 3th quantile by at least 1.5 times the interquartile range (IQR), or smaller than 1st quantile by at least 1.5 times the IQR.

Changes to manuscript: We will adapt the colorscheme of the figures in the revised manuscript