We would like to thank the referee for taking the time to read and comment on our manuscript, which will help us improve the abstract and the overall manuscript.

Answers to individual points raised can be found below.

Referee 2

Overall, I thoroughly enjoyed this manuscript. However, I am concerned that much of the language in the discussion is speculative in nature.
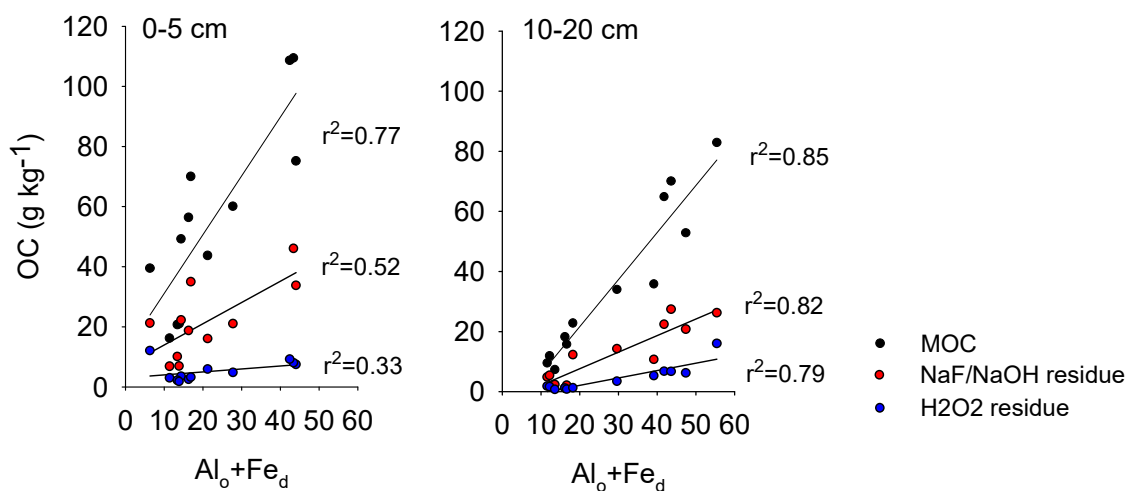We agree with the referee that the discussion contains some speculation. The reason is that many results did not meet the expectations and that we therefore offer alternative hypotheses to explain them. We will make this point clearer and reduce the degree of speculation in the revised version.

I do not believe that reasonable speculation should not be allowed, as true statistical replication in soil studies is sometimes difficult to achieve due to measurement cost (NMR) or sample availability. But the number of statements basing interpretation of mechanism or process which are based on data from one sample or site seems very high in this manuscript.

We agree with the referee that more real replicates would help to underline the obtained results and others might prove us wrong in the future. However, for now we have to interpret the pattern that we see in the obtained result. With five study sites from different places in Europe and true replicates within the sites we are also not at the lower end of sample numbers being analysed in experimental studies considering topsoils and subsoils. The main result of our study is anyhow that the sites have more in common than we expected and so we focused the discussion rather on common patterns than on differences among individual sites. We agree that with only one sample being analysed for the NMR results of bulk MOC and MOC residues, the observed differences can only hint at potentially underlying mechanisms and mostly serve a basis for developing new hypothesis to be further tested in the future. The similarity of the spectra of extracted OC among sites and the uniform extractability suggest, however, that the idea of similar trends at other sites is not fully unrealistic. Nevertheless, we will reduce the degree of speculation and highlight the new hypotheses more clearly in the revised version. Besides the costs, also technical problems with low signal-to-noise ratios in C-poor and oxide-rich samples reduced the broader application of NMR for more samples in our study.

The authors could focus solely on interpretation of how extraction chemistry influences the portion and composition of extractable organic matter, and that would be sufficient. The comments linking NaF-extractable C with Al/Fe oxyhydroxides don't seem well supported by the data. Figure A2 illustrates a significant relationship between H2O2 residues and extractable Al/Fe. Why are there not similar plots for NaF residues?

Since we already showed in Schrumpf et al. (2013) that HF-OC was closely related to Al/Fe oxyhydroxides and NaF/NaOH extractable OC represented roughly 60% of the total HF-OC across sites, we thought it would not be necessary to show that plot. However, we are happily willing to add the graph shown below.

0-5 cm

OC (g kg^{-1})

r^2=0.77

r^2=0.52

r^2=0.33

120
100
80
60
40
20
0

0 10 20 30 40 50 60

Al$_o$+Fe$_d$

10-20 cm

r^2=0.85

r^2=0.82

r^2=0.79

120
100
80
60
40
20
0

0 10 20 30 40 50 60

Al$_o$+Fe$_d$

- MOC
- NaF/NaOH residue
- H2O2 residue

From the abstract it isn't clear from the abstract why these two particular extractants were chosen. This may be a major gap in my knowledge base, and if that is the case, please ignore. I have never previously seen an experiment that utilized a mixture of NaOH and NaF to evaluate SOM extractability/solubility. Is there a reference that can be associated with this method? If the abstract included an explanatory statement, it would be very helpful.

First of all, we thank the referee for the critical look at the abstract and are happy to extend the description and motivation of the methods further if the editor agrees that such an extension is not too detailed for an abstract and would make it overly lengthy.

We assume that the referee refers to the combination of NaF and NaOH as extractants since we thought that the reason for using NaF/NaOH and H_2O_2 was explained in the abstract. Both, NaOH alone and in combination with NaF have been used and published in the literature as extractants for OC before (see e.g. Möller et al. 2000 Aust. J. Soil Res., Kaiser et al. 2007 SSSAJ). The reasons for using NaOH were that it has the power to remove much of previously sorbed dissolved organic carbon from mineral surfaces, and because the high pH increases the deprotonation, and thus, the solubility of organic acids. Since NaF is more strongly competing for binding sites while less dispersive than $Na_4P_2O_7$, the combination of NaF-NaOH was used to study a potential maximal desorption. However, since it is not common to have references in the abstract, we would keep those in the main manuscript, but add the following sentence explaining the reason for using this extract in the abstract:
"The combination of NaF-NaOH was used because F$^-$ is a strongly sorbing anion capable to replace anionic organic molecules from mineral surfaces and the high pH of the extract additionally supports desorption and solubility of MOC."

It sounds as if H2O2 is meant to represent oxidizable SOM, which the authors may be arguing is not very representative of the portion of SOM which is readily exchangeable under field conditions. In comparison, a NaOH and NaF solution presumably extracts SOM which is actually bound on the exchange surfaces of minerals. If this is indeed the case, it needs to be stated more explicitly in the abstract. Maybe something similar to the statement on lines 102-104 of the introduction.

Actually, we already wrote in lines 13-15 in the abstract: "Therefore, we determined the extractability of MOC into a mixture of 0.1 M NaOH and 0.4 M NaF as a measure for maximal potential desorbability, and compared it with maximal potential oxidation in heated H_2O_2." We believe that this sentence already demonstrates the main intention for using the two different treatments. It can be complemented by stating that they address different aspects of MOC stability: H_2O_2 addresses the chemical stability of the molecules, and NaF-NaOH the possible displacement and mobilization of the molecules from mineral surfaces by competing ions.

In the abstract, the language describing trends in radiocarbon data uses all the following terms: 14C, 14C content, MO14C, 14C-depleted, older, and age. It would improve the clarity of the results if the language was more consistent when describing this data.
Since depletion and enrichment are relative terms (depleted in comparison to what?), it seems necessary to include actual values with associated error terms when describing significant differences between extractions and residues.

Following the recommendation of the referee, we will add absolute ^{14}C values and error measures to the abstract and harmonize the language used to describe the results.

The abstract fails to mention that the soils were density fractionated prior to conducting the extractions.
To fulfil this request, we will change the sentence in lines 15ff. in the following way:
"We selected MOC samples (>1.6 g cm^3) obtained from density fractionation of samples from three soil depth increments (0-5 cm, 10-20 cm, 30-40 cm) of five typical soils of the mid-latitudes, differing in contents of clay and pedogenic oxides, and being under different land use."

It is unclear in the abstract how the experimental results could lead to the conclusion that MOC was dominated by OC interactions with pedogenic oxides. This seems like a complete non sequitur.

The conclusion on the role of pedogenic oxides in lines 36ff of the abstract is just repeating what was stated and introduced before. We would kindly like to draw the attention of the referee to lines 16 and lines 22-23 of the abstract. In line 16, we mentioned that sites were chosen to have some spread of pedogenic oxide contents. In lines 22-23, we stated: "Total MOC amounts were linked to the content of pedogenic oxides across sites, independent of variations in total clay. The uniform MOC desorption could therefore be the result of pedogenic oxides dominating the overall response of MOC to extraction." We can make this point more clear by adding that also MOC residues were correlated to pedogenic oxides. Therefore, one explanation for the uniform extractability of MOC is that MOC is dominated by interactions between OC and pedogenic oxides and that clay minerals were less important.

Throughout the manuscript, it would increase the clarity of the arguments and results if the language regarding radiocarbon analysis was cleaned up. I hate to argue semantics, but I believe the terms 'stability' and 'lability' are now often rejected by the community due to lack of specificity. Colleagues have been insisting on using the term 'persistence' as far as I

know. I also believe 'older', 'younger', and 'age' are not appropriate for use in a manuscript, unless in reference to a 'mean system age'.

We are not convinced that there is consensus in the community when it comes to terminology and on whether or when to use the term persistence instead of stability. The term "persistence" only became popular after its usage in the title of the Schmidt et al. (2011) paper and therefore stands maybe for replacing the old chemistry-based "stability" (=recalcitrance?)-paradigm by a new one. In that sense, using persistence in our manuscript might be appropriate but the term has not been clearly defined in the Schmidt et al. paper. We think it is reasonable to state that OC can become temporarily stabilized, e.g. by adsorption to mineral surfaces, and that it might be re-mobilized later on. Also, we think it makes more sense to directly refer to "stabilization processes" instead of "processes increasing the persistence of OC soils". Both terms might have their justification and will re-check the manuscript for their correct use.
Regarding the ^{14}C results, we did not calculate any absolute ages, mean ages, age distributions or transit times of OC in samples, since this exercise requires models with specific assumptions. Irrespective of the reference system, it is correct to state that carbon in samples with smaller ^{14}C contents or more negative Δ^{14}C values than in other samples had been isolated from the atmosphere for a longer time and can accordingly be referred to as being older. We will add a respective sentence to the methods part to make clear what we mean when referring to younger or older carbon.

In hypothesis #5, it is unclear what the term 'organic acids' is referring to.

Hypthesis 5 will be reworded in the revised version following the recommendations of referee 1 and will now read:
"MOC should be less prone to desorption and accordingly of older ^{14}C age, when organic molecules forming MOC have many carboxyl groups that enable strong bonds with minerals surfaces."

In the figures, it would be helpful to readers if the soils were referred to according to their dominant physicochemical attribute rather than the site name. Hesse = loess/Cambisol; Laqueuille = grassland/Andisol: : : or something similar. That way the reader can immediately draw inferences based on differences in soil physicochemical characteristics.

We thank the referee for this suggestion and will complement the figure caption with a brief soil/site description.

Again in section 3.3, using site names gives little information, and forces the reader to continually check back to table 1 for context.

We agree and will provide the respective context in the revised version.

What was the pH of the NaF/NaOH solution (13-14, correct?), and what influence do the authors think this had on the amount and characteristics of the extracted C? Perhaps the effect of using such a strongly basic solution masked any influence of F on the amount of exchangeable C displaced on mineral surfaces.

The referee is right that besides F⁻ also OH⁻ anions from NaOH in the extraction solution will compete with and remove organic molecules from mineral surfaces. The higher pH will also lead to a deprotonation of organic acids. Since both, NaF and NaOH affect desorption in the same way and direction (also NaF alone would lead to an increase in soil pH), their relative importance should be unimportant for the overall result.

Butnor et al., 2017 (doi: 10.1016/j.foreco.2017.01.014) is an additional reference that explores H2O2 residues and Fe oxyhydroxides.

We thank the referee for making us aware of that study. Using a combination of 20% H_2O_2 and 0.33 $\underline{M}$ HNO_3, Butnor et al. 2017 also found increasing oxidation resistance of OC with soil depths, and a relation between oxidation resistant OC and extractable Fe (Mehlich-3) and clay. We will add the reference to line 389 of the manuscript, where we are discussing the effect of pedogenic oxides for the oxidation resistance of OC.

It would offer additional insight if the authors cold measure _13C and _15N for the extracts and residues to see if there were significant differences in the degree of microbial processing. Perhaps it would be acceptable to use the 13C values from the AMS measurements? I know these aren't typically considered publishable due to possible fractionation effects, but it seems as though they could be used for intercomparison purposes if it can be assumed that all samples experienced similar degrees of fractionation during AMS measurement?

We thank the referee for this suggestion. We actually already tested the potential application of the OC-to-N ratio for a similar reason but did not obtain consistent, and thus, easily interpretable results, also because N concentrations in (subsoil) residues were sometimes very small. ^{15}N analyses cannot be added any more now because for many fractions the sample material was little and used up during analyses. The same applies for additional ^{13}C analyses. For the reason mentioned by the referee, we typically do not use the ^{13}C values of the AMS but will evaluate their application for the revised version and add them if appropriate. Unpublished results from a previous experiment showed that both, oxidation and NaOH extraction residues were depleted in ^{13}C (and ^{14}C) relative to the original sample, indicating that resistant OC was rather less microbial processed than the extracted or oxidized fractions.

Additional NMR data would also be interesting. Would it be possible to measure the H2O2 residuals?

From a conceptual point of view we agree that it would be great to do that. Unfortunately, it is neither technically feasible nor reasonable. First, because of the small OC concentrations in the HF-residues, where the noise produced by the mineral background will impede the interpretation of the spectra even with long measurement times, and also because the H_2O_2 treatment probably also chemically altered the residues, so that it would be difficult to tell if the molecular structures observed now are no artefacts cause by the strong oxidation.