We would like to thank the editor for taking the time to handle our manuscript and for finding three very constructive reviewers. We also want to thank all reviewers for taking the time and reviewing our manuscript to help improve its quality. We are grateful for the honest and thorough feedback. The suggestions were highly useful and provided us with information, where misunderstandings could be possible and where we needed to make our message clearer and to discuss the limitations of the DSI in more detail. They helped to further improve the quality of this manuscript and we hope that we addressed concerns to a satisfying extent. Our comments to the reviewers in the following are in blue color. We made use of the constructive criticism and altered the text of the manuscript, where applicable. We added screenshots of alterations in the text related to the comments. These are displayed in green color.

Comments of Reviewer 3: Lauric Cécillon

"Reservation on the rationale of the DRIFTS stability index of soil organic matter (SOM) in mineral soil, and its use for partitioning the C kinetic pools of SOM dynamics models" This draft by Laub and colleagues describes a method to divide soil organic matter(SOM) into fast and slow cycling C pools in the soil organic module of the DAISY model. This method is based on the characterization of bulk mineral soil samples using mid-infrared diffuse reflectance spectroscopy (DRIFTS). DRIFTS spectra of bulk mineral soils are used to compute the "DRIFTS stability index" of SOM, defined as the ratio of aliphatic C-H (2930 cm-1) to aromatic C=C (1620 cm-1) stretching vibrations.

The DRIFTS stability index was previously published by Demyan et al. (2012) in the European Journal of Soil Science.

The development of routine and operational method to initialize the relative size of C kinetic pools from SOM dynamics models is a very important and timely topic. Indeed, the accuracy of the simulations of SOM evolution in mineral soils by current models is strongly questioned, notably because a poor initialization of the size of C kinetic pools. The method proposed by Laub and colleagues, using the DRIFTS stability index to divide soil organic matter (SOM) into fast and slow cycling C pools in the soil organic module of the DAISY model is original and very interesting, and their draft is well structured and written. However, I have a major concern regarding the rationale of the DRIFTS stability index of SOM in mineral soil, and its use for partitioning the C kinetic pools of SOM dynamics models. In this review, I will only discuss this concern, though this stimulating and timely work would deserve many other comments, as highlighted by the two other reviewers of this draft. First, I would like to come back on the justification of the DRIFTS stability index by Demyan and colleagues in their 2012 paper. Demyan et al. (2012) searched for information related to SOM in DRIFTS spectra of bulk mineral soils, and its link to SOM stability as assessed by a SOM density fractionation scheme. In their search for SOM information in DRIFTS spectra of bulk mineral soils, they discarded "wavenumbers of functional groups associated with non-organic compounds such as silicates and alumino-iron oxides". For them, "these criteria removed the peaks <1000 cm−1 and the peaks at 1980, 1870, 1792 and 1390 cm−1", but not the 1620 cm−1 peak. For them, "the [DRIFTS] peak at 1620 cm−1 was assigned to predominantly aromatic C =C stretching and/or asymmetric–COO−stretching but possibly also C = O vibrations". Demyan et al. (2012) show that "a positive relationship was found between the ratio of the peaks at 1620 and 2930 cm−1 (1620:2930) and the ratio of stable C (sum of C contained in clay and >1.8 g cm−3 fractions) to labile C (amount of C in the <1.8g cm−3 fraction) (R2= 0.62, P = 0.012)." For the authors, this result justifies that the DRIFTS stability index can reliably be "taken as an indicator of SOM stability" (Demyan et al., 2012).

We originally stated (line 369ff) that the peaks were selected in order to have limited mineral interference (e.g. Demyan et al., 2012). In their original publication only soils from the same field experiment with the same texture and mineral background were taken as additional measure of caution. As this approach showed potential for the site at Bad Lauchstädt, we hypothesized that this could justify evaluating the use of the DRIFTS 1620:2930 ratio as a more general stability index. We are aware of the mineral signal in the vicinity of the 1620 cm^{-1} peak and this fact was also acknowledged in the original publication of Demyan et al. (2012). By carefully selecting the integration limits, it was possible to minimize the mineral interference and get a general applicable stability index (see evidence below). In the current study, we aimed to combine several sites with differing textures and mineralogies to have several test cases. The reason for the statistical analysis of the model error was exactly that we wanted to test whether the DSI is a useful proxy across a range of sites. We state some further reasoning below why we think the 1620 cm^{-1} peak and the specific peak limits that we have used (1660 – 1580 cm^{-1}) is representative of aromatic carbon and what was changed in the main text.

However, a short look at the literature on DRIFTS of soils show that the 1620 cm-1 peak in bulk mineral soils cannot be exclusively assigned to absorption from SOM functional groups (C = C or C = O) as claimed by Demyan et al. 2012. I will only cite two important papers: Nguyen et al. (1991) and Reeves (2012).Nguyen and colleagues, based on DRIFTS spectra of pure mineral compounds and various soil samples demonstrated that "The DRIFT spectra of soils containing organic matter show considerable overlap of the silicate combination bands in the 2000-1600 cm-1 region". I provide here the Figure 1 modified from Nguyen et al. (1991) showing the DRIFTS spectra of quartz (pure or diluted in KBr), highlighting the strong absorption of quartz at 1620 cm-1 (for the DRIFT spectra of pure quartz). They suggested that "Spectral subtraction techniques or prior chemical treatment may thus be required to resolve these peaks." (Nguyen et al., 1991).

Reeves (2012) based on works similar than Nguyen et al. (1991), concluded that "With the exception of the bands at 2930and 2850 cm−1 due to aliphatic CH [when the soil does not contain carbonates, added by me] and the large OH band spanning most of the region between 2700 and 3500cm−1, there is little that is obviously due to OM in the soil spectra". Regarding the 1620cm-1 DRIFTS peak, he suggested, following Nguyen et al. (1991) that "the region between 1750–1600 cm−1 can be interpreted, despite the presence of strong silica bands, because silica can be ash subtracted quite well". But he also concluded his paper with this warning regarding spectral subtraction: "It will detect not only whether your sample is changed by 0.1% at some point in time, but will also seem to detect the phases of the moon and the mood you were in while you were measuring the data."(Hirschfeld, 1984; cited by Reeves, 2012).I deduce from this short literature survey that in their 2012 paper, Demyan et al. incorrectly assigned to SOM compounds (C = C, C = O) exclusively the 1620 cm-1 DRIFTS peak of bulk mineral soils, as this peak is also due to mineral compounds such as quartz (but also to water in some phyllosilicates).

It is not correct that we claimed an "exclusive" assignment of the 1620 cm^{-1} peak to SOM functional groups, but rather that by carefully selected integration limits, the delimited area of the 1620 cm^{-1} is mostly representative of those organic groups.

In fact, the different spectra of soils before and after ashing or pyrolysis (as the example below taken from the supplementary material of Nkwain et al. (2018)) demonstrate that a considerable part of the delimited 1620 cm^{-1} peak is lost. Demyan et al. (2013) found a decrease in absorbance intensity at 1620 cm^{-1} with maximum losses occurring between 400-500°C (Figure S8, Left) for bulk soils. In the same study separated fractions were also analyzed, with a similar 1620 cm^{-1} peak loss found for particulate organic matter (POM) that was assumed to be mineral free. These consistent findings of the organic contribution to the 1620 cm^{-1} peak from both rapid pyrolysis

and in situ thermal monitoring of soil samples up to 700 °C where also found when pretreating bulk soil or fractions with NaOCl (Yeasmin et al., 2017).
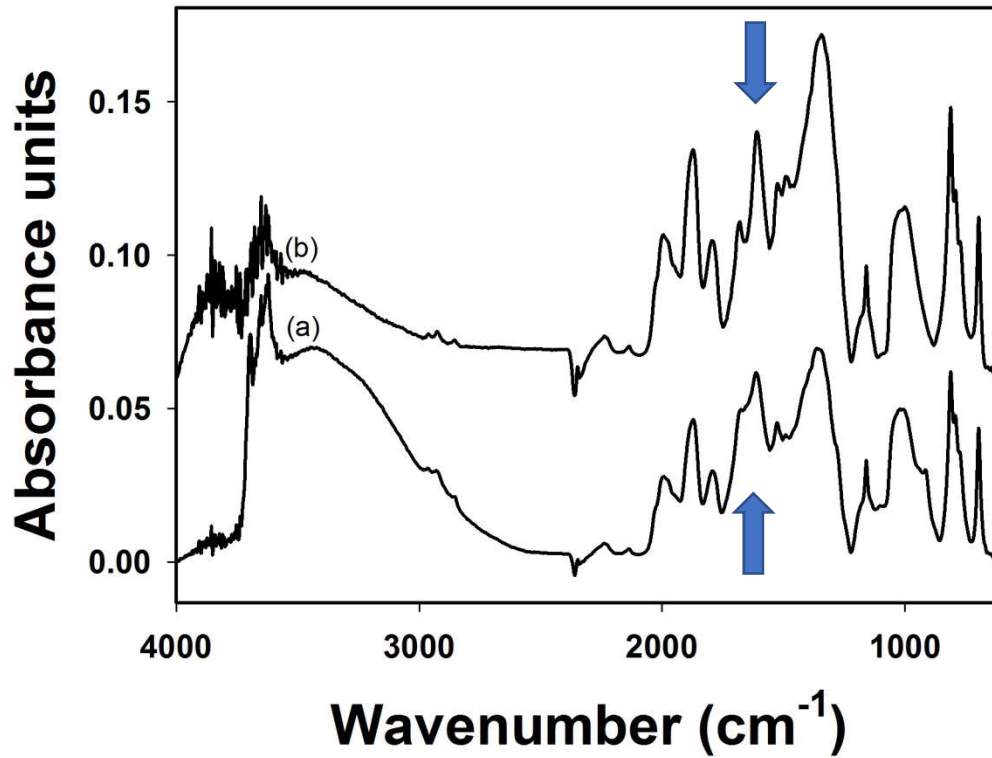

Figure S7. DRIFTS spectra of (a) unpyrolyzed soil and (b) pyrolyzed soil from Bad Lauchstädt (FYM). From (Nkwain et al., 2018)
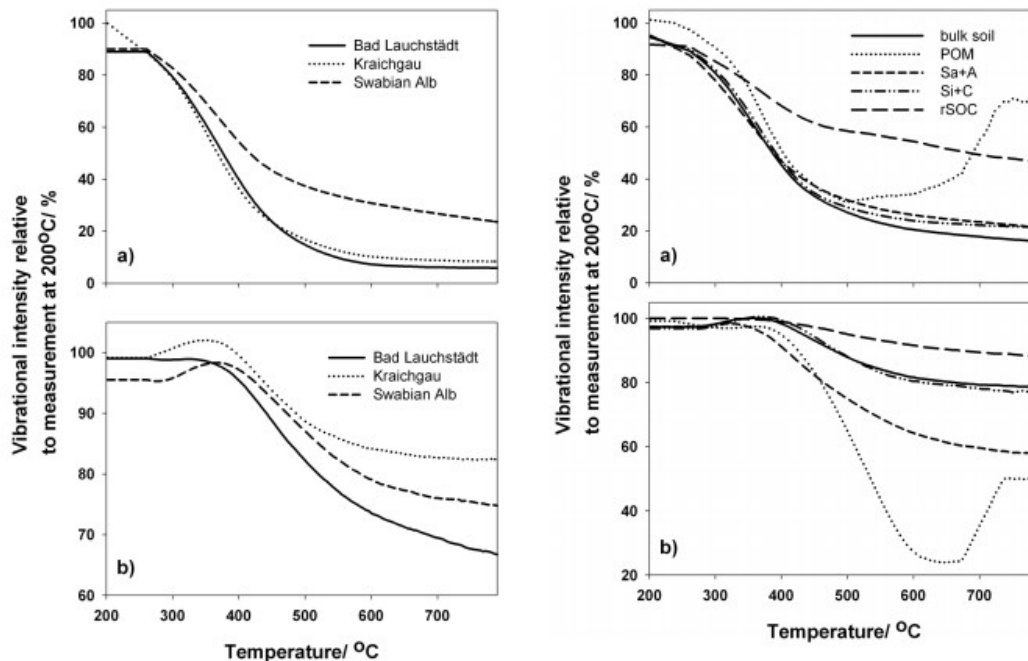

Figure S8 Left: (a) Change of C-H (2930 cm^{-1}) and (b) C = O/C = C (1620 cm^{-1}) vibrations with heating as measured by in situT DRIFTS of bulk soil samples from Bad Lauchstadt, ¨ Kraichgau, and Swabian Alb. Right: (a) Change of C-H (2930 cm^{-1}) and (b) C = O/C = C (1620 cm^{-1}) vibrations measured in bulk soil and fractions of soils from Kraichgau and Swabian Alb (Demyan et al., 2013). *POM-particulate organic matter, Sa+A-sand and stable aggregates, Si+C-silt and clay, rSOC-resistant soil organic carbon.

We would like to draw the attention to the fact that by a careful selection of the integration limits, we only take the top of the larger 1620 cm^{-1} peak (which in our samples made up 15 to 33% of the whole peak area). As the three examples above demonstrate, this is mostly the part, which is removed by burning, pyrolyzing or NaOCl treatment. This is the same principle as used for the aliphatic peak area at 2930 cm^{-1}, which is on top of a larger OH peak and to our knowledge, there is little debate about using this approach for the 2930 cm^{-1}. See the picture below for typical peak areas from our samples.
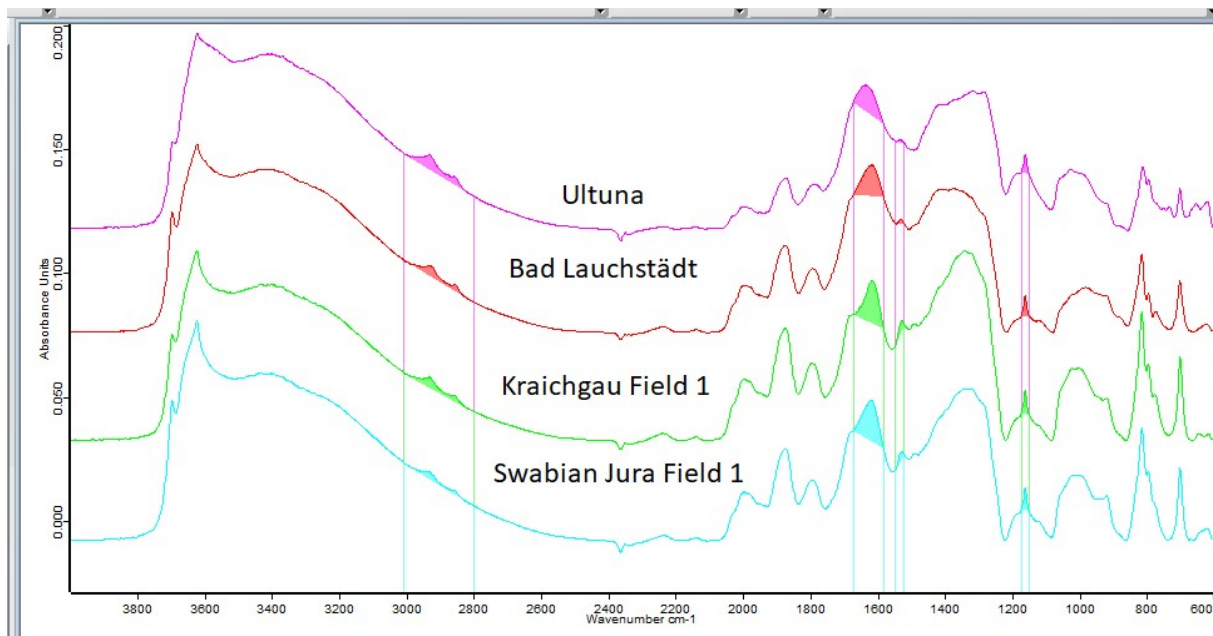
While we certainly do not claim that we can completely eliminate mineral interference, we think that the specifically delimited 1620 cm^{-1} peak that we use mostly consists of aromatic carbon i.e. the part of the peak that is selected is the part that disappears with the mentioned methods of SOC destruction. The finding, that it really is a meaningful proxy for carbon quality or stability is corroborated by the strong correlation (0.84) between the DSI and the percent of CPsoc, as was suggested to be computed by Lauric (new Figure S1 and comment below). As we recently demonstrated (Laub et al., 2019), and further found in the current study, the 2930 cm^{-1} peak is also subject to interference even in non-carbonate containing soils. This is mostly by water, which can partly be removed by higher drying temperatures. So, in summary we believe that there is sufficient evidence that, even though there is noise in the DSI at both peaks, DSI is still a meaningful and useful proxy, which is highly correlated to other measures of SOC composition but has the advantage of being cost/time effective to measure.

To further illustrate how the 1620 cm-1 DRIFTS peak of bulk mineral soil is poorly related to SOM compounds, I provide the Figure 2 based on published and unpublished data from the paper of Barré et al. (2016) in Biogeochemistry showing the non-parametric Spearman's Rho coefficient of DRIFTS spectra from soils coming from the Ultuna Fame trial, one site that was used in this reviewed work by Laub and colleagues, with SOC concentration. In Figure 2, we clearly see the strong and positive Rho coefficient of the 2900 cm-1 spectral region with SOC concentration while the 1620 cm-1spectral region show a Rho coefficient with SOC concentration close to 0, suggesting (though not demonstrating) that other compounds that organic matter absorb energy in the 1620 cm-1 spectral region of DRIFTS spectra, when scanning bulk mineral soils. From the above-mentioned information, I

therefore question the rationale of the DRIFTS stability index of soil organic matter (SOM) in mineral soil samples.

> The result is strongly affected by the delineation of the peak area. We thus agree if the whole 1620 cm-1 peak area (ca. 1755-1555 cm^{-1}) is taken results may not be reliable.

My interpretation is that this index is dividing a quantity that is highly correlated to SOC concentration (the 2900 cm-1 spectral region), by a quantity that is weakly changing when SOC concentration is modified (the 1620 cm-1 spectral region, provided a similar mineral composition). The DRIFTS stability index may thus show an increased SOC lability when SOC concentration is increased. I thus hypothesize that the DRIFTS stability index, as proposed by Demyan et al. (2012) and Laub and colleagues in this reviewed draft, may provide some information that is basically the same (though with added noise) than a variable much simpler than their index: total SOC concentration.

> We agree with the interpretation that the DSI is "dividing a quantity that is highly correlated to SOC concentration by a quantity that is weakly changing when SOC concentration is modified", and as we demonstrate above, both quantities are linked to forms of SOC. The fact that the selected subregion of the 1620 cm^{-1} peak does not change strongly with SOC content, while, as destructive techniques demonstrate, it is still consisting mostly of aromatic carbon compounds (according to our integration limits), is exactly the reason why it is a very suitable proxy for slow turnover SOC.

It is well documented that an increase in SOC concentration is associated with an increased in the labile/stable SOC ratio, and all proposed indicators of SOM stability should be compared to SOC concentration, the most simple and straightforward indicator of SOM stability (though not very accurate).

What is the Spearman's Rho coefficient of the DRIFTS stability index with SOC concentration in the dataset of Laub and colleagues?

> We calculated the Pearson`s correlation coefficient -0.57 and Spearman`s rank correlation coefficient to be -0.68 between OC content and the DSI (as in formula 2) for the whole dataset (n=50). See the plot below
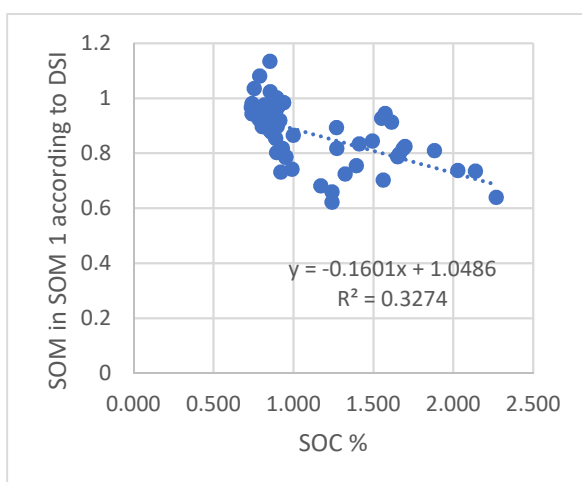
> We think, that the nonlinearity of the relationship between the DSI and the SOC content, as indicated by a higher rank correlation coefficient, points towards the possibility, that as SOC increases, most of the carbon is added to the fast turnover pool and that this could potentially be lost rather fast again.

I suggest that the authors (rather than using the spectral subtraction technique suggested by Nguyen et al., 1991 or Reeves, 2012), (i) test a soil dilution in KBr to reduce mineral artifacts in the 1620 cm-1 spectral region of neat DRIFTS, (ii) or test attenuated total reflectance mid-infrared spectroscopy (MIR-ATR) as an alternative technique.

Dilution with KBr (1:3 and 1:100) had been tested by Demyan et al. (2012), mainly to determine whether there was specular reflectance in the 1159 cm^{-1} region, but was not found to yield better performances for deriving the DSI. Using neat samples avoids hygroscopic KBr which would have the potential to absorb water interfering with the 2930 cm^{-1} and 1620 cm^{-1} peaks and other non-desired interactions with the sample. We think that the major advantage of the DSI and DRIFTS PLSR is that it is possible to use undiluted bulk soil samples, and that it is nondestructive (cost effective and other analysis can be done with the same samples). We see the major advantage also in large scale applications, such as regional simulations, where other techniques are either too expensive or time consuming.

Indeed, MIR-ATR is a technique where the 1620 cm-1 peak region seems to be much less affected by quartz and other minerals that neat DRIFT signal, as illustrated in Figure 3 (Cécillon, Unpublished data).

From our understanding, the issue with ATR usually is that the signal throughput to the detector is weaker, thus the overall spectral features stand out less and are dominated by the silica vibrations at <1500 cm^{-1}, which is also shown in Figure 3 of Lauric Cécillon´s comment. The maximum absorbance in the DSI wavenumbers is almost an order of magnitude lower as compared to DRIFTS. If you zoom in on the figure, you can also see a small peak probably around 1620 cm^{-1} in the silica sample, so it seems not to be free of mineral interference.

It might be possible that MIR-ATR is an alternative to DRIFTS, if it can reduce mineral interference at the 1620 cm^{-1}, but given the less strong signal of organic peaks it might be of limited use in low C soils. It could be worthwhile to do further research towards that direction and we think that this could be the content of another future publication.

Finally, as Laub and colleagues benefit from soil samples from two long-term bare fallow sites in Europe, I suggest that they compute the Spearman's Rho coefficient of their DRIFTS stability index with the proportion of centennially persistent soil organic carbon(CPsoc), that may be derived from the SOC evolution in the bare fallow plots, as shown by Cécillon et al. (2018).

A higher Spearman's rho coefficient of the DRIFTS stability index with CPsoc than the Spearman's rho coefficient of SOC concentration with CP-soc, would suggest an added value of the index compared to SOC concentration, in its current state.

Thanks for the discussion on this comment. We have now computed %CPsoc with the value of 6.95 g kg^{-1} CPsoc from Ultuna derived by Cécillon et al. (2018) and 16.0 g kg^{-1} CPsoc from Franko and Merbach (2017) for the bare fallow data we have available. As shown below, when combining the two datasets of Bad Lauchstädt and Ultuna the correlation between SOC and CPsoc across sites is poor. This shows that SOC alone is not a sufficient indicator for SOC quality. The correlation between the DSI and CPsoc on the other hand is quite strong (0.84), which according to Laurics comment is a strong indicator of its added value.

We think that it would be highly interesting to test this for other long-term bare fallows, where CPsoc could be mathematically derived (needing probably 30+ years of fallow) and this might

help to optimize the DSI further. We think that a future publication could go into this direction and are excited about this finding.
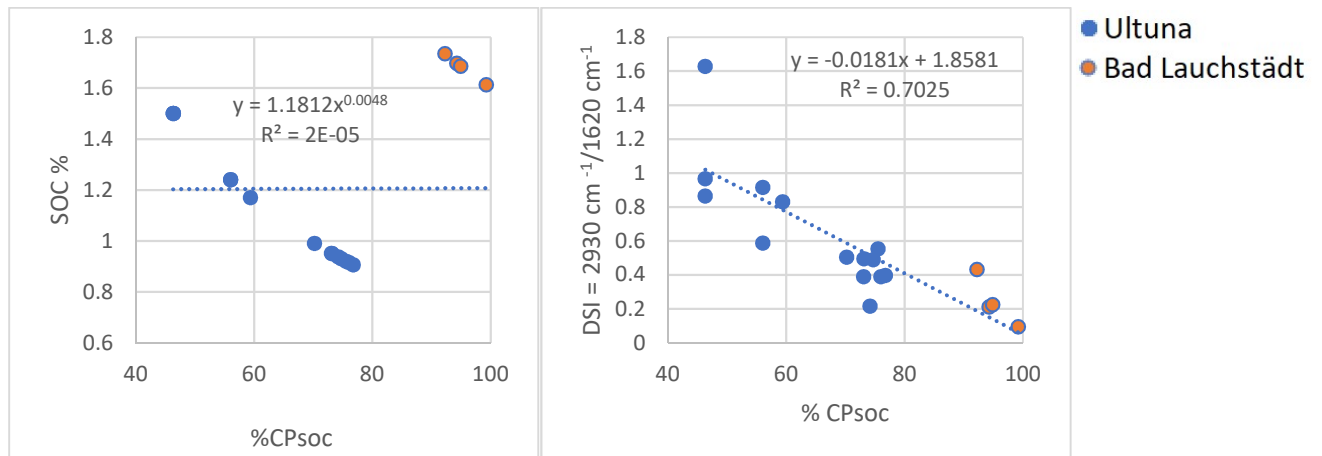
As the reasoning behind CPsoc comes from RothC type models, which assume that there is only one actively decomposing SOC pool and another passive or inert SOC pool which is NOT subject to decomposition, this could mean that the DSI might also be useful for these types of models. In this study, we worked with DAISY, which is a CENTURY type model, that has a fast and slow SOC pool, both subjected to decomposition. We think that this is in agreement with the principle behind DRIFTS, and that microorganisms primarily target high energy aliphatic SOC, but aromatic SOC is also decomposed at a much slower rate, probably as a byproduct of enzyme release.

Overall, we very much acknowledge the issue of mineral interference addressed by the reviewer (see line 369 in the original manuscript) and the new addition:

information on SOM quality (Giacometti et al., 2013; Margenot et al., 2015). ~~Demyan et al. (2012)~~While both peaks are subject to interference (2930 cm^{-1} mainly from water and 1620 cm^{-1} mainly from minerals (Nguyen et al., 1991)), it should be possible to limit the interference using subregions of the peaks with carefully selected

85   integration limits, only selecting the specific peak area of interest. Indeed, Demyan et al. (2012) found ~~aliphatics~~aliphatic compounds to be enriched under long-term farmyard manure application and depleted in mineral fertilizer or control treatments, and showed that the ~~ratio of the 2930 cm^{-1} to 1620 cm^{-1} peaks had a~~ ~~significant positive correlation with the ratio of labile to stable SOM obtained by size and density fractionation.~~ ~~Hence, we hypothesise~~dratios of the 1620 cm^{-1} to 2930 cm^{-1} peak had a significant positive correlation with the

90   ratio of stable to labile SOM obtained by size and density fractionation. It was further corroborated that the specific integration limits of the peaks they used, which mainly selected the top subregion of the peak areas, are lost during combustion (Demyan et al., 2013). Hence, we hypothesized that the ratio of the aliphatic to aromatic DRIFTS

We have addressed this issue mainly by carefully delaminating the integration area and now have more clearly pointed to this in the methods:

peak areas of the four subsamples averaged after that. The local baselines were drawn between the intersection of the spectra and a vertical line at the integration limits (3010 – 2800 cm⁻¹ for the aliphatic C-H ~~streching, 1660 – 1580 cm⁻¹ for aromatic C=C streching vibrations).~~ stretching, 1660 – 1580 cm⁻¹ for aromatic C=C stretching vibrations). Example spectra and integrated peak areas are displayed in **Figure S 1**. These carefully selected integration limits were critical to reducing signal interference from water and minerals. Particularly, the mineral interference close to the 1620 cm⁻¹ peak makes accurate selection of integration limits necessary, so that only its top part (assumed to consist mostly of aromatic carbon) is selected. In the case of our samples, the selected specific peak area of the 1620 cm⁻¹ peak accounted for approximately 10 to 30 % of the total peak area (ca. 1755-1555 cm⁻¹), and roughly corresponds to the peak portion that is lost with combustion or chemical oxidation (Demyan et al., 2013; Yeasmin et al., 2017). A strong correlation between the DSI and the percentage of centennially persistent SOC (r = 0.84) from the combined long term experiments used in this study (using values of centennially persistent SOC from Cécillon et al., 2018; and Franko and Merbach, 2017), showed that the DSI selected in this manner did in fact explain a large portion of the SOC quality change across sites (**Figure S 2**).

We have further added a more detailed discussion on open questions of the DSI in the new manuscript version and finalize the section with the limitation that DSI should be tested before used with different soil types.

Compared with the other proxies for SOM quality discussed above, the measurements by DRIFTS are inexpensive, relatively simple, and the equipment of the same manufacturer is standardized. This should also constrain variability between different laboratories and be attractive for large-scale applications with large sample ~~size, for example to initialize simulations at the regional scale~~ numbers, for example to initialize simulations at the regional scale. However, for standardization of the DSI for model initialization one needs to address how the type of spectrometer (e.g. detector type) influences the spectra, if water and mineral interferences (Nguyen et al., 1991) in the spectra can be further reduced and if a mathematical standardization of the spectra and DSI (across instruments and water contents) is possible. While a complete elimination of mineral interference is not possible, a careful selection of integration limits and the use of a local baseline minimizes mineral interference of DRIFS spectra from bulk soils. This mostly selects the top part of the 1620 cm⁻¹ peak, which corresponds to the part that is reduced or completely lost when SOC is destroyed (Demyan et al., 2013; Yeasmin et al., 2017). Other approaches such as spectral subtraction of ashed samples or HF destruction of minerals prior DRIFTS analysis have been developed

in the attempt to obtain spectra of pure SOC. All are rather labor intensive and still produce artifacts, as it is not possible to destroy only the minerals or only the SOC without altering the respective other fraction (Yeasmin et al., 2017). Hence, we think that the selected integration limits might represent at this point the most feasible option for obtaining a robust and cost-effective proxy of SOC quality for modeling. The strong correlation of DSI and centennially persistent SOC as well as the model results of this study seem to corroborate this. The method of DSI estimation might be improved by a study of the best integration limits optimizing the fit of the DSI and centennially persistent SOC, which would require more bare fallow experiments than in this study. It could be worthwhile to use a purely mineral peak to correct for the mineral interference at 1620 cm⁻¹ similar to what was done to correct for carbonates in the 2930 cm⁻¹ peak by Mirzaeitalarposhti et al. (2016). The recent coupling of pyrolysis with

In the quest to find measurable fractions for model pools, we think that the DSI is a useful proxy (carefully selected integration limits, nonlinear relation with SOC, evidence that our 1620 cm⁻¹ is mostly from carbon, drying at 105 °C to reduce water interference at 2930 cm⁻¹). We first and foremost consider the DSI as a potential proxy to help initializing two pool SOM models, and our

question was, whether it was useful for this purpose or not, compared to steady state initializations. We think the value of this publication is to establish that the DSI has the potential to be a measurable fraction as a model pool proxy and thereby reducing model uncertainty, and show this to the scientific community. As any research this opens new questions which could lead to further development and refinement of the DSI. We think, that our study could demonstrate the DSI's usefulness and that it might be worthwhile to put further efforts and research towards its validation, use or optimization, especially because of its ease of use and inexpensive nature.

References:

Demyan, M. S., Rasche, F., Schulz, E., Breulmann, M., Müller, T. and Cadisch, G.: Use of specific peaks obtained by diffuse reflectance Fourier transform mid-infrared spectroscopy to study the composition of organic matter in a Haplic Chernozem, Eur. J. Soil Sci., 63(2), 189–199, doi:10.1111/j.1365-2389.2011.01420.x, 2012.

Demyan, M. S., Rasche, F., Schütt, M., Smirnova, N., Schulz, E. and Cadisch, G.: Combining a coupled FTIR-EGA system and in situ DRIFTS for studying soil organic matter in arable soils, Biogeosciences, 10(5), 2897–2913, doi:10.5194/bg-10-2897-2013, 2013.

Franko, U. and Merbach, I.: Modelling soil organic matter dynamics on a bare fallow Chernozem soil in Central Germany, Geoderma, 303(May), 93–98, doi:10.1016/j.geoderma.2017.05.013, 2017.

Laub, M., Blagodatsky, S., Nkwain, Y. F. and Cadisch, G.: Soil sample drying temperature affects specific organic mid-DRIFTS peaks and quality indices, Geoderma, 355, 113897, doi:10.1016/j.geoderma.2019.113897, 2019.

Nkwain, F. N., Demyan, M. S., Rasche, F., Dignac, M.-F., Schulz, E., Kätterer, T., Müller, T. and Cadisch, G.: Coupling pyrolysis with mid-infrared spectroscopy (Py-MIRS) to fingerprint soil organic matter bulk chemistry, J. Anal. Appl. Pyrolysis, 133(April 2017), 176–184, doi:10.1016/j.jaap.2018.04.004, 2018.

Yeasmin, S., Singh, B., Johnston, C. T. and Sparks, D. L.: Evaluation of pre-treatment procedures for improved interpretation of mid infrared spectra of soil organic matter, Geoderma, 304, 83–92, doi:10.1016/j.geoderma.2016.04.008, 2017.