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 predominately 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 with the same mineral background were taken as additional measure of caution. As this approach showed potential for Bad Lauchstädt, we thought that this could justify trying to use DRIFTS as a general stability index. This was the reason for conducting this study. We are aware of the mineral signal at the 1620 cm^{-1} region and this fact was also acknowledged in the original publication of Demyan et al. (2012). By carefully selecting the integration limits, it should be possible to minimize the mineral interference and get a general applicable stability index. We aimed to combine several sites 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 as we selected 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).
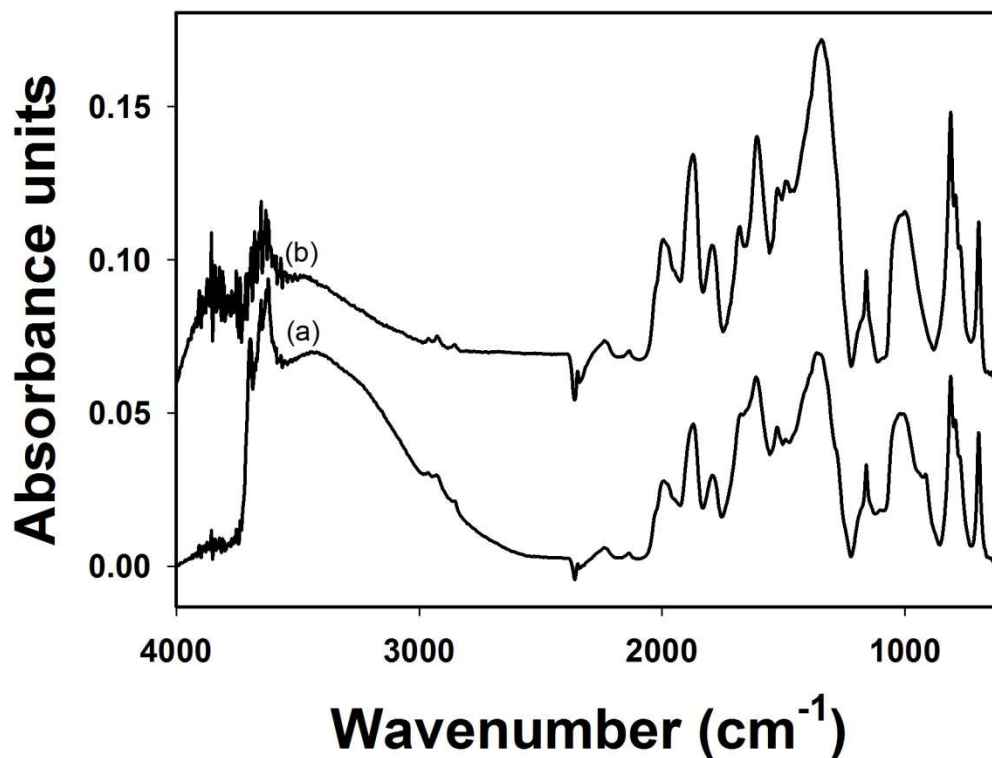
Figure S7. DRIFTS spectra of (a) unpyrolyzed soil and (b) pyrolyzed soil from Bad Lauchstädt (FYM). From (Nkwain et al., 2018)
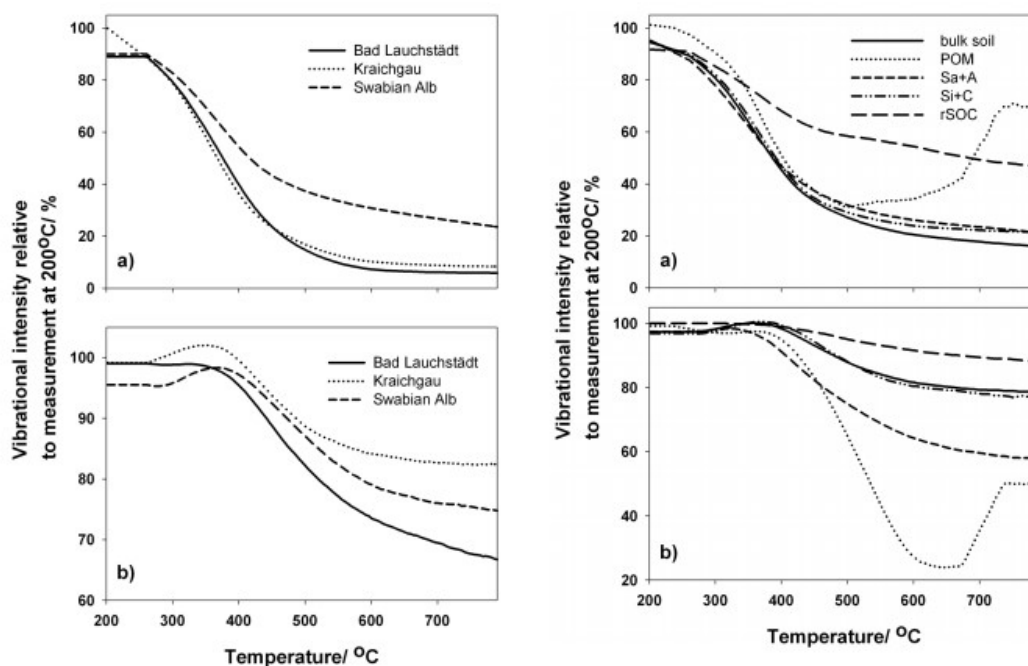

Figure S8 Left: Change of C-H (2930 cm−1) and (b) C = O/C = C (1620 cm−1) vibrations with heating as measured by In situT DRIFTS of bulk soil samples from Bad Lauchstadt, ¨ Kraichgau, and Swabian Alb. Right: Change of C-H (2930 cm−1) and (b) C = O/C = C (1620 cm−1) vibrations measured in bulk soil and 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 way we selected the integration limits. We only take the top of the larger 1620 cm^{-1} peak. As the three examples above demonstrate, this is mostly the part

which is removed by burning, pyrolyzing or NaOCl treatment. See the picture below for typical peak areas from our samples.
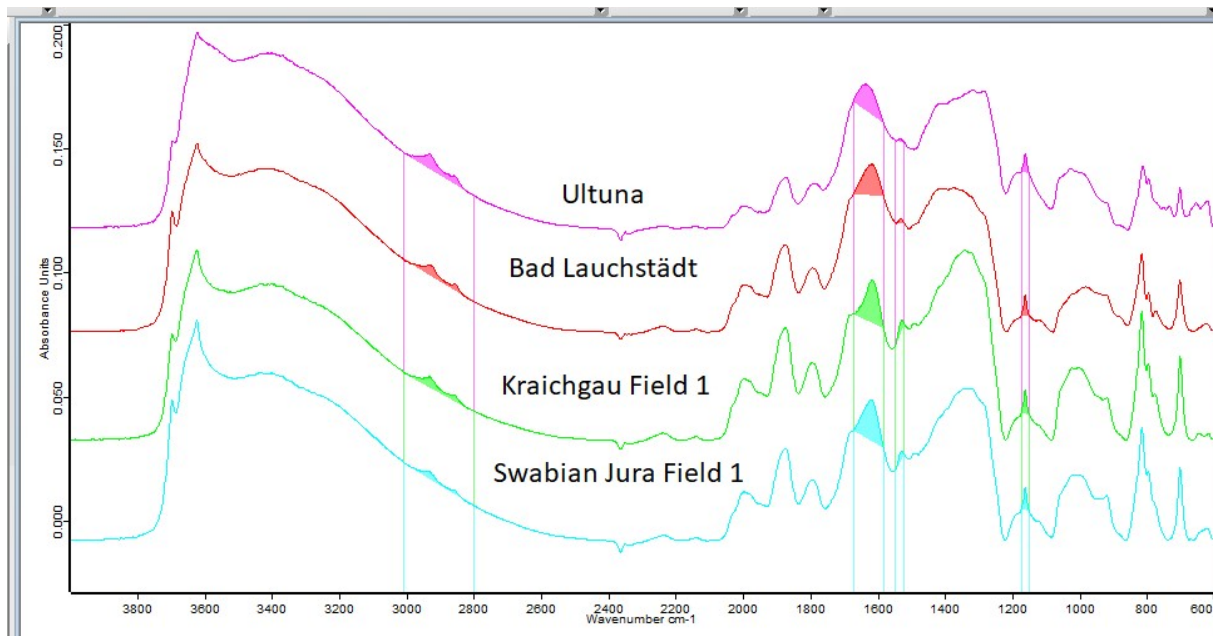
However, as we recently demonstrated (Laub et al., 2019), and further found in the current study, the 2930 cm^{-1} peak is also subject to interference even in non-carbonate containing soils. This is mostly by water, which can partly be removed by higher drying temperatures.

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.

Actually, we found that the 1620 cm^{-1} was mostly negatively correlated with TOC, but in our recent publication, we showed that there was a slight positive correlation with TOC, if the 1620 cm^{-1} peak was normalized (divided by the 1880 cm^{-1} quartz/ silicates Peak) (though not significant for the small number of 21 archive samples we used within this study. See the supplementary material of Laub et al., (2019).

| correlation of with	TOC 2930 cm^{-1}	TOC n2930 cm^{-1}	TOC 1620 cm^{-1}	TOC n1620 cm^{-1}	TOC 1530 cm^{-1}	TOC n1530 cm^{-1}	TOC 1159 cm^{-1}	TOC n1159 cm^{-1}
temperature								
32	0.77	0.85	-0.69	0.14	-0.70	-0.85	-0.60	0.18
55	0.73	0.77	-0.64	-0.01	-0.66	-0.83	-0.56	0.08
65	0.75	0.81	-0.63	0.06	-0.68	-0.84	-0.55	0.11
75	0.76	0.82	-0.62	0.12	-0.68	-0.80	-0.56	0.18
85	0.74	0.79	-0.63	0.09	-0.67	-0.81	-0.56	0.15
95	0.72	0.80	-0.63	0.19	-0.68	-0.84	-0.57	0.24
105	0.76	0.83	-0.62	0.16	-0.68	-0.81	-0.58	0.20

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. That the 1620 cm^{-1} 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
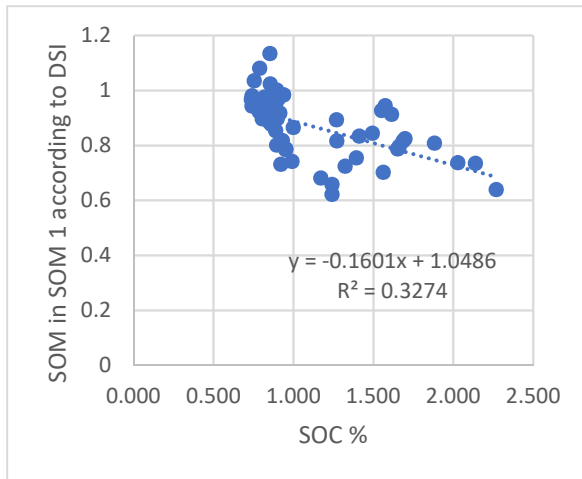
We think, that the nonlinearity of the relationship between the DSI and the SOC content points towards the possibility, that as SOC increases, most of the carbon is added to the fast turnover pool and that this could potentially be lost rather fast again.

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} peak 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 organics peaks it might be limited 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.

We cannot follow the reasoning here. We computed CPsoc with the value of 6.95 g kg^{-1} from Ultuna derived by Cécillon et al. (2018). As shown below, by definition of CPsoc, Spearman`s correlation coefficient (or in our graph an R^2 of an exponential equation, which is effectively the same) between SOC% and CPsoc of SOC is perfect (1!) if we only use data from one long term site. To resolve this, one would need data for the CPsoc and DSI from many long term experiments.
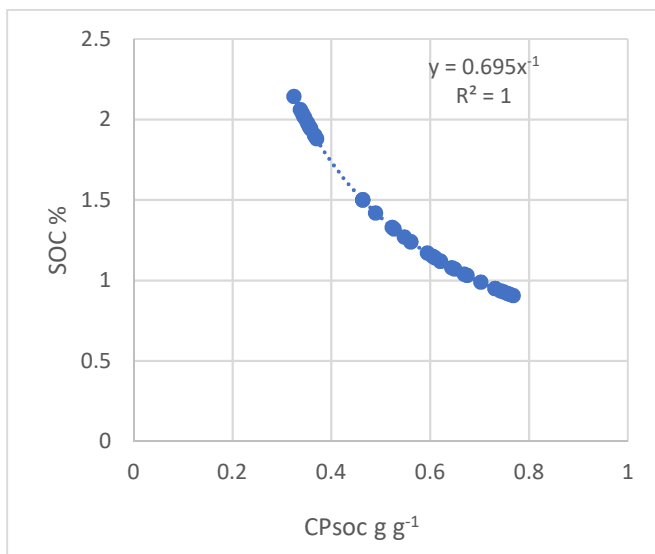
Additionally, the reasoning behind CPsoc is only valid for RothC type models, which assume that there is only one actively decomposing SOC pool and another passive SOC pool which is NOT subject to decomposition. In this study, we worked with DAISY, which is a CENTURY type model, that has a fast and slow 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.

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:

> Margenot et al., 2015). While both peaks are subject to interference (2930 cm^{-1} from water and 1620 cm^{-1} from minerals mainly (Nguyen et al., 1991), it should be possible to limit the interference with carefully selected

> integration limits. Indeed, Demyan et al. (2012) found aliphatics 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. It was further corroborated that the part of the peaks they used mainly select the top part of the peak areas, which are lost during combustion (Demyan et al., 2013). Hence, we hypothesised that

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

the aliphatic C-H streching, 1660 – 1580 cm^{-1} for aromatic C=C streching vibrations). The carefully selected integration limits, as well as using only the peak area on top of the local baselines were critical to reducing signal

140 inteference from water and minerals. Additionally, soils from the experiments in Kraichgau and Swabian Jura

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.

There are some remaining questions that should be answered to standardize the application of the DSI for model initialization. Those are related to how the type of spectrometer influences the spectra, as well as how water and mineral interferences (Nguyen et al., 1991) in the spectra can be eliminated or at least be further reduced. We had the experience, that spectra and therefore peak areas vary to some degree between different spectrometers

435 (mostly due to different detectors, types of detector cooling and resolution). Hence, it will be necessary to either use the same spectrometer, or to develop techniques to standardize spectra across a large number of instruments. While a careful selection of integration limits and the use of a local baseline minimizes mineral interference of DRIFS spectra from bulk soils, and mostly selects the part of the 1620 cm^{-1} peak which gets usually lost when SOC is destroyed (Demyan et al., 2013; Yeasmin et al., 2017), it is not possible to completely eliminate mineral

440 interference with DRIFTS. The recent coupling of pyrolysis with DRIFTS (Nkwain et al., 2018) might be a further advancement of the DSI, as it overcomes mineral interferences in the spectra. However, this technique is more complex due to a larger number of visible organic peaks, including CO$_2$ that develops from the pyrolysis, which makes it not easily applicable to established two-pool models such as DAISY. In addition, a considerable portion (30 – 40 %) of SOM is not pyrolyzed and therefore not recorded in the spectra. So iIn summary, even

445 despite of the acknowledged shortcomings, , it was found that the DSI can be directly used to distribute SOM between pools in two pool models, though there is some mineral interference. Furthermore, DSI was suitable for a wide range of soils and improved model performance. Hence, DSI seems to be a more robust proxy for pool initialization then methods such as steady state or long-term spin-up runs which rely on strong assumptions to which they are very sensitive though there is very limited data to prove them. It would still be advisable to test

450 the DSI before using it for mineralogy that differs considerably from the soils of this study.

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.

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.