The manuscript by Demyan et al. reports on a new method for the characterization of soil organic matter (SOM) composition and stability, which represents a potentially important advancement of our knowledge of SOM dynamics. The application of thermal analysis techniques to the characterization of SOM has gained increasing traction, but many problems with the technique remain unresolved. The coupling of EGA and in-situ DRIFTS represents a potential means of addressing some of these problems, particularly the attribution of thermal reactions to particular chemical components of SOM. While the manuscript is of high quality and appropriate for the journal, I found that it lacked the reporting of some important results. The authors report on the optimized results, but as a method development manuscript, more results from the non-optimal conditions should also be reported for the reader to be able to make their own judgments. Also, I found the most novel part of the work to be the attribution of thermal reactions to chemical components, but the manuscript lacked most of the DRIFTS results, which should instead be highlighted. I think the manuscript needs some major revisions before I would recommend it for full publication. I provide specific guidance through the comments below.

p.15384, ln.2-3: The citation Blumstein (1965) is not in the reference list.

p.15384, ln.7: The method of DTA is not the measure of differential mass loss, that it TG. Instead, DTA measures a differential in the temperature between a sample and reference.

p.15384, ln.15: This is a very important point, and one of the major challenges and opportunities in thermal analysis for SOM characterization. I will refer to this point in later comments.

p.15385, ln.9-13: I would refer the authors to a recent article in ES&T (Fernandez et al. 2012 ES&T 46:8921) where we report on the coupling of TG-DSC with EGA by infrared gas analysis for CO2. This approach is similar in part to the one being proposed, and addresses some of the same issues.

p.15386, ln.12 & 16: Refrain from using "and/or". In terms of logic, "or" is not exclusive and therefore includes the possibility of "and". As such, using only "or" is sufficient.

Section 2.1: Given that this work is method development, and that one of the major issues with thermal analysis is the role of the mineral matrix, I was somewhat disappointed by the limited number and range of soils used in the study. I understand the rationale of using multiple treatments or fractions to characterize SOM variability within a given soil, but I think the method development would have been much more robust by applying it to a much broader range of soils. Has the method been tested more extensively? Given that the study includes a rather limited number of soils with limited variability in their properties, it is imperative to include a description of the mineralogy of these soils to partly address the issue of how the mineral matrix affects the results.

p.15387, ln.16: When was the bare fallow treatment initiated? That is, how long had the soil been bare before sampling?

p.15388, ln.1: Typo. Replace "3%" with "30%".

p.15388, ln.4-6: This sentence should be moved to the previous section as it describes pre-treatment of the soil in general rather than steps in the fractionation procedure.

p.15388, ln.26: Is the term "rSOC" really necessary? Would hypochlorite-resistant C not be sufficient? I am reluctant to see new and unnecessary acronyms used.

p.15388, ln.27: By "twice", do the authors mean that two separate aliquots were fractionated to generate pseudo-replicates, rather than a single sampling between fractionated twice in sequence?

p.15389, ln.2: I am not familiar with the "Scheibler method" and am not sure how well know it is in general. Please elaborate with one sentence to describe what the method entails.

Section 2.3: I understand the rationale of the incubation experiment, but it should be stated more explicitly here. Much of the second paragraph can be omitted to reduce the emphasis on this experiment because many of the results are either not reported, or are not highly relevant to the study. For instance, I'm not convinced of the value of the microbial biomass and metabolic quotients. I'll refer to this again in the discussion section. My recommendation would be to simplify this section to report how the incubation was performed, and then report only the cumulative amount respired as a proportion of initial C to give an indication of C loss during the incubation, and the thermal results (Fig 6).

p.15390, ln.26: Here and several other places in the manuscript, I would recommend reporting the data that is "not shown" in some way. As a methods development better, the results obtained under non-optimal conditions can be as useful as those that the authors have decided are indeed optimal. For this reason, I would like to see these data even if they are in some kind of supplemental file.

p.15391, ln.12 – p.15392, ln.2: This section reports on the post-processing of data acquired, and therefore I recommend that it be moved to and merged with section 2.4.2.

p.15392, ln.3-7: Again, I would strongly recommend reporting these results in a figure showing how the curves differ by heating rate, similar to what we reported in Fernandez et al. (2011, JTAC).

p.15392, ln.22-29: I would recommend separating this section out and moving it to a new section before 2.1 Soils. The analysis of model materials is an important first step in the method development, I would suggest that moving this paragraph would generate a better logical order to the methods, that would then be reflected in the order in which the results are reported.

p.15392, ln.28: What was the rationale for the use of quartz as the inert material as opposed to the more frequently used calcinated kaolin?

p.15394, ln.4: The Demyan et al. (2012) citation appears to be quite important. Unfortunately it is not listed in the references.

Section 2.6: This is perhaps the most novel and important component of the work, but unfortunately, I found it difficult to follow. Perhaps it can be restructured and revised to improve clarity. My paraphrasing of this section is that DRIFTS data were used to determine the presence/absence of components within the CO_2 evolution curve. These components (identified by changes in portions of the DRIFTS spectra) were used to guide the deconvolution of the CO_2-EGA curves. Part of the confusion

from the current structure is that the DRIFTS and FTIR-EGA are somewhat muddled together and it is difficult to tease them apart.

p.15394, ln.13: Perhaps it is terminology, but it is not clear to how the Area$_{wt}$ output is used "for" curve fitting. Is it not more correct that the Area$_{wt}$ output is "subjected to" curve fitting/deconvolution? I am also not clear as to how the Area$_{wt}$ data forms a curve. Is Area$_{wt}$ not a single value for each sample? The FTIR-EGA curves reported in the figure are for Absorbance. Please clarify.

p.15394, ln.16: When the authors refer to "*a* vibrational organic functional group", are they referring to "each", "any" or "all" groups? Please clarify.

Section 3.1: I don't believe that the C concentration data of the soils and fractions are significant enough to warrant a separate results section, particularly since these samples come from well documented field experiments. Instead, my recommendation would be to fold the C concentration data into the background characterization of the soils in the methods section. Also, as noted above, I don't think the microbial biomass and metabolic quotient is relevant to the main story of the manuscript and should be omitted (also see comments below for p.15402 and Table 1).

Section 3.2 and 3.3: The titles for these sections do not appear to be completely appropriate. The title for Sec 3.2 should be limited to "organic substances" only and not include "soils", while the title for Sec 3.3 referring to "soils and fractions" is indeed appropriate. This structure for reporting the results for the model organic materials before the results for the soils and fractions is very good and should be reflected in the structure of the methods section as I noted above.

p.15396, ln.8: Again, I think these would be important data to show for method development.

p.15396, ln.25-28: This is a very important result and should be better highlighted. We reported similar incomplete yield of C in our IRGA-based EGA in our ES&T paper. We have been pursuing this line of investigation to determine the potential causes for incomplete recovery and to determine if the variability in the yield is systematic. Perhaps this is something the authors should elaborate on for their EGA method.

p.15397, ln.1-11: I would suggest this paragraph is more appropriately located in the methods section and should be moved there.

p.15397, ln.8: Why was the pre-heating treatment of the sand performed up to 600°C instead of 700°C like the samples or hotter?

Section 3.3: The structure of this section should be revised to parallel the previous section on the model organic materials by describing the curve shapes and C yield, then CO2$_{max}$. For instance, what was the C yield for the soil samples, and how did it compare to the model organic materials? It would also improve the structure and clarity to separate the results from each soil in separate paragraphs. In addition, it is essential that the various FTIR-EGA curves be reported as figures. These are some of the most important results to be reported. See additional comments below concerning Table 2 & 3 and Fig 5. Lastly, it does

not appear that the results of the various "thermal characteristics" described in Sec2.4.2 are actually reported.

p.15398, ln.9-14: The results of the cumulative CO_2 evolved (perhaps as a proportion of initial C) should be reported here along with the comparison of the pre- and post-incubation thermograms. I would also recommend moving this paragraph ahead of the previous one to keep all the results of bulk soils together first, before moving on to the results of the fractions.

Section 3.4: These results are the most novel part of the study, and therefore warrant better reporting. Again, instead of "data not shown", I would strongly recommend reporting these results in some form or another. The authors need to show more of these results to demonstrate the diversity of results among samples. See comment below in reference to Fig 7.

Section 3.5: It is not clear to me why the peak-fitting appears to only have been performed on two soils/samples. The authors should consider a different comparison. Comparing the SOM composition/stability in two different soils is not a straight-forward exercise because of the large number of confounding variables that might affect the results of the thermal analysis such as mineralogy. Instead, I strongly recommend that the authors expand the analysis to the other samples and consider alternative comparisons. Better comparisons would be among treatments or fractions within a soil. Or has the method not proven to be sufficient sensitive to detect these differences? See also the comment below in reference to Table 4.

p.15401, ln.12-28: This is great and really hits right at the heart of the underlying assumptions for using thermal analysis for SOM characterization. The distinction between results of the model organic materials, litter, and POM fractions, and the results of mineral-associated C fraction is very important has not yet been properly highlighted in the literature. We have unpublished data that show similar trends that the authors report here. I recommend this finding be better highlighted, and better placed in the context of previously published results on litter, compost and soils (e.g., papers by Rovira, DellAbatte, etc.).

p.15401, ln.14: Should "but CO_2" refer instead to $CO2_{max}$?

p.15402, ln.6-8: This sentence is not necessary in the discussion. It should instead be in the methods or results.

p.15402, ln.21-26: I am not at all convinced that the differences in microbial biomass are sufficient to be expressed in the thermal analysis. Biomass was reported to be 0.6 mgC/g, which represents approximately 3% of the total organic C in the bulk soil (which is reported to be 19.4 mgC/g). I believe this is a red herring unless there is substantial supporting evidence. My recommendation, as I have noted above, is to omit the microbial biomass and metabolic quotient components of the experiments, and limit the incubation study results to the total cumulative CO_2 respired as an index of C loss during the incubation. More important to the discussion, however, is to highlight that the newly developed coupled FTIR-EGA and DRIFTS method provides a mechanistic/concrete foundation for the deconvolution of the thermogram. I previously published some peak deconvolution of DSC data (Plante

et al. 2005 Geoderma 129:186), but have become highly reluctant to continue to do so become of the problem of attribution. Peak deconvolution is strictly a statistical method that has no theoretical or mechanistic foundation with other data to guide it. The authors have provided some of the first data using their DRIFTs analysis that can provide this guidance. This, in my opinion, is the most important outcome of the new method, and should be highlighted.

Section 4.4: I am not sure this is the best position for this section. It does not provide the best bookend for the discussion, which should instead end with the broadest implications for the study. I would recommend moving this section between 4.1 and 4.2, or perhaps even merging it with Sec4.1.

Conclusions: The impact of the conclusion can be greatly improved by shortening it. The conclusion section can be easily shortened by omitting much of the summary or repetition of the results and discussion. Rather than a summary, the conclusions should instead highlight the most salient outcomes of the study and provide guidance about future developments and applications of the method.

Table 1a: Which treatment(s) are represented in the TOC data? Is it the mean of all treatments? I would actually recommend omitting the TOC data from this table, and limiting this table to basic characterization data. For instance, in addition to %clay, pH and mineralogy could be reported.

Tables 1b and 2: Instead of reporting the TOC concentrations in Table 1a, I would recommend creating a new, merged table that reports only the TOC concentrations for all soils, treatments/years and fractions. As it is now, the TOC data is spread over three different data, and I do find that do be very effective. It would be a better presentation of the background information about C concentrations it were all compiled into a single table.

Table 1b: Why are TN data reported? How are they relevant? If the authors consider them important, they need to be explained in the methods and described in the results text. I would recommend that they be omitted as they have little bearing on the central issues of the study.

Table 2: Why are the TOC data reported as the mean of the three years? Did the C concentration not change through time? My guess would be that they did not actually change much because the standard errors appear to be quite low. In any case, this should be explained in the results. Also, as I described above, I recommend a single table to report the TOC data. However, the authors might elect to report peak temperatures along with the TOC data or as a separate, similarly-structured table reporting only peak temperatures.

Table 3: I strong recommend that the authors replace this table with a figure showing the shapes of the thermograms. This would be much more impactful for the reader. It is difficult to imagine the shape of the thermograms based on a listing of peak and shoulder temperatures.

Table 4: Was peak deconvolution performed only for these two samples? I found this to be a rather weak comparison, and as I noted in a comment above, I recommend that this table be expanded to include more samples provide a better comparison (i.e., among treatments or fractions within a soil). Or is the method not sufficiently sensitive to detect these differences?

Figure 5: Which treatments are represented here? As noted above, I strongly recommend that this figure be expanded to illustrate the thermograms of many more of the soils, treatments and fractions analyzed.

Figure 6: Rather than a three-panel figure, I would suggest that the figure would be more effective if the difference curves for each soil were included in each of panels A and B. Having the third panel as is suggests that the comparison of the difference curves between the two soils is important. I'm not convinced that a comparison between the soils is intended or meaningful.

Figure 7: As noted above, the figure illustrating the DRIFTS data needs to be expanded to include more samples.