We thank the anonymous referees for their comments. Our responses are noted here in blue text; revisions in quoted text are indicated with bold text. We reference page/line numbers in the original manuscripts as 'OM' and in the revised manuscript as 'RM'.

Referee #1

This paper presents detailed measurement of semi- and intermediate-volatility S/I-VOC compounds from laboratory burns of a range of wildland fuels using a novel approach, which employs Teflon membrane in front of solid-phase extraction (sorbent-infused) filters to capture particle and vapor-phase fractions, respectively. Derivativized extracts from these media are analyzed using high-resolution 2-D GC-MS and the resulting chromatograms are analyzed in several ways to provide insights into similarities and differences between I/SVOC profiles from the different fuels. For example, analysis points out that different benzenediol isomers (that would not be separately resolved by other analytical techniques) have distinctive emission profiles across the different fuels/burns. The authors also explore the volatility distribution of the analyzed compounds. This is an interesting and well written paper, and I believe it should be published. I include some minor comments/clarifications below. My main concern is that this is more of a 'methods' paper (and so perhaps more appropriate for AMT than ACP) as it explores the application and some evaluation of a new and interesting method to a range of test burns. It does include a 'survey' of analysis across a few dimensions but does (and perhaps can't) go as far as actually yielding quantitative estimates of emissions the may be of most interest to the ACP audience. The main contribution is to show the promise of this new measurement approach and give a 'sampling' of some of the insights that its application yields. In short, while it is definitely an important contribution, it is more exploratory than broadly quantitative. The paper itself generally makes this orientation clear, but perhaps the title and abstract could be slightly reframed to emphasize the nature of the study.

We generally agree with the referee's assessment of the scope of the manuscript and also considered whether to submit it to AMT or ACP. Ultimately, we decided that the lack of a fully optimized, quantitative analysis protocol meant this work was better suited to ACP as the majority of the discussion focuses on the chemical insights gained from the filter analyses. Although qualitative, many of the findings, such as isomer-dependent emissions and partitioning of compound classes in biomass burning samples, have not been previously reported and will be of interest to the ACP community.

We feel that the title is a concise and accurate representation of the work that does not overstate the scope of the manuscript; thus we have opted not to revise it. However, we have revised the abstract in response to this referee's comments. We have added a note that the analysis is qualitative in the following sentence: "Emitted I/SVOCs were collected onto Teflon filters and solid-phase extraction (SPE) disks to **qualitatively** characterize particulate and gaseous I/SVOCs, respectively" (Page 1, Line 20, RM). We have also modified the objectives listed in the introduction to better emphasize the qualitative nature of the study "(1) **demonstrate the potential of** SPE filters for the untargeted analysis of compounds with a wide range of volatilities; (2) **qualitatively** investigate the **diversity** of I/SVOCs emitted from biomass burning across a range of fuel types…" (Page 3, lines 16-17, RM)

Specific points
P1, Line 23 – The abstract discusses 'speciation profiles', but the paper later specifies (P. 8, L25) that it will not provide speciation profiles (at least as this term is typically applied). As noted above, the abstract could be re-worked to clarify the scope/intent of the study.

We have changed 'speciation profiles' in the abstract to '**observed chromatographic** profiles' (page 1, line 23, RM) to make it clearer that we are considering the differences in observed chromatographic peaks between samples rather than identified compounds. We agree that is an important distinction.

P7, L20 – 'error estimated to be approximately a factor of two' – this phrasing doesn't make that much sense, especially given that you reference RSDs of ~80%. Is it really indending to say an uncertainty of 200%?

A factor of 2 uncertainty is effectively equivalent to 100% RSD, which is consistent with our conservative estimate based on maximum RSDs of 80%. To avoid further confusion, we have revised the text to:

"Further, the relative standard deviation (RSD) for each compound was conservatively estimated to be approximately **100 %** based on the maximum RSDs determined in the recovery tests (**80 % and 64 %** for the PTFE and SPE tests…)" (Page 7, Lines 33-34, RM).

P8, L25 – This framing, 'investigating the diversity of emissions' – should be more emphasized early in the paper, or else some effort should be made to present profiles (even in relative/approximate form). In general, I think there would be interest in having access to actual profiles in a supplement, as this is the key product that might be applied by the interested research community (apart from demonstrating the method). Are there are least some compounds for which you could provide emission ratios or C2 factors?

We have now used the preferred framing when introducing the goals of the study on page 3, line 17 (RM): "(2) **qualitatively** investigate the **diversity of** I/SVOCs emitted from biomass burning across a range of fuel types."

We have posted all PTFE and SPE filter data to the NOAA FIREX archive that is now publicly available. A note has been added describing the data availability (Page 17, lines 19-20, RM).

Considering the analysis issues affecting some samples and the lack of recovery and internal standards noted throughout the text, we prefer to remain conservative and provide qualitative, rather than potentially incorrect quantified data.

P10, L33 – Why is this difference likely too large? Need to justify statement.

This response assumes the referee is referring to Pg 12, line 33 (OM): "…for comparison, $C^*_{\mathrm{Cat}}/C^*_{\mathrm{HQ}}$ = 51 using the Verevkin and Kozlova (2008) ΔHvap data, which is likely too large of a volatility difference between the two isomers." We have added the following justification (page 13, lines 11-12, RM): "which is likely too large of a volatility difference between the two isomers **based on a comparison with hexane- and cyclohexane- diols (Capouet and Muller, 2006)."**

P13, L29-32 – I'm not sure this statement is supported. Doesn't this depend on the lifecycle of the VOC? Couldn't either of these be mostly reacted in the gas-phase if they don't find themselves around liquid water?

We have toned down this discussion and now suggest that differences in oxidation phase may, rather than will, be different between the two isomers. The text now reads: "Overall, the partitioning estimates calculated in this work support the conclusions of Smith et al. (2015) that **aqueous/condensed phase** SOA production mechanisms **are more likely to occur for hydroquinone than catechol**, for which gas-phase oxidation will dominate SOA production **based on its low condensed-phase fraction** (Figure 5)…" (Page 14, Lines 7-9, RM)

P14, L14-15 – How is it possible that particles in only one or other phase should be in a middle volatility bin? This needs justification.

We have added the following explanation: "It is possible that such compounds could belong in one of the middle volatility bins **if, for example, the compound was actually present in both phases, but was below detection limit/peak selection criteria in either the PTFE or SPE filter sample."** (Page 14, Lines 25-26, RM)

P14, L24-31 – It would make sense to renormalize your distribution to the same C* range when making this comparison – it is good to show your entire distribution, but doing the higher volatility bins are presumably 'invisible' in the May distribution, so I would advise a version of this plot making the comparison with a version of your distribution normalized to the same range as the May et al. distribution. Also, note that the fuel and burn conditions are different, and this is another source of variability (in addition to different analytical approaches). This discussion also falls into the category of 'descriptive/exploratory'.

We disagree that the bins should be rescaled to the May et al. volatility distribution, which was derived based on particle-phase measurements alone. A key motivation for this work was to measure compounds with volatilities intermediate to those accessible using the commonly applied particle- and gas-phase measurements. Rescaling our data to particle-phase measurements would diminish those efforts. We have modified the following sentence to make that point: "The trends observed in bins $\log C^*$ = 2-4 are nominally similar between the two studies, although the fractions are lower in this work due to allocation of some mass into bins >4 and **illustrates that the SPE filters enabled**

measurement of higher volatility compounds, although the relative abundances of such compounds are likely underestimated" (Page 15, lines 8-9, RM)

We have also added that the fires represented in the two distributions are different: "the agreement is reasonable considering the two very different analytical approaches… **and the different fires included in each distribution.**" (Page 15, lines 12-13, RM)

Minor/technical corrections
P9, L32 – minor quibble, but using 'fuels' where you mean 'burns' should be avoided for clarity. . .

We agree and have made the suggested change.

P10, L12 – I believe this should be 10ˆ3, not 10ˆ-3?

We thank the referee for catching this typo (on Pg 12, L12). We have corrected the text.

Referee 2:
This manuscript describes the analytical methodology associated with a relatively novel sampling strategy, analyte extraction/recovery, and analysis by two-dimensional gas chromatography for biomass combustion aerosol sampled at the Fire Science Laboratory. The authors have apparently studied a wide number of samples using a new analytical method without sufficient preliminary validation to understand the potential method errors associated with intermediate- and semi-volatile organic compounds (I/SVOC). For example, samples were dried under N2 to exchange solvents one or two times and the authors noted sample losses during storage. These are very significant errors leading to incorrect quantities of I/SVOC components. Some of these errors are discussed, but the associated limitations in the results interpretation is not sufficient. Overall, the paper is quite long including 3 major goals: (1) test the application of untargeted analysis to the new sampling strategy, (2) compare the I/SVOC speciation profiles, and (3) assess the accuracy of the measurements. Unfortunately, I did not understand how the accuracy was assessed. Considering the method errors, it is very difficult to have a meaningful discussion of the I/SVOC profiles, which I thought was the main point of this paper. I think the authors could do a better job of eliminating the analytes must susceptible to evaporative losses, and then proceed with the comparison of the profiles and their correlation. I would like to see the paper strongly revised with careful attention to the limitations associated with the method errors/assumption. At the same time, I think the paper should be substantially shortened to focus on the most valuable results.

Regarding the comment: "…without sufficient preliminary validation to understand the potential method errors associated with intermediate- and semi-volatile organic compounds (I/SVOC)." We have added to the Supplement (page S6, RM) a discussion of additional quality control tests we performed to investigate potential volatilization losses, which demonstrate that the variability is likely within the factor of 2 uncertainty we proposed on page 7, line 20 (OM). With those additional tests, we have provided method validation in the form of (1) potential recovery biases associated with the SPE filters using a broad range of standards (Section 3.1, page 6, OM); (2) breakthrough/collection efficiency of the SPE filters (page 8, lines 12-17, OM); and (3) the impact of sample mass/absorbing phase on volatilization losses (page S6, RM), where (1) and (2) were addressed in the original manuscript and (3) was added in response to the referee's comments. We believe this is sufficient validation to demonstrate the feasibility of this relatively novel method; we have been clear in the original and revised manuscripts that the results are qualitative rather than quantitative and the limitations of our conclusions have been clearly reported. We have refrained from calling this a "methods" paper and submitting to a methods journal, as we entirely agree that further method optimization and validation are needed to achieve quantitative analysis, and have stated as such throughout the text (e.g., see response to Referee #2's comment #10 below). Furthermore, we were very conservative in presenting only results that likely were not significantly impacted by the noted errors and uncertainty. When introducing each analysis of the fire samples, we discussed the justification and/or potential errors associated with that specific analysis (e.g., page 8, lines 30-34, OM; page 11, lines 3-5, OM; and page 14, lines 5-11, OM).

Regarding the comment: "…3) assess the accuracy of the measurements. Unfortunately, I did not understand how the accuracy was assessed," as written the referee has significantly broadened the scope of objective 3 relative to that stated in the manuscript. Our objective 3 was specific to investigating gas-particle partitioning (page 3, lines 18-19, OM): "(3) assess the accuracy of phase-separated SPE measurements to predict gas-particle partitioning of compounds". We have done that through a detailed comparison of the filter-based benzenediol (Figure 5, Section 3.2.1, OM) and levoglucosan (page 15, lines 12-17, OM) volatility estimates with available thermodynamic/volatility data. These two examples reflect compounds with volatilities differing by several orders of magnitude. We have added 'specific' before 'compounds' in objective 3 to further make this clear (pg 3, line 19, RM). We additionally acknowledged that further optimization would improve the partitioning measurements (page 12, lines 7-8, OM): "Future use of recovery standards and further method optimization will help to correct for any sample-to-sample inconsistencies in the extraction efficiency and ultimately better constrain the partitioning estimates."

Regarding the manuscript length, from a general perspective, the length is well within limits of previously submitted manuscripts by Hatch and/or Barsanti to ACPD. The current paper, as formatted for ACPD, is 30 pages; other manuscripts, also as formatted for ACPD and published post 2015 since formatting has changed, range from 22-36 pages. Two recent FIREX publications by other authors are 34 and 44 pages in ACPD, neither of which introduces a new method. Further, referee #2 has provided no specific comments in their review related to which results s/he found to be most valuable vs. confusing/unnecessary. The justification for substantially shortening the manuscript is not clear, as all of the specific comments seek additional information regarding the methods. We believe the manuscript in its current form satisfactorily addresses each of our stated objectives: demonstrating the feasibility of a new analytical approach and presenting results on gaseous emissions profiles as a function of fuel type and partitioning of I/SVOCs in smoke, which are of general interest to the atmospheric community. We have made an effort to be as concise as possible in presenting our approach and findings, while also addressing current uncertainties and limitations (which we agree are a critical component), but have not shortened the manuscript by any significant measure.

Specific comments:
1) Please specify how the blank subtraction was performed.

The background/blank subtraction was explained on page 5, lines 3-6 (OM): "Background correction was performed separately on the PTFE and SPE filters. Within each filter group, the maximum signal of a given peak observed on any *blank or background sample* was multiplied by 2 to ensure that all artifact peaks arising from the extraction and derivatization procedures were entirely removed; that value was subsequently used for background correction of all corresponding fire samples." As the referee did not specify which additional details s/he is looking for, we have attempted to provide further clarity by revising the sentence to (page 5, lines 3-7, RM):

"Within each filter group, the maximum **peak area** of each peak observed **across the full set of corresponding blank and background samples** was multiplied by 2 to ensure that all artifact peaks arising from the extraction and derivatization procedures would be entirely removed. **Those adjusted values were subsequently used to background correct all associated peaks in the corresponding fire samples; only 16 (PTFE) and 32 (SPE) peaks in the fire samples required background correction**."

2) MeOH extraction was performed prior to BSTFA derivatization and BSTFA derivatizes hydroxyl groups. Unfortunately, this can lead to unexpected variations in the amount of BSTFA that is available for reaction with the analytes, since the relative reaction rates are expected to be much slower for carboxylic acids compared to primary alcohols. This limitation should be noted.

We assume the referee is suggesting that the methanol solvent consumed some of the derivatization agent. However, the extracts were fully dried down prior to derivatization; thus we do not believe availability of BSTFA was a limiting factor in the derivatization of carboxylic acids. We have also revised the text accordingly (Page 4, lines 19-20, RM):

"For derivatization, each aliquot was dried completely under nitrogen **to avoid reactions between the solvent and derivatizing agent**."

3) Drying extracts leads to substantial losses of components depending on their specific volatility and can increase the possibility of solvent reaction artifacts (e.g., ester and hemiacetal formation). This limitation must be made clear along with appropriate caveats in the interpretation of the results.

We are aware of this limitation and discussed the potential for variable volatilization losses in the Supplement (Page S2, OM): "As seen in all figures (especially Figures S1 and S2), the recovery from both PTFE and SPE filters decreases with decreasing retention time (increasing volatility), likely due to losses during the drying periods. In both the PTFE and SPE tests, recovery increases for later-eluting compounds, although in the "transition" region where recovery gradually increases for each subsequent compound (e.g., tetradecane – heptadecane, Figure S1; guaiacol, Figure S3), *the measured recoveries were highly variable, as indicated by the large error bars. Despite such variability, we have retained these compounds in the biomass burning dataset and assume a factor of 2 uncertainty to account for this observation*".

We also provided relevant caveats in the main text when introducing relative abundance (Page 8, lines 29-35, OM): "The peak area fractions are used as an indication rather than an absolute measure of the % abundance *because the individual compounds have not been corrected for recovery or differences in instrument response.* This approach is reasonable for the qualitative assessment sought here because *many of the most volatile compounds were still observed with high relative abundance*. For example, cresols likely exhibited poorer recovery (<0.3) than guaiacol (Figure S3) due to their higher volatilities, yet were still among the most abundant of the observed compounds. However, *the relative abundance for volatile compounds should be considered a lower limit."*

In the revised manuscript, we have added the following additional discussions and caveats regarding volatilization losses:
-Page 6, lines 15-19, RM: "**However, there is no clear relationship between the poor recovery of these four compounds and oxygenated compounds that elute in the same primary retention time window. For example, heptanoic acid displayed significantly better recovery in the PTFE test (0.86±0.06, Figure S4) than tridecane (0.14±0.07, Figure S1) despite eluting significantly earlier (1045 s vs. 1190 s). As a result, the poor recovery of the volatile alkanes/naphthalene was not extrapolated to other relatively polar oxygenated compounds present at similar retention times in the fire samples**"
-Page 14, line 31-32, RM: "**Additionally, the relative fractions of the higher volatility bins (i.e., logC^* bins ≥5) would likely increase following correction of the relatively poor recovery exhibited by such compounds."**
-Page 15, lines 8-9, RM: **"…which illustrates that the SPE filters enabled measurement of higher volatility compounds, although the relative abundance of such compounds are likely underestimated."**
-Page 16, lines 26-29, RM: "However, **SPE-induced** extraction biases were generally consistent for each compound class and could be corrected, **although recovery was low for the most volatile compounds due to volatilization during blow down.** Future improvements in the extraction protocol, **specifically testing additional solvents and drying time, in addition** to the use of recovery, internal, and external standards, should enable quantification of compounds collected onto SPE filters."

We have also added to the revised Supplement a discussion of additional tests that were conducted to further probe the recovery of the relatively volatile compounds encountered in the SPE samples. We find that the variability in the recoveries of volatile compounds to be within the uncertainty stated above (Pages S6-S8, RM).

In the original Supplement (page S4, OM), we had also discussed extensively the observation of methyl ester derivatives that formed by reaction of methanol with aldehydes during storage of the standard mixtures. We further assessed the likelihood that such reactions impacted the biomass burning samples with the following discussion:

"Because acids were observed in the SPE and PTFE fire samples, extraction and subsequent storage of methanol extracts may have caused some conversion of aldehydes to methyl esters, although we expect this artifact to be relatively minor. For example, the peak area of methyl vanillate was ~3.5% that of vanillin in the dung SPE sample. Artifacts resulting from methanol extraction have also been previously observed (Sauret-Szezepanki and Lane, 2004). Therefore, further optimization of the extraction solvent should be conducted to limit potential reaction products, as well as to improve the recovery of hydrocarbons, as discussed in the main text (Section 3.1)."

However, we agree that the potential for analyte-solvent reactions was not clearly addressed in the main text. We have therefore added the following discussion to the main text (Page 7, lines 12-16, RM):

"**Because other artifacts resulting from methanol extraction have been previously observed (Sauret-Szezepanki and Lane, 2004), further characterization of methanol-extraction-related artifacts and/or further optimization of the extraction solvent may be needed for studies seeking to accurately identify/quantify certain compounds. However, given the consistency of the observed HMF reactions, potential solvent-analyte interactions likely did not substantially impact the observed differences in the chromatographic profiles of the biomass burning samples discussed below**."

4) An interesting point of discussion that could be elaborated here relates to the meaning of untargeted and targeted analysis. What is the actual certainty associated with the NIST library matches? I would expect them to have variations, especially with consideration to derivatized components. What happens to the analyte signals without library matches? What fraction of the overall signal do they represent?

We assume the referee is referring to the following text on page 5, lines 10-13 (OM): "For other peaks, a compound name was assigned and considered *tentatively identified* if the match similarity with the NIST library hit was >800 in any sample that contained that peak and the identified structure was consistent with the derivatization approach used here (i.e., contained trimethylsilyl groups where applicable)."

A recent paper assessed the accuracy of NIST library matches and found that the likelihood of an incorrect match was ~30% in the match range of 800-900 and even lower for matches >900. We have incorporated these details and the reference into the manuscript: "**Worton et al. (2017) have recently demonstrated that the probability of an incorrect NIST library match is ~30% for matches between 800-900 and 14% for matches >900"** (page 5, lines 14-15, RM). We emphasize that these compounds are considered tentatively identified, as we specified in the original manuscript (see quote above).

We are not entirely clear what the referee means by “variations” in the library matches. If s/he is referring to variability in the similarity values, such variability could arise depending on the quality of the analyte spectra (e.g., influence of background signal in low abundance peaks). For that reason, we allowed a tentative identification for a given peak if the library match was >800 in any sample. If s/he is referring to variability in the actual compounds matched, that could occur if there are a number of library compounds available with comparable match similarity values, in which case the best match may differ from sample to sample due to slight variations in the individual similarity values. We did not find that to be the case for the few tentatively identified compounds noted in the text. Because electron ionization is performed at an industry standard -70 eV specifically to minimize variability in the generated mass spectra, there shouldn’t be any additional variability introduced by derivatization. However, compounds for which TMS-derivatives are not included in the library will not be identified, which thus explains the large number of unknown compounds within the fire samples: of the 1488 compounds observed across the full set of samples, only 24 were tentatively identified and 66 were positively identified. Because we have noted relatively few tentatively identified compounds in the text, and in the interest of keeping the text concise, we do not elaborate on the potential error in the tentative identifications.

The NIST library match was used to tentatively identify a few additional compounds (beyond those for which standards were available); the identification is independent of the analyte signal. Compounds for which the NIST match similarity was <800 were still retained, they are just considered as unknown identity. The NIST match has no influence on the analyte peak area.

5) Extraction efficiency tests include a test named PTFE test, but this test doesn’t have anything to do with the PTFE filter. Please rename the tests. See also figure 1, where PTFE recovery is used to assess the SPE recovery. What is the slope of the fit? What are the outliers in the plot with red symbols?

When introducing these tests on page 5, lines 19-21 (OM) we were clear that the “PTFE test” was meant as a proxy for the PTFE filters and provided the reasons those tests were conducted as they were: “standard was spiked into 7 mL of methanol (HPLC grade) as a *proxy* for the PTFE extractions (hereafter called “PTFE test”; *standards were not spiked directly onto a PTFE filter due to the potential for rapid evaporation of the relatively volatile standard constituents*).” This approach implicitly assumes no recovery bias was introduced by the PTFE filter, which we have added to the sentence:

“…standards were not spiked directly onto a PTFE filter due to the potential for rapid evaporation of the relatively volatile standard constituents (Dhammapala et al., 2007); **this approach assumes no recovery biases are introduced by the PTFE filter**;” (Page 5, line 25, RM)

Because the PTFE test was used as the benchmark for comparison with the SPE filters, that assumption would overestimate rather than underestimate potential biases introduced from the SPE filters.

The regression parameters of the fit in Figure 1a were given on page 6, lines 14-16 (OM): “Linearity with a *correlation coefficient* (R^2) *of 0.72* was observed between the SPE tests and PTFE tests (Figure 1a); however there is clearly a negative bias in the extraction from SPE filters compared to the PTFE tests, *with a slope of ~0.75.”* We have added these details to the Figure 1 caption in the revised manuscript so they are easier to find.

There are two red outliers in Figure 1a. The outlier with the square marker near the 1:1 text is one of the carboxylic acids that exhibited artifacts from the extraction/derivatization procedure, which was noted on page 6, line 25 (OM). The other red outlier with a circle marker at ~1.0 along the x-axis is an alkane, for which we discussed the under-recovery from the SPE filters on Page 7, line 5 (OM).

6) Please discuss the possibility of positive and negative sampling artifacts with respect to your sampling scheme.

We assume the referee is referring to filter adsorption artifacts. We pointed this out on page 12, lines 27-28 (OM): “…considering that the estimated Ci* values are more likely to be underestimated rather than overestimated due to adsorption of I/SVOCs to the PTFE filter (Mader and Pankow, 2001).” We also discussed the potential for sample-line losses on page 13, lines 14-24 (OM).

7) The threshold for poor recovery is quite low at 20%; I recommend raising it toward 50% for the discussion.

For the qualitative demonstration of the SPE method sought here, we have opted to use a relatively low threshold in order to illustrate the types of compounds that can be collected on the SPE filters. We have stated this rationale on page 7, lines 28-29 (RM): “As the interest at this stage is in assessing the full diversity of I/SVOCs in smoke **and demonstrating the range of compounds collected by SPE filters,** all observed compounds (except those explicitly noted above) were retained without correction for recovery.”

We have tried to be very clear that the recovery of relatively volatile compounds is low (e.g., page 8, line 34, OM: “the relative abundance for volatile compounds should be considered a lower limit.”). We have also adopted relatively large error (100% RSD) to account for the potential variability in the recoveries of the relatively volatile compounds; all tests that we have conducted indicate such uncertainty is reasonable, as shown/discussed in the Supplement, as well as on Page 7, lines 20-24 (OM). Additionally, relatively few standard compounds included in the SPE tests exhibited recovery below 50 % (Figures S1-S5) and thus we do not feel that we have enough information to adequately map the recoveries of peaks that eluted early in the chromatogram (where recoveries were poorest), as retention time alone is not a sufficient proxy. Thus any changes in the compounds included in the analysis would largely be arbitrary without analyzing a wider range of standards. We have added some discussion to that effect (page 6, line 15-19, RM):

"However, there is no clear relationship between the poor recovery of these four compounds and oxygenated compounds that elute in the same primary retention time window. For example, heptanoic acid displayed significantly better recovery in the PTFE test (0.86±0.06, Figure S4) than tridecane (0.14±0.07, Figure S1) despite eluting significantly earlier (1045 s vs. 1190 s). As a result, the poor recovery of the volatile alkanes/naphthalene was not extrapolated to other relatively polar oxygenated compounds present at similar retention times in the fire samples."

8) Several specific compounds were poorly recovered from the SPE tests. To what degree are these recovery issues related to the various method errors (solvent evaporation steps, sampling artifacts, or derivatization artifacts).

In the original manuscript, we pointed out examples of specific compounds/classes that exhibited clear experimental bias and the likely method error responsible for each bias. On pages 6-7 (OM), we discussed the following specific cases:

a) Poor recovery of the most volatile compounds that "likely volatilized during the drying periods." (Page 6, line 10, OM)
b) Palmitic and stearic acid were over-recovered and "are likely an artifact of the extraction/derivatization procedure." (Page 6, line 26, OM)
c) Maltol was significantly under-recovered from the SPE filters and state "The reason for this large SPE bias for maltol is currently unknown." (Page 6, line 30, OM)
d) 5-Hydroxy methyl furfural reacted during extraction in both the PTFE and SPE tests. (Page 6, line 32, OM)
e) All hydrocarbons were poorly recovered compared to oxygenates and suggest that "it is possible that hydrocarbons were more strongly bound to the non-polar octadecane-based sorbent material than more polar compounds resulting in lower recovery." (Page 7, line 6, OM)
f) Octadecane was over-recovered in the SPE tests "likely due to degradation of the SPE sorbent material during extraction" (Page 7, line 9, OM)

We additionally offered further discussion of the recovery results in the supplement (Section S1, OM) and therefore believe we have fully explained the results related to the observed SPE biases. As the referee did not specify which other compounds require explicit discussion or which explanation(s) s/he found inadequate, we cannot provide further clarity.

9) Extracts for I/SVOC analysis should never be allowed to dry. Practically speaking, there are variations in the N2 flow rate and there are sample matrix differences that affect the rate of solvent evaporation. The two practical considerations contribute to variations in time the dried solute is exposed to the N2 thus varying the extent of analyte losses. This is a very important limitation that is poorly addressed.

In this case, it was necessary to completely dry the solvent to avoid reactions between methanol and the BSTFA derivatization agent. As the referee noted in comment #2, such reactions could limit the amount of BSTFA available to derivatize analytes. We agree that drying to completeness is not ideal for maximizing recovery of I/SVOCs and should be one variable considered in future method optimization, which would necessarily require testing different non-protic solvents or other derivatization agents. We have added that as an additional optimization parameter for future experiments on Page 8, Lines 30-31 (RM): "**For example, other non-protic solvents could be tested to enable derivatization without the need to completely dry the samples, which would help to minimize volatilization losses**."

To further address the volatilization losses, additional quality control tests are now discussed in the Supplement (pages S6-S8, RM). Although the goal of the tests was to assess the role of the amount of absorbing phase on the recovery of relatively volatile compounds, the results provide some insights into the effects of drying time and matrix effects. Because of the different volumes used for those tests (100-500 uL), the drying time varied considerably (although was not controlled). We did not observe variability outside the factor of 2 (100% RSD) that we assumed for the samples on page 7, line 20 (OM), as long as some absorbing phase was present, as would be the case in the biomass burning samples. We did observe evidence for potential matrix effects wherein a polar matrix reduced the recovery of an aliphatic hydrocarbon. However, we do not expect this to introduce major errors for this qualitative analysis. For example, we find much higher fractions of aliphatic compounds in peat smoke compared to sagebrush smoke (i.e., the example shown in Figure 4c). Based on complementary measurements from the same fires using sorbent tube collection followed by thermal desorption-GCxGC analysis, which are not subject to the same matrix effects (unpublished data, but similar protocol to that described in Hatch et al. 2015), we also find negligible aliphatic compounds in sagebrush compared to peat. We have added this discussion to the revised Supplement (Page S8, RM).

10) Were any internal standards used to monitor analyte losses?

No. This was noted several times throughout the manuscript in the context of future needs for quantitative analysis, specifically:

- Page 6, line 17 (OM): "Such bias could be corrected in future work through application of recovery standards."
- Page 7, lines 12-13 (OM): "Further method optimization (e.g., use of recovery, internal, and external standards) would be necessary for quantification."
- Page 12, lines 7-8 (OM): "Future use of recovery standards and further method optimization will help to correct for any sample-to-sample inconsistencies in the extraction efficiency and ultimately better constrain the partitioning estimates."
- Page 16, lines 12-13 (OM): "Future improvements in the extraction protocol, specifically the use of recovery, internal, and external standards, should enable quantification of compounds collected onto SPE filters."

11) Considering the method limitations, please justify how the analyte recovery from the SPE filters was determined to be satisfactory? For which analytes?

We have conducted a number of tests to characterize the recovery for a broad range of compounds from the SPE filters, as outlined in the main text (Section 3.1, OM) and Supplement (Section S1, OM). We found that several compounds exhibited very poor recovery regardless of filter type due to substantial volatilization losses, as discussed on page 6, lines 8-12 (OM) and possible activity effects, as now discussed on page 6, lines 12-14 (RM). We noted specifically which compounds were omitted from the analysis on page 6, lines 11-12 (OM) and page 7, line 10 (OM). For the other standard compounds, our tests indicated that the variability in recovery was within a factor of 2 (and significantly better for most compounds), which we noted on page 7, lines 20-24 (OM). We deemed this factor of 2 uncertainty as sufficient for the qualitative analysis sought here. We acknowledge that the absolute recovery is not consistent across all compounds and further characterization tests would be needed to adequately correct for that in order to achieve quantitative analysis. We stated as such on page 7, lines 13-15 (OM): "Characterization of the recoveries for more standards would also be necessary to adequately correct the recoveries for the wide range of compounds observed in untargeted analysis of biomass-burning smoke (including many unknown compounds), especially the most volatile observed compounds."

Please also see the responses to Referee #2's comments 3, 7, and 9 for additional discussion of the method limitations and the associated caveats we have noted in the text.

12) Again, the assessment of diversity in the I/SVOC is dependent on the assumption that the samples did not have losses associated with the sample preparation. What if some of the variation you see is due to solvent evaporation during storage or during the solvent exchange steps?

All samples were stored identically; we did not observe obvious solvent losses during storage. We noted volatilization losses in some samples during derivatization due to faulty vial caps and have attempted to correct for that artifact (Table S1, OM). We looked for and found no clear association between those losses and potential errors in the gas-particle partitioning estimates where such artifacts would have the greatest impact among the analyses presented in the manuscript. As we've noted above, we believe the variability in compound recovery is on the order of a factor of 2 or less. In the discussion of the results, we have specifically emphasized results that fall outside this factor of 2 uncertainty (e.g., in the discussions of Figure 3 and in Figures 4b,c).

13) Please define what is meant by composite chromatograms?

For additional clarity, we have revised the text to:
"Figure 2 shows a comparison of composite SPE and PTFE chromatograms **wherein peaks from all SPE and all PTFE samples are represented** (GasPedal, Decodon Gmbh, Greifswald, Germany) (Schmarr and Bernhardt, 2010)" (page 8, lines 5-6, RM). Readers are referred to the cited reference for additional information.

14) Please define the RI values ~ 1200-2100.

Retention index is a standard and commonly used chromatography parameter, and as such does not require further definition here. We have additionally provided the retention times for comparison to Figure 2 to put the noted RIs in context.

15) References on page 8 did not simply use PUF for SVOC collection. XAD has been frequently used and XAD is an SPE material.

We had noted as such on page 2, lines 30-31 (OM): "All of the off-line I/SVOC characterization studies cited above collected samples using a filter and polyurethane foam (PUF) plug (or PUF/*XAD-sorbent*/PUF sandwich". We agree that XAD should be noted on page 8 (OM), as well (page 9, line3, RM). The text now reads: "samples are largely consistent with the compounds reported from PUF/**XAD** measurements".

16) Figure 4 is quite complex. The effect of levoglucosan on the correlation is quite strong. Since levoglucosan is non-volatile, it could suggest that the more volatile components were more highly variable even between replicate burns. But, is this related to the sampling strategy, sample prep, or actual burn conditions? This makes me curious about other volatility bins and their correlations.

We investigate the referee's questions using the peat-sagebrush comparison shown in Figure 4c for visual reference and because sampling/analysis artifacts are more likely to lead to poorly correlated sample pairs than well correlated sample pairs. We have separated the data into two fractions-- compounds eluting before and after levoglucosan-- to investigate the correlations among different broad volatility classes. Of the 358 compounds present in only one of those fires, approximately 40% (by compound number and relative abundance) are attributed to compounds that elute after levoglucosan, indicating that there were significant differences in the compounds that are generally less volatile than levoglucosan. Additionally, although the correlation coefficient is very poor for both compound fractions, it is significantly lower among the less volatile compounds: 0.075 for compound eluting before levoglucosan and 1.7e-5 for compounds eluting after levoglucosan ($R^2 = 0.057$ including all data except levoglucosan). Thus the correlations are not driven solely by variability in the recoveries of volatile compounds. Furthermore, of the compounds that could be identified (positively or tentatively), the differences are consistent with previous studies. For example, the fact that peat smoke is uniquely rich in aliphatic compounds has been shown previously, as we stated on Page 10, lines 19-22 (OM): "Peat smoke was also poorly correlated with most other biomass-burning emission samples, largely due to a high abundance of aliphatic compounds, including *n*-alkanes and 1-alkenes, as has been observed previously (George et al., 2016; Hatch et al., 2015; Jayarathne et al., 2018)." In summary, we find no significant indication that the correlations in Figure 4 are dominated by sampling errors rather than real differences in the smoke samples.

17) What is the purpose/significance of figure 5?

The entirety of section 3.2.1 is devoted to discussion of the significance of the results shown in Figure 5. In particular, we illustrate the fuel dependence of the benzenediol isomer speciation (Figure 5a), which would not be observed by many of the instruments commonly used these days that are unable to resolve isomers (e.g., chemical ionization). We additionally used the benzenediols as an example to demonstrate the ability of the filters to capture phase separation and estimate saturation concentrations, C*. In Figure 5b, we show a comparison of the filter-based C* estimates with available thermodynamic data, as a means of testing the accuracy of the method, which was one of the objectives of the paper as outlined in the Introduction. Figure 5b further illustrates the significant differences in gas-particle partitioning between the two isomers, for which we have discussed the implications (i.e., potentially underestimating emission factors, differences in sample-line losses, and available reaction pathways) on page 13, lines 5-8, 29-32 (OM). As the referee did not specify what s/he found confusing or inadequate within the ~3 page discussion related to Figure 5, we cannot provide further clarity regarding the significance of that figure.