

We would like to thank all reviewers for their time and especially for the detailed comments provided by Reviewers 1 and 3. As a result of these comments we have made substantial changes to the structure of the paper, placing some of the more technical aspects into Appendices. We have gathered together (under their own sub-headings) thoughts on the potential sampling biases, difficulties of estimating uncertainties and the likely applicability of the emission factors that we report to other fires in the region. We have deleted or corrected passages where needed and believe that these changes will improve the accuracy and readability of the paper. We have also sought to address all the individual comments of the reviewers making corrections and clarifications where necessary. Below we have reproduced the reviewers' comments, numbering the comments that require a reply and/or change to the manuscript. In blue font below each numbered comment is our reply.

Reviewer 1

The manuscript by Paton-Walsh et al. is a well-written and clearly articulated manuscript that provides emission factors for gases from fires in Australian temperate forests. The authors provide plenty of reasons for the necessity of these emission factors, not just because of the direct impact of fire emissions on atmospheric chemistry but also because of the paucity of available data for which to use in air quality modelling. This is surprising for a country as fire prone as Australia and the data presented in this manuscript is, therefore, sorely needed

The manuscript is very detailed in its methodology and I appreciate its honesty in describing potential caveats and how these were overcome. Although, I personally have not used FTIR as a measurement technique, the detailed description of the instruments used and how the spectral data were interpreted give me confidence in the measurement technique used and the emission ratios/factors that they have produced.

- (a) My main concern with the manuscript is based on the treatment of the uncertainties. The authors have treated the overall effect of many of the measurement uncertainties by summing in quadrature. However, no mention is made of whether they have accounted for the co-variance in some of the factors? For example, the authors mention that air temperature has an effect on both spectra and on air density. Therefore, the error in the spectra and density are not independent of each other and will exhibit co-variance. This needs to be accounted for in the error propagation calculations.

The first point raised by Reviewer 1 is that co-variance between uncertainties needs to be accounted for. We have done this for the main case where we expect there to be co-variance (the spectral and density temperature errors). This is mentioned already on lines 20 – 23 on page 24, but obviously this point was not made clearly enough in the original manuscript.

The uncertainty analysis has now been moved to an Appendix B but this point is now clarified with the following text:

“The different temperature errors are obviously correlated and so the combined uncertainty was determined by some basic sensitivity studies. In the case of CO₂ the density and spectral temperature errors are in the same direction giving a combined error of 15.3% for an underestimation of the temperature of 20°C, whereas for CO they partially compensate giving an overall temperature error of 5.1%. These combined temperature uncertainties are added in

quadrature to those resulting from uncertainties in the assumed background mole fractions, spectral fitting errors and errors in the HITRAN lines used in the retrieval.”

We have also replaced the separate temperature errors in what was Table 3 (now Table 7) with a column for “combined temperature error”.

This issue is also relevant when estimating uncertainties in the emission ratios. However the difficulty is that the cross-sensitivities are extremely complicated. As the reviewer points out later, this can be achieved by a Monte Carlo analysis. However the software used to retrieve concentrations from the spectra is not able to run this way as the different gases require different input files and different retrieval runs. It is well out of the scope of this study to re-write the software to enable a Monte Carlo analysis. Also such an effort is not commensurate with the likely return because of the other difficulties in making an accurate appraisal of the measurement uncertainties.

The paper now contains a paragraph outlining a number of issues in determining the measurement uncertainties and the detailed discussion has been moved to Appendix B. The following paragraph has been added to that Appendix:

“Component values for the uncertainties arising from spectral temperature sensitivities are estimated by assuming a maximum 20° C temperature error and adding sensitivities of target gas and reference gas in quadrature. (Whilst these errors are clearly not uncorrelated, the true correlation of the sensitivities of each gas are complicated by strong non-linearity and by feedbacks into the spectral fitting errors. Thus adding in quadrature provides a very approximate estimate commensurate with the other difficulties in obtaining a good estimate of our true measurement uncertainty).”

- (b)** Indeed, there is considerable use of error propagation calculations throughout the manuscript that are then used to assign an error to the emission factors for the five sites. As the authors ask the reader to accept their mean emission factors as default emission factors for Australian temperate forests, even though they were measured from fires within one State (NSW), I would expect a more probabilistic approach to the uncertainties, such as a Monte Carlo type simulation. This would provide the reader with a probability distribution over which we could expect the emission factors to come from. It would also allow the readers to ascertain whether the mean emission factors reported are negatively or positively biased and whether median values need reporting instead.

Reviewer 1 suggests that a Monte Carlo approach would be a good idea - however as well as being extremely difficult to implement, the authors do not believe that it would yield the information suggested (i.e. whether or not the mean emission factors reported are negatively or positively biased). This is because our uncertainty estimate only yields information about the measurement uncertainty and contains no information about the true underlying natural variability in the emission factors from the fires, nor any potential sampling bias. In our revised paper we have included a section (4.4) pointing out limitations in our ability to accurately estimate uncertainties. The detailed process of reaching the uncertainties is now moved to an Appendix B.

- (c) The authors report the mean emission factors from the five sites and report the variability of the emission factors with a standard deviation. Normally, I would accept that as a method of reporting. However, these values are then compared against the measurement uncertainties (see p.4357 L1-10). The standard deviation is only informing me about 68% of the variability in the values from the mean of the sample population. The measurement uncertainties are reported as absolute limits so the comparison the authors make against the standard deviations is not valid. So the authors cannot state that “the variability of the fuel carbon in this ecosystem is less than the overall uncertainty assigned”. As a crude check, the error on the mean, using the quadrature method, gives an uncertainty of approximately 5%, which is the same as the uncertainty assigned to the carbon content. Alternatively, the 95% confidence interval of the mean value is approximately 3.9% of the mean and the 99% confidence interval is 8.8% of the mean value. Therefore, I feel that the authors need to reconsider what the data is telling us here. This also applies to the second paragraph on page 4358 with regards to the other trace gases and the second paragraph of the summary and conclusions.

Reviewer 1 points out that our estimates of measurement uncertainty are reported as absolute limits and so some of the inferences made from comparison to the standard deviations in emission factors from the different fires are not valid (e.g about variability of the fuel carbon in the ecosystem). This is a good point and the authors have deleted these comments in the revised manuscript.

Minor Comments

1. Some of the descriptive information about the sites in the method section (e.g fuel loads) is replicated in Table 2. Include it in one or the other.
The duplicated information has been removed from the methods section text and retained in the Table.
2. P. 4335 L8. Space needed between ‘it’ and ‘decreases’.
Fixed
3. P. 4343 L26. Reference is made to a manuscript in preparation. I can’t check this manuscript out, therefore, include the details in the current paper or remove this paragraph.
Paragraph removed.
4. P4351 L11. Volkova and Weston use the value of 0.47 for one fuel fraction (trees, I think) but actually measured C content for other fuel fractions. Unfortunately, they don’t report the measured values.
Rephrased to correct this as suggested.
5. P4352 L20-21. This sentence repeats what was said at P4352 L4-5.
Repetition deleted.
6. Figure 2. Units missing on axes (e.g. cm-1 and arbitrary units).
Done.
7. Figure 2. In panels b and h I found some of the ‘red’ colours hard to distinguish.
Figure 2 has been remade with different colours – now clearer.

Reviewer 3

This is a good paper, the data is useful, and it should be published. However the presentation can be improved in my view. Currently about 29 pages (including 6 tables) are devoted to methods, error, fire description, and some speculation about things outside the scope of this experiment. Then the discussion about the representativeness, novelty, uses, etc of the final results is only three pages.

Due to this and other comments by this reviewer we have made a significant restructure of the paper, including putting some of the more technical details into Appendices.

- (a) For instance, the authors recommend that their data should replace data currently used for Australian forest fires when some users might prefer to include this new data weighted into an evolving literature average. Or some may prefer to use the data for modeling small hazard reduction burns."

These data represent the first directly measured emission factors for many trace gases from Australian forest fires, however it is true that this does not automatically mean that all users will wish to use them as is. Whilst eucalyptus trees might be expected to produce different emissions to other forest types, the variability of different ecosystems is not truly known amongst other variabilities that affect emissions. We have added in a section to discuss the applicability of the emission factors to address this and other such comments from this reviewer.

- (b) I offer more thoughts on discussion items in the detailed comments below. My thought would be to shift the balance by reducing the first 29 pages (delete some stuff and move some stuff to supplement) and then flesh out the discussion by examining the work in a larger context. In my view, the detailed description of the spectral analyses is too long. Much of it has been in the literature before and most could only be followed by a few specialized readers and may be better in a specialized AMT paper, supplement, or appendix. I leave this to the authors and editor to decide.

This methods section serves both this paper and part B. In our view past papers have sometimes lacked detailed descriptions of the spectral analysis - and in particular as pertains to a detailed uncertainty budget, probably because of the difficulties in doing this well. Thus we wanted to present a full description as part of this pair of papers. Nevertheless, we accept this point and have moved the most technical material to appendices as suggested.

- (c) For a general audience it would be of more widespread interest to describe the overall challenges, trade-offs, and innovations associated with the measurement configuration in 1-3 paragraphs written in not overly technical fashion. By way of explanation, most scientists do spectroscopy when the optical path is all at a single easily-measured temperature and pressure. A smaller subset do atmospheric total column spectroscopy when slowly-changing a-priori profiles for concentration and temperature can be assigned to a system modeled as a "stack of layers." In the present work, the fire essentially produces "layers" of potentially very different, unknown temperature and concentration that are rapidly changing in position, extent, and content. Each packet of emissions has a different, unknown vertical velocity that

may be needed to properly calculate a flux of emissions. There may be rising, hot CO₂ and falling, cold CO in the same measurement path? This challenge may be highest if one targets flaming emissions that normally account for most of the emissions with an optical path in/near the flames. If one targets more cooled/mixed emissions with the measurement path to minimize the (unmeasured?) variability along the optical path, it seems to come at the cost of relying on the complex fire-side meteorology to drive representative emissions sideways to the optical path rather than in "the normal" upward direction. Thus, the measurement geometry seems to present non-trivial challenges based on first-principles alone. Despite the lengthy technical section, I did not feel like these basic concerns were completely dismissed.

We have added a section specifically on the difficulties of obtaining a representative sample with different geometries. We have pulled all the existing comments together and also expanded our discussion of this topic.

- (d) The authors do cite Smith et al who evidently explored the effect of assuming an incorrect single temperature for the whole optical path, but maybe the major source of uncertainty is temperature variation along the optical path: potentially from 300-1100+ C? It was not specifically stated that the study of Smith et al. addressed all the issues that might be encountered with this measurement geometry. Having offered these concerns, it is important to state that in my view the authors are perhaps more qualified than anyone else to deal with these challenges and enact innovative solutions. Further, a very important point is that the authors obtained results within the envelope of previous work and so maybe I've overestimated the problems. In any case, it would be of widespread interest to summarize how these challenges are addressed in a manner accessible to a wide audience and how they impact true overall uncertainty for the individual fires sampled.

It is true that there may be some really large temperature differences along the measurement path and we have added some comments about this in the new paragraph that discussed the difficulties of estimating uncertainties. We also point out that huge discrepancies in the assumed temperature will likely be evident from examining the spectral fits obtained in the retrievals.

- (e) At the next level up, a discussion item is essentially: fire to fire variability vs individual fire uncertainty. In airborne studies of well-mixed samples, the uncertainty in individual fire EF is almost always smaller than the uncertainty in the study-average EF. In this work, the uncertainty in the study mean is sometimes smaller than the uncertainty in the individual fire value, but in general the study standard deviation is larger reducing the need for discussion of error in individual fires and making the possibility of systematic bias more of a concern to address. Real variability, possible bias, and measurement representativeness are important for modelers using emission factors at any scale. Sample size (this study is a rather small sample of five fires) is an important aspect of assessing those issues, but is not yet discussed. An additional discussion item, perhaps the single most important issue, involves going from uncertainty in these measurements to representativeness for regional modeling when the majority of biomass burned in Australian forest fires is consumed in very large, intense, uncontrolled fires as opposed to the small planned fires the authors current approach is designed for. The authors have done a great job

of comparing in the limited fashion possible to emission ratios in major Australian wildfires that were measured using their solar-tracking FTIR system when smoke was transported over the group's lab. The good comparison for fires burning on different scales should be pointed out more clearly even though the result may not be as good in future comparisons of fire emissions at different intensity scales. The authors already compare to airborne forest fire measurements, from the rest of the world, which is also good, but they point out the possibility of regional differences. Given the small sample size in their study and the large natural variability observed in all work, it may still be TBD if regional differences or if other factors (fuel moisture, season, etc) drive EF variability. Some papers that explore "non-geographic" or "non-vegetation" sources of variability include:

van Leeuwen, T. T., et al.: Dynamic biomass burning emission factors and their impact on atmospheric CO mixing ratios, JGR, 2013.

Akagi, S. K., Yokelson, R. J., Burling, I. R., Mainardi, S., Simpson, J., Blake, D. R., McMeeking, G. R., Sullivan, A., Lee, T., Kreidenweis, S., Urbanski, S., Reardon, J., Griffith, D. W. T., Johnson, T. J., and Weise, D. R.: Measurements of reactive trace gases and variable O₃ formation rates in some South Carolina biomass burning plumes, Atmos. Chem. Phys., 13, 1141-1165, doi:10.5194/acp-13-1141-2013, 2013.

<http://www.atmos-chem-phys.net/13/89/2013/acp-13-89-2013.html>

We have added a new section to more thoroughly address the likely applicability of the emission factors that we report to other fires in the region.

- (f) Some of the speculation that is beyond the scope of this study could be minimized throughout. For example, the authors discuss airborne vs ground-based measurements. While I think such a discussion doesn't need to be in this paper, I reluctantly offer some thoughts that may have been overlooked in the author's discussion mostly to illustrate that the issues may be more complex than recognized in the discussion paper. Ground-based measurements can only sample at the edges of a fire whereas airborne platforms can sample over the whole fire (which is often inhomogeneous with different fuels in the interior). Aircraft can sample the explosive blow-ups that may account for most of the fuel consumption. In particular, the airborne observer has the perfect view of the whole fire and can usually command the aircraft to all parts - including the most active parts. The ground-based observer is confined to fire-lines that are normally constructed well away from the heaviest fuel accumulations to make control easier and they cannot see most of any moderately-sized fire because of intervening smoke and vegetation. If the ground-based observer does see other important areas, they may be unable to move the pre-ordained fixed path. In Akagi et al. (2014) the fire was ignited on the perimeter and burned towards the interior and away from the immobile ground-based measurement path. Thus the aircraft was still sampling weakly-lofted, mostly-smoldering emissions from the interior of the unit after the ground-based measurements had ceased. Ground-based measurements can only be used for prescribed fires, can't be used to search for fires, and can't probe plume evolution in the described configuration. Both airborne and ground-based have some sensitivity to both flaming and smoldering. The ground deployment gets some flaming if/when fortuitous wind gusts direct smoke sideways instead of up. Airborne deployments get some smoldering because the intense convection column entrains smoldering emissions and even debris from

the site. On real fires a moving flame front will often persist throughout most of the fires occurrence rather than as a brief initial burst at a point. Airborne sampling favors smoke from high fuel consumption rates, while ground-based sampling favors smoke from low fuel consumption rates, thus airborne sampling may often be inherently more representative of whole fire emissions. An exception would be fires with little or no flaming combustion. Ground-based sampling would be inherently more representative of fires that burn almost entirely by smoldering, where some subset of peatland fires may fall in the category. Lab fires are the only way to capture all the smoke and they usually produce the highest flaming to smoldering ratio of all three main deployment platforms, but this is probably an artifact related to dryness and lack of wind which makes burning wetter fuels harder and may elevate MCE. See discussion in link above. Speculating about "thermal mass," large metal blocks used in "green buildings," is not tested and may be irrelevant.

The discussion of airborne versus ground-based measurements is relevant to this paper in that it puts into context the biases of the methods used in this study compared to other possible methods of sampling fires. Nevertheless the reviewer makes some valid points about the advantages of different methods and we have rewritten much of this discussion and removed it from parts of the text where it was ill-placed and put it instead in its own sub-section entitled “**Sampling Geometries and Potential Biases**”.

Specific comments.

The paper could be re-focused to some extent so here I just list possible errors and offer ideas on potential revisions.

- 1) P4328, L14: "given" to "measured locally"
Done.
- 2) P4328, L17-20: Include realistic uncertainties in the EF.
Done.
- 3) P4329, L27: "are large uncertainties" to "is large natural variability" to minimize implication of measurement error?
Changed to “there is a large range” – since if we include the geometry of measurement then probably a combination of natural variability and measurement uncertainty is responsible for the large spread of values.
- 4) P4330, L10 and 15: Actual large-scale fires in the field don't really burn in flaming and smoldering stages that often. It's more that there is a site-specific, potentially dynamic mix of flaming and smoldering combustion. Most other comments on how to sample fires are contained in general comments above. E.g. smoldering emissions are also lofted by convection, flaming may be long-lasting (for days); not "rapid." 125: Re: "very limited temporal coverage"- flaming is not just an "initial" phenomenon on real fires, where the flames can constantly move into new fuels.

OK – this has been rephrased to: “At any single location within the fire combustion can be thought of as progressing from a flaming stage through to a smoldering stage.” This also sets the context for the comment on line 15.

- 5) P4331, L4: What ground-based geometry works when the flame lengths are 50-200 feet as is common on Australian forest fires?

A geometry looking up at a steep angle up an escarpment would work. Obviously the larger the flames the greater the distance and angle required. Nevertheless a ground-based geometry does have the “potential” to operate in geometries that capture flaming emissions as stated.

- 6) L8: Akagi et al. (2013) cited but not in references.

Reference in place in revised version.

- 7) L4334, L15: Eqn 5 is not actually needed to get EF since that is already covered by eqns 1 and 2. I don't think it is ever made clear if there is a need for eqn 5. Just getting a full set of ER to CO₂ or CO by the sum or plot method and then converting to EF is simplest. The idea to use ER times EF of a reference species just seems to add extra work and error since the ER are included in the EF of reference species. Perhaps what the authors are getting at is the ER plots minimize effect of low S:N for some species?

Equations 1 and 2 requires a knowledge of the “background concentration” of each trace gas – and this is not always well known. Using Equation 5 removes this requirement. This is explained in lines 10-13 and reiterated later when discussing why we chose this method.

- 8) P4335. L8: "ltdecreases"
Corrected.

- 9) P4336, L6: "often" only applies to lab work, for ground-based it is "on rare occasions" and in the air "never"- see also comments on Fig 4 at end.

Replaced “often lies” with “can lie”

- 10) P4336, L21-25: How could Wooster et al get fire-averaged EF that resembled those at high MCE, but a medium fire-average MCE of 0.91?

I don't really understand what the reviewer is getting at here. This depends upon how you weight the measurements to get the “fire-averaged” value. The point is that in this paper they occasionally had fire radiative power measurements that they used to weight the “fire-averaged” emission factors. This produced a result that was closer to the EF measured during the flaming-dominated stages.

- 11) P4337, L7: The EF from the summation method is indeed better than the time average EF, but it ignores potentially higher vertical velocity of emissions during flaming. In the field, the flaming emissions (and entrained smoldering emissions) may "surge" into the atmosphere at 10 m/s while any un-entrained smoldering emission may linger at ground level or only rise weakly. Thus the same mixing ratio can represent different production of products. The fix for this in lab studies is to use a constant entrainment flow that is much larger than the fire-driven flow so that the summation method is rigorous, but this is not feasible in field.

The reviewer raises a question about the ability of the measurements to obtain a representative sample – due to the possibility of flaming emissions moving past the line of sight faster than smouldering emissions. Whilst the point is in itself valid (you can only be sure of getting a really representative sample in a laboratory fire with a controlled flow), the exact same issue applies regardless of how you subsequently treat the data. i.e the summation method does not suffer this problem any more or less than any other technique to derive an emission factor.

We have moved the discussion of sampling issues from the description of the summation method to a more appropriate part of the paper – a new section that discusses the potential biases from viewing geometry and we have addressed this point more explicitly.

12) P4339, L13: does unmodulated radiation cause a baseline (zero) offset? If so, this can impact retrievals.

Using unmodulated radiation means that the detector is also sensitive to the radiation from the surrounding environment. This needs to be subtracted or else there is a zero-offset that can affect the retrieval of some gases to a large extent. In our configuration there is no sensitivity to the surrounding environment (because the radiation from the surroundings does not get modulated and is therefore not seen in the spectrum (because it is outside the bandwidth of the detector)).

13) P4339, L16-18: This doesn't sound realistic. How is the 1100 C temperature of flames accounted for when sampling "thru flames" or is there a filtering process by which this data can actually be rejected without impacting representativeness?

As explained above, in our configuration there is no sensitivity to any radiation that does not come from the spectrometer because it does not get modulated and is therefore outside the bandwidth of the detector and therefore not seen. This applies to flames regardless of how hot they are.

14) Related question re P4341, L13: How would you know "true amounts" along a complex open path other than in well-mixed, independently-measured background air?

In the study by Smith et al, 2011 described in these lines of text, true amounts were known because they filled a large gas cell with calibration gases and then measured the concentrations using an open-path arrangement to determine the accuracy of the technique used here. This has been clarified in the text so that the reader does not need to read the paper referenced to understand this.

15) P4339, L23: 7 seconds for a 1.0 cm^{-1} scan sounds like a slow mirror speed and there can be effects of concentrations changing during the scan.

The scan speed is not adjustable in the spectrometer we used. We did not need to co-add 3 spectra – but this is a trade-off between signal-to-noise and changing concentrations.

16) Sect 3.2: presents an overly long discussion of error that may largely omit what might be the largest source of error; temperature?

Section 3.2 describes quantitative analysis using MALT (not errors). A detailed discussion of errors/uncertainties is given in Section 5.2 (including temperature uncertainties). Nevertheless we have cut some of this section and put it instead in an appendix.

17) Sect 3.3, specifically P4341, L2 – P4342, L2: These spectral "windows" (i.e. wave-length regions) could be helpful starting points for other novice workers, but I would disagree with recommending these spectral windows for "any other users." The best windows depend on RH, resolution, path length, concentration range, what interferences are present (it varies), instrument function, etc. and each study must optimize for their conditions to ensure good results. One way to test open-path systems is by using cal gases and permeation tubes with closed cell systems at similar path/concentrations and with an identical instrument function, etc as in Akagi et al (2013). Examples of deviating from the authors "recommended" windows follow. In Brazil, at H₂O mixing ratios of 2-3% and with a 100 m path, the region above 3500 cm⁻¹ can have too much water absorption for some detectors. Akagi et al., 2013 found excellent agreement with a suite of directly-introduced NIST-traceable CO/CO₂ standards that covered the range of field observations when using a spectral window from 2040 to 2100 cm⁻¹. That window keyed on the CO₂ feature at 2077 cm⁻¹, which is outside the CO₂ region in Table 1. The alternate region has lower H₂O interference and it out-performed the region in Table 1 in the Akagi et al tests for their specific application. It appears the N₂O retrievals may not have been challenged with calibration gases yet and Griffith et al. (1991) detected no significant N₂O enhancements in smoke with a longer path. The window from 920-1000 cm⁻¹ also contains CO₂ hatbands which can be prominent in smoke spectra though maybe not important in this work. There are other species sometimes analyzed for not in the author's windows. A useful QC check is that when a species appears in more than one window it should ideally have the same value in all windows. For instance, the CO₂ in the C₂H₂ window should agree with the CO₂ in the "main window" (first and last windows in authors table). This ensures that the interference is fit properly. Of course this requirement can be relaxed if the interfering peaks are very small compared to the target analyte. It's also sometimes more accurate to average multiple windows. For many reasons such as these cited above, I believe a lengthy discussion of windows is of limited value to the general readership and should be condensed to less than 5% of its current length. For research applications, it's important that each practitioner optimize for their specific application and not simply accept a prescription.

This section has been largely moved into an Appendix and these micro-windows are now recommended as a starting point rather than as a recipe with the paragraph:

“Optimised spectral windows are dependent upon many factors including path-length, spectral resolution, humidity and the concentration range of the species retrieved and any interfering species. The spectral regions and fitting parameters presented here would be a useful starting point for new groups employing open-path FTIR spectrometry to measure in similar highly polluted atmospheric environments. “

- 18)** Section 4. This could be too much information on the fires if the rest of the paper is not shortened. It could be described in a paragraph or put in a supplement. Or OK to keep all this detail if the preceding material is condensed.

This has been trimmed slightly by removing from the text information that can be found in the Table.

- 19)** Section 5.1. Now we are 22 pages into the paper and again presenting detailed methods as results. Eqns 7 and 8 apparently explicitly repeat earlier equations, which were already a bit much in my view. Sect 5.1-5.3 place a burden on the reader when most just want the mean, a realistic estimate of the uncertainty in simple tabular form, and then some interpretation.

Equations 7 and 8 have been removed and much of the discussion that relates to calculating uncertainties has been moved to Appendix B. We have placed this discussion in an Appendix rather than remove it or place it into Supplementary materials because we believe that an examination of the uncertainties inherent in the measurement technique is a worthwhile exercise (even when there are difficulties in doing it) and worth reporting in the paper. Many papers neglect to assess the measurement uncertainties (probably because it is fraught with difficulties) and so they are simply disguised amongst the natural variability of the emissions.

- 20)** P4350, L14-15: If most of the absorbing gas is in a small region of elevated temperature, then does it matter what the average temperature is? L26. I don't see a discussion of S:N:R, which is a common element of uncertainty. It's not 100% clear that all the known and unknown uncertainties are easily estimated, but even 20-30% uncertainty in many trace gases and more for some is very useful and if this is realistic for this experiment it could be communicated as a short, useful estimate rather than a lengthy, complex, and potentially over-optimistic discussion.

Signal-to-noise ratio is one component of the spectral fitting error. This is now stated explicitly. Most of this is now in the Appendix, and the point about temperature is also spelt out in more detail now:

“The temperature was measured at a single point close to the spectrometer but in reality the temperature may vary substantially across the spectrometer’s line of sight with a significant probability of increased temperatures where there are enhanced amounts of trace gases absorbing such as in the flaming emissions from the fire. However, flames are unlikely to cover the whole path-length and so we estimated that the temperature error was likely to be approximately 20°C. This estimate is meant to account for the possibility that a large temperature error for a fraction of the path may be more significant than a smaller temperature error over the entire path. (Very large errors in the temperature are expected to produce large errors in the spectral fits that would be easily identified).”

- 21)** P4351: The uncertainty in fuel carbon is estimated at 0.5 ± 0.05 , which is probably a 2-sigma uncertainty in view of published carbon analyses for forest fuels. Meanwhile, maybe the other uncertainties are one-sigma? Best to adopt and specify same "number of sigma" throughout. Also 10% may be high, but regardless the uncertainty in EFCO₂ is also 10%, which may be optimistic given the many difficult sources of error.

This uncertainty (like many others) is an approximate estimate. Studies where the fuel carbon is actually measured usually have 1-sigma uncertainties of ± 0.03 . In our study we are simply assuming that the fuel carbon will have been similar to previous studies (we have not measured it) and so we have estimated a significantly larger 1-sigma uncertainty..

- 22) L19:** Reason to switch from sum method here? It was never clear to me why the authors don't just get their best estimate of how the carbon is partitioned and then put that in the carbon mass balance method; once.

The reason for this is that it avoids biases introduced by assuming the wrong background values. This is clarified now when first mentioned:

“We have used the whole fire summation method to obtain our best estimate for the emission factors of CO₂ and CO, using the background mole fraction values for these gases measured before ignition of the fires (which can be measured relatively easily). For all other gases (where the background values are often closer to the quantitation limits of the measurement technique), we have used emission ratios via equation 5 and the emission factors for CO₂ and CO calculated for the fire via the summation method. This produces lower uncertainties than using the summation method for all gases because it does not rely on poorly defined background values for many of the trace gases.”

Also the emission ratios are useful in themselves -and can be compared to those measured in other ecosystems and those measured in smoke from Australian wildfires. Since the correlations are typically very strong, with either CO or with CO₂, these emissions ratios are not sensitive to sampling biases that can significantly influence the measured MCE and hence the emission factors.

- 23) L27:** "reson"

This typo has been corrected now.

- 24) P4352: L26-27:** Temperature affects density, but may also impact the relative line strength of different gases differently (via the Boltzmann factor), or different parts of the path may have different gases and different temperatures, so temperature errors may not completely cancel in ratios.

This is true. That is why the temperature uncertainty has been explicitly separated into a density effect (that will cancel out) and a “spectral” uncertainty that will almost certainly not cancel out and in fact (for the reasons that the reviewer points out) the way that this part of the temperature uncertainty will effect different gases is complex and difficult to predict, since the combined effect could partially cancel or could enhance the uncertainty. Given these difficulties in predicting the true uncertainty, we have taken the sensitivity values from Smith et al, or from the sensitivity of the line-strengths and made the approximation that adding in quadrature will give us a reasonable first order estimate of the combined uncertainty.

Given comments on the treatment of uncertainties in this paper by both reviewers, we have added a section that points out some of the difficulties in making good estimates

of the uncertainties, added a few relevant caveats to the approach we have chosen and moved the details of our method of uncertainty estimation to an Appendix.

- 25)** P4353, L21-26: Maybe not useful, but worth checking is the CH₄/H₂O window keyed on the CH₄ q-branch at 1306 cm⁻¹.

Yes – useful! Excellent suggestion because actually the two windows are consistent with one another (although with ~10% bias between them) for all the data *except* that they start to diverge in the Alford's Point data where we see our odd non-linearity. So we now believe that this is a spectral artefact caused by a temperature error. The 1306cm⁻¹ region shows much greater linearity with CO but simulations show this is also temperature sensitive. In the revised manuscript we have simply excluded the anomalous spectra and used the remaining spectra to obtain the Emission Factor for CH₄. We suspect that excessive temperature errors are also the cause of the anomaly in NH₃ and so we have taken the same approach for the NH₃ EF from Alford's Point (see comment 27 below).

- 26)** P4353, L28: Again, is assuming a single temperature relevant?

This sentence has been removed in the rewrite.

- 27)** P4354 and Fig 6: NH₃/CO lower for low MCE is counterintuitive since NH₃ is a smoldering compound. Might be fuel N and also worth checking if it is an artifact of not having CO₂ in the window.

We tried refitting with CO₂ also fitted in the window but this made no significant difference. However we also believe that this is caused by a spectral artefact because if we remove the spectra that show anomalous CH₄ (and therefore we believe are at significantly enhanced temperatures) then the remaining NH₃ emission ratio is constant and is more consistent with those measured at the other fires. This is the emission ratio that we have used in the revised manuscript.

The best way to deal with this temperature problem is to fit the temperature alongside the concentration of the trace gases absorbing. However (although this is theoretically possible) it is computer intensive since the cross-sections need to be recalculated at every iteration of the fit. For this reason it has never been implemented in MALT and so it is not possible to do this for this study.

- 28)** P4355, L13: NH₃ decreasing with aging is well-known and seen in Goode et al. (2000) and Akagi et al. (2012).

Akagi, S. K., Craven, J.S., Taylor, J. W., McMeeking, G. R., Yokelson, R.J., Burling, I.R., Urbanski, S. P., Wold, C. E., Seinfeld, J. H., Coe, H., Alvarado, M. J., and Weise, D. R.: Evolution of trace gases and particles emitted by a chaparral fire in California, Atmos. Chem. Phys., 12, 1397-1421, doi:10.5194/acp-12-1397-2012, 2012.

References added.

- 29)** P4355, L18: Has the 2005 value been corrected for the large change in HCOOH cross-

section in HITRAN? Factor of 2.

No – thank you – this point now added.

- 30)** P4355, L21-22: The Goode et al value should be adjusted to reflect the change in HCOOH cross-section since that paper was published as explained in Akagi et al., (2011). The Goode et al. (2000), Akagi et al. (2012), Yokelson et al. (2009), and others all show HCOOH/CO increasing with aging as measured by FTIR or CIMS and this is by now a well-known phenomenon.

Yokelson, R.J., Crouse, J.D., DeCarlo, P.F., Karl, T., Urbanski, S., Atlas, E., Campos, T., Shinozuka, Y., Kapustin, V., Clarke, A. D., Weinheimer, A., Knapp, D. J., Montzka, D. D., Holloway, J., Weibring, P., Flocke, F., Zheng, W., Toohey, D., Wennberg, P. O., Wiedinmyer, C., Mauldin, L., Fried, A., Richter, D., Walega, J., Jimenez, J. L., Adachi, K., Buseck, P. R., Hall, S. R., and Shetter, R.: Emissions from biomass burning in the Yucatan, *Atmos. Chem. Phys.*, 9, 5785-5812, doi:10.5194/acp-9-5785-2009, 2009.

This point is now clarified with the suggested references added.

- 31)** P4356, L3-4: The original HCOOH data in Andreae and Merlet 2001 is about a factor two too high due to the old incorrect HITRAN cross-section. The corrected values are in Akagi et al. (2011).

OK – this part removed in new discussion as per previous comment.

- 32)** P4356, L6: The citation to Akagi et al 2013 should be to Akagi et al 2014. Also Burling et al., 2011 (ACP) report a range of study-average EF_{HCOOH} of from 0.11- 0.57 in the same nominal ecosystem so a similar average but high variability illustrating the importance of a large sample size.

OK – corrected.

- 33)** P4356, L16: Rapid drop in C₂H₄/CO also seen in Akagi et al. (2012).

Added the sentence “Akagi et al, [2012] also report a rapid drop in C₂H₄ to CO ratios as the smoke ages.”

- 34)** P4357: LS-7: I agree that the variability in fuel carbon (at one-sigma) is less than assumed, but it is not clear how a comparison of standard deviations would suggest that. It seems the authors get about the same EFCO₂ every time because they assume the same ρ/C every time and the CO₂/CT is not varying a lot. CO₂/CT not varying is also essentially why the MCE are all similar. But if CO₂/CT was the same and the real fuel carbon was unknowingly only 25% you would get same EF.

This discussion point has been removed in the revised paper.

- 35)** P4357, L11-15: Not sure how you conclude the sample is representative from the uncertainties? That is often assessed by looking at sample size ("n") rather than standard deviation of mean. A low range of MCE for different geometries doesn't prove

representativeness if the real MCE's varied. It's interesting that the MCE from all "fire-side deployments of FTIR" from 1991 thru 2014 have almost the same MCE near 0.9 On the other hand, airborne, lab, and tower-based (e.g. Ward et al. 1992 JGR) measurements always show a large range of MCE. I'm not sure what that means, but perhaps worth discussing? Again, capturing some flaming and some smoldering is not necessarily the same as capturing the right relative amounts.

This has been rewritten in the revised paper and a discussion of MCE and how representative the measurements are likely to be has been added.

- 36)** P4358, 8-9: I think these data are valuable to use as is, or averaged with other available data, which may make more sense for some users. The choice could depend on several factors: such as model specificity or whether one chases to believe five data points represent real differences in Australia or just another data set to add to the pile and maybe shift the overall average. For instance, Akagi et al. (2013) report very different EF results than Burling et al. (2011) in the same nominal ecosystem using the same measurement approach, but one in spring and one in fall. If the authors repeat this experiment in other Australian temperate forests and get similar values, then more meaningful evidence for a regional difference is starting to accumulate. In any case, it's not apparent what the logic is here: the logic unintendedly comes across as "since we don't understand this we recommend you use it." I think the intent was to say the variability doesn't correlate with something like MCE. Normally fire EF correlate with MCE, but in this small study the MCE are essentially all the same so there is no way to isolate any MCE dependence. Finally, a lot of important gases are missing from this data set so literature values from elsewhere still need to be used for those.

This has been rewritten in the revised paper and a more measured recommendation made for the usefulness of the data.

- 37)** P4359, L10-16: It's great to recommend using this data, but I caution against recommending that the exact methodology be used in other future work. Future workers may find that different spectral windows or different analysis methods work better for their data.

See previous comment.

- 38)** P4360, L3-5: Most of the paper is about uncertainties, but they are dropped at the end of the text Standard scientific practice for many reasons is to include the uncertainties for any recommended values.

Uncertainties now included.

- 39)** References: randomly noticed that Yokelson et al 1999a and b appear to be same paper?

OK –this was a mess up with the endnote library – now sorted.

- 40)** Fig 4: Circular CO versus CO₂ plots like this are the norm for lab fires and it is just a time series starting with post-ignition flaming progressing up the low leg and then returning to origin via the high leg during smoldering. This indicates the ground-based system measured the history at a location, whereas airborne data in fresh

smoke above small hazard reduction burns is normally a series of similar well-mixed samples with gradually increasing CO/C02 ratios

Agreed. This is clarified in the text as per comment 4.