Review of Molecular Composition of Particulate Matter Emissions from Dung and Brushwood Burning Household Cookstoves in Haryana, India, by Fleming et al.

By Bob Yokelson

The authors have successfully identified numerous individual chemical species that absorb UV light in an important, but under-studied type of biomass burning aerosol (wood and dung cooking fires). The authentic cooking fire samples are difficult to obtain and the author's samples were subsequently analyzed off-line with a unique, extremely powerful array of coupled optical and mass-spectral techniques. The team has a great deal of relevant expertise in all aspects of the study and it includes leading practitioners of these advanced analysis techniques. The large body of data represents a significant investment of effort and the work should definitely be published. With some straightforward improvements the paper could be very good. I provide an overview followed by specific suggestions referenced to page and line number in order of appearance.

Overview:

More information is needed about the sampling, calculations, and error budget in the paper. A few additional sampling details and equations are given in the SI, but more are needed and should be in the main paper. Currently, the emission factors are said to be "orders of magnitude" upper limits, but it's proposed via limited comparisons that the MAC are "OK" – though they are given without an error estimate. The final MAC seem potentially "high" to me and their application should be clarified. Along with more detail on the sampling, the associated uncertainty in each step should be discussed near the beginning of the paper so the context and limitations are clear as the discussion is read, rather than reader wondering and only potentially finding out later in the paper.

I suggest a clear and commonly-used format would be to clarify how each step is done and discuss/estimate individually the different sources of error of each of these steps in order. For instance:

1. What were the inlet positions w.r.t. to the fires? Flaming produces a greater vertical velocity and flux of products than smoldering. Gas-particle partitioning is proportional to particle concentration and is also temperature dependent. Further from the source the smoke has diluted and cooled to some degree and vertical velocity doesn't need to be measured. A detailed diagram with distances, temperatures, concentrations (description of any dilution?), etc should be given in the paper.

2. What were the relative positions of the canister and filter inlets? Were they close or in well-mixed smoke?

3. What was the relative timing of the canister and filter sampling? Canisters tend to fill quickly at a non-constant rate while the filters were acquired over the whole fire.

4. When the results from two filters are coupled, what was the spatial and temporal overlap of the filter collection?

5. The authors evidently measured extraction efficiency once at 50%, but did not specify 50% of what (PM2.5, OA, etc)? Nor is it stated if the extraction efficiency is the same for all chromophores. Later in paper it's stated that the extraction efficiency was "lower sometimes." How much lower? Maybe give a range and consult some studies where extraction efficiency has been estimated by comparison to some familiar term e.g. OA measured by AMS or conventional OC analyses.

6. Ionization efficiency is maybe stated near the end of the paper to be higher for polar compounds, but some polar compounds that are ubiquitous in BBOA like levoglucosan were not seen in some samples. Can this be explained?

7. Given the realistic limitations of any technique, in the introduction or after the fleshed-out experimental/error section then maybe a brief summary of the strengths and weaknesses of various approaches to understanding aerosol optical properties is in order to provide context for readers. For instance, my understanding is that laser-based techniques would be well-suited for measuring the overall absorption of real aerosol (after drying) that contains BC, BrC, and other species at specific wavelengths, but non-power law features are an issue in fitting the cross-section at unmeasured wavelengths, they certainly cannot identify individual compounds, and they have some uncertainty in differentiating between BrC and coating effects (e.g. Pokhrel et al., 2017). Extractions (I think) eliminate BC and coating effects so that only BrC is probed. Following with off-line analysis by broadband UV absorption, retention times, and exact mass is very powerful for measuring the true BrC spectral shape and compound identification. But quantification and how the MAC relate to real-world aerosol is not clear from the paper now.

8. In light of above, provide some at least brief, rough guidance on how the overall optical properties of real BC-containing aerosol (coating effects and all) could be estimated from the extraction results that are presented. It seems possible using independently measured BC emissions, but I did not find that in the paper.

9. Make sure to specify what is being measured throughout (e.g. the MACs are absorption of what per mass of what?) and provide uncertainties.

10. BBOA almost certainly contains 1000's of species. The mass spectra obtained depend on variable detection limits and loading and natural variability makes the number of samples important. Thus, throughout the paper, I would refer to "observed" complexity, with actual complexity beyond the capabilities of current instruments.

11. Some new species are observed, but many species commonly observed by other techniques are missing some or all of the time. It would be helpful to clarify which off-line compound identification techniques access which types of chemical space well and, if possible, the relative mass contributions to total OA measured by various techniques. This could be a sentence or two in the paragraph recommended in point 7 above.

12. Tracers and markers are not the same thing and these terms are frequently misused. Tracers are emitted in a narrow, well-characterized ratio to the observable of interest and can be used to quantify impacts at receptor sites. Markers are useful but typically emitted in highly variable amounts by unique sources and useful qualitatively.

Comments on text in order of appearance:

P1, L13: "organic particles" should be "organic matter in particles" or "organic aerosol" to allow for the possibility of internally mixed particles.

P1, L20: "selected "extractable"" or similar seems appropriate before "compounds" or "numerous" instead of "selected"

P1, L21: The fact that many of these species are newly observed will make the reader curious if they are newly observed because they are relatively rare or because most of the organic matter in BBOA was previously un-speciated. If that question can be answered it would be great to do so in the paper somewhere. A related fine point is that if most of the observed compounds are newly seen despite the fact that numerous other compounds have already been seen in BBOA, then it is unlikely that this study probes the true molecular complexity. This is an empirical, observed molecular complexity impacted by detection limits and loading.

P1, L22: "stove-specific combustion conditions" might be better than "stove" if the stoves impact emissions by impacting the mix combustion processes?

P1, L23: "emission factor" and "observed molecular complexity"

P1, L25-7: These MACs seem ~2x too high compared to other studies if they are referenced to g of PM2.5, especially as a lower limit? And it should be clear what they are for; real aerosol with BC included or just the BrC? If BC is added back to represent real aerosols, how is that done?

P2, L12: "depends to some extent" since amount of PM and individual susceptibility to various toxins matter a lot.

P2, L20: lifetime of BrC also important.

P2, L24: probably don't need same reference twice?

P2, L25: "a photoacoustic spectrometer" should be "photoacoustic extinctiometers (PAX)" (throughout) and "895" should be "870", L26: "cook" > "cooking"

P2, L26-29: The EFs for BrC quoted from Stockwell et al are actually their EF Babs at 405 nm (not the same thing as explained next). These include absorption by BC and BrC. Different EFabs, tentatively for just the BrC are also provided though.

Some relevant background on Stockwell et al 2016 from the corresponding author:

Lack and Langridge (2013, Table 1) recommended an MAC for “BrC,” but actually meant an MAC referenced to the mass of “BrC-containing OA.” They clarified that their MAC was an average value for BBOA but that the MAC can range a lot. The OA MAC was later found to depend on BC/OA by Saleh et al 2014. We used the concept of an EF for “BrC” based on the Lack and Langridge MAC in Stockwell et al, but have since abandoned that terminology as we think it is too easy to misinterpret. Meanwhile, in the Stockwell paper cited here we tried to present qualifying text as follows: “The BrC mass calculated this way is considered roughly equivalent to the total organic aerosol (OA) mass, which as a whole weakly absorbs UV light, and not the mass of the actual chromophores. The MAC of bulk OA varies substantially and the BrC mass we calculate with the single average MAC that we used is only qualitatively similar to bulk OA mass for “average” aerosol and even less similar to bulk OA for non-average aerosol (Saleh et al., 2014). The BrC mass estimated by PAX in this way was independently sampled and worth reporting, but the filters and mAMS provide additional samples of the mass of organic aerosol emissions that have lower per-sample uncertainty for mass. Most importantly, the optical properties from the PAX (SSA, AAE, and absorption EFs calculated as detailed below) are not impacted by MAC variability or filter artifacts.”

Thus, our EF BC and the values listed just above are the best to compare to in Stockwell et al. Thanks to the authors, we rechecked our cooking fire table, found one error that only impacts the SSAs, and have posted a corrigendum. (We discovered that the SSA labels in Table 4, but not Table S8 had been reversed: it should be SSA 870 and then SSA 405 below that in Tab 4.)

Then finally, the more robust estimates of speciated PM2.5 mass from the same study that we alluded to in Stockwell et al., 2016 can now be found in Jayarathne et al. (2017) and are useful for comparisons with this work as will be pointed out.

Jayarathne, T., Stockwell, C. E., Bhave, P. V., Praveen, P. S., Rathnayake, C. M., Islam, Md. R., Panday, A. K., Adhikari, S., Maharjan, R., Goetz, J. D., DeCarlo, P. F., Saikawa, E., Yokelson, R. J., and Stone, E. A.: Nepal Ambient Monitoring and Source Testing Experiment (NAMaSTE): Emissions of particulate matter from wood and dung cooking fires, garbage and crop residue burning, brick kilns, and other sources, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-510, in review, 2017.

P2, L31-32: Just an observation that the dung is expected to have lower MAC than wood consistent with lower BC/OA per Saleh et al 2014.

P3, L13: Should “The” which could imply “all” be “Many”?

L17-19: This second half of the sentence starting with "while" is unclear in and also ceanothus is from US NW.

P3, L21: N accounts for a small mass fraction of BB PM2.5 so N-containing organics are likely a small fraction of the total BBOA?

P3, L25: eliminate "The most" since that will quickly be dated?

P3, L32: "more" or "additional"?

P4, L4-5: I'm sure this is worthwhile data regardless if it is the first or last data, but it seems unlikely this would be the first detailed study of brushwood or dung smoke since these sources have been studied intensively for > twenty years. A recent example that contains examples of historical references is Jayarathne et al. (2017). Simoneit et al have been characterizing smoke for > 20 years.

P4, L23-30: The pictures are useful, but Fig 1A currently shows a concoction of unidentified tubing and a diagram is also needed. Questions arise with some also in the overview. How did the authors ensure representative sampling of well-mixed emissions for all devices? For instance, was the data corrected for the different vertical velocity in the smoke column above the fire? This is important because during flaming, the flux of emissions can be much greater than during smoldering. What were the flow rates and residence times in inlets, was any dilution used, were the downstream filters side by side or in series, were the pumps downstream of the filters, how were filters stored during the 25-30 hours not at -80 C (in a cooler with dry or blue ice or at ambient T, what was ambient T if relevant?), were backgrounds or field blanks taken, error in gravimetric analysis, etc?

P4, L24: "BBOA" > "PM2.5"

P5, L1-9: The next step after filter collection and before mass spec is filter extraction, i.e. how was extraction done, what is extraction efficiency compared to total BBOA and is it the same for all the species detected, etc? As written it sounds as if the filter is inside a capillary. It may not be possible to estimate extraction efficiency if done by a droplet flowing over the filter surface? But the bottom line should be clear.

P5, L6-7: Why positive mode and are there species only seen in negative ion mode?

P5, L7 What are MRFA and Ultramark, what is the relevance of the calibration species, how were the cal results applied?

P5, L8-10: Filter deposits may not be uniform. What percent of peaks were seen in only one sample or had S:N < 3. This is useful context that helps relate reported complexity to observed complexity, which is itself a subset of actual complexity.

P5, L12: Does Kendrick analysis with CH2 and "H2" base units help with O-containing species?

P5, L13: “mass-calibrated”

P5, L16: About what percent of signal or number of peaks was above m/z 350?

P5, L20: K is normally the most abundant alkali metal in BB-PM (not Na). S and Cl can be high in dung (Hosseini et al and numerous other papers). Brief explanation of why not included?

P5, L22: No “O” in the formula, but it is in the explanation?

P5, L27: Why is the extraction solvent mixture different here, i.e. not containing water? What fraction of mass is extracted and of what types of compounds (in summary form)?

P6, L16: It should be specified that this is an MAC for the extracted BrC only. Are the units (cm) consistent with the reported MAC (m)? “Cmass” is the “solution mass concentration” of what and how measured?

P6, L17-18: Another filter collected where and when? Here is an example why the spatial and temporal overlap of the filter collection is important to describe in experimental section. How are PM2.5 on the other filter and “Cmass” connected?

P6, L20: So is the target mass reference for the MAC PM2.5 mass then?

P6, L20-21: How do they do know the mass extraction efficiency was < 50% sometimes and why would a lower mass extraction mean the MACs are a lower limit? What if the extraction got all the chromophores, but only half the total mass, then would the raw MAC be a factor of two high? It seems the impact on the MAC would depend on the relative extraction efficiency of the chromophores and other constituents.

P6, L21: Again it seems the AAE are for the BrC component only? Can the authors add a sentence on how their AAE and MAC can be adjusted to represent real aerosol that also contains BC?

P6, L27: Abundance is usually used to indicate how many there are of an item. Peak area or height is usually used to estimate the signal strength or amount of compound. Should “abundance” be replaced by “area” or something else?

P6, L27-29: So if I understand this and the SI right, the EFPM(total) was estimated separately (in a sparsely described experiment) and then EFPM(total) was partitioned among the peaks observed on the MS according their relative intensity. If so this should be stated as a sentence in the main text. It’s a concern that the real EF could be ten times or even several “orders” of magnitude lower. Should the authors reconsider even reporting EF directly as such? Maybe it’s safer to label them as “upper limit EF in the tables and figures to spare potential future misinterpretation?

Also, can the error be reduced or the error budget be tightened up? For instance, the authors may detect some species that have been better quantified from these sources in other studies? Can they use their ratios to any overlap compound with the extensive quantitative analyses reported in Jayarathne et al 2017 and many others? Ratios to levoglucosan or PAHs for instance may be helpful to constrain the "EFs" to a realistic range?

P6, L28-30 and Figure 2 comments: It's true that authors show non-identical spectra from the sources even though the two dung spectra look pretty similar. But proof of distinct signatures requires enough samples to quantify the variability in each source. Right now it's not clear if the method uncertainty is larger or smaller than natural variability or what the observed variability is. The authors should try to characterize observed source to source variation with some metric (# or % of unique peaks) and this is done to some extent later in the paper, which is good. As noted above, the y-axis label should be "Upper limit EF" or "approx. relative abundance" Relevant to earlier comments, the spectra are too simple to represent all the components of BBOA although more details might be revealed with a log scale.

P7, L8: Dung has much higher N-content than wood and that makes excellent sense based on known plant and animal physiology. The Gautam reference has much higher N than normal for wood. See Stockwell et al 2016 for dung and Coggon et al Fig 3 for wood or many other sources referenced in these papers.

P7, L8-11: This sentence doesn't quite make sense. Typo? Punctuation?

P7, L12: Could provide a few key citations on lignin pyrolysis – there are many. It would be interesting if the authors could show that the cellulose preferentially ends up in the gas phase? But maybe the method has low sensitivity for sugars (from cellulose), which are usually abundant in BBOA (Christian et al., 2010; Jayarathne et al., 2017)?

P7, L 14: Dung and embedded grasses are both high in Cl content (Stockwell et al., 2016 and references there-in, especially Lobert et al., 1999) so expect high Cl from cooking with dung and ag residues as in Stockwell et al. 2014-2015 ACP papers.

P7, L15-16: Clarify if these large peaks were included when the EF was partitioned?

P7, L21-26: K is well-known to be enhanced in biomass (Table 1 in Hosseini et al., 2013) and K has a very long history as a biomass burning "tracer" (e.g. Sullivan et al., 2014 and references there-in). It is well-known that K is primarily emitted by flaming while levoglucosan is primarily emitted by smoldering. If the stoves had different flaming/smoldering ratios that could explain variability in K production.

Hosseini, S., Urbanski, S., Dixit, P., Li, Q., Burling, I., Yokelson, R., Johnson, T., Shrivastava, M. K., Jung, H., Weise, D., Miller, W., and Cocker III, D.: Laboratory characterization of PM

emissions from combustion of wildland biomass fuels, J. Geophys. Res., 118, 9914–9929, doi:10.1002/jgrd.50481, 2013.

Sullivan, A. P., May, A. A., Lee, T., McMeeking, G. R., Kreidenweis, S. M., Akagi, S. K., Yokelson, R. J., Urbanski, S. P., and Collett Jr., J. L.: Airborne characterization of smoke marker ratios from prescribed burning, Atmos. Chem. Phys., 14, 10535-10545, doi:10.5194/acp-14-10535-2014, 2014.

P7, L27-30: There is a difference between a tracer and a marker and levoglucosan (LG, a cellulose pyrolysis product) is the latter. If LG was a good tracer then it would be in all the samples and at reproducible amounts. LG is in fact normally found to be a major, but variable component of BBOA (Sullivan et al 2014; Jayarathne et al 2017; Christian et al 2010). If LG is missing from many of the samples that suggests detection issues that could cause some of the BBOA species to escape the analysis procedures used.

P7, L31 - P8, L2: There were more than three spectra of each stove/fuel combo, but three were chosen how? Then peaks not appearing in all three spectra of a combo were discarded why? Then the remaining peaks were scaled and the spectra for each stove/fuel combo were averaged together. I think that is the right order, which may be jumbled in the text?

P8, L6-12: First compounds found from all three cooking types are listed. Then compounds only found from brushwood are discussed. That I can follow. Then the compounds that are found in all dung cooking. But not unique to dung cooking? It’s not clear here what the difference between sections 3.5.1 and 3.5.2 is. I.e. what is the difference between “common to” and “detected in all”?

P8, L11-12: If the stove material caused the emissions it potentially might not matter what the fuel is since all fires heat the stove.

P8, L14-17: The true number of constituents for all the PM types is much greater than observed so maybe just say ~we saw more peaks from A than B.

P8, L15-30: Can this overview of the results section be re-phrased or re-organized to make it easier to follow?

P9, L4-5: “detected elemental” and the relative insensitivity for sugars, or anything else needs to be discussed earlier – maybe in the introduction or at latest the experimental section.

P9, L6-8: I don’t believe the %N in Gautam et al, it goes against all the other studies I’ve seen dating back to Susott et al., 1996, unless Gautam et al included the foliage with their wood.

P9, L17: All biomass is ~25% lignin, 25% hemicellulose, and ~50% cellulose polymers though the monomer units differ.

P9, L18: delete “at”

P9, L18-19: by “commonly detected” do they mean by their group or are there references to other groups? There are numerous studies that characterized BBOA.

P9, L20-21: “found” to “observed” better?

P9, L25: “coniferyl” alcohol and that comes only from conifers whereas the others are unique to hardwoods or grasses, which are probably more relevant in India.

P9, L30: In the experimental overview, the range of ionization efficiencies could be provided?

P9, L31: Jayarathne et al also discuss species unique to dung

P10, L1: important to qualify “the observed chemical …was far more complex” etc

P10, L2: By “reproducibly” it means “seen” in all “n” samples, but not in the same ratio to total PM2.5? That should be clear if so.

P10, L5-6: This is a little hard to follow as the authors use “found in all dung …” in the section header, then next mention “detected exclusively” in one type or another, and then “combine all…” Maybe change “Hereafter” to “However” would help with the transition?

P10, L10 - P11, L6; Figures 5 -7: Jayarathne et al 2017 and references there-in quantified numerous PAHs for similar fuels, which may be helpful to compare to.

P11, L8: At the outset useful to state what these MACs represent: absorption due to extractables only per PM2.5 mass? I think the mass reference is not extract mass or mass of chromophores. It’s important in the larger context if they are MAC that don’t include the BC component or any un-extracted chromophores.

P11, L8-11: “browner” OA from wood cooking makes sense empirically given the higher EC or BC to OA ratio for these fires per Saleh et al., 2014, Jayarathne et al., 2017. The latter reference also reports higher PM emissions from dung cooking than wood cooking in agreement with cited previous work. Their measured and cited EFPM2.5 from South Asia for wood and dung cooking may be useful to compare to.

P11, L12: The MAC is the absorption coefficient per mass of PM. In order to convey the absorption per unit fuel consumption, “coefficient by” should probably be replaced by “emission factor of”

P11, L 11-13: On the overall absorption per unit fuel consumption. This was evaluated by Stockwell et al., (2016) for particles containing BC and BrC and, with higher uncertainty, for just the BrC component, with the following results (in m^2/kg) (variability also shown in reference).

	Wood	Dung
EF Babs 405	10.6	5.85

EF Babs-405-BrC	8.40	5.43
EF Babs 870	1.04	.197

Wood-burning aerosol absorbed about 5 x more per kg burned at 870 because of the higher BC emissions, but just 2 x more at 405 for the overall particles. More BrC absorption per unit fuel consumption was observed on average for wood, but within variability the amounts overlapped. Can the authors use BC/EC EFs for these fire types to get their own estimates for absorption EFs for real fires?

P11, L14-21: Throughout this section, for this work and other work, the mass reference (PM2.5, OA, etc?) should be rechecked and specified and uncertainties for the MACs should be provided.

It makes sense that the brushwood MAC is larger than the dung MAC given the BC/OA dependence of MAC described by Saleh et al. (2014) as noted above.

At a cursory glance, it seems like the MAC values at ~400 nm that the authors selected for comparisons are on average higher than values in Lack and Langridge, the extensive tables in Olson et al 2015, or Bluvshtein et al., (2017). It could be helpful if the authors could include these studies in their discussion and comment on any implications there may be.

Olson, M. R., Garcia, M. V., Robinson, M. A., Van Rooy, P., Dietenberger, M. A., Bergin, M., and Schauer, J. J.: Investigation of black and brown carbon multiple-wavelength-dependent light absorption from biomass and fossil fuel combustion source emissions, J. Geophys. Res.: Atmos., 120, 6682-6697, doi:10.1002/2014JD022970, 2015.

P11, L22-26: If my understanding is correct, one advantage of extraction techniques is that the BC is eliminated along with uncertainties in BrC attribution due to BC coatings or AAE. If that is right, the authors should clarify this is an AAE for the BrC only. Also, the AAEs can depend on the choice of wavelengths fit with a power law and this study has much better wavelength coverage than just the 2-3 wavelengths commonly used in optical in-situ approaches. The overall AAE with BC included is also important to describe real in-situ aerosol. The AAE with BC included may be lower since the AAE of BC is near 1. Stockwell et al obtained AAEs of 3 and 4.6 for wood and dung, respectively using an optical in-situ approach on aerosol containing BC. Can the authors estimate overall absorption values for real intact BC-containing aerosol from their extract values? It seems doable using EFBC and BC optical properties.

P11, L25: Why is BC mentioned? Per above, isn't it eliminated in the extraction process? It's not clear why the authors say the following: "However, the observed absorption can be definitively attributed to BrC since AAE values of 2 or greater indicate that light absorption comes from BrC as opposed to BC (Kirchstetter et al., 2004; Laskin et al., 2015)." Pure uncoated BC has an AAE near 1, but any aerosol AAE can have some contribution from BC if BC is present. I think the definitive attribution comes from the fact that BC and BC coating effects are eliminated in the extraction.

P11, L27: Move first sentence to P12, L17?

P12, L4: “lignin-derived”?

P12: L3-16: This is great stuff. I wonder if retention times were measured for standards to support identifications.

P12: L22-28: On line 24, why were the early eluters ignored? Combining line 22 and line 26, is the conclusion that polar compounds tend to be detected more efficiently, but elute earlier and may be ignored? Minor tweaks to text could likely clarify this section.

P12, L29 – P13, L2: It’s neat that absorption features can be used with mass and retention times to support compound identification. In the example given, the results were inconclusive. When two components contribute to the observed absorption one could theoretically resolve that if the absolute cross-sections are known. Figure 10 legend and trace colors should be consistent. I don’t see a red trace.

P12, L33 “that peaks”

P13, L15: “observed chemical complexity” – in general measuring true complexity is likely beyond scope of any one study? The main benefit of this study is a wealth of new chemical information and tying that to absorption. For the former point, if not already done, it might be worth flagging which peaks are new. Since many commonly observed species were not seen, but some new ones were, perhaps a good topic for the conclusions is if this approach occupies a unique niche in “chemical space”?

P13, L31-32: This could be taken as: this study is the first to see non-lignin-derived entities in BBOA? Seems unlikely, clarify?

On SI:

Position/timing of cans is not clarified, etc.