Review of Jayarathne et al, Nepal Ambient Monitoring and Source Testing Experiment.

Summary

This paper describes particulate matter and its composition from several sources in the Kathmandu Valley of Nepal. In this paper, “composition” includes elemental and organic carbon, water-soluble inorganic ions and metals, and single organic species useful for speciation. The information presented here is relevant for atmospheric chemistry and will serve to provide inputs to atmospheric models. The measurement methods are competent and consistent with the state of the science. The paper is well written and the organization is clear. I support publication of this work after attention to some of the issues raised here.

Major issues.

The authors have gathered a lot of information in this measurement campaign, NAMASTE. They have chosen to publish several papers and to organize those papers by measurement type (gas phase in a different paper also published in Stockwell et al 2016, ACP- 16-11043-2016; particles in this paper, other papers promised.) This arrangement seems unavoidable or at least I can’t think of a better one. I understand the need to divide the information into multiple presentations for tractability. However, I find that it leads to a somewhat haphazard feel and some repetition as each paper walks through a number of different emission sources and yet doesn’t provide an overall integrated understanding of any single emission source. In this review I have some comments on the integration of this paper with the earlier paper. Although the earlier paper is already published, I hope that these comments can be useful to frame this and future publications.

Another major issue is the very small number of samples for each source and the implication, given in the rationale, that these are representative of South Asia. If sources are different worldwide, then one might not expect sources in Kathmandu to be similar to those in India—why should they be? I understand that source testing always provides data from a small number of sources, relative to the total population. But the sample size and representativeness has to be discussed in the context of natural variability within the source population, and (if known) the causes of that variability. The paper has a good discussion of why the source categories were chosen, but hardly any discussion of why the individual units were selected or what they represent. This discussion could be improved. For example:

- two generators (one gasoline, one diesel?) are described only as "old" and "a size that is commonly used"—what does this designation entail? What power output and capacity factor? I assume these are four stroke engines but it's not stated.
- two diesel groundwater pumps: again what size? How were they operated? (This information is given later in the paper; should be in the Methods)
- motorcycles are said to be different in Kathmandu because of "steep gradients, congested traffic, low vehicle speeds, high altitude, and frequent re-starting" yet these motorcycles were measured at idle, capturing only the altitude—why? Are these two stroke or four stroke engines?

- biofuel stoves "brought a pot of water to boil"; is this Water Boiling Test with hot start, cold start simmer or is it a different sequence?

In most of the descriptions no mention is made of the operating conditions: power levels, acceleration or steady state, fuel quality, analysis, moisture content (for solid fuel), sulfur content. For some sources, emissions vary during the course of operation, such as garbage burning, field burning, kiln operation. It's not stated whether the emissions were measured from beginning to end—probably not—or whether a fraction of the time was measured, which fraction was selected and why. The exception is "5 hours" for the brick kiln. How long were the samples? These factors affect emissions, representativeness, and comparability. When papers are given on individual sources this information would be expected. The presentation of just a small number of many different sources doesn't relieve this responsibility. The information may be in the earlier paper, but I would rather see repetition of this important data, rather than a repetition of the reasons for sampling.

Finally, there could be more comparison with constraints. For example authors have both EC measurement (thermal optical) and BC measurement (PAX), but these are never compared except when one of them does not yield a result. Since both are employed to infer model inputs, this comparison should be discussed. There is a good discussion of this only for the pump results, and that one is rather inconclusive. Likewise, there are measurements of both SO2 and SO4, which should be possible to compare with fuel sulfur content.

Specific comments.

Page 7 Sample collection. How were the capture points of the probes aligned and how were they chosen? How was it ensured that a representative portion of the plume was captured?

Page 10 Elemental and organic carbon "adjusted for positive sampling artifacts." How was this done? Denuders or subtracting loading of quartz filters behind Teflon filters?

Page 10-11 Field blanks were subtracted, but there isn't a mention of how large the field blank correction was. Is it significant relevant to average concentrations—especially for individual organic species? This could be a measure of contamination under the challenging field conditions.

I don’t have many comments on the chemical measurement methods. They seem competent. The carbon-balance method is commonly used for sources where plumes are hard to capture.

Zig Zag Kiln, Comparison with previous measurement (Weyant and Christian): This seems important because these are the only 3 measurements existing, to the best of my knowledge, and the results seem very different in this study. This paper states that “measurements were sampled within the stack at higher temperatures” but the Weyant paper described dilution to cool the sample stream before measuring, not measuring at stack temperature. Does the reasoning still apply? If the other measurements were cooled and diluted (but not diluted to ambient concentration), then in fact the gas-to-particle partitioning (without chemical conversion) would

favor higher emission factors from the Weyant measurements, wouldn't it? "2-3 m downwind" (page 7) doesn't allow a lot of time for cooling. Authors finally compare the total carbon measurements as similar (for all measurements: Christian, Weyant and these) and attribute the difference in this study to sulfur and bound water. So it seems that there is a second hypothesis, difference is likely due to the SO_2 conversion to SO_4.

Zig Zag Kiln, Sulfate and Water: The sulfur appears very important in the emission factor. Is the emitted sulfur (gas and particle) consistent with the amount of sulfur in the fuel? Also, is reporting bound water consistent with other measures of PM emission? For use in modeling, the report of bound water would seem to give a high bias for atmospheric PM concentrations since the models also account for water uptake.

Zig Zag Kiln, "EC was not detected by thermal-optical analysis, and thus the optically-determined EFBC at 0.112 g kg^{-1} for this source (Stockwell et al., 2016) is recommended to estimate the soot component of the smoke." This statement seems arbitrary, unless authors can show that this level of EC would be undetectable with this method.

Clamp Kiln, composition and closure: Since the clamp kiln PM had 20% sulfate, why is the particle bound water not also contributing to an underestimate of total PM when reconstructing total from the species? It seems this PM behaves quite different from the zig zag emitted PM.

Low EF of levoglucosan (page 16) Is the fraction of PM (not absolute emission factor) also compared to wood?

Garbage burning: Comparison between EC emission factor and PAX based BC emission factor. They seem very different. 7.4 g/kg (PM) x 2.6% EC = 0.19 g/kg. Whereas PAX BC is 0.56 g/kg (wet) or 6 g/kg (dry). This should be explained.

Garbage burning: emissions of PAH. Again totals are given in mg/kg, and it would also be useful to identify whether the PAH/PM ratio is relatively high.

Diesel generator: EFPM are compared with US EPA, but "generator" is not a unique class. There is a lot of speculation in this discussion, e.g. that generators would have lower emissions if better maintained.

Diesel generator, composition: I am surprised by the low EC content, particularly since there appears to be little oil contribution to the PM. In that case where does the OC come from? Authors cite another study that finds mostly OC in emissions for a high-sulfur fuel. Was no sulfur detected in this PM?

Pumps OC-BC split discussion. This discussion is not strong and leads to question about the work presented. Information is presented from AMS data which haven't been published yet. A method is used to divide the PM that is not discussed in the methods. Authors point out that the different measurement methods were measuring at different times and over different conditions (e.g. another method included start-up and high black smoke emission while this one does not.) They then propose applying composition from other measurements to these emission factors

after having just explained that the emissions were different. One gets the impression that the other measurements are better and these shouldn't even have been reported. Perhaps this impression could be improved with a better presentation.

Motorcycles: Although the motorcycles were measured only during idle, interesting results about the change in PM emission and composition with servicing are presented here. It is stated that results are compared with start-up emissions. Why is that condition comparable with idle emission?

However, it seems unlikely that the emissions profile here is representative, since only idle was included.

Biofuels: Authors find that (1) field samples EF are higher than previous reported EF, (2) PM2.5, OC, and EC were not significantly different between the field and laboratory samples. Does this mean their laboratory EF are higher than previous lab EF? Is the finding of no significant difference between lab and field the same for the other reported species?

Although authors have discussed some previous literature, they have missed comparison with some other studies. The following 3 studies measured many types of household stoves in laboratory setting. Are the relative comparisons in this study (which stoves are better) similar?

Smith, K. R., R. Uma, V. V. N. Kishore, J. Zhang, V. Joshi, and M. A. K. Khalil (2000), Greenhouse implications of household stoves: An analysis for India, *Ann. Rev. Energy Environ.*, *25*, 741-763.
Jetter, J. J., and P. Kariher (2009), Solid-fuel household cook stoves: Characterization of performance and emissions, *Biomass & Bioenergy*, *33*(2), 294-305, 10.1016/j.biombioe.2008.05.014. –
Jetter, J., Y. X. Zhao, K. R. Smith, B. Khan, T. Yelverton, P. DeCarlo, and M. D. Hays (2012), Pollutant Emissions and Energy Efficiency under Controlled Conditions for Household Biomass Cookstoves and Implications for Metrics Useful in Setting International Test Standards, *Environ. Sci. Technol.*, *46*(19), 10827-10834, 10.1021/es301693f.

The following 2 studies compared laboratory and field emissions. Are the comparisons similar to those found here?
Johnson, M., R. Edwards, C. A. Frenk, and O. Masera (2008), In-field greenhouse gas emissions from cookstoves in rural Mexican households, *Atmos. Env.*, *42*, 1206-1222.

Roden, C. A., T. C. Bond, S. Conway, A. B. O. Pinel, N. MacCarty, and D. Still (2009), Laboratory and field investigations of particulate and carbon monoxide emissions from traditional and improved cookstoves, *Atmos. Env.*, *43*, 1170-1181.

Heating fire Page 29: it is stated that this single source profile (one measurement) is "representative of open co-burning of dung and fuel wood under smoldering conditions in the Tarai." This seems like a strong statement without any support. It is also stated that "the high OC:EC ratio (~150) is also characteristic of smoldering combustion conditions" – it seems that a large number of high OC:EC ratios could be "characteristic" of smoldering conditions. I think

that authors should be more careful of using the words "representative" and "characteristic" without being able to support the use of those words.

Uncertainties: Throughout the paper uncertainties are given. However I began to wonder what these uncertainties represent. Are they uncertainties in method, obtained from field blanks, and do they also represent natural variability among members of a source class? I began to suspect that the natural variability was not represented when an uncertainty was also given for the single heating fire. This should be clarified.

Conclusion: This paper is rather long and it is sometimes difficult to extract the authors' contributions to the field. The paper concludes with a statement that the measurements will be useful. I think it would be very helpful to the reader for the authors to make a short list of the specific new information. A few examples: (1) I don't know other measurements of groundwater pump emissions, and so they could indicate that these are some of the first measurements. (2) In other cases the authors added to the database of emission factors for total PM, e.g. for kilns, or stoves. They could say that this is an addition, whether the measurements are higher, or lower. (3) There were several discussions of source markers; some were identified as unique markers and some were dismissed. These could be summarized.

Editorial comments

"Terai" is misspelled throughout

page 19 "a bit damp" is not professional language

Title 3.7 "of" should be removed