ARCTAS characterized the trace gas composition of the biomass burning impacted boreal atmosphere in NH summer 2008. The total number of pyrogenic gases measured during ARCTAS greatly exceeded that in previous field campaigns in the boreal region with the instrumental improvements since NASA’s last boreal forest fire mission (ABLE) being nothing short of spectacular. The measurements covered an extensive geographic range and are of high accuracy. The data should definitely be published. This paper is one of at least four papers that summarized the trace gas data from multiple instruments on the DC-8. All were multi-investigator, but none were comprehensive in including all the data. In Simpson et al the fresh plumes without signs of aging were carefully selected and used to compute initial emission ratios (ER) and emission factors (EF) as model input. They unfortunately did not include the TOGA gases though. In contrast, this paper presents the average normalized excess mixing ratios (NEMR) observed for each non-methane organic compound (NMOC) to CO (for fresh and aged plumes) over the course of the campaign, which depends on the average plume age and species lifetime. Because of the extensive geographic range sampled, these average NEMR give some idea of how polluted the boreal atmosphere was due to fresh or aged smoke. This paper however only includes a few of the WAS gases and not the HCHO. It’s unclear to me if all the PTR-MS data was published or how much overlap the above papers have with papers by Singh et al and Vay et al. Since outsiders are too busy to read all the papers, my review will reflect that knowledge gap and it would be great if future missions could strive to produce comprehensive papers organized by themes.

I have some general comments on a few concerns I have with the paper, followed by a summary of the many things I like. I should state at the outset, that most of my concerns are largely an artifact of the flight planning, which was likely beyond the authors control. I discuss these concerns at some length, not to lecture the present authors, but hopefully to impress upon the entire ARCTAS team that these issues may need to be acknowledged throughout all the ARCTAS publications.
It has been clear since the beginning of biomass burning research that it is highly variable – not unlike many natural phenomena. The highest variability may be seen during a single fire as a function of time because the flaming to smoldering ratio evolves causing high variability in individual samples or “intercepted plumes” as demonstrated at least all the way back to Lobert et al 1991 and Yokelson et al 1996 in lab studies. Real fires also normally flare-up and die-down as they burn in response to many different environmental variables, which changes their flaming/smoldering ratio. In fact, as the figures in Yokelson et al., (2008) show, the full range of variability cannot even be measured from a single platform in the field, with most of the range only being accessible to ground-based sampling.

If multiple samples of a single fire are acquired and one estimates fire average values, then one can examine fire-fire variation. Most papers provide the average EF for NMOC and the 1-sigma uncertainty in the mean, often in a vegetation classification scheme. The variability is often prominently displayed as a function of MCE or the CO/CO2 ratio, which are both proxies for the flaming to smoldering ratio. This can be seen in Figure 3 of Yokelson et al., (2009), but has been ubiquitous in the fire literature since the early 1990s. Reporting the mean and uncertainty for a group of fire-average values is not done to be “variability deniers,” but simply to provide the data in compact, convenient form. It's widely assumed that, in large-scale applications, the modelers will probably want to apply an average EF value to all detected fires in an ecosystem rather than having to generate a distribution of EF about the mean and applying that distribution of EF to a set of hotspots, or whatever. Based on standard statistical theory, it’s understood that 95% of the observations would likely fall within 2 standard deviations of the mean. Since the average literature value for one standard deviation is typically ~40% for an NMOC, the range would usually be at least +80% or a factor of 9. However despite this high variability there are reproducible ecosystem effects on average values. E.g. taking CH3OH as a generic smoldering NMOC, the study-average CH3OH EF for savanna fires is lower than the study average CH3OH EF for tropical deforestation fires. Since vegetation maps are ubiquitous, this is an accessible approach to increase the spatial representativeness of the model. But in some applications, such as modeling individual plumes, the limitations of using average values may be severe and should be acknowledged.

Finally some work has looked at the variation in study average values for individual ecosystems. The compendium of Andrea and Merlet (2001), and the updated compendium of Akagi et al (2011a) both report high average study variability and even high variability study to study in some cases.

This variability discussed above means that in studies that randomly sample different plumes of different ages, the photochemical changes would be expected to be masked to a large degree by the inherent variability. Even in studying a single plume, the photochemistry can be impossible to quantify if the source is changing. In particular, an aircraft is much faster than the wind, so sequential (long axis) sampling, if e.g. started at the source and then pursued downwind is likely to compare source smoke to downwind smoke emitted much earlier. This can mask the photochemical changes unless it can be shown that the source was burning in a fairly constant manner as in Yokelson et al 2009, or if the sampling is Lagrangian or pseudo Lagrangian as in Hobbs et al 2003 or Akagi et al 2011b.

There is an additional problem that as plumes evolve the variability may increase. In any case, the NEMRs can be significantly distorted due to changes in the background air that the plume dilutes with. One can easily demonstrate in a spreadsheet that diluting a plume with constant background air will not effect the dX/dY values, the NEMRs in the plume. But, if X/Y changes in the entrained background air, then dX/dY no longer reflects the original source characteristics. This is seen in a few of the plumes in the present work that have non-physical MCE that are > 1.

The uncertainty in sample ages is very significant and requires some subtle analysis. In the best of circumstances, in a single plume the 1-sigma variability in the windspeed is usually ~50% of the mean. When the plume source is physically located and sam-
samples in an isolated plume are aged based on distance downwind and windspeed, the uncertainty in age is quite high due to the uncertainty in the windspeed. However, because there are unlikely to be any “passing lanes” in a single smoke plume, at least the age order corresponds to the distance order and so the relative ages of samples are well-known. When sampling different plumes this advantage is lost. A plume thought to be 4 h old may only be 2 h old due to uncertainty in windspeed and a different plume thought to be 2 h old may really be 3 h old. Thus, the order of samples is no longer confidently known and the age uncertainty needs to be prominently acknowledged.

Another large source of age uncertainty that is hard to quantify likely exists in the ARCTAS data if samples were aged by running back trajectories back to active fire detections. Many fires are not detected as hotspots due to size or cloud cover (Yokelson et al., 2011, Burling et al 2011), with the latter being ubiquitous during ARCTAS. In addition, plume injection heights are largely unknown. So a sample may not be emitted at a hotspot that the back trajectory passes over, but instead at a different hotspot or at a fire with no hotspot.

In summary, to do aging best, one needs lagrangian or pseudo-lagrangian flight plans or some evidence of a steady source and an isolated, individual plume that can be verified not to have it's photochemistry “reset” by mixing with other fresh plumes downwind.

If the goal is to instead compare average observed NEMRS, a best effort is needed to compare NEMRs for reactive species in similarly-aged samples.

Because of variability in the NEMRs and uncertainty in age, it is unclear to me how well-constrained the measurement model comparison is. I’m not sure what the take-home message of the modeling portion is as explained in more detail later. It surprises me that two key photochemical intermediates, HCHO and OH, were measured on the DC-8, but not discussed in this paper. These species are critical parameters to get right in models (Alvarado and Prinn, 2009, JGR). Thus, e.g. the model OH should be compared, even in a sentence, to the measured OH.

The disconnect between the mission goals and what the flight plans could actually deliver seems to have led to some intellectual confusion in this and many ARCTAS papers. In this paper, the abstract may be an example of this. In outline form it seems to say that (1) variability is huge and (2) the variability has been underestimated, (3) our variability agrees well with previous studies, (4) the model agrees well with the hugely variable data. On first principles one would wonder if the latter means the line bisects the scatter or that the model can reproduce the scatter.

Some other key species that are not mentioned were not measured in ARCTAS, e.g. organic acids, but their role should be acknowledged anyway. Further many NMOC still cannot be measured with current technology (Christian et al., 2003; Karl et al., 2007; Warneke et al., 2011). The uncertainty due to both these factors should be acknowledged.

The authors do a good job of comparing to a lot of studies, but a great deal of relevant previous work in the boreal ecosystem is unmentioned. E.G. Goode et al., 2000 (JGR), the entire NASA ABLE mission that studied boreal fires, Nance et al., 1993 (JGR), etc. The Goode et al paper and references there-in, may help the authors track down papers that provide important context and relevant info! Also modeling papers that are relevant are not cited as I point out below.

It would be helpful if the authors could clarify even in a few sentences the rationale behind the TOGA species selection – could it be expanded to include more pyrogenic species?

Terminology issues: if EMR stands for “normalized excess mixing ratio” then what would you use to abbreviate excess mixing ratio? Since the NEMR abbreviation has been in the literature since Hobbs et al., 2003 it may be wise to stick with this admittedly clumsy term. I am trying to get people to use NMOC instead of VOC as a general term for organic gas-phase emissions from fires since about half of the organic gases are
not VOC, but in fact “SVOC.”

I am not 100% sure the intercomparison was done right as explained in more detail below.

LRT happens at all levels of atmosphere including BL too, just typically at different speed and direction if in BL instead of FT.

Some positive comments. In general, the improvements in instrumentation are mind-boggling since the last NASA boreal forest fire study (ABLE). A lot of good information can be extracted from the author's measurements. The authors deserve a great deal of credit for emphasizing variability rather than succumbing to the more common tendency to oversimplify things. With a little work to make the discussion of variability more quantitative and specific (see above) this aspect of the paper will be useful.

Specific:

1. The authors could provide a very useful list of EF for the TOGA species that were not included in Simpson et al. The formula is easy to implement in a spreadsheet. There is no need to redo the whole carbon mass balance as involving CO2, CO, CH4, and the TOGA NMOC will get close enough to total carbon to get a great approx to the EF. However, the authors should be careful to select only fresh plumes (with MCE < 1) for EF for the reactive species.

2. HCN and CH3CN ER to CO – this section is already good – although the authors might explore the impact of eliminating plumes with MCE >1 and discuss how their results compare to those of Simpson et al., (2011).

3. It's useful to see TOGA confirmation of the biomass burning source for infrequently observed emissions also reported in Simpson et al such as ethanol and the discussion could be slightly expanded to compare to that paper.

4. The intercomparison between TOGA, PTR-MS, and WAS may need to be tweaked (vide infra), but is of great value.

5. The authors do not extrapolate Boreal results to global scale! This shows excellent judgment since boreal fires and tropical fires differ in many important ways.

6. The comparison of observed NEMRs is useful, but it should be clear what is being compared and when needed, the authors should tighten up or clarify the comparison of NEMRs for reactive species to literature values to ensure that they are comparing fresh to fresh or “old to old” i.e. similar ages.

7. The authors work to compare the ratios of NMOC to each other instead of CO - with age is not a panacea for eliminating variability, but definitely has high potential for useful simplifications since similar compounds are often produced by similar combustion processes and “track” with each other (Yokelson et al 1996). As time allows, the authors may wish to explore this further.

8. The variability is not news and as community we strive to see the underlying patterns despite variability. Thus, perhaps it’s worth a few minutes to try binning or other standard mathematical approaches to discerning underlying trends in variable phenomena?

9. Is there any potential that the body of ARCTAS data, properly de-resolved, could be used to validate global model output or satellite retrievals (especially the many new products from ACE, IASI, TES, but also MODIS AOD, etc)?

Specific comments in order of appearance “P” is followed by the “last two” page numbers and “L” indicates the line on that page.

P29, L17-21: Usually the abstract highlights new findings. My comments on the abstract are previewed in my general comments above. Here I point out specifics. The long established inherently variable nature of fires is presented in the abstract, first as if it is new finding and then it is stated that this was already in the literature.

P129, L21-3: “However, this variability is often diluted in the literature when individual observations are averaged to generate an overall regional EMR from a particular study.”
1. Most studies report the stdev of the mean, which is a fairly standard way to report variability. Are they saying it is important to instead report the full range instead of letting people calculate the range from ± 2-3-sigma. 2. The averaging is done because modelers generally ask for a best estimate of the average. Are they saying models should instead use random number generators and generate a large range of EF to apply randomly to inferred fire locations — and should those random numbers then average to the central estimate and would it make it any difference in the model result? interesting q. 3. Some regions have very diverse ecology or land-use. Thus the trend has been to develop average EF for ecosystems or land-cover types rather than “regions.” In summary, what specific concrete point are the authors trying to make here?

P29, L24: “generally consistent within a given region” . . . what is meant by “generally consistent?” I think previous studies demonstrated that there are some differences in the average EF values between gross ecosystem types. I don’t think the authors of most previous works intended any more than that.

P29, L25: What is meant by “earlier assumptions”? Do they mean earlier measurements are in error or are they referring to some other “assumptions?” this needs to be specified.

P30, L3-5: The oxidation products of NMHCs and OVOCs are typically OVOCs so the sum of all OVOC/CO likely can’t decrease over 2.5 days although individual precursors often do. Unless what you are saying is essentially that most of the NMOC (and their oxidation products) are very short lived. Can the authors tell us what the model says happened to the oxidation products and what they were? Unless the modeling analysis included organic acids and the unidentified species as input and products the statement might need qualification by referring to “measured” or individual species.

L4-7: During 2008 there were a lot of fires in northern California in high elevation coniferous forests that are similar to boreal forests in many aspects. The authors should check on the likely fuels for their samples (including for Asia) – if the samples being compared had similar age, the similarities observed could either reinforce or downplay the idea of characteristic averages for ecosystems

L20, NMOC probably a better term as many of the emissions are SVOC?


P31, L3: The new fire emissions inventory draws on a review by Akagi et al., 2011a, which explains the NMOC differences in detail – it’s subtle. Not sure this sentence got it exactly right.

P31, L9: NOAA had a P3 in ARCPAC. Was the NASA P3 in ARCTAS then?

P32, L8-9: The authors point out that plumes are extremely variable and then state that by randomly intercepting fresh and aged plumes they studied plume aging in an “ideal manner.” I think my general comments above show why the ARCTAS data is not well-suited for studying plume evolution. It may also be helpful to see the discussion of plume evolution in section 3.5 of Akagi et al., 2011a.

L19-26: An unusually extensive section on a paper in preparation, which could be pared down to customary length.

P33, L19-26: TOGA undoubtedly made high quality measurements of the species indicated. It is interesting to me that the species list is a small fraction of the species that would likely have appeared in the raw chromatograms or that have been purportedly measured by GC-MS in other studies. Is it possible to give a sentence explaining how this specific suite of compounds was chosen? At this point it seems a bit eclectic or random and not targeted at the main emissions of biomass burning such as CO2, CO,
CH4, C2H4, C3H6, etc? (I realize that ARCTAS had multiple objectives and that the reported species are very interesting and important!)

P34, L3-7: The PTR-MS did not monitor a lot of significant BB species that it could have monitored as can be seen by referring to Karl et al. 2007 or Yokelson et al., 2008. I am not sure why and not criticizing the selection, but perhaps something could be inserted to say ~ “full mass scans were not implemented to increase time resolution and S:N for the following species selected because . . . .” The reduced suite of species may need to be acknowledged when drawing conclusions from the data.


P36, L3: Might want to mention here that Mason et al (2001 and 2006) used the NCAR Master Mechanism at least twice to investigate BB plumes previously. Daniel Jacob modeled boreal fire emissions in ABLE and Alvarado did in ARCTAS. Stu McKeen and Mike Trainer also modeled boreal forest fire plumes in JGR and Science, respectively. Yokelson et al., (1999) showed simple demo of including HCHO in a BB plume model.

P36-P38, section 3.1, Figs 1 and 2: As the authors themselves seem to point out, it’s potentially misleading to compare absolute amounts between instruments that sample at different times in a structured environment. The text seems to imply that the largest differences are due to timing issues, so does that obscure detection of the actual instrumental differences present? There is a simple, standard method to compare instruments in a heterogeneous environment that works well though. For continuous instruments, one compares the integrated excess amounts across the whole plume (as seen by each instrument) to compensate for response time differences as explained most recently by Karl et al., 2007 or Yokelson et al., 2009. For intermittent instruments there are two cases. If the plume transit time is much longer than the gap between samples and the concentrations are varying slowly one again computes the integrated excess amounts and compares those. If the plume width and sampling intervals are comparable one typically compares dX/dCO from each instrument where the CO is carefully matched temporally to the measurement interval for X. For instance the excess TOGA acetone divided by the excess TOGA-merge CO should be compared to the excess WAS acetone divided by the excess WAS-merge CO (or better yet, the WAS CO and TOGA CO if they are available). This can be tedious, but it yields a meaningful comparison. An idea: something like dAcetone/dMeOH (TOGA) versus dAcetone/dMeOH (WAS) vs dAcetone/dMeOH (PTR) etc might help and be quicker? It might be useful to try this for a few of the outliers and see if it affects the results?

P36: Is there a reference for exactly how the merges are done? Is the data carefully lined up with response times in mind or just “overlapped” Is it better to compare dX/dY for the PTR-MS, TOGA, or WAS in some of the ways I propose just above?

P36, L17-18: Are these comparing all data or just data in BB plumes?

P36, L26: If this is the final best result to partition furan and isoprene in boreal BB plumes the authors could compare to the isoprene plus furan and the split found to that in Karl et al. 2007 and Christian et al. 2003 (JGR).

P38: Again, because of these timing issues it’s not that meaningful to compare absolute amounts, as explained above.

P38, L22-24: So in the end all the instruments are assigned equal accuracy – sounds fair enough for now.

P39, L19: It’s a good idea to eliminate mixed plumes from an analysis of fire emissions and the authors likely did a good job of this, but I’m not sure NOx is an anthropogenic tracer? BB produces a large fraction of global NOx

P39, L23-24: A lot of chemical evolution can happen in 2 hours see Reid et al. 1998; Goode et al., 2000; Yokelson et al. 2003, 9; Hobbs et al. 2003; Alvarado et al. 2010; Akagi et al. 2011b. Also, note as a general aging issue that 2 hours flying down the long axis of a plume can represent almost 1000 km of plume length. At a windspeed of 10 km/hour this would represent the fire output over a time span of 100 hours or ~
4 days.

P39, L25: As explained above, there will be considerable uncertainty in the plume age due to missing hotspots or incorrect guesses at the injection altitude, except in the case of the McKay Lake fire where the source was actually located and sampled. Note, that fire exhibited high source variability (Alvarado et al., 2010).

P39, L26-27: the word “similar” is vague. Past studies tended to look out the window and see what was burning and classify the fires by vegetation type. In this study what are the new classification parameters and the category limits? Is there a connection between region and land cover type? Also, what is the logic for the grouping? 1) What are they trying to achieve by groups? 2) How does the grouping strategy accomplish the goal? Slightly more detail on the grouping strategy is given in the conclusions, but this is good place to clarify what is meant by “similar” or “chemical composition.”

P40, L3: I don’t think this study measures “fire-averages” I think the term here would be “group average” or “regional average”?

P40, L9: Table 1 has numerous MCE > 1, which is problematic. As explained above, a simple calculation shows that you can dilute a plume with background air and dX/dY is preserved as long as the background air is constant in X and Y. Once X/Y in the background air changes, dX/dY no longer reflects the original characteristics of the plume. For diluted aged plumes, the effect can be huge. Since MCE cannot be > 1 by definition at the source of a fire, and the authors define plumes as having positive dCO, an MCE > 1 means the dCO2 is a negative number smaller in absolute value than dCO. Since CO2 is the main emission of fires the negative dCO2 can only arise as an artifact due to changes in local background CO2 during transport and thus CO2/CO in the background likely changed during transport. That means X/CO likely changed as well where X is any random analyte. The original fire ratios have probably been corrupted in the NEMR from these plumes, but they are still valuable measurements of how BB affects the atmosphere. Thus, the results are worth presenting, but it might be best to eliminate them from any derivation of the attributes of fires.

P40, L17: “long-lived” better term than “useful”?

P40, L19: This was confusing because I thought the authors were using EMR to stand for “excess mixing ratio” but I guess its “normalized excess mixing ratio.” They are not the same thing.

P40, L21 and L24: maybe add an “N” to “EMR” throughout as this is confusing otherwise and excess mixing ratios don’t compare to ER.

P40, L26: This is one of several places where I wonder why the OH measurements were not used.

P40, L27: Again at this point in a first reading I get confused as it seems you are now changing the definition of EMR? Or add the word “normalized” and an N to acronym. One needs to read carefully to get this straight. I recommend reading Akagi et al., 2011a (ACP) as a quick course in fire terminology (and plume aging).

P41, L2: 0.2 days as in 4.8 hours is too long for a NEMR to be an ER for some species (see the OH on the ARCTAS archive and think about the rate constants for e.g. isoprene. See the decay plots in Hobbs et al., 2003 (JGR).

P41, L15-19: Initially one wonders what the selection criteria and purpose are for Figure 6 since it is presented without discussion, but then it does come up again later. At some point it might be instructive to compare e.g. the toluene data to the toluene lifetime at the measured OH (~2 days?) to illustrate how the variability masks the photochemistry? It could be made slightly clearer how the colors line up with plume groups 1-40.

P45, L2: “there is”

P45, L27: Is production the right word here? It seems odd that the values which are lowest by far would be associated with production.
there's the expected factor of $\sim 10$, which is also not too badly approximated by the authors mean $\pm 2$ quoted standard deviations. E.G. four standard deviations is "48" whereas the observed range is "49."

Yokelson et al., (2009) report CH$_3$OH/CO (Table 2) ranging from 12.5 - 48.1 mmol/mol, where those are fire-average values. The study-average (not including one point from C-130) is 27.3 +/- 10.2 (n=17). The high variability is shown to be correlated with MCE in Figure 3 in agreement with general comments above.

Discussion of ethanol is interesting and useful and can be expanded to compare to Simpson et al.

Different OH may be the main factor, but may want to recognize a possible contribution from potentially different spring vs summer fuels and/or fire behavior in Asia.

Yokelson et al., (2011) (ACPD) on isoprene. I would eliminate "our group" and just cite Apel et al 2011, but add it to the references as well.

Here is a good opportunity for a comparative discussion regarding the findings of Simpson et al who concluded there was a characteristic ethane/CO ratio for boreal fires? It would be useful to clarify the different assumptions that lead to different conclusions?

A fresh BB plume, or a combustion plume of any kind with MCE>1 is not physically realistic. Plume 28 is also an outlier for MEK. Plumes such as these should probably be weeded from the analysis. This could help make the paper shorter and less complex. Unless this outlier is telling us something important? Sometimes the outlier is the best point of all!

Another place where bringing in the OH could be useful.

Good place to acknowledge SVOC by switching to NMOC, especially since the SVOC are more likely to become particulate organic matter.
P53, L1-2: If Figure 6 is retained (which I think it should be) here would a great place to indicate the expected decay of these species based on the measured OH for comparison to the scattered observations.

P53, L11: My first guess would be that in real plumes the variability in NEMR would increase with aging due to different OH, light intensity, background mixing etc. Of course, the reactive species can all go very quickly to very small values that look similar on a plot with a large y-scale?

P53, L13-15: It seems surprising that the outcome is largely independent of the input and it may be a coincidence in that increasing NOx increases O3 to a point and then further NOx increases “poison” the O3 chemistry as discussed in Mason et al., 2001. Are the author’s model results consistent with the message of the well known, traditional NOx/NMOC O3 isopleths?

P53, L13-20: It would be helpful to summarize what the message is here. It’s probably not that it doesn’t matter what emissions are used in atmospheric models, but it’s hard to be sure as written.

Section 3.5: This is very promising. The ratios of NMOC to each other as a function of plume age may take out some of the variability from the analysis and could be explored further as a way to extract important new findings or enhanced comparisons to the model. E.G. it is known that the alkanes are mostly made by glowing, many of the OVOC are mostly made by pyrolysis, etc (Yokelson et al 1996, 97). Thus the NEMRs between NMOC that are similar to each other in structure should be less scattered in many cases than the NEMRs of NMOC to CO, but nonetheless, the NMOC within a “similar origin group” can have different OH rate constants and so the evolution of their ratios may be a useful window into the aging!

P55, L14-16: This is an important result of the paper, the measurements of rarely measured species.

P55, L19-20: Actually literature ranges in NEMR are well over an order of magnitude. For instance, Yokelson et al., (1999) cite a range of 600 for acetic acid EFs. I agree with the authors that there surely are some papers in the literature that underestimate variability, but they should be identified specifically rather than referring to the “literature” as a whole.

P55, L21-25: I think I disagree with the message here, but the overall message of the conclusions is vague. Here the authors seem to claim that fires are more variable than recognized in the literature (inaccurately I think), but in lines 13-14, the “good agreement” with the literature is emphasized. The authors should decide if they made quantitative progress in determining the seasonal, spatial, temporal, chemical, etc, resolution needed to adequately represent BB emissions, and if so give the results in specific, quantitative language. E.g. see Korontzi et al., (2003) or van Leeuwen et al., (2011).

P55, L23: A paper with this “perception” should perhaps be quoted. Also a region and a fuel type are not similar concepts.

P55, L25-6: “confirmed” more accurate than “shown.” Regarding the next sentence: using the mean can’t over or under estimate the mean, unless the mean is wrong. But of course, in any particular plume the “mean” may not be observed. So this is like the weather vs climate.

P56, L3-8: “The Californian BB plume was significantly more polluted than the Canadian BB plume, with $10\times$ greater NOx, twice the O3 concentrations, and $4\times$ greater VOC and CO concentrations.” Are the concentrations referred to here the background, excess, or absolute in plume? Unless the differences are in the background it’s hard to believe they make no difference in the outcome. The implications of that would be enormous.

Fig 7: maybe expand the caption or refer to discussion in text?
Fig 8: Change “VOC” to “Measured NMOC” to acknowledge that not everything was (or can be) measured. (Akagi et al. 2011a; Warneke et al., 2011; Christian et al., 2003; Karl et al., 2007).

Selected additional references:


Yokelson, R. J., Burling, I. R., Urbanski, S. P., Atlas, E. L., Adachi, K., Buseck, P. R., C5978


Trentmann, J., R.J. Yokelson, P.V. Hobbs, T. Winterrath, T.J. Christian, M.O. Andreae, and S.A. Mason, An analysis of the chemical processes in the smoke plume from a


Interactive comment on Atmos. Chem. Phys. Discuss., 11, 14127, 2011.