Dynamic savanna fire EFs by Vernooij et al

Review by Bob Yokelson

As described in a separate group of papers, this team has recently completed a very seasonally and geographically extensive series of measurements of savanna-fire, fire-average MCE and emission factors (EFs) for CO_2, CO, CH_4, and N_2O; along with appropriate related factors like weather, fuel moisture, etc. In this paper, the authors focus mainly on:

1) How well their field-measured environmental factors can be used to calculate their field-measured EFs.

2) How widely-available remote sensing products (aka "features") correlate to their field-measured environmental factors.

3) How well the remote sensing products can be used to predict the field-measured EFs.

4) Using the remote-sensing proxies to calculate a new global emissions database that is sensitive to local spatial and temporal variability in the environment.

Not surprisingly, given the coefficient of variation (COV) (aka relative standard deviation (RSD)) of the EFs in both the literature and the author's work, local variability in emissions was found to be high and the authors work provides a way forward for those wanting to account for variability at a finer scale than is currently common. Also interesting, the +/- variability tends to cancel in their total global emissions, which is encouraging for researchers focused on larger-scale averages. This is important, creative, work that should be published. I have some minor suggestions that I think could strengthen the paper as summarized next.

1/ A few more sentences describing the sampled fires and data reduction would be helpful. I glanced at the previous publications and did not quickly find all the common or potentially useful details. For instance:

a/ Were the fires all prescribed?

b/ How big were they?

c/ Were they detected from space as hot-spots or burned areas?

d/ Were they all lit the same way? (In Brazil we noted that fires were often lit on opposing sides and the flame-fronts burned together. Fires were sometimes lit at night after wind died down.)

e/ What, in a nutshell, was the sampling strategy?

f/ Were RSC samples collected when relevant?

g/ How were the data processed into emission ratios (ERs) and EFs? To clarify last question, Yokelson et al., 1999 compared the impact of processing grab samples into ERs and EFs with several different justifiable approaches. Without proving one approach was best, they found only small differences among approaches. Similarly, regarding the authors work, I don't plan to critique their approach, but it's useful for posterity to specify the approach used (see below on RSC for more).

2/ The paper would be easier to comprehend the first time thru with slightly more plain language and consistent terminology in describing the statistical analysis.

3/ The discussion on possible future applications is nice. Perhaps one other addition would be to identify which environmental variables might be available in timely enough fashion and have enough predictive power to improve air quality forecasts. I.e. could current or forecast temperatures from the global weather services help predict how fires will burn in near real time?

I also have a few other more focused general comments.

Overview of value:

Somewhat related to #3 above, can the computational burden be specified of using the author's full-scale approach or partial implementation? How much easier and how relatively accurate is simply using EDS and LDS EFs?

What is the error in the satellite proxies and how does propagated error in the dynamic EF compare to the impact of switching to dynamic EF?

During a recent field campaign, we found that one of the global vegetation products mapped a pine forest and an alpine wilderness area to savanna and agriculture respectively. Simple added info would be useful such as: do all the author's savanna fires show up as being in a savanna in the remote-sensing products? There is also considerable difficulty/uncertainty in field-measured fuel consumption, etc. Easier than adding many columns for uncertainties would be at least generic uncertainties in the table explaining the data set. The error bars in the figures do look generous to the author's credit. Again, it might be worth stating how the local variability compares to full, propagated uncertainty?

CH_4 is exceptionally dependent on MCE, but not all important emissions are as seen in Yokelson et al. (2003) and other work including Andreae 2019.

Yokelson, R.J., I.T. Bertschi, T.J. Christian, P.V. Hobbs, D.E. Ward, and W.M. Hao, Trace gas measurements in nascent, aged, and cloud processed smoke from African savanna fires by airborne Fourier transform infrared spectroscopy (AFTIR), J. Geophys. Res., 108, 8478, doi:10.1029/2002JD002322, 2003.

Questions on data:

It seems the Excel spreadsheet is giving time as local time?

The spreadsheet seems not to include background samples. ERs and background values can be derived from slopes and intercepts, respectively. By subtraction of the "_em" column from the "_abs" column, it appears there was a fixed background for each fire. These backgrounds are interesting in themselves. For instance, one fire had a background of 0.17 ppm CO, which is pretty low compared to the 1-5 ppm CO background that can occur during regional smoke episodes during peak fire season. As we also see by FTIR (but don't report), there were negative

N_2O emissions and EF at times. How were these negative emissions handled in further data processing?

There are a number of non-physical values in the spreadsheet easily found by plotting the columns in a line chart. E.g., rows 2209-2211, 2353, and especially 2382 and 3116. These data were presumably not used in the training or validation and might be removed?

The letter and number convention for the sample names, does it have any significance that should be explained?

Why are EF calculated for the cals?

Why is no date/time given for the cals?

Why are the cals not all the same or nearly the same? Were there different calibration mixtures or does the scatter reflect the precision?

I was surprised that field-measured temperature had poor correlation with the satellite temperature in Table 4. Then I noticed in the spreadsheet that the temperatures measured on the drone correlate with CO_2. In general, the temperature, RH, and VPD seem to be measured in the convection column at times where they would reflect the heat and water production of the fire, rather than an ambient air value that would influence fire behavior. If this is the case, I suggest replacing sample-specific values from the drone with one best ambient value per fire and (if not already done) seeing how that correlates with measured EF and remote-sensing products. Or did the authors use pre-fire met data measured differently or on the drone during the pre-fire cal and that data is available somewhere else?

For example. Picking one fire randomly, EDS19_3 on a June afternoon in Mozambique, one notices that Tsat is close to the climatological average high for June in Maputo (26 C), but is well below the lowest Tdrone (33.57 C). Is that a shade versus sun-exposed thing? Was there a Tdrone during a cal or background that is more appropriate? Further, VPDsat is only close to VPDdrone at minimum Tdrone suggesting combustion products make VPDdrone not representative of ambient VPD unless a VPDdrone measured in background air was actually used? Likewise the RH comparison reveals differences.

In the LGR N_2O-CO instrument, the N_2O data needs to be corrected for CO and the correction only works up until 5ppm CO. This is because at high CO values, the CO line broadens enough to interfere with the N_2O line. In general, the strongest N_2O band is overlapped by water, CO_2, and CO (and other gases). The CO values in the author's spreadsheet are in the 100s. The manufacturer of the author's N_2O instrument (AERIS) product literature claims to use an interference-free, but unspecified, alternative spectral region and have an upper limit of 500 ppm for some unspecified molecule (probably CO?). Kudos to the authors for not using LGR for N_2O, but I am curious if the authors have any evidence against or for CO interference in their N_2O data? I am not assuming issues exist, but if they can be ruled out, it would be worth mentioning as N_2O is an important, but undersampled fire emission.

Variability of emissions during fire:

Emissions can vary significantly fire to fire, but also during a fire and there are both lofted (at a range of velocities) and unlofted emissions. Atmospheric chemists tend to focus on chemical completeness and pack a plane full of many instruments and target vigorously lofted emissions at the peak of fire season. This approach is not sensitive to the un-lofted emissions from residual smoldering (aka “post frontal”) combustion (RSC, Bertschi et al., 2003). In the author’s approach, they were ground-based, but had a drone that could presumably sample both unlofted and lofted emissions, though not the vertical velocity of the emissions. They also measured fuel consumption of the fine fuels (whose emissions are mostly lofted) and the heavy fuels (more prone to RSC). This is a very powerful sampling strategy (for a limited selection of gases) to be commended. The challenges for representative sampling of fires are outlined elsewhere (Bertschi et al., 2003; Akagi et al., 2013) and especially with the timeline given for access by different sampling platforms to the same fire in Akagi et al. (2014). Thus, it would be helpful if this paper gave a brief narrative of when and where samples were collected with respect to the course of the fire and how the data were processed to get fire-averaged EFs. I did not find this info in the previous papers so just a summary here would be great.

Bertschi, I.T., R.J. Yokelson, J. G. Goode, D.E. Ward, R.E. Babbitt, R. A. Susott, and W.M. Hao, Trace gas and particle emissions from fires in large diameter and belowground biomass fuels, J. Geophys. Res., 108, 8472, doi:10.1029/2002JD002100, 2003.

Akagi, S. K., Yokelson, R. J., Burling, I. R., Meinardi, S., Simpson, I., 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.

Akagi, S. K., Burling, I. R., Mendoza, A., Johnson, T. J., Cameron, M., Griffith, D. W. T., Paton-Walsh, C., Weise, D. R., Reardon, J., and Yokelson, R. J.: Field measurements of trace gases emitted by prescribed fires in southeastern US pine forests using an open-path FTIR system, Atmos. Chem. Phys., 14, 199-215, doi:10.5194/acp-14-199-2014, 2014.

Line by line

I expand on, or overlap with, these overview comments in my detailed line-by-line remarks. The format is: page number/line number(s)

1/21 change to: “… the breakup of the constituents of the fuel …”? Elements cannot be broken down via chemistry.

1/22 “ … other things …”

1/27 “collected” > “made”

1/28 delete “EF”, add “EFs” after “N2O”. Also “also” before “measured.”

1/31 awesome data set! “85 savanna fires”, delete “known” change “listed” to “provided”

2/1 Not 100% sure what is meant here. It almost reads like the biome average EF is 60-85% off on average. I think you mean e.g. if a measured fire had an EF 10% below the biome average EF, the satellite-based recalculation of the EF would be ~6-8.5% below the biome average?

2/2 “total global savanna fire emissions estimates”

2/3 change in CO_2 totals? (expect small)

2/2-3 It’s amazing that the global totals based on average biome EFs were within 1.8 to 18% of global totals using dynamic EFs. The difference is much smaller than the uncertainty in almost any other thing. However, it should be clear what biome average EFs are employed here. Probably the old literature average? Also, is good agreement seen every year or just for the 14 year total?

Ultimately, the paper could compare the old literature average EFs to the evolved literature average EFs that include the author’s new data, and the average EFs based on just the authors new work. I.e. how much impact does this study have on averages? Finally, in addition to predicting measured EF better, it would be interesting to know if the use of dynamic EFs also better predicts downwind impacts, but that might be another paper.

2/5-6 Did not the authors observe that CO and CH_4 EFs decreased with drying in xeric grasslands, but increased with drying in mesic woody savannas? Also “… annual average savanna fire …”

2/7 Are there just reductions? There is good agreement on totals so there should also be localized increases. In general, from the 1-sigma standard deviation in literature EFs we expect +/- 40% variation in EFs fire-to fire 1-sigma.

2/12 Throughout paper it should be “emissions inventories”

2/15 60% of net emissions? Could deforestation and peat be more important in the C-cycle if minimal regrowth?

2/26-28 There are many direct field measurements and they quantify overall variability, but previously we could not account for the total variability with quantitative contributions from very many specific factors. Previous studies targeted the average and variability, but not the causes of variability.

2/30 change “closed” to “open”? At least in Akagi et al there was a split between dry forest and woody savanna at 60% canopy coverage.

3/4-5 How about “, a series of savanna burning experiments measuring EFs using unmanned aerial systems (UAS) has resulted in a large amount of new data …”

3/6-7 How about “… the variability in over 4500 individual bag-measured EFs of CO_2, CO, CH_4 and N_2O covering 129 fires.”

3/25-26 Could other real-time data besides that from satellites be useful?

3/31 “savanna fires” and no comma after “fires”. It wasn’t immediately obvious that the cited references contained much detail about the fires, sampling strategy, and data reduction. Maybe I missed it in a supplement?

3/32 “… at altitudes between …”

3/36 Maybe I missed it somewhere, but worth repeating here as no page limit. Include averaging/plotting schemes, RSC sampling? Fire descriptions? How met data collected, etc. As detailed in overview

4/10 “sometimes prevented”?

4/15 The fire dates and coordinates are always “known” but not always “provided” in the paper, though perhaps available from authors. This illustrates value of describing fires in some detail.

4/21 collected > measured

4/23 The pre-fire met data mentioned here, where is it? The spreadsheet has non-useful met data collected in the fire convection column.

4/30 Okay this provides some of details requested above.

4/35 One naturally wonders here if the authors field environmental data can be used for insight into the accuracy of the global satellite products and were their fires detected by the satellite products GFED4s uses?

4/36 built

5/7 Impressive set of products. Is it easy to explain why no VIIRS or geostationary? Not available as long? Useful going forward?

5/8 Were all the samples of a fire usually in the same feature pixel?

5,/15 Is it easy to explain why not using historic NDVI range?

5, 25-26, TRMM useful for rainfall?

5/30-31 risk or behavior or both? Are any ideas in the “hot dry windy index” useful as predictors here?

5/34-35 Is the daily cycle of fine fuel moisture captured? Was FFMC compared to the author’s field-measured fine-fuel moisture data?

6/6 How does spatial resolution of the fire severity proxies (dNDVI etc.) compare to the size of fires? If the fire is smaller, then is the signal diluted? Would a small severe fire look like a larger less severe fire? Did the authors expect better correlation of scorch and char height with the severity proxies?

6/17 modes or models?

6/18 all in-situ or 70% as on line 20?

6/19-20 What is “a measurement with a missing value of an included feature”? Do you mean you did not use EF measurements if even a single associated satellite product out of the whole set was missing?

6/21-22 What does “resampled using ten-fold cross validation while allowing sample replacement (i.e., bootstrap method)” mean? Can a simple plain language explanation be added?

6/22-23 Explain that “hyper parameter” refers to the most influential parameters?

6/28-30 This is hard to follow. How would an EF require a resolution and how would that be computed? Overlap is within or between features? Do you mean some fires were bigger than or occupied more than one grid cell in the original feature (note we have slipped into calling remote-sensing proxies “features” for short), so you averaged, or extrapolated, or built a new grid for each fire such that the fire was centered in a single grid cell? Sometimes a few extra words can help a lot!

6/32 How can an EF have a temporal resolution? Are the EFs referred to fire-average or sample-specific? Is the daily cycle in RH and fine fuel moisture considered?

6/40 “… savanna fire emissions …” Were the dynamic EFs calculated using global products and RF? Change “dry matter emissions” to “dry matter consumption” also at 7/8.

7/15-16 How were samples with negative N_2O emissions treated when calculating fire-average N_2O emissions?

7/19-20 Clarify this is the Andreae 2019 average and not the average of the 85 measurements used from other groups? Otherwise, how do you get locations for study-average or vegetation average emissions (unless one fire in study)?

7/23 substantial variability in fire-averages or samples?

7/24-25 The higher CO and CH_4 EFs in woody savanna is supported in previous literature at least once, e.g. Sinha et al., (2004).

Sinha, P., P.V. Hobbs, R.J. Yokelson, D.R. Blake, S. Gao, and T.W. Kirchstetter, Emissions from miombo woodland and dambo grassland savanna fires, J. Geophys. Res., 109, D11305, doi:10.1029/2004JD004521, 2004.

FWIW, the Miombo fire was included in the tropical dry forest category in Akagi et al, but it was also a small part of a savanna fire study-average used in the savanna category.

7/25 Taking this to mean the authors study-averages were lower than previous literature averages.

7/29 by “seasonally inundated grasslands” do you mean aka dambos?

8/2-3 Any benefit to comparing the authors fuel measurements to similar measurements by Shea et al (1996) and Hoffa et al (1999) and others?

Shea, R. W., Shea, B. W., Kauffman, J. B., Ward, D. E., Haskins, C. I., and Scholes, M. C.: Fuel biomass and combustion factors associated with fires in savanna ecosystems of South Africa and Zambia, J. Geophys Res., 101(D19), 23551–23568, 1996.

8/4 What is meant by “corresponding mixtures of fuel age”? In Table 3, why was a higher percent of the heavy fuels consumed in the EDS in Australia, unlike elsewhere; maybe lit more aggressively?

8/10-13 I’m pretty sure that increased RSC and increased CO and CH4 EFs in the LDS in wooded savannas is already in the literature, but haven’t found the reference. Maybe Hoffa or Korontzi?

8/16-28 good summary.

8/32 For Table 4, clarify which field-measured met data were compared to satellite met data, preferably NOT drone data in fire-processed air! However, Table 4 seems to specify that T and RH from the drone were used, which could be okay if NOT when drone was above the fire, but instead in ambient (background) air. Then again, currently, it’s odd that the satellite temperature and drone temperature are weakly positively correlated at 0.18 while satellite temperature is most strongly correlated with field measured nitrogen content in the grass (perhaps a seasonal coincidence?).

8/31-33 This text and Table 4 could be clarified with slightly more precise and consistent terminology. I think Tab 4 shows how the *field measured-ecosystem attributes* correlate with the field-measured MCE and EFs and also how the *field-measured ecosystem attributes* correlate with the satellite products, but NOT how satellite products correlate with field-measured EF or how anything correlates with model-calculated EF? At this point in the paper, evidently, calculated EFs vs measured EFs and the sensitivity of calculated EFs will be discussed elsewhere.

8/40 strongest > most strongly

9/4-5 I’m taking this to mean that 70% of field-measured EF were used with “features” to train the RF model and the RF model then used features to predict EF for the other 30% of field measured EF (out of sample means fires not in training set) and performed well in terms of r-squared. Could give the slope too? Is the training set randomly selected or varied run to run?

9/5-7 Is there a simple way to connect feature importance and the concept of hyper parameters? Is “impurity decrease” essentially a fraction of total variability?

9/8 The red line in Fig 4 is useful for comparing the range of EF to the old literature average. But later in paper, the effect of dynamic EF should perhaps be compared to the biome average based just on the field data used by the authors, which could be shown with a second vertical line. Then recalculate MAE and improvement %. Currently, the comparison is “apples to oranges” in that “improvement” is based on a difference resulting partly from incorporating new data and partly from a change in approach.

9/9 replace “data” with “features”

9/10 change “recalculate the MCE” to “predict the field-measured MCEs of the fires in the validation set”

9/11 “static” > “old literature”

9/13-20 This is a nice exploration of simplifying the RF approach. Can the authors explain why VPD is the most important feature in the small subset of features, despite having a low rank in the full set of features? Any estimate of reduced computational burden?

9/24-25. Does this mean you ran the RF model once to get MCE and then used the MCE as a new feature in a re-run of the RF model?

9/28 130 or 129? Including fires with negative N2O emissions?

9/35-38 It would be interesting to see the study-average EFs vs the former literature average EFs and then also what the new literature averages are including this study, all in a little 3x4 table.

10/2 How common are mixed biome grid cells? Percentage of total? Is the most common type of mix with tropical dry forest? Is there a percent tree cover or canopy closure that defines the boundary between what the authors consider savanna and something else?

10/11-12 What is “annual effective EF”? An annual global savanna-fire average EF for each compound? This is also saying the year to year variation in global average EFs is small?

10/14-15 averaged over what time and space? I.e. the daily average over all areas occupied by the indicated vegetation class? Fig 7 doesn’t seem to show much or any EFCO increase in woody savanna as the fire season progresses? Does this figure clash with previous text? What is “typical savanna”?

10/30-34 Interesting, shows the RF model may have value to at least partially correct sampling bias in a field campaign!

11/2 Just to be clear, the N is in the foliage of the trees, not the wood itself

11/15 “CO-”?

11/24-26 Did Hoffa and Korontzi predict higher MCE in LDS?

11/30 The Eck trend in SSA is averaging over all sub-Saharan Africa AERONET sites?

11/32 References that support an increase in SSA as MCE decreases include Liu et al and Pokhrel et al and probably many others

Liu, S., Aiken, A. C., Arata, C., Dubey, M. K., Stockwell, C. E., Yokelson, R. J., Stone, E. A., Jayarathne, T., Robinson, A. L., DeMott, P. J., and Kreidenweis, S. M.: Aerosol single scattering albedo dependence on biomass combustion efficiency: Laboratory and field studies, Geophys. Res. Lett., 41, 742–748, doi:10.1002/2013GL058392, 2014.

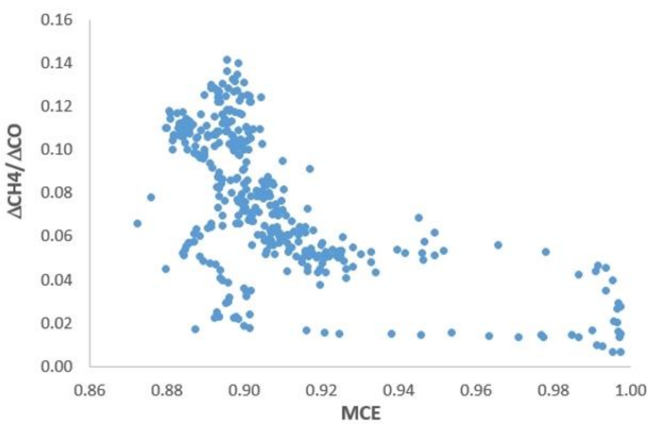
Pokhrel, R. P., Wagner, N. L., Langridge, J. M., Lack, D. A., Jayarathne, T., Stone, E. A., Stockwell, C. E., Yokelson, R. J., and Murphy, S. M.: Parameterization of single-scattering

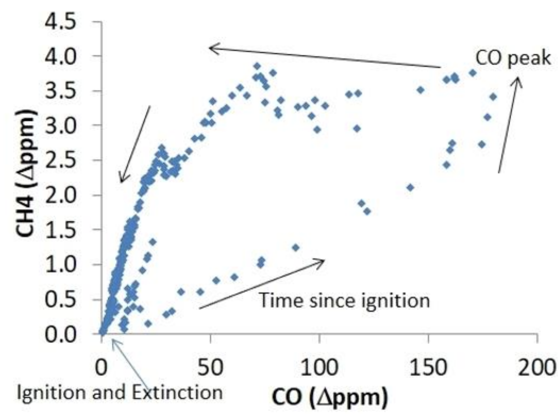
albedo (SSA) and absorption Ångström exponent (AAE) with EC / OC for aerosol emissions from biomass burning, Atmos. Chem. Phys., 16, 9549-9561, doi:10.5194/acp-16-9549-2016, 2016.11/37 linear or non-linear?

11/40-12/1 Not sure about the interpretation here. Does CH4/CO vary with MCE? CO is not technically independent of MCE since MCE has CO in its definition.

12/2-3 interesting

12/5-9 This discussion could be misleading in a subtle way. I think the effect seen here is probably because the other studies compared to are plotting the fire-average EFCH4 versus the fire-average MCE, while the authors are plotting EFCH4 vs MCE for "snapshot grab samples" that could include samples during flaming that may have much higher MCE than the fire-average MCE for typical useful real-world fires. We've seen this often over the years. To illustrate we can revisit the comparison to the Selimovic et al lab fire study. If one plots the instantaneous EFCH4 vs instantaneous MCE for these typical lab fires you often get "curvature" at high MCE values during "pure flaming" and other effects. The ERCH4 vs MCE can also be non-linear at high MCE or have interesting other interesting patterns with time. The plots show this for the 1-s data from randomly selected Fire #74 on the NOAA FIREX-Firelab archive (https://esrl.noaa.gov/csd/groups/csd7/measurements/2016firex/FireLab/DataDownload/). Fire #74 is one of the fires in the linear plot of fire-integrated EFCH4 vs MCE in Selimovic et al. (2018). Interesting topic but variability during a fire is a level of detail large-scale models can't cope with yet. Thus, in providing guidance for large-scale models it may be best to stick to fire-average data.





12/11-22 Since Brazil has a lower elevation than Zambia it may have higher temperatures and more evaporation than Zambia making it more xeric at the same total rainfall? In addition, African sites may have keystone grazing species that encourage forest cover by reducing grass fuels and thus fire intensity and tree mortality?

12/25-26 Think you mean "This is the first study to quantify the spatial distribution of GHG EFs over the entire savanna biome by using both field measurements from a variety of savanna ecosystems and their relation to global data mainly from satellites". I.e. the field measurements have gaps as explained in the following lines, but by connecting the measurements to features you have a new way to get a useful global savanna estimate!

13/11 The idea of a gross underestimate here is worrisome. How well do the authors think GFED4s accounts for fires too small to show up in their burned area product? Worth mentioning here?

14/7 change "propose" to "developed"?

14/9 "modelled"> "model-produced". Also "significant fire-specific improvements"

14/10 Here I think it's important to preserve the idea that you have not concluded the biome averages have large errors, just that fire to fire variability is large and is better accounted for by using a more sophisticated model. Also + and – local errors tend to cancel. It worries me that someone reading quickly may think you mean that global CO and CO2 emissions from savanna fires are off by ~80%.

14/21 Remind reader that the N2O decrease is a combined effect of new data and the dynamic approach.

14/31 I did not check the zenodo link. If it is different from spreadsheet, I could check it by request.

Fig 1. Andreae 2019 is not indicated in yellow as caption suggests?

Fig 7. Why do "typical savanna" fire emissions peak earlier than all the subtypes?