Comment on amt-2021-223
Anonymous Referee #2

The study uses TROPOMI observations to derive biomass burning NOx emissions. The authors apply two methods: flux and EMG, and suggest EMG is a better approach. The authors further evaluate satellite derived NOx emissions with aircraft measurements. Given the uncertainties of satellite retrievals, deriving NOx emissions from fire plumes is not an easy job. The authors have done a lot of work putting together satellite, models and aircraft measurements. I was expecting this study will represent a significant contribution to literature from the abstract, but I was a little disappointed after reading the whole manuscript.

Overall, I feel the authors made a lot of assumptions that are not justified or evaluated carefully. While the authors show their approach can somehow agree with measurements, I’m not really convinced whether these methods can be applied widely to other fires. Especially since this manuscript is under review at AMT, developing a solid, justified and widely applicable method is the key. I strongly recommend the authors carefully evaluate each of their assumptions. Since the authors put together a large dataset from aircraft campaigns. I think this paper will be useful if they can justify their assumptions with the aircraft measurements. Below are my detailed comments.

1. The novelty of this study is the explicit aerosol correction for AMF calculation, but I’m not convinced with the methods. There are many issues:

1) It’s not clear to me why the authors only account for aerosol scattering effect. What about the aerosol absorption effect? The emissions of black carbon from fires should not be small. Aerosol scattering and absorption will influence AMF differently (Lin et al., 2015).

2) It’s not clear why the authors decide to use a constant profile shape for NO2 and aerosol. The authors simply assume constant NO2 and aerosol between surface and 2.5 km. Is this really a good assumption? Shouldn’t the vertical profiles of NO2 and aerosol vary with meteorology, topography, fire injection height etc.? For aerosols, depending on the height of aerosol (above or below cloud), its impact on AMF should also be different.

3) Equation 3 is also confusing. First, why use 3.5km as the cut off? Shouldn’t this vary with boundary layer height? Second, VCD_KNMI uses NO2 profile from TMS simulations,
but VCD_freetrop uses NO2 profile from GEOS-Chem profile. How could the difference between VCD_KNMI and VCD_freetrop be used to scale the NO2 profile? Without further justifications, I feel such approach is arbitrary.

(4) Does the monthly GEOS-Chem simulations include fire emissions? If not, what if the fire emissions are released into the free troposphere?

(5) The authors assume clear sky inside the plume, but outside the plume, they account for cloud conditions. I’m not convinced how good the assumption is here. I think this may lead to some inconsistency in the AMF or derived NO2 VCD within and outside the plume if the cloud and aerosols affect AMF differently. On the other hand, if clouds and aerosols affect AMF in the same way within the plume, why do the authors use explicit aerosol correction rather than implicit correction?

(6) The authors use NO2 VCD as a proxy for AOD, which is also confusing to me. They simply assume a constant relationship between NO2 and AOD, and the relationship is not at all evaluated in literature (Bousserez 2009 is not a peer-reviewed paper).

2. I’m not convinced that the assumption of constant lifetime and plume spread is valid. A recent study shows large variation of NOx lifetime in fire plumes (Jin et al., 2021). The spread of the fire plume should also vary with wind speed and fire intensity. Figure 3c clearly shows how the emissions would change with different assumptions of lifetime. While assuming constant lifetime is fine for the flux method, isn’t the main idea of EMG method is to derive emissions and lifetime simultaneously while accounting for variation in plume spread (Lu et al., 2015)? If lifetime and spread is considered constant, EMG is essentially a smoothed exponential decay function, which is mathematically similar to the flux method. What’s the motivation of using two methods then?

3. The evaluation with aircraft measurements is new, but the comparison is overall limited to the statistics. For example, it’s interesting to see statistically EMG performs better than flux methods, but why? Since the authors made the same assumptions with lifetime, I’m curious what factors could lead to such differences. Also, I guess the difference between TROPOMI and aircraft emissions is related to the short-term variability of fire emissions, which however is not discussed. These aircraft measurements may also help assess the assumptions made in AMF calculation, and I don’t see any discussions on this.

4. A lot of details are missing in terms of how the authors perform EMG. The authors simply listed a number of equations, but I’m not sure how to interpret these equations. What does each equation and parameter mean? How is implemented for each fire? I notice there is large data gap in Figure 4. How would this influence the EMG method?

Specific comments:

Page 3 Line 24: You already mentioned TROPOMI in previous paragraph.

Page 12 Line 30: Here you mentioned using TROPOMI aerosol layer height for wind speed, but why not use this information in aerosol correction for AMF?

Page 13 Figure 2: Here VCD_EC seems to be much smaller than VCD_KNMI, but Figure 8 shows the opposite. I understand that aerosol may influence AMF differently. To avoid confusion, I’d suggest the authors either limit to one fire case (the case with aircraft), or explain under which conditions aerosol corrections lead to higher VCD and vice versa.

Page 14 Line 20: Why did you choose 20 km for box size? It seems that the fire plume goes much further than 20 km in Figure 3?
Page 17 Line 20: Did you look same fires for TROPOMI and GEM-MACH? Why NO2 lifetime is shorter in model than observations? Maybe it’s due to different resolutions? Also, what’s the chemical lifetime of NO2 in GEM-MACH?

Page 18 Line 33: The errors of satellite retrievals are not necessarily random. Studies have reported low bias of TROPOMI NO2.

Page 23 Line 2: Do you assume constant lifetime and spread here? If so, why not try relaxing these assumptions for synthetic observations, and see whether original EMG method still works?

Table 1: I think there are other sources of uncertainties not discussed here. Just to name a few: 1) uncertainties of your AMF method (especially with prior); 2) uncertainties of the aerosol information; 3) biases in satellite retrieval of NO2 columns; 4) uncertainties in the plume injection height.

Figure 8c: It looks like there is large gradient of NO2 at low altitude, which differs from the interpolated profile. This again made me doubt about the validity of your assumption with the NO2 profile. Also, it’s better to present vertical profiles in pressure gradient, which could better show the vertical gradient of NO2 at lower altitude.

Page 26 Line 11: While NO2 columns downwind are less important for overall enhancement, this would impact on the lifetime of NO2.

Page 27 Line 3: Did you account for diffusion when calculating NOx emissions from aircraft measurements?

Page 29 Line 7: Please provide justification for the threshold.

Figure 10: The plot looks messy. Why not just show the mean and standard deviation of NO2 column from CU-DOAS?

Figure 10: I suggest include TROPOMI VCD_KNMI here.

Page 29 Line 18: Not clear what you mean here. Emissions = columns x wind speed? It doesn’t sound right to me.

Page 32 Line 13: It’s not clear where the scale factor of 1.3 to 1.5 comes from. Table 2 shows large difference between TROPOMI EMG and aircraft derived emissions, and the difference also varies a lot from fire to fire. I don’t think it’s correct to suggest a universal scale factor.

Page 32 Line 15: I’m confused here. Didn’t you assume constant lifetime for EMG approach (Page 15 Line 5)?

References:

